

# 基于乘性规则的支持向量域分类器

陆从德 张太镒 胡金燕

(西安交通大学电信学院信息与通信工程系 西安 710049)

**摘要** 该文提出了一种基于支持向量域描述(SVDD)的学习分类器。在两类样本分类中,该算法在训练时通过对 1 类样本的描述求取包含 1 类样本的球形边界,然后通过该边界对两类样本数据进行分类,并且在求取边界的优化问题中,采用乘性规则来直接求取 Lagrange 乘子,而不是用传统的二次优化方法。该文所获得的学习算法和支持向量机(SVM)与序列最小优化(SMO)算法相比,不仅降低了样本的采集代价,而且在优化速度上有了很大提高。通过 CBCL 人脸库的仿真实验,将该算法和 SVM、SOM 算法的实验结果进行对比,说明了该学习算法的有效性。

**关键词** 支持向量域描述;支持向量域分类器;乘性规则;支持向量机;序列最小优化

中图法分类号 TP18

## Support Vector Domain Classifier Based on Multiplicative Updates

LU Cong-De ZHANG Tai-Yi HU Jin-Yan

(Department of Information and Communication Engineering, Xi'an Jiaotong University, Xi'an 710049)

**Abstract** This paper proposes a support vector domain classifier based on multiplicative updates. In two-class problem, through description of the training samples from one class, this algorithm obtains the sphere boundary containing these samples, and then uses this boundary to classify the test samples. In addition, instead of the traditional quadratic programming, multiplicative updates is used to solve Lagrange multiplier in optimizing the solution of boundary. Compared to Support Vector Machine(SVM) and Sequential Minimal Optimization(SMO) algorithms, the learning algorithm shown in this paper not only decreases the collecting cost of samples, but also improves greatly the computational speed of optimization. The experiment on CBCL face database illustrates the effectiveness of this algorithm in comparison with SVM and SMO.

**Keywords** support vector domain description; support vector domain classifier; multiplicative updates; support vector machine; sequential minimal optimization

## 1 引言

支持向量机(Support Vector Machine, SVM)是近几年来的研究热点,并且使用 SVM 进行识别和分类获得了极大的成功<sup>[1~3]</sup>。SVM 算法是一种监督学习的两类分类器,在训练时需要知道两类样本

的类别。然而在很多分类和识别问题中,能获得已知类别的数据相对来说是非常少的,即一般得到的大部分数据都是不知道类别的数据,因此这时基于学习的两类分类器就不是很适用了。

TAX 等人在 SVM 的基础上提出了一种数据描述算法——支持向量域描述(Support Vector Domain Description, SVDD)<sup>[4]</sup>,主要用来进行数据描述和剔

收稿日期:2003-01-06;修改稿收到日期:2004-03-11。陆从德,男,1972 年生,博士研究生,主要研究方向为统计学习理论、核方法、神经网络、模式识别、数字图像处理。E-mail: congdelu@mailst.xjtu.edu.cn。张太镒,男,1944 年生,教授,博士生导师,主要研究方向为新一代移动通信技术、软件无线电、无线宽带接入技术、模式识别、数字图像处理和数字音视频压缩编码。胡金艳,女,1973 年生,博士研究生,主要研究方向为分形理论与应用、音视频信号处理、多媒体通信等。

除奇异点或出格点(outliers). SVDD的主要思想是通过计算包含一组数据的最小超球形边界来对该组数据进行描述. 该算法给我们的启发是: 在对两类数据进行分类时, 如果我们只知道一部分样本的类别是1类, 而其它样本的类别是未知的, 则我们可以在SVDD的基础上设计一个学习分类器. 该分类器在训练时只需要对已知类别的数据进行描述, 然后得到该类数据的描述边界, 根据此边界就可以对未知的两类数据进行分类了.

