

文章编号:1001-9081(2011)12-3297-05

doi:10.3724/SP.J.1087.2011.03297

# 求解双边加权模糊支持向量机的序贯最小优化算法

李艳,杨晓伟

(华南理工大学 理学院, 广州 510641)

(xwyang@scut.edu.cn)

**摘要:**高的计算复杂度限制了双边加权模糊支持向量机在实际分类问题中的应用。为了降低计算复杂度,提出了应用序贯最小优化算法(SMO)解该模型,该模型首先将整个二次规划问题分解成一系列规模为2的二次规划子问题,然后求解这些二次规划子问题。为了测试SMO算法的性能,在三个真实数据集和两个人工数据集上进行了数值实验。结果表明:与传统的内点算法相比,在不损失测试精度的情况下,SMO算法明显地降低了模型的计算复杂度,使其在实际中的应用成为可能。

**关键词:**序贯最小优化;双边加权模糊支持向量机;支持向量机;模糊支持向量机

**中图分类号:**TP301.6    **文献标志码:**A

## Sequential minimal optimization algorithm for bilateral-weighted fuzzy support vector machine

LI Yan, YANG Xiao-wei

(School of Sciences, South China University of Technology, Guangzhou Guangdong 510641, China)

**Abstract:** High computational complexity limits the applications of the Bilateral-Weighted Fuzzy Support Vector Machine (BW-FSVM) model in practical classification problems. In this paper, the Sequential Minimal Optimization (SMO) algorithm, which firstly decomposed the overall Quadratic Program (QP) problem into the smallest possible QP sub-problems and then solved these QP sub-problems analytically, was proposed to reduce the computational complexity of the BW-FSVM model. A set of experiments were conducted on three real world benchmarking datasets and two artificial datasets to test the performance of the SMO algorithm. The results indicate that compared with the traditional interior point algorithm, the SMO algorithm can reduce significantly the computational complexity of the BW-FSVM model without influencing the testing accuracy, and makes it possible for the BW-FSVM model to be applied to practical classification problems with outliers or noises.

**Key words:** Sequential Minimal Optimization (SMO); Bilateral-Weighted Fuzzy Support Vector Machine (BW-FSVM); Support Vector Machine (SVM); Fuzzy Support Vector Machine (FSVM)

## 0 引言

支持向量机(Support Vector Machine,SVM)是模式识别和机器学习领域的一种很重要的分类和非线性函数估计方法,其主要的缺点是标准的支持向量机模型对噪声和孤立点是敏感的<sup>[1]</sup>。针对这一问题,Lin等人<sup>[2]</sup>在2002年提出了模糊支持向量机(Fuzzy Support Vector Machine,FSVM),Jayadeva等人<sup>[3]</sup>在2005年提出了模糊近边界支持向量机,Tao等人<sup>[4]</sup>在2004年提出了一种基于加权间隔的模糊支持向量机。

模糊支持向量机的关键是如何设置训练样本的模糊隶属度。针对这个问题,2005年,Lin等人<sup>[5]</sup>通过引入置信因子和无用因子提出了模糊隶属度的自动生成方法。2006年,Jiang等人<sup>[6]</sup>基于高维特征空间样本与类中心的距离提出了一种新的模糊隶属度函数。2004年,基于模糊C均值聚类和模糊if-then规则,Leski<sup>[7]</sup>提出了 $\varepsilon$ 间隔非线性分类器来解决带孤立点或噪声点的分类问题,并提出了迭代设置样本权重和集成学习的策略。

在以上提出的模型中,模糊隶属度 $s_i$ 是对应样本点属于

某一类的程度,而 $1-s_i$ 是无意义的程度。考虑到在实际的分类问题中,同一个样本点可能属于多个类,Wang等人<sup>[8]</sup>分别提出了双边加权模糊支持向量机模型和它的最小二乘版本<sup>[9]</sup>来评估信贷风险。2008年,Jilani等人<sup>[10]</sup>将双边加权模糊支持向量机应用在多分类问题中。对于一个具有 $l$ 个样本点的训练集,双边加权模糊支持向量机模型需要解 $2l$ 个变量的二次规划问题。在先前的研究中,这个模型运用传统的优化算法来求解,其计算复杂度为 $O(8l^3)$ 。当训练样本上万时,如果计算机没有足够的内存,用传统的优化算法来求解双边加权模糊支持向量机模型是不现实的,这就限制了其在实际中的应用。目前,如何降低计算复杂度是双边加权模糊支持向量机模型的关键问题之一。在本文中,我们主要处理这个问题。

在支持向量机领域中,分解算法是处理实际分类问题的主要算法之一<sup>[11-13]</sup>。其中序贯最小优化(Sequential Minimal Optimization,SMO)算法的应用最广泛<sup>[14]</sup>。SMO算法将整个二次规划问题分解成一系列规模为2的二次规划子问题,然后解这些二次规划子问题,这使得SMO算法解决大规模分类

收稿日期:2011-06-24;修回日期:2011-08-05。

基金项目:国家自然科学基金资助项目(61070033);广东省自然科学基金重点资助项目(92510090001000005)。

作者简介:李艳(1985-),女,河南开封人,硕士研究生,主要研究方向:支持向量机; 杨晓伟(1969-),男,河南平顶山人,教授,主要研究方向:优化计算、数据挖掘、自适应算法。

问题成为可能。目前,SMO 算法在大规模的分类问题中已有应用<sup>[15~20]</sup>。为了降低计算复杂度,本文用 SMO 算法求解双边加权模糊支持向量机(Bilateral-Weighted Fuzzy Support Vector Machine,BW-FSVM)模型,使得该模型在实际分类问题中的应用成为可能。

## 1 双边加权模糊支持向量机模型

对于二分类问题,给定如下的训练数据集

$$T = \{(\mathbf{x}_1, y_1), (\mathbf{x}_2, y_2), \dots, (\mathbf{x}_l, y_l)\} \quad (1)$$

2005 年,Wang 等人<sup>[8]</sup>将训练集扩展如下:

$$T_b = \{(\mathbf{x}_1, +1, m_1), (\mathbf{x}_1, -1, 1 - m_1), (\mathbf{x}_2, +1, m_2), \\ (\mathbf{x}_2, -1, 1 - m_2), \dots, (\mathbf{x}_l, +1, m_l), (\mathbf{x}_l, -1, 1 - m_l)\} \quad (2)$$

