第 38 卷 第 4 期
Vol.38 No.4

计 算 机 工 程
Computer Engineering

2012 年 2 月
February 2012

·人工智能及识别技术· 文章编号：1000—3428(2012)04—0152—03 文献标识码：A 中图分类号：TP18

线性支持向量机多类分类器几何构造方法

唐 英，李应珍

(北京科技大学机械工程学院，北京 100083)

摘 要：针对支持向量机多类分类问题，根据样本点集凸包找寻模式类间隙，通过提取模式类间隙多边形中轴线构造多类分类边界。当基本支持向量机扩展为多类分类问题时，该方法克服了 OAO 和 OAA 等传统方法存在的决策盲区和类别不平衡等缺陷。基于仿真数据集的实验结果表明，构造的分类边界在保证分类精度的同时，能够使分类空隙最大化，实现对线性可分多类数据的最优分类。

关键词：支持向量机；最优分类线；点集凸包；Delaunay 三角剖分；多边形中轴线；多类分类

Geometric Construction Method of Linear SVM Multi-class Classifier

TANG Ying, LI Ying-zhen

(School of Mechanical Engineering, University of Science and Technology Beijing, Beijing 100083, China)

【Abstract】A new method to construct multi-class classifier based on linear SVM is proposed in the paper. Its major procedures include: to form interval space polygon among point sets by subtracting operation of convex hulls, to extract polygon axes and then extend to construct the classification boundaries. The method can avoid problems like blind area in decision-making and imbalance data sets like traditional multi-class classification ways of One-Against-All(OAA) and One-Agianst-One(OAO) encounter. Simulation test results show that classification boundaries constructed by the method can realize the minimum risk and the maximum interval space among point sets, thus can be seen as an embodiment of the optimal classification lines of multi-class point sets.

【Key words】Support Vector Machine(SVM); optimal classification line; convex hull; Delaunay triangulation; polygon axis; multi-class classification

DOI: 10.3969/j.issn.1000-3428.2012.04.049

1 概述

支持向量机(Support Vector Machine, SVM)是一种基于统计学习理论的机器学习方法，具有推广能力强、维数不敏感等优良性能，广泛应用于模式识别、函数回归、故障诊断等。SVM 在本质上解决的是 2 类分类问题，而实际应用中的模式识别往往是多类分类问题。如何将 SVM 推广至多类别分类领域是机器学习的研究热点之一[1]。

有通过改变 SVM 目标函数使 SVM 本身成为多类分类器的方法，当训练样本数目相对较大时，存在运算时间长且误差较大的问题。

也有将多类分类问题转化为 2 类分类问题，如 OAA(One-Against-All)和 OAO(One-Agianst-One)等。这些方法对于多类分类问题需要依次两两分类下去，导致出现决策盲区和类别不平衡等缺陷[2]。

为了克服上述缺陷，本文根据 SVM 几何学的观点，基于最优分类面概念，提出线性 SVM 多类分类器构造方法。

2 SVM 最优分类线的几何求解法

SVM 从线性可分情况下的最优分类面发展而来。如图 1 所示的二维 2 类线性可分情况，图中实心点和空心点分别表示 2 类训练样本，H 为把 2 类没有错误地分开的分类线，$H1$、$H2$ 分别为过各类样本中离分类线最近的点且平行于分类线的直线，$H1$ 和 $H2$ 之间的距离叫做 2 类的分类间隔或分类空隙。所谓最优分类线，就是要求分类线不但能将 2 类样本正确地分开，而且要使 2 类的分类空隙最大。前者保证经验风险最小，后者使推广性的界中的置信范围最小，即真实风险最小。推广到高维空间，最优分类线就成为最优分类面。

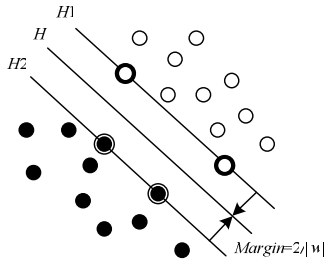
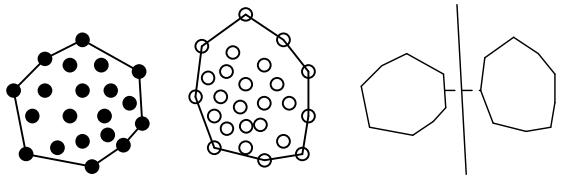


图 1 最优分类线示意图

构造最优分类面的方法有平分最近点法和最大间隔法，它们求解得到的是同一个超平面。如果把分类中的每一类都看成一个凸包集合，根据 SVM 工作原理，如图 2 所示，要将图 2(a)中的 2 类模式分开，即寻找一个超平面，该超平面为两数据类的凸包之间最近 2 个点的连线的平分线，如图 2(b)所示，其核心是寻找最大空白边缘超平面，等价于寻找 2 个相应凸包的最近点[3]。