在SVM和SVDD算法中都把原问题转化为对偶问题(二次规划)进行求解, 而在对偶问题的优化计算上存在着一些问题, 如训练算法速度慢、算法复杂而难以实现以及测试阶段运算量大等. 传统的利用标准二次型优化技术解决对偶问题的方法可能是训练算法慢的主要原因. 首先, SVM方法需要计算和存储核函数矩阵, 当样本点数目较大时, 需要很大的内存, 其次, SVM在二次型寻优过程中要进行大量的矩阵运算, 多数情况下, 寻优算法是占用算法时间的主要部分.

近年来人们针对SVM方法本身的特点提出了许多改进算法来解决对偶寻优问题, 例如SMO(Sequential Minimal Optimization)算法<sup>[5]</sup>等. 大多数改进算法的一个共同的思想就是循环迭代: 将原问题分解成为若干子问题, 按照某种迭代策略, 通过反复求解子问题, 最终使结果收敛到原问题的最优解.

在本文所提出的学习分类器中并不采用常规的优化技术, 而是采用乘性规则(multiplicative updates)<sup>[6]</sup>. 乘性规则给出了一种新的解决二次优化问题的解法, 该方法不需要像SVM中那样仔细选择每一步的学习速率, 也不需要固定工作子集, 所有的变量可以并行迭代, 因此它提供了一种直接优化的方法.

本文第2节对SVDD进行回顾并且在此基础上设计出支持向量域分类器(Support Vector Domain Classifiers, SVDC); 第3节对乘性规则进行阐述并且把该乘性规则应用到支持向量域分类器中, 得到基于乘性规则的支持向量域分类器(Support Vector Domain Classifiers based on Multiplicative updates, SVDCM); 第4节给出了SVDC, SVDCM, SVM和SMO方法的对比实验结果并进行了必要的分析; 第5节在分析的基础上给出结论.

## 2 支持向量域分类器(SVDC)

### 2.1 支持向量域描述(SVDD)<sup>[4]</sup>

在对一组数据进行描述时, SVDD算法就是通

过找到能包含该组数据的最小半径的超球体以便能对其进行描述. 对于奇异点来说, 它应该在该超球体的外面. 为了减少奇异点的影响, 使用松弛变量 $\xi_i$ 把该奇异点排除在超球体的外面. 最小化超球体的体积就是二次规划问题, 即在约束条件

$$(\mathbf{x}_i - \mathbf{a})^\top (\mathbf{x}_i - \mathbf{a}) \leq R^2 + \xi_i \quad (1)$$

下最小化

$$\min [R^2 + C \sum_i \xi_i] \quad (2)$$

其中 $\xi_i \geq 0$ , 并且C是一个常数. 这个优化问题的解是由下面的拉格朗日(Lagrange)泛函的鞍点给出的:

$$L(R, a, \alpha_i, \xi_i) = R^2 + C \sum_i \xi_i - \sum_i \alpha_i (R^2 + \xi_i - (\mathbf{x}_i \cdot \mathbf{x}_i - 2\mathbf{a} \cdot \mathbf{x}_i + \mathbf{a} \cdot \mathbf{a})) \quad (3)$$

其中 $\alpha_i \geq 0, \beta_i \geq 0$ .

求式(3)的最小值可变成了求其对偶问题的最大值:

$$W(\alpha) = \sum_i \alpha_i K(\mathbf{x}_i, \mathbf{x}_i) - \sum_{i,j} \alpha_i \alpha_j K(\mathbf{x}_i, \mathbf{x}_j) \quad (4)$$

其中对 $\alpha_i$ 的约束是 $\sum_i \alpha_i = 1$  和 $0 \leq \alpha_i \leq C$ , 并且已经用核函数代替了内积.