其中输入数据  $\mathbf{x}_n \in \mathbb{R}^p$ , 对应的类标  $y_n \in \{-1, +1\}$ ,  $m_n$  是  $\mathbf{x}_n$  属于 +1 类的模糊隶属度。

基于训练集  $T_b$ , 双边加权模糊支持向量机的原问题<sup>[8]</sup>如下:

$$\min_{w, b, \xi_n, \xi_n'} \frac{1}{2} \|w\|^2 + C \sum_{n=1}^l (m_n \xi_n + (1 - m_n) \xi_n') \quad (3)$$

$$\text{s. t. } w \cdot \varphi(\mathbf{x}_n) + b \geq 1 - \xi_n; n = 1, 2, \dots, l,$$

$$w \cdot \varphi(\mathbf{x}_n) + b \leq -1 + \xi_n'; n = 1, 2, \dots, l,$$

$$\xi_n, \xi_n' \geq 0; n = 1, 2, \dots, l,$$

式中:  $w$  是超平面的法向量,  $b$  是偏项,  $\varphi(\mathbf{x}_n)$  是将  $\mathbf{x}_n$  映射到高维特征空间的非线性映射,  $\xi_n, \xi_n'$  是松弛变量,  $C$  是控制模型复杂性和训练误差的正则化常数。

原始优化问题的对偶问题为:

$$\min_{\alpha_i, \alpha'_i} \frac{1}{2} \sum_{n=1}^l \sum_{t=1}^l (\alpha_n - \alpha'_n)(\alpha_t - \alpha'_t) k(\mathbf{x}_n, \mathbf{x}_t) - \\ \sum_{n=1}^l (\alpha_n + \alpha'_n) \quad (4)$$

$$\text{s. t. } \sum_{n=1}^l (\alpha_n - \alpha'_n) = 0,$$

$$0 \leq \alpha_n \leq Cm_n; n = 1, 2, \dots, l,$$

$$0 \leq \alpha'_n \leq C(1 - m_n); n = 1, 2, \dots, l,$$

其中  $k(\mathbf{x}_n, \mathbf{x}) = (\varphi(\mathbf{x}_n), \varphi(\mathbf{x}))$  是 Mercer 核函数。

求解上面的优化问题,从而可以得到如下的决策函数。

$$f(\mathbf{x}) = \sum_{n=1}^l (\alpha_n - \alpha'_n) k(\mathbf{x}_n, \mathbf{x}) + b \quad (5)$$

从对偶问题(4)可知,对于  $l$  个样本的训练集,双边加权模糊支持向量机模型需要求解  $2l$  个变量的二次规划问题。如果用传统的优化算法来求解此模型,其计算复杂度  $O(8l^3)$  限制了其在实际分类问题中的应用。在后边的研究中,考虑到 SMO 算法在实际分类问题中的广泛应用,我们给出求解对偶问题的 SMO 算法。

## 2 BW-FSVM 模型的 SMO 算法

本章首先给出对偶问题(4)的最优条件,其次给出选择工作集的策略,然后给出更新拉格朗日乘子的公式,最后给出 SMO 算法的步骤。

### 2.1 对偶问题的最优条件

对偶问题(4)的拉格朗日函数如下:

$$L_D(\alpha_n, \alpha'_n, \beta, \pi_n, \psi_n, \delta_n, \eta_n) =$$

$$\frac{1}{2} \sum_{n=1}^l \sum_{t=1}^l (\alpha_n - \alpha'_n)(\alpha_t - \alpha'_t) k(\mathbf{x}_n, \mathbf{x}_t) - \\ \sum_{n=1}^l (\alpha_n + \alpha'_n) + \beta \sum_{n=1}^l (\alpha_n - \alpha'_n) - \sum_{n=1}^l \pi_n \alpha_n - \sum_{n=1}^l \psi_n \alpha'_n -$$

$$\sum_{n=1}^l \delta_n (Cm_n - \alpha_n) - \sum_{n=1}^l \eta_n [C(1 - m_n) - \alpha'_n] \quad (6)$$

式中:  $\beta, \pi_n \geq 0, \psi_n \geq 0, \delta_n \geq 0$  和  $\eta_n \geq 0$  是拉格朗日乘子。

令

$$F_n = - \sum_{t=1}^l (\alpha_t - \alpha'_t) k(\mathbf{x}_t, \mathbf{x}_n); n = 1, 2, \dots, l \quad (7)$$

对偶问题(4)的最优条件如下

$$\frac{\partial L_D(\alpha_n, \alpha'_n, \beta, \pi_n, \psi_n, \delta_n, \eta_n)}{\partial \alpha_n} = \\ - F_n - 1 + \beta - \pi_n + \delta_n = 0 \quad (8)$$

$$\frac{\partial L_D(\alpha_n, \alpha'_n, \beta, \pi_n, \psi_n, \delta_n, \eta_n)}{\partial \alpha'_n} = \\ F_n - 1 - \beta - \psi_n + \eta_n = 0 \quad (9)$$

$$\pi_n \alpha_n = 0; n = 1, 2, \dots, l \quad (10)$$

$$\psi_n \alpha'_n = 0; n = 1, 2, \dots, l \quad (11)$$

$$\delta_n (Cm_n - \alpha_n) = 0; n = 1, 2, \dots, l \quad (12)$$

$$\eta_n [C(1 - m_n) - \alpha'_n] = 0; n = 1, 2, \dots, l \quad (13)$$

$$0 \leq \alpha_n \leq Cm_n; n = 1, 2, \dots, l \quad (14)$$

$$0 \leq \alpha'_n \leq C(1 - m_n); n = 1, 2, \dots, l \quad (15)$$

$$\pi_n \geq 0, \psi_n \geq 0, \delta_n \geq 0, \eta_n \geq 0; n = 1, 2, \dots, l \quad (16)$$

为了方便讨论,先给出下面的定理。

**定理 1** 当对偶问题取得最优解时,集合  $\{n | \alpha_n = 0, \alpha'_n = 0\}, \{n | \alpha_n = 0, 0 < \alpha'_n < C(1 - m_n)\}, \{n | 0 < \alpha_n < Cm_n, \alpha'_n = 0\}$  和  $\{n | 0 < \alpha_n < Cm_n, 0 < \alpha'_n < C(1 - m_n)\}$  是空集,其中  $0 < m_n < 1$ 。