(a)凸包 (b)最优分类线

图 2 最优分类线的几何求解法

作者简介：唐 英(1968－)，女，副教授、博士后，主研方向：人工智能，故障诊断；李应珍，硕士研究生

收稿日期：2011-08-11 **E-mail：**tangydl@public3.bta.net.cn

对于线性不可分情况，待分类样本可以通过选择适当的非线性变换$\phi(x)$映射到某个高维的特征空间，使得在目标高维空间这些样本线性可分。这个非线性映射函数$\phi(x)$称为核函数[4]。将实际问题通过非线性转化到高维特征空间，在高维空间中构造线性函数实现原来空间中的非线性，这一特殊性质能保证机器有较好的泛化能力，同时还巧妙地解决了维数灾难问题，使其算法复杂度与样本维数无关。构造线性 SVM 的空间称为判别域空间。本文研究判别域空间的多类数据最优分类线构造问题。

3　线性 SVM 多类分类器几何构造方法

c 类样本数据在二维判别域空间的分布表征为判别域平面的离散点集。如图 3 中 c=3。为了在某个空间中进行分类，通常假设同一类的各个模式在该空间中组成一个紧致集[5]。

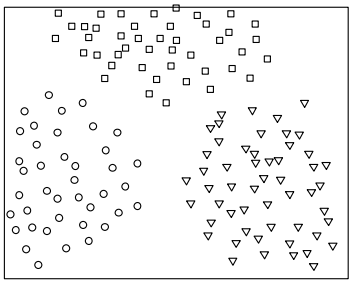


图 3　3 类样本数据的二维判别域空间分布

对于 c 类离散点集的分类，构建最优分类线即决策线的流程图如图 4 所示，包括以下 4 个主要步骤：(1)生成离散点集凸壳；(2)寻找判别域内模式类间空隙多边形；(3)提取间隙多边形中轴线；(4)延展中轴形成多类分类决策线。

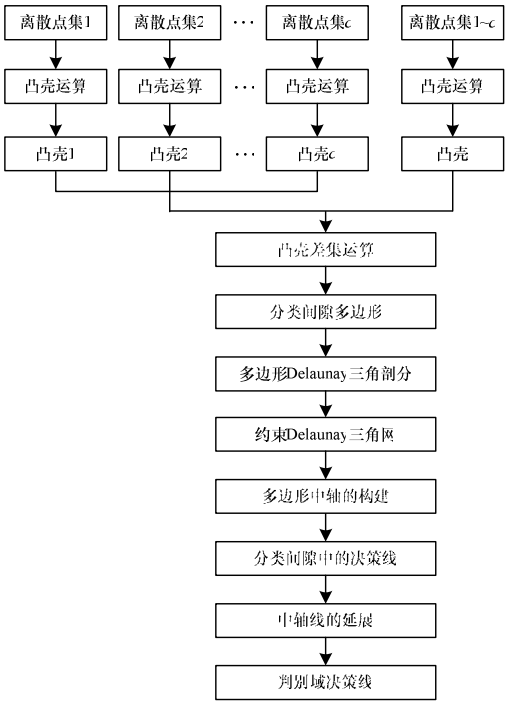
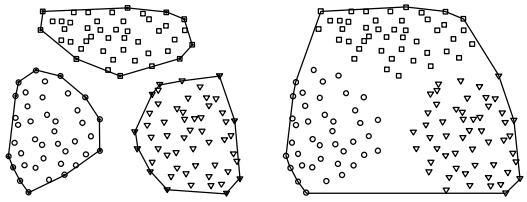


图 4　线性 SVM 多类分类线构建流程

3.1　凸壳的生成

凸壳也称最小凸包，是包含集合 S 中所有对象的最小凸集。平面点集 S 的凸壳是包含 S 中所有点的最小凸多边形，其顶点为 S 中的点。目前，计算平面点集凸壳的算法主要有卷包裹法、格雷厄姆算法、分治算法和 Z3-8 算法等[6]。本文以格雷厄姆算法为基础生成离散点集的凸壳。

图 5(a)是对图 3 的 3 个离散点集分别进行凸壳运算生成的子类凸壳，图 5(b)是混合判别域平面上所有离散点集后进行凸壳运算求得的整体凸壳。

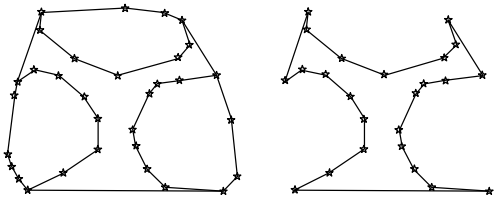


(a)各子类凸壳　　(b)整体凸壳

图 5　凸壳的生成

3.2　模式类间空隙多边形的生成

模式类间空隙多边形是通过整体凸壳与子类凸壳之间的差集计算获得的，利用了多边形差集的布尔运算概念[7]。针对图 3 的 3 类样本数据，整体凸壳与子类凸壳之间的差集计算获得的模式类间空隙多边形如图 6(b)所示。找到此分类间隔多边形，即找到了决策线构建的区域。