如果一个测试点 $\mathbf{z}$ 满足如下条件的话, 则这个测试点就被接受, 否则就被拒绝.

$$K(\mathbf{z}, \mathbf{z}) - 2 \sum_i \alpha_i K(\mathbf{x}_i, \mathbf{z}) + \sum_{i,j} \alpha_i \alpha_j K(\mathbf{x}_i, \mathbf{x}_j) \leq R^2 \quad (5)$$

### 2.2 支持向量域分类器(SVDC)

SVDD算法主要是对1类数据进行描述, 但现在我们把它推广到两类分类的情况. 考虑两类样本 $(\mathbf{x}_i, y_i), i=1, 2, \dots, l, l+1, \dots, n$ , 且

$$y_i = \begin{cases} 1, & i = 1, \dots, l \\ -1, & i = l+1, \dots, n \end{cases}$$

现在我们对 $y_i = 1$ 的样本建立超球体, 而不考虑 $y_i = -1$ 的样本, 则所要求解的问题可以写为

$$\min [R^2 + C \sum_{i=1}^l \xi_i] \quad (6)$$

其约束条件为

$$y_i (R^2 - (\mathbf{x}_i - \mathbf{a})^\top (\mathbf{x}_i - \mathbf{a})) \geq \xi_i, \quad i = 1, 2, \dots, l \quad (7)$$

其中 $\xi_i \geq 0, y_i = 1$ , 并且C是一个常数. 同理, 这个优化问题的求解由下面的Lagrange泛函的鞍点给出:

$$L(R, a, \alpha_i, \xi_i) = R^2 + C \sum_{i=1}^l \xi_i + \beta_i \xi_i - \sum_{i=1}^l \alpha_i (y_i (R^2 - (\mathbf{x}_i - \mathbf{a})^\top (\mathbf{x}_i - \mathbf{a}))) \quad (8)$$

其中  $\alpha_i \geq 0, \beta_i \geq 0$ . 在式(8)中对  $R, \alpha, \xi$  求偏导, 然后再把所求出的  $R, \alpha, \xi$  的表达式代回式(8), 则可得到关于参数  $\alpha$  的表达式如下:

$$W(\alpha) = \sum_{i=1}^l \alpha_i y_i K(\mathbf{x}_i, \mathbf{x}_i) - \sum_{i,j=1}^l \alpha_i \alpha_j y_i y_j K(\mathbf{x}_i, \mathbf{x}_j) \quad (9)$$

其约束条件为

$$\sum_{i=1}^l \alpha_i y_i = 1, \quad 0 \leq \alpha_i \leq C$$

于是可以设计出分类器为

$$f(x) = \text{sgn}\left(R^2 - K(\mathbf{x}, \mathbf{x}) + 2 \sum_i \alpha_i y_i K(\mathbf{x}, \mathbf{x}_i) - \sum_{i,j} \alpha_i \alpha_j y_i y_j K(\mathbf{x}_i, \mathbf{x}_j)\right) \quad (10)$$

其中

$$R^2 = \frac{1}{k} \sum_{\{k|y_k=1, k \in \{\text{SV}\}\}} \left( K(\mathbf{x}_k, \mathbf{x}_i) - 2 \sum_i \alpha_i y_i K(\mathbf{x}_k, \mathbf{x}_i) + \sum_{i,j} \alpha_i \alpha_j y_i y_j K(\mathbf{x}_i, \mathbf{x}_j) \right) \quad (11)$$

在式(11)中,  $\mathbf{x}_k$  为支持向量,  $k$  为支持向量的个数.

在进行两类分类时, 首先, 由于 SVM 算法<sup>[1~3]</sup>要最大化两类数据的间隔, 因此在训练时需要两类数据. 在式(6), (7)中, 由于要求的是包含 1 类 ( $y_i = 1$ ) 数据的最小超球体边界, 于是在训练时只需要 1 类数据, 因此 SVDC 算法在训练时将极大减少了计算量. 假如 SVM 算法和 SVDC 算法在训练时使用样本数相等的两类样本数据, 那么 SVDC 算法从理论上将减少一半的计算量, 因为它在训练时只需要 1 类数据样本, 而训练时间却又是整个算法时间的主要部分. 其次, 在 SVM 算法中, 支持向量是那些过两类样本中离分类面最近的点且平行于最优分类面的超平面上的训练样本, 而在 SVDC 算法中, 支持向量是位于最小超球体边界上且  $y_i = 1$  的训练样本, 因此, SVDC 算法中的支持向量数要比 SVM 算法中支持向量数要少, 这就导致在样本测试阶段, SVDC 算法所需的计算量要少于 SVM 算法的计算量.