证明 假设集合  $\{n | \alpha_n = 0, \alpha'_n = 0\}$  非空。由式(8),(9),(12),(13),(16)可得:

$$- F_n - 1 + \beta \geq 0 \quad (17)$$

$$F_n - 1 - \beta \geq 0 \quad (18)$$

由式(17),(18)可得

$$F_n + 1 \leq F_n - 1 \quad (19)$$

这是不可能的。故  $\{n | \alpha_n = 0, \alpha'_n = 0\}$  是空集。

同理可证  $\{n | \alpha_n = 0, 0 < \alpha'_n < C(1 - m_n)\}, \{n | 0 < \alpha_n < Cm_n, \alpha'_n = 0\}, \{n | 0 < \alpha_n < Cm_n, 0 < \alpha'_n < C(1 - m_n)\}$  是空集。

接下来,简要讨论  $m_n = 0, m_n = 1$  和  $0 < m_n < 1$  三种情况下的最优条件。

1)  $m_n = 0$ 。由式(14)可知,  $\alpha_n = 0$ , 对于不同的  $\alpha'_n$ , 由对偶问题的最优条件可得:

当  $\alpha'_n = 0$  时,  $F_n - 1 \geq \beta$ ; 当  $\alpha'_n = C$  时,  $F_n - 1 \leq \beta$ ; 当  $0 < \alpha'_n < C$  时,  $F_n - 1 = \beta$ 。

2)  $m_n = 1$ 。由式(15)可得,  $\alpha'_n = 0$ , 对于取不同值的  $\alpha_n$ , 可以得到最优条件为:

当  $\alpha_n = 0$  时,  $F_n + 1 \leq \beta$ ; 当  $\alpha_n = C$  时,  $F_n + 1 \geq \beta$ ; 当  $0 < \alpha_n < C$  时,  $F_n + 1 = \beta$ 。

3)  $0 < m_n < 1$ 。考虑到  $\alpha_n$  和  $\alpha'_n$  分别有三种不同的取值,因而  $(\alpha_n, \alpha'_n)$  有 9 种组合。但由定理 1 可知,其中  $(\alpha_n, \alpha'_n)$  的 4 种组合是不可能的,因此只有下面 5 种可能的组合:

当  $\alpha_n = 0, \alpha'_n = C(1 - m_n)$  时,  $F_n + 1 \leq \beta$ ; 当  $\alpha_n = Cm_n, \alpha'_n = 0$  时,  $F_n - 1 \geq \beta$ ; 当  $\alpha_n = Cm_n, \alpha'_n = C(1 - m_n)$  时,  $-1 \leq F_n - \beta \leq 1$ ; 当  $\alpha_n = Cm_n, 0 < \alpha'_n < C(1 - m_n)$  时,  $F_n - 1 = \beta$ ; 当  $0 < \alpha_n < Cm_n, \alpha'_n = C(1 - m_n)$  时,  $F_n + 1 = \beta$ 。

定义下面的指标集

$$I_{01} = \{n | \alpha_n = 0, \alpha'_n = 0, m_n = 0\}$$

$$I_{02} = \{n | \alpha_n = 0, 0 < \alpha'_n < C, m_n = 0\}$$

$$\begin{aligned}
I_{03} &= \{n \mid \alpha_n = C, \alpha'_n = 0, m_n = 1\} \\
I_{04} &= \{n \mid 0 < \alpha_n < C, \alpha'_n = 0, m_n = 1\} \\
I_{05} &= \{n \mid \alpha_n = 0, \alpha'_n = C, m_n = 0\} \\
I_{06} &= \{n \mid \alpha_n = Cm_n, \alpha'_n = 0, 0 < m_n < 1\} \\
I_{07} &= \{n \mid \alpha_n = 0, \alpha'_n = 0, m_n = 1\} \\
I_{08} &= \{n \mid \alpha_n = 0, \alpha'_n = C(1 - m_n), 0 < m_n < 1\} \\
I_1 &= \{n \mid \alpha_n = Cm_n, \alpha'_n = C(1 - m_n), 0 < m_n < 1\} \\
I_2 &= \{n \mid \alpha_n = Cm_n, 0 < \alpha'_n < C(1 - m_n), 0 < m_n < 1\} \\
I_3 &= \{n \mid 0 < \alpha_n < Cm_n, \alpha'_n = C(1 - m_n), 0 < m_n < 1\}
\end{aligned}$$

设

$$\begin{aligned}
\overline{\overline{F}}_n &= \begin{cases} F_n + 1, & n \in I_{03} \cup I_{04} \cup I_1 \cup I_3 \\ F_n - 1, & n \in I_{01} \cup I_{02} \cup I_{06} \cup I_2 \end{cases} \\
\overline{F}_n &= \begin{cases} F_n + 1, & n \in I_{04} \cup I_{07} \cup I_{08} \cup I_3 \\ F_n - 1, & n \in I_{02} \cup I_{05} \cup I_1 \cup I_2 \end{cases}
\end{aligned}$$

那么,最优条件可以改写为

$$\begin{cases} \beta \leq \overline{\overline{F}}_n, n \in I_{01} \cup I_{02} \cup I_{03} \cup I_{04} \cup I_{06} \cup I_1 \cup I_2 \cup I_3 \\ \beta \geq \overline{F}_n, n \in I_{02} \cup I_{04} \cup I_{05} \cup I_{07} \cup I_{08} \cup I_1 \cup I_2 \cup I_3 \end{cases}$$

设

$$b_{up} = \min \{ \overline{\overline{F}}_n, n \in I_{01} \cup I_{02} \cup I_{03} \cup I_{04} \cup I_{06} \cup I_1 \cup I_2 \cup I_3 \} \\
b_{low} = \max \{ \overline{F}_n, n \in I_{02} \cup I_{04} \cup I_{05} \cup I_{07} \cup I_{08} \cup I_1 \cup I_2 \cup I_3 \}$$

则最优条件满足当且仅当

$$b_{low} \leq b_{up} \quad (20)$$

表 1 指标集和要优化的变量对之间的对应关系

| 指标集                                                                                            | 要被优化的变量对                                                                                     |
|------------------------------------------------------------------------------------------------|----------------------------------------------------------------------------------------------|
| $i \in I_{01} \cup I_{03} \cup I_{06} \cup I_1, j \in I_{05} \cup I_{07} \cup I_{08} \cup I_1$ | $(\alpha_i, \alpha_j), (\alpha'_i, \alpha'_j), (\alpha'_i, \alpha_j), (\alpha_i, \alpha'_j)$ |
| $i \in I_{01} \cup I_{03} \cup I_{06} \cup I_1, j \in I_{04} \cup I_3$                         | $(\alpha_i, \alpha_j), (\alpha'_i, \alpha_j)$                                                |
| $i \in I_{01} \cup I_{03} \cup I_{06} \cup I_1, j \in I_{02} \cup I_2$                         | $(\alpha'_i, \alpha'_j), (\alpha_i, \alpha'_j)$                                              |
| $i \in I_{02} \cup I_2, j \in I_{05} \cup I_{07} \cup I_{08} \cup I_1$                         | $(\alpha'_i, \alpha'_j), (\alpha'_i, \alpha_j)$                                              |
| $i \in I_{02} \cup I_2, j \in I_{04} \cup I_3$                                                 | $(\alpha'_i, \alpha_j)$                                                                      |
| $i \in I_{02} \cup I_2, j \in I_{02} \cup I_2$                                                 | $(\alpha'_i, \alpha'_j)$                                                                     |
| $i \in I_{04} \cup I_3, j \in I_{05} \cup I_{07} \cup I_{08} \cup I_1$                         | $(\alpha_i, \alpha_j), (\alpha_i, \alpha'_j)$                                                |
| $i \in I_{04} \cup I_3, j \in I_{04} \cup I_3$                                                 | $(\alpha_i, \alpha_j)$                                                                       |
| $i \in I_{04} \cup I_3, j \in I_{02} \cup I_2$                                                 | $(\alpha_i, \alpha'_j)$                                                                      |

由表 1 可知,在优化过程中,仅仅需要优化其中的 4 个变量对  $(\alpha_i, \alpha_j), (\alpha'_i, \alpha'_j), (\alpha'_i, \alpha_j)$  和  $(\alpha_i, \alpha'_j)$ 。

设

$$(\alpha_i - \alpha'_i) + (\alpha_j - \alpha'_j) = (\alpha_i^{\text{old}} - \alpha'^{\text{old}}_i) + (\alpha'_j - \alpha'^{\text{old}}_j) = \gamma \quad (26)$$

$$\eta = k(\mathbf{x}_i, \mathbf{x}_i) + k(\mathbf{x}_j, \mathbf{x}_j) - 2k(\mathbf{x}_i, \mathbf{x}_j) \quad (27)$$

$$\varphi_t = \sum_{n=1}^l (\alpha_n^{\text{old}} - \alpha'^{\text{old}}_n) k(\mathbf{x}_n, \mathbf{x}_t) + b; t = 1, 2, \dots, l \quad (28)$$

$$S_w = \{i, j\} \quad (29)$$

将式(4)的目标函数  $\frac{1}{2} \sum_{n=1}^l \sum_{t=1}^l (\alpha_n - \alpha'_n)(\alpha_t - \alpha'_t) k(\mathbf{x}_n, \mathbf{x}_t) - \sum_{n=1}^l (\alpha_n + \alpha'_n)$  展开可得

$$\begin{aligned}
&\frac{1}{2} \sum_{n=1}^l \sum_{t=1}^l (\alpha_n - \alpha'_n)(\alpha_t - \alpha'_t) k(\mathbf{x}_n, \mathbf{x}_t) - \\
&\sum_{n=1}^l (\alpha_n + \alpha'_n) = \frac{1}{2}(\alpha_i - \alpha'_i)^2 \eta + \\
&(\alpha_i - \alpha'_i)[(\varphi_i - \varphi_j) - \eta(\alpha_i^{\text{old}} - \alpha'^{\text{old}}_i)] - \\
&\sum_{n \in S_w} (\alpha_n + \alpha'_n) + C_1
\end{aligned} \quad (30)$$

其中  $C_1$  是常数。

## 2.2 选择工作集

设满足下面的两个条件之一的训练样本对为  $(i, j)$

$$i \in I_{01} \cup I_{02} \cup I_{03} \cup I_{04} \cup I_{06} \cup I_1 \cup I_2 \cup I_3, j \in I_{02} \cup I_{04} \cup I_{05} \cup I_{07} \cup I_{08} \cup I_1 \cup I_2 \cup I_3 \text{ 并且 } \overline{\overline{F}}_i < \overline{\overline{F}}_j \quad (21)$$

$$i \in I_{02} \cup I_{04} \cup I_{05} \cup I_{07} \cup I_{08} \cup I_1 \cup I_2 \cup I_3, j \in I_{01} \cup I_{02} \cup I_{03} \cup I_{04} \cup I_{06} \cup I_1 \cup I_2 \cup I_3 \text{ 并且 } \overline{F}_i > \overline{F}_j \quad (22)$$

那么这样的训练样本对  $(i, j)$  定义了一个矛盾对。

设

$$i_{up} = \operatorname{argmin}_{n \in I_{01} \cup I_{02} \cup I_{03} \cup I_{04} \cup I_{06} \cup I_1 \cup I_2 \cup I_3} \overline{\overline{F}}_n \quad (23)$$

$$i_{low} = \operatorname{argmax}_{n \in I_{02} \cup I_{04} \cup I_{05} \cup I_{07} \cup I_{08} \cup I_1 \cup I_2 \cup I_3} \overline{F}_n \quad (24)$$

$$\text{并且 } b_{low} > b_{up} \quad (25)$$

那么训练样本对  $(i_{low}, i_{up})$  定义了一个最大的矛盾对,为了提高优化过程的速度,我们采取两个策略来挑选工作集:一个是最小矛盾对策略,也就是说将式(23)和(24)定义的最大矛盾对  $(i_{low}, i_{up})$  挑选到工作集内进行优化,另一个策略是全部违反策略,将所有满足式(21)或(22)的矛盾对都放到工作集里进行优化<sup>[21]</sup>。