### 3 基于乘性规则的支持向量域分类器(SVDCM)

#### 3.1 乘性规则<sup>[6]</sup>

首先研究非负二次规划的一般问题. 考虑二次目标函数的最小化

$$F(\mathbf{v}) = \frac{1}{2} \mathbf{v}^\top \mathbf{A} \mathbf{v} + \mathbf{b}^\top \mathbf{v} \quad (12)$$

约束条件是  $v_i \geq 0, \forall i$ . 假定矩阵  $\mathbf{A}$  是对称的和半正定的, 因此对目标函数  $F(\mathbf{v})$  的优化是凸优化. 然而, 由于非负的限制, 所以不存在一个分析工具去求全局最小, 因此需要一种迭代的方法.

在式(12)中, 让  $\mathbf{A}^+$  和  $\mathbf{A}^-$  表示非负矩阵:

$$\begin{aligned} A_{ij}^+ &= \begin{cases} A_{ij}, & A_{ij} > 0 \\ 0, & \text{其它} \end{cases}, \\ A_{ij}^- &= \begin{cases} |A_{ij}|, & A_{ij} < 0 \\ 0, & \text{其它} \end{cases} \end{aligned} \quad (13)$$

于是可得  $\mathbf{A} = \mathbf{A}^+ - \mathbf{A}^-$ . 根据这些非负矩阵, 则乘性迭代规则如下:

$$v_i \leftarrow v_i \left[ \frac{-b_i + \sqrt{b_i^2 + 4(\mathbf{A}^+ \mathbf{v})_i (\mathbf{A}^- \mathbf{v})_i}}{2(\mathbf{A}^+ \mathbf{v})_i} \right] \quad (14)$$

可见, 该式很明显实现起来比较简单. 首先, 上式描述了一个乘性规则, 通过这个规则,  $\mathbf{v}$  的第  $i$  个元素由  $\mathbf{b}, \mathbf{A}^+ \mathbf{v}, \mathbf{A}^- \mathbf{v}$  的第  $i$  个元素来确定. 第二, 因为  $\mathbf{v}, \mathbf{A}^+, \mathbf{A}^-$  是非负的, 上式中右边的所有因子乘上  $v_i$  也是非负的. 因此, 这些迭代规则并没有脱离非负的限制.

由下面的定理可知, 乘性规则能收敛到全局最小点.

**定理 1<sup>[6]</sup>.** 利用等式(14)中的乘性规则, 式(12)中的目标函数可以单调下降到它的全局最小值.

#### 3.2 支持向量域分类器的乘性规则

我们现在考虑计算 SVDC 的 Lagrange 乘子问题. 式(9)的损失函数是式(12)中的特殊情况并且  $A_{ij} = y_i y_j K(\mathbf{x}_i, \mathbf{x}_j)$  和  $b_i = y_i K(\mathbf{x}_i, \mathbf{x}_i)$ . 因此, 在支持向量域分类器中的乘性规则如下:

$$\alpha_i \leftarrow \alpha_i \left[ \frac{y_i K(\mathbf{x}_i, \mathbf{x}_i) + \sqrt{(y_i K(\mathbf{x}_i, \mathbf{x}_i))^2 + 4(\mathbf{A}^+ \alpha)_i (\mathbf{A}^- \alpha)_i}}{2(\mathbf{A}^+ \alpha)_i} \right] \quad (15)$$

其中,  $\mathbf{A}^+$  和  $\mathbf{A}^-$  由式(13)定义.

传统的二次规划由于涉及到大量的矩阵运算, 所以运算速度慢成为了 SVM 的最大缺点. 通过使用式(15)的乘性规则代替传统的二次规划, 将可以避免这些缺点, 使得优化速度有了极大的提高, 这从后面的实验中可以得到验证.