## 2.3 更新拉格朗日乘子 $\alpha_i, \alpha'_i, \alpha_j$ 和 $\alpha'_j$

由指标集的定义可知,由于  $i$  和  $j$  属于不同的集合,这样就有九种集合的组合方法来确定要优化的变量对,具体组合见表 1。

如果待优化的变量对是  $(\alpha_i, \alpha_j)$  和  $(\alpha'_i, \alpha'_j)$ ,设定  $s = 1$ ,否则  $s = -1$ ,由式(30)可知,对偶问题(4)等价于下面的优化问题:

$$\begin{aligned}
&\min \frac{1}{2} (\alpha_i - \alpha'_i)^2 \eta + (\alpha_i - \alpha'_i)[(\varphi_i - \varphi_j) - \\
&\eta(\alpha_i^{\text{old}} - \alpha'^{\text{old}}_i)] - (\alpha_i + \alpha'_i)(1 - s) \quad (31)
\end{aligned}$$

$$\text{s. t. } \alpha_i \in [L, H], \alpha'_i \in [L', H']$$

其中:  $H$  和  $L$  分别是  $\alpha_i$  的上、下界,  $H'$  和  $L'$  分别是  $\alpha'_i$  的上、下界。关于  $\alpha_i$  和  $\alpha'_i$  的无约束最小优化问题如表 2。

表 2 变量对和无约束值  $\alpha_i$  或者  $\alpha'_i$  之间的关系表

| 变量对                      | 无约束值 $\alpha_i$ 或者 $\alpha'_i$                                                                                                          |
|--------------------------|-----------------------------------------------------------------------------------------------------------------------------------------|
| $(\alpha_i, \alpha_j)$   | $\alpha_i^{\text{unconstrained}} = \alpha'_i + (\alpha_i^{\text{old}} - \alpha'^{\text{old}}_i) - \eta^{-1}(\varphi_i - \varphi_j)$     |
| $(\alpha_i, \alpha'_j)$  | $\alpha_i^{\text{unconstrained}} = \alpha'_i + (\alpha_i^{\text{old}} - \alpha'^{\text{old}}_i) - \eta^{-1}(\varphi_i - \varphi_j - 2)$ |
| $(\alpha'_i, \alpha_j)$  | $\alpha_i^{\text{unconstrained}} = \alpha_i - (\alpha_i^{\text{old}} - \alpha'^{\text{old}}_i) + \eta^{-1}(\varphi_i - \varphi_j + 2)$  |
| $(\alpha'_i, \alpha'_j)$ | $\alpha_i^{\text{unconstrained}} = \alpha_i - (\alpha_i^{\text{old}} - \alpha'^{\text{old}}_i) + \eta^{-1}(\varphi_i - \varphi_j)$      |

$\alpha_i$  和  $\alpha'_i$  的最终的计算公式分别如下:

$$\alpha_i^{\text{new}} = \begin{cases} H, & \alpha_i^{\text{unconstrained}} \geq H \\ \alpha_i^{\text{unconstrained}}, & L < \alpha_i^{\text{unconstrained}} < H \\ L, & \alpha_i^{\text{unconstrained}} \leq L \end{cases} \quad (32)$$

$$\alpha_i^{\text{new}'} = \begin{cases} H', & \alpha_i^{\text{unconstrained}'} \geq H' \\ \alpha_i^{\text{unconstrained}'} , & L' < \alpha_i^{\text{unconstrained}'} < H' \\ L', & \alpha_i^{\text{unconstrained}'} \leq L' \end{cases} \quad (33)$$

表 3 要优化的变量对和对应的上下界之间的关系表

| 要优化的变量对                  | 变量对对应的上下界                                                                                                        |
|--------------------------|------------------------------------------------------------------------------------------------------------------|
| $(\alpha_i, \alpha_j)$   | $L = \max(0, \gamma + \alpha_i' + \alpha_j' - Cm_j), H = \min(\gamma + \alpha_i' + \alpha_j', Cm_j)$             |
| $(\alpha_i, \alpha_j')$  | $L = \max(0, \gamma + \alpha_i' - \alpha_j'), H = \min(\gamma + \alpha_i' - \alpha_j + C(1 - m_j), Cm_i)$        |
| $(\alpha_i', \alpha_j)$  | $L' = \max(0, -\gamma + \alpha_i - \alpha_j'), H' = \min(-\gamma + \alpha_i - \alpha_j' + Cm_j, C(1 - m_i))$     |
| $(\alpha_i', \alpha_j')$ | $L' = \max(0, -\gamma + \alpha_i + \alpha_j - C(1 - m_j)), H' = \min(-\gamma + \alpha_i + \alpha_j, C(1 - m_i))$ |

## 2.5 更新 $F_n$

在优化过程中,每一步优化之后,  $F_n$  需要进行更新,更新  $F_n$  的公式如下:

$$F_n^{\text{new}} = F_n^{\text{old}} - \sum_{t \in S_w} [\alpha_t^{\text{new}} - \alpha_t^{\text{old}} - (\alpha_t^{\text{new}'} - \alpha_t^{\text{old}'})] k(\mathbf{x}_t, \mathbf{x}_n); \quad n = 1, 2, \dots, l \quad (34)$$

## 2.6 双边加权模糊支持向量机的 SMO 算法

基于上面的分析,我们给出双边加权模糊支持向量机的 SMO 算法如下:

步骤 1 输入训练数据集和超参数的值

步骤 2 设定拉格朗日乘子  $\alpha_n$  和  $\alpha_n'$  的初始值,  $n = 1, 2, \dots, l$ 。

步骤 3 利用式(23)和(24),计算  $i\_low, i\_up$ 。

步骤 4 检查  $i\_low$  和  $i\_up$  是否违背式(25),如果违背了,就优化对应的拉格朗日乘子  $\alpha_{i\_low}, \alpha_{i\_low}', \alpha_{i\_up}$  和  $\alpha_{i\_up}'$ ,然后返回步骤 3;否则继续。

步骤 5 遍历所有的训练样本  $i$ ,利用式(21)和(22)检查是否存在与  $i$  匹配的  $j$ ,如果存在,优化对应的拉格朗日乘子  $\alpha_i, \alpha_i', \alpha_j$  和  $\alpha_j'$ ,返回步骤 3。如果所有的训练样本都不违背式(21)和(22),就转到步骤 6。

步骤 6 输出支持向量和对应的拉格朗日乘子。

对于双边加权模糊支持向量机的对偶优化问题,SMO 算法在每一步优化中仅仅挑选两个拉格朗日乘子去优化,解析地求解这两个拉格朗日乘子,避开了整体优化数值二次规划;另外,SMO 算法也不需要额外存储矩阵,这些优良的特征使得双边加权模糊支持向量机在带噪声和孤立点的分类问题中的应用成为可能。双边加权模糊支持向量机模型中涉及到  $2l$  个变量,因此其时间复杂度为  $O((2l)^{2,2})^{[14]}$ 。

## 3 数值实验和讨论

在三个实际数据集和两个人工数据集上进行实验来测试 SMO 算法的性能。为了比较双边加权模糊支持向量机模型和标准支持向量机模型的性能,我们也用 SMO 算法<sup>[18]</sup>求解标准的支持向量机模型。为了表明 SMO 算法大大降低了计算复杂性,给出了利用预测一校正算法(Predictor-corrector algorithm,PrCo)<sup>[22]</sup>即传统的内点算法来求解双边加权的模糊支持向量机模型的结果。

### 3.1 实验环境和数据集

实验中,采用高斯核函数,通过网格剖分方法来寻找近似的最优超参数,剖分的网格为  $\sigma = [2^{-4}, 2^{-3}, 2^{-2}, \dots, 2^5]$  和  $C = [2^0, 2^1, 2^2, \dots, 2^9]$ ,程序运行的硬件环境是拥有英特尔双核处理器、最大内存 3.25 GB,CPU 为 3.16 GHz 的 PC,软件环境是 Windows XP,编程语言是 C++,编译器是 VC++ 6.0。

Letter 和 Statlog 数据集来自于 <http://archive.ics.uci.edu/ml>,并经过了如下的预处理使之变为二类数据集:Letter 是一个 26 类的数据集,将类标为 {A, B, ..., M} 的看成正类,类标为 {N, O, ..., Z} 看成负类;Statlog 是一个 6 类的数据集,将类标为 {1, 2, 5} 的看成正类,类标为 {3, 4, 6} 看成负类。数据集的其他详细信息如表 4。

基于式(26),我们能够计算  $\alpha_i$  和  $\alpha_i'$  最终的值。

### 2.4 确定 $\alpha_i$ 和 $\alpha_i'$ 的上下界

由等式(14)、(15)、(26)可以得到待优化的变量对与其对应的上下界的关系如表 3。

表 4 数据集的详细信息

| 数据集     | 样本数   |      | 合计    |
|---------|-------|------|-------|
|         | 训练集   | 测试集  |       |
| Ripley  | 250   | 1000 | 1250  |
| Letter  | 16000 | 4000 | 20000 |
| Statlog | 4435  | 2000 | 6435  |
| 人工数据集 1 | 279   | 281  | 560   |
| 人工数据集 2 | 344   | 336  | 680   |

### 3.2 产生模糊隶属度

数据集 Ripley,人工数据集 1,人工数据集 2 的模糊隶属度利用下面的方法设置。

设在高维特征空间中正类和负类的中心分别为

$$\varphi_+(\mathbf{x}) = \frac{1}{l^+} \sum_{i=1}^{l^+} \varphi(\mathbf{x}_i)$$

$$\varphi_-(\mathbf{x}) = \frac{1}{l^-} \sum_{i=1}^{l^-} \varphi(\mathbf{x}_i)$$

其中  $l^+$  和  $l^-$  分别为正负类的样本数。

训练获得的分类超平面为

$$\mathbf{w} \cdot \varphi(\mathbf{x}) + b = 0$$

其中  $\mathbf{w} = (\varphi_+(\mathbf{x}) - \varphi_-(\mathbf{x}))$ ,  $b = [-(\varphi_+(\mathbf{x}) - \varphi_-(\mathbf{x})) \times (\varphi_+(\mathbf{x}) + \varphi_-(\mathbf{x}))]/2$ 。

如果  $y_i = +1$ ,那么

$$m_i = \min \left( 0.5 \times \left( 1 + \frac{\mathbf{w} \cdot \varphi(\mathbf{x}_i) + b}{\mathbf{w} \cdot \varphi_+(\mathbf{x}) + b} \right), 1.0 \right)$$

$$\text{否则 } m_i = \max \left( 0.5 \times \left( 1 - \frac{\mathbf{w} \cdot \varphi(\mathbf{x}_i) + b}{\mathbf{w} \cdot \varphi_-(\mathbf{x}) + b} \right), 0.0 \right).$$

如果  $m_i < 0$  或者  $m_i > 1$ ,那么就删除训练集中对应的样本点。

对于 Letter 和 Statlog 数据集,用 Keller 和 Hunt 提出的策略产生模糊隶属度<sup>[24]</sup>。

$$m_i = \begin{cases} 0.5 \times \left( 1 + \frac{\exp(C_0(d_-(\mathbf{x}_i) - d_+(\mathbf{x}_i))/d) - \exp(-C_0)}{\exp(C_0) - \exp(-C_0)} \right) & y_i = +1 \\ 0.5 \times \left( 1 - \frac{\exp(C_0(d_+(\mathbf{x}_i) - d_-(\mathbf{x}_i))/d) - \exp(-C_0)}{\exp(C_0) - \exp(-C_0)} \right) & y_i = -1 \end{cases} \quad (35)$$

其中  $d_+(x_i) = \|\varphi(x_i) - \varphi_+(x)\|$ ,  $d_-(x_i) = \|\varphi(x_i) - \varphi_-(x)\|$ ,  $d = \|\varphi_+(x) - \varphi_-(x)\|$ 。 $C_0$  是控制隶属度函数的参数。在实验中,通过在网格  $C_0 = [-100, -90, \dots, 90, 100]$  中搜索寻找到最优值  $C_0 = -100$ 。