### 4 实 验

为了验证 SVDC 和 SVDCM 的性能, 本文的实验是对人脸和非人脸进行分类. 实验采用的是 MIT

的CBCL人脸库,该人脸库已经被归一化为 $19 \times 19$ 的标准图像,其部分样本参见图1。为了说明SVDCM的有效性,在实验中分别对它和SVM和SMO算法进行了比较。在SVM和SMO的实验中,训练(或学习)的人脸样本数是200,非人脸样本数也是200;测试中使用的人脸样本数是100,非人脸样本数也为



图1 CBCL库的归一化人脸样本

为了得到明显的对比实验效果,根据分类时所需要的特征的不同,分别完成三个不同的实验:(1)分类所使用的特征直接使用图像的像素点。使用像素点作为特征,然后分别使用SVDC,SVM,SMO和SVDCM分类器进行分类汇总,分别获得的支持向量数、优化时间以及识别率如表1所示;(2)使用线性主成分分析(Linear Principal Component Analysis,LPCA)<sup>[7]</sup>提取的特征作为分类的特征,然后使用SVM,SMO,SVDC及SVDCM进行分类所获得的结果见表2;(3)使用基于核的主成分分析(Kernel based Principal Component Analysis,KPCA)<sup>[8]</sup>进行特征提取,然后使用SVM,SMO,SVDC和SVDCM进行分类所得到的实验结果如表3所示。

表1 像素点作为分类特征的实验结果

	SV数	优化时间(s)	正确率(%)
SVM	54(13.5%)	164.8	95.5
SMO	56(14.0%)	129.6	95.5
SVDC	28(14.0%)	91.8	93.5
SVDCM	28(14.0%)	70.5	93.0

表2 采用LPCA提取的特征作为分类特征的实验结果

	SV数	优化时间(s)	正确率(%)
SVM	57(14.3%)	125.6	94.5
SMO	58(14.5%)	108.4	94.5
SVDC	29(15.0%)	75.2	92.0
SVDCM	29(15.0%)	55.3	91.0

表3 采用KPCA提取的特征作为分类特征的实验结果

	SV数	优化时间(s)	正确率(%)
SVM	54(13.5%)	150.7	94.5
SMO	54(13.5%)	125.1	94.5
SVDC	27(13.5%)	85.4	91.5
SVDCM	27(13.5%)	63.3	91.0

从表1可看出:(1)SVM,SMO算法中支持向量几乎是SVDC和SVDCM算法的两倍,这是因为SVM和SMO算法在训练时需要计算人脸图像和

100.在SVDC(采用传统的优化方法)和SVDCM(采用乘性规则)中,训练的人脸样本数是200个,即在式(9)中令人脸样本的 $y_i=1$ ,并且训练时不考虑非人脸样本。测试的时候使用了100个人脸样本和100个非人脸样本。

非人脸图像,也就是样本数是SVDC算法所使用的样本数的两倍。(2)在优化时间上,SMO和SVM相比确实有了很大的提高,约提高了27%,这说明SMO算法的速度比SVM算法有优势。(3)由于SVDC和SVDCM使用的样本数较少,所以其优化的时间SVM及SMO要少得多,这也是本文所推荐的学习算法的优点。(4)由于SVDC和SVDCM在训练时只使用1类样本数据,所以支持向量少于SVM和SMO所获得的支持向量,所以提高了测试速度,这和理论分析是符合的。(5)也是最重要的是,SVDCM的优化时间比SVDC的优化时间要少,即使用乘性规则的优化时间更低,这说明乘性规则的优越性,它极大地提高了算法的运算速度。最后,从识别率来看,SVM,SMO具有一样的识别率,而SVDC和SVDCM算法的正确识别率要比它们低2%~3%,但是因为该算法极大地提高了学习速度,综合考虑,我们认为是值得的。