表 5 五个数据集上各种方法的测试精度、测试时间和最超参数比较

| 数据集     | 算法             | 测试精度   | 测试时间/s     | $\sigma$ | $C$ |
|---------|----------------|--------|------------|----------|-----|
| Ripley  | SVM            | 0.9070 | 0.016      | 0.5      | 1   |
|         | BW-FSVM (SMO)  | 0.9140 | 0.016      | 0.25     | 1   |
|         | BW-FSVM (PrCo) | 0.9140 | 19.625     | 0.25     | 1   |
| Letter  | SVM            | 0.9845 | 273.578    | 2        | 2   |
|         | BW-FSVM (SMO)  | 0.9845 | 634.969    | 2        | 2   |
|         | BW-FSVM (PrCo) | —      | —          | —        | —   |
| Statlog | SVM            | 0.9845 | 36.078     | 32       | 4   |
|         | BW-FSVM (SMO)  | 0.9850 | 46.625     | 32       | 4   |
|         | BW-FSVM (PrCo) | 0.9850 | 118708.203 | 32       | 4   |
| 人工数据集 1 | SVM            | 0.8292 | 0.015      | 2        | 2   |
|         | BW-FSVM (SMO)  | 0.9146 | 0.672      | 0.5      | 128 |
|         | BW-FSVM (PrCo) | 0.9146 | 41.890     | 0.5      | 128 |
| 人工数据集 2 | SVM            | 0.9554 | 0.031      | 4        | 16  |
|         | BW-FSVM (SMO)  | 0.9643 | 0.078      | 2        | 512 |
|         | BW-FSVM (PrCo) | 0.9643 | 72.312     | 2        | 512 |

由表 5 可知,与 PrCo 算法相比,SMO 算法大大降低了双边加权支持向量机模型的计算复杂度,例如,对于 Statlog 和 Letter 数据集,所提出 SMO 算法仅仅分别花费 46.625 s 和 634.969 s。然而对于 Statlog 数据集,PrCo 算法用了 118708.203 s,对于 Letter 数据集,由于没有足够的内存从而不能进行训练。另一方面,如果模糊隶属度设置得合理,双边加权模糊支持向量机模型比标准的 SVM 模型能获得更好的性能。例如,对于数据集 Ripley、Statlog、人工数据集 1 和人工数据集 2,双边加权支持向量机模型的测试精度高于标准的 SVM 模型;对于 Letter 数据集,双边加权支持向量机模型的测试精度与标准的 SVM 模型的测试精度一样。

#### 4 结语

对于双边加权模糊支持向量机模型的高计算复杂度问题,本文运用 SMO 算法来求解。实验结果表明:与 PrCo 算法相比,SMO 算法大大降低了模型的计算复杂度,使得双边加权支持向量机模型在带噪声和孤立点的实际分类问题中的应用成为可能。尽管 SMO 算法只在二分类问题上进行了实验,但是它也能够很容易应用到多分类问题中。

在以后的工作中,需要继续探索设置双边加权模糊支持向量机模型的隶属度的算法,进一步研究 SMO 算法在实际分类问题中的应用。

#### 参考文献:

- [1] ZHANG X G. Using class-center vectors to build support vector machines[C]// Proceedings of the 1999 IEEE Signal Processing Society Workshop. New York: IEEE, 1999: 3 – 11.
- [2] LIN C F, WANG S D. Fuzzy support vector machines[J]. IEEE Transactions on Neural Networks, 2002, 13(2): 464 – 471.
- [3] JAYADEVA A, KHEMCHANDANI R, CHANDRA S. Fast and robust learning through fuzzy linear proximal support vector machines [J]. Neurocomputing, 2004(61): 401 – 411.
- [4] TAO QING, WANG JUE. A new fuzzy support vector machine based on the weighted margin[J]. Neural Processing Letters, 2004, 20(3): 139 – 150.
- [5] LIN C F, WANG S D. Fuzzy support vector machines with automatic membership setting[C]// Studies in Fuzziness and Soft Computing. Berlin: Springer, 2005: 233 – 254.
- [6] JIANG X F, YI Z, LV J C. Fuzzy SVM with a new fuzzy membership function[J]. Neural Computing and Applications, 2006, 15(3/4): 268 – 276.
- [7] LESKI J K. An  $\varepsilon$ -margin nonlinear classifier based on fuzzy if-then rules[J]. IEEE Transactions on Systems, Man, and Cybernetics Part B: Cybernetics, 2004, 34(1): 68 – 76.
- [8] WANG Y Q, WANG S Y, LAI K K. A new fuzzy support vector machine to evaluate credit risk[J]. IEEE Transactions on Fuzzy Systems, 2005, 13(6): 820 – 831.
- [9] HUANG W, LAI K K, YU L, et al. A least squares bilateral-weighted fuzzy SVM method to evaluated credit risk[C]// Proceedings of the 4th International Conference on Natural Computation. Washington, DC: IEEE Computer Society, 2008: 13 – 17.
- [10] JILANI T A, BURNEY S M A. Multiclass bilateral-weighted fuzzy support vector machine to evaluate financial strength credit rating [C]// International Conference on Computer Science and Information Technology. Washington, DC: IEEE Computer Society, 2008: 342 – 348.
- [11] DONG J X, KRZYZAK A, SUEN C Y. Fast SVM training algorithm with decomposition on very large data sets[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2005, 27(4): 603 – 618.
- [12] JOACHIMS T. Making large-scale SVM learning practical[C]// Advances in Kernel Methods-Support Vector Learning. Cambridge: MIT Press, 1998: 169 – 184.
- [13] OSUNA E, FREUND R, GIROSI F. Training support vector machines: an application to face detection[C]// Proceedings of 1997 IEEE Computer Society Conference on Computer Vision and Pattern Recognition. Washington, DC: IEEE Computer Society, 1997: 130 – 136.

(下转第 3317 页)

变能力差, 预测曲线平缓, 收敛速度慢。因此, 当足球运动状态发生突变时, 基于 STF 和  $H^\infty$  滤波算法对足球位置的预测要明显优于  $H^\infty$  滤波算法。

此外, 在 Intel Core i3-350M(2.26 GHz), 2 GB 内存的笔记本电脑 Windows 平台下, 测得两种滤波算法的平均处理时间如表 1 所示。

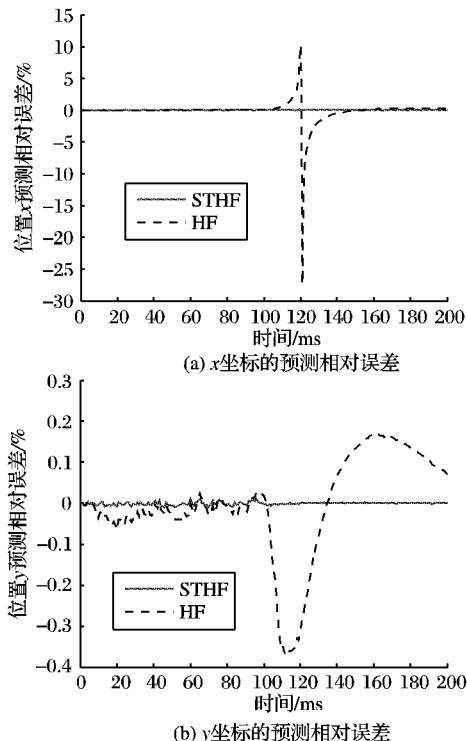


图 2 两种滤波算法的足球位置预测相对误差

表 1 两种滤波算法的平均处理时间

| 算法                     | 平均处理时间/ms |
|------------------------|-----------|
| $H^\infty$ 滤波          | 0.3       |
| 基于 STF 和 $H^\infty$ 滤波 | 0.4       |

根据表 1, 基于 STF 和  $H^\infty$  滤波算法的平均处理时间虽然稍长于  $H^\infty$  滤波算法, 但提高了预测精度, 且仍满足比赛过程中实时性高的要求。

(上接第 3301 页)

- [14] PLATT J C. Sequential minimal optimization: a fast algorithm for training support vector machines [C]// Advances in Kernel Methods. Cambridge: MIT Press, 1998: 185–208.
- [15] CHEN P H, FAN R E, LIN C J. A study on SMO-type decomposition methods for support vector machines [J]. IEEE Transactions on Neural Networks, 2006, 17(4): 893–908.
- [16] KEERTHI S S, GILBERT E G. Convergence of a generalized SMO algorithm for SVM classifier design [J]. Machine Learning, 2002, 46(1/3): 351–360.
- [17] KEERTHI S S, SHEVADE S K. SMO algorithm for least-squares SVM formulations [J]. Neural Computation, 2003, 15(2): 487–507.
- [18] KEERTHI S S, SHEVADE S K, BHATTACHARYYA C, et al. Improvements to Platt's SMO algorithm for SVM classifier design [J]. Neural Computation, 2001, 13(3): 637–649.
- [19] KNEBEL T, HOCHREITER S, OBERMAYER K. An SMO algorithm for the potential support vector machine [J]. Neural Computation, 2008, 20(1): 271–287.
- [20] LIN C J. Asymptotic convergence of an SMO algorithm without any assumptions [J]. IEEE Transactions on Neural Networks, 2002, 13(1): 248–250.
- [21] SHEVADE S K, KEERTHI S S, BHATTACHARYYA C, et al. Improvements to SMO algorithm for SVM regression [J]. IEEE Transactions on Neural Networks, 2000, 11(5): 1188–1193.
- [22] MEHROTRA S. On implementation of a primal-dual interior point method [J]. SIAM Journal on Optimization, 1992, 2(4): 575–601.
- [23] RIPLEY B D. Pattern recognition and neural networks [M]. Cambridge: Cambridge University Press, 1996.
- [24] KELLER J M, HUNT D J. Incorporating fuzzy membership functions into the perceptron algorithm [J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 1985, 7(6): 693–699.

## 4 结语

本文根据 STF 算法将渐消因子引入  $H^\infty$  滤波, 提出了基于 STF 和  $H^\infty$  滤波的算法, 并将该算法应用于机器人足球比赛中足球位置的预测。实验结果表明, 与  $H^\infty$  滤波算法相比, 在足球运动状态发生突变时, 该算法能够更准确地预测足球的位置, 具有更高的滤波精度和更快的收敛速度。

## 参考文献:

- [1] FREESTON L. Applications of the Kalman filter algorithm to robot localization and world modeling [R]. Newcastle: University of Newcastle, 2002.
- [2] HEINEMANN P, PLAGGE M, TREPTOW A, et al. Tracking dynamic objects in a robocup environment: the attempto tubingen robot soccer team [C]// Robot Soccer World Cup VII. Berlin: Springer-Verlag, 2004: 1–6.
- [3] RUIZ-DEL-SOLAR J, VALLEJOS P A. Motion detection and tracking for an AIBO robot using camera motion compensation and Kalman filtering [C]// RoboCup 2004: Robot Soccer World Cup VIII. Berlin: Springer-Verlag, 2004: 619–627.
- [4] 王文学, 王强, 孙萍. 卡尔曼滤波在机器人足球比赛系统中的应用 [J]. 机器人, 2006, 28(4): 410–414.
- [5] LI XIANG, ZELL A.  $H^\infty$  filtering for a mobile robot tracking a free rolling ball [C]// RoboCup 2006: Robot Soccer World Cup X. Berlin: Springer-Verlag, 2006.
- [6] 付梦印, 邓志红, 吴莉萍. Kalman 滤波理论及其在导航系统中的应用 [M]. 2 版. 北京: 科学出版社, 2010: 95–107.
- [7] HASSIBI B, SAYED A H, KAILATH T.  $H^\infty$  optimality of the LMS algorithm [J]. IEEE Transactions on Signal Processing, 1996, 44(2): 267–280.
- [8] SHEN X, DENG L. Game theory approach to discrete  $H^\infty$  filter design [J]. IEEE Transactions on Signal Processing, 1997, 45(4): 1092–1095.
- [9] 周东华, 席裕庚, 张钟俊. 非线性系统带次优渐消因子的扩展卡尔曼滤波 [J]. 控制与决策, 1990, 5(5): 1–6.
- [10] 周东华, 席裕庚, 张钟俊. 一种带多重次优渐消因子的扩展卡尔曼滤波器 [J]. 自动化学报, 1991, 17(6): 689–695.
- [11] 范文兵, 刘春风, 张素贞. 一种强跟踪扩展卡尔曼滤波器的改进算法 [J]. 控制与决策, 2006, 21(1): 73–76.