从表2及表3,我们得到的结论和表1得到的结论几乎是相同的,这说明不管是采用图像像素点,还是使用主成分分析所提取的特征作为分类特征,SVDCM都体现出了它良好的性质,因此,基于乘性规则的支持向量域分类器是一种良好的学习分类器,它在保持识别率较高的情况下,极大地提高了其运算速度,当然也同时降低了图像采集的成本或代价。

## 5 结 论

在两类样本分类情况下,如果采用基于两类学习的分类算法,那么将在训练时同时需要知道两类样本的类别,这不仅增加了训练时间,而且也增加了图像样本的采集代价。本文在SVDD的基础上提出

了一种学习分类器——SVDCM, 通过和学习算法 SVM, SMO 的理论分析和实验比较, 表明它具有以下优点:(1)由于训练时只需要 1 类样本, 因此降低了学习样本的采集成本;(2)由于训练时只使用 1 类样本, 因此提高了运算速度;(3)通过引进乘性规则进行优化计算, 使得其优化速度得到了极大的提高. 另外, 可以推广 SVDCM 去解决多类分类情况, 即对于每一类样本分别训练并且对每一类样本建一个分类器. 这种思想是符合人的学习过程的, 即人类是一点点地学习并且进行知识积累, 然后在积累的基础上进行决策. 这种思想和王守觉教授提出的“仿生模式识别”<sup>[9]</sup>思想是一致的.

## 参 考 文 献

- 1 Vapnik V.. Nature of Statistical Learning Theory. New York: Springer-Verlag, 2000
- 2 Vapnik V.. Statistical Learning Theory. New York: John Wiley, 1998
- 3 Burges C. J. C.. A tutorial on support vector machines for pattern recognition. Data Mining and Knowledge Discovery, 1998,
- 4 Tax D., Duin R.. Support vector domain description. Pattern Recognition Letters, 1999, 20(11~13): 1191~1199
- 5 Platt J. C.. Sequential minimal optimization: A fast algorithm for training support vector machines. In: Scholkopf B., Burges C., Smola A. eds. Advances in Kernel Methods: Support Vector Machines. Cambridge, MA: MIT Press, 1998
- 6 Sha F., Saul L. K., Lee D. D.. Multiplicative updates for non-negative quadratic programming in support vector machines. Neural Information Processing System, 2002, 15. Available from: <http://www-2.cs.cmu.edu/Groups/NIPS/NIPS2002/NIPS2002proceedings/papers/AA71.html>
- 7 Kung S. Y., Diamantaras K. I., Taur J. S.. Adaptive Principal component EXtraction(APEX) and applications. IEEE Transactions on Signal Processing, 1994, 42(5): 1202~1217
- 8 Scholkopf B., Smola A., Muller K. R.. Nonlinear component analysis as a kernel eigenvalue problem. Neural Computation, 1998, 10(5): 1299~1319
- 9 Wang S. J.. Bionic (topological) pattern recognition: A new model of pattern recognition theory and its applications. Acta Electronica Sinica, 2002, 30(10): 1417~1420(in Chinese)  
(王守觉. 仿生模式识别(拓扑模式识别)——一种模式识别新模型的理论与应用. 电子学报, 2002, 30(10): 1417~1420)



**LU Cong-De**, born in 1972, Ph. D. candidate. His current research interests include statistical learning theory, kernel method, neural network, pattern recognition, digital image processing.

**ZHANG Tai-Yi**, born in 1944, professor and Ph. D. su-

2(2): 1~47

- pervisor. His current research interests include advanced mobile communications, software-defined radio, wireless broadband communications, pattern recognition, digital image processing and data compression.
- HU Jin-Yan**, born in 1973, Ph. D. candidate. Her current research interests include fractal theory and its application, audio and video signal processing, multimedia communication, etc.

## Background

The research project is supported partly by Xi'an Flying Technology company, aimed at developing new method for detecting the facial changing of driver with fatigue, including development of new theories and techniques of detection and recognition based on support vector machine or neural net-

works. The research group has developed some new algorithms on the detection of face and iris. The paper introduces multiplicative updates on support vector domain classifier for improving the computational speed of optimization.