(a)凸壳　　(b)凸壳差集

图 6　模式类间空隙多边形的生成

3.3　间隙多边形中轴线的提取

多边形中轴也叫多边形的对称轴、骨架线或中线，具有中间性、连续性和唯一性的特点。本文正是利用中轴线的这 3 个特点，提取模式类间空隙多边形的中轴线。中轴线将模式类间空隙多边形分割成无偏倚的 c 个区域(c 为模式类的数目)时，中轴线的连续性避免了决策盲区；中轴线的中间性使每个模式类得到的分类间隙最大，体现了最优分类线的特点，避免了类别不平衡的缺陷；中轴线的唯一性保证所构成的多类分类线存在且是唯一的。本文模式类间空隙多边形中轴线的提取是在任意多边形 Delaunay 三角剖分的基础上，采用非约束边中点法的中轴构建方法实现的。

(1)任意多边形 Delaunay 三角剖分

对多边形进行中轴线的构建，首要步骤是对多边形进行 Delaunay 三角剖分，即将多边形拆分为满足 Delaunay 三角剖分原则的三角网。Delaunay 三角剖分具有唯一性、最优性、区域性等优异的特性[8]。

本文研究的任意多边形的 Delaunay 三角剖分算法，在构网时加入约束边，而且在生长法的基础上附加约束条件，从而实现对任意多边形的 Delaunay 三角剖分[9]。用于图 6(b)空隙多边形的 Delaunay 三角剖分，结果如图 7 所示。

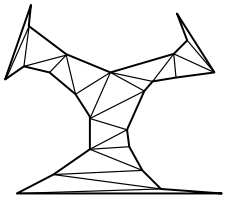


图 7　模式类间空隙多边形的 Delaunay 三角剖分

(2)多边形中轴的构建

常用的提取多边形中轴线的方法有质心法、中线中点法、

非约束边中点法和 Voronoi 图法。相比其他方法，非约束边中点法得到的中轴线过渡更柔滑，用于区域划分时更具有优越性而被本文采用。具体方法是：对给定的多边形 Delaunay 三角网，考察网络中的三角形与约束边之间的关系，可以将网络中的三角形分为 3 类：2 边为约束边的Ⅰ类三角形，只有一边为约束边的Ⅱ类三角形，3 边中没有约束边的Ⅲ类三角形。对于Ⅰ类三角形，取三角形中非约束边中点和及其所对应的三角形顶点的连线；对于Ⅱ类三角形，取三角形中两非约束边中点的连线；对于Ⅲ类三角形，取三角形质心和其 3 边中点的连线。将这些连线按照其拓扑关系连接成一个整体，即为所求中轴线。

应用于图 7 的 Delaunay 三角剖分后的多边形进行中轴构建，结果如图 8 所示。该中轴线将多边形分为 3 个部分，即是将 3 类模式之间相关联的公共区域划分成了 3 个部分，中轴起到了决策线的作用。但是，此时的决策线只是各类之间分类空隙内的。实现整个判别域的划分，需延展中轴，将整个判别域划分成互不连通的 3 个区域。

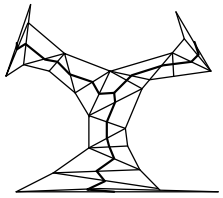


图 8 多边形中轴线的提取

3.4 中轴线在模式空间的延展

中轴线在模式空间延展时，延展线以中轴线的端点为起点，还需确定延展线方向。方法如图 9 所示，在 II 类三角形 △ABC 中，非约束边 AB 的 2 个端点 A 和 B 分别为不同模式类 $\omega 1$ 类与 $\omega 2$ 类的点，AB 的中点为中轴线的端点，则以该点为起点，沿平行于角 $\angle ACB$ 平分线的方向延伸中轴线至判别域边界。

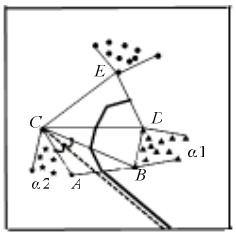


图 9 延展线构建示意图

采用上述方法，延伸图 8 中的中轴线至判别域边界，如图 10 所示。将判别决策线抽取出来还原到判别平面中，如图 11 所示。可以看到，该判别域决策线具有封闭性及连续性的特点。

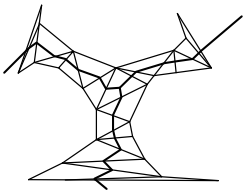


图 10 中轴线的延展

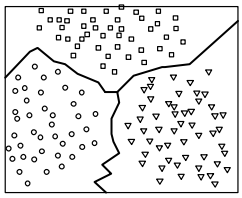


图 11 样本数据的分类线

4 结束语

本文基于 SVM 几何学的观点，提出了一种多类分类新方法，即通过样本点集凸包运算，找寻模式类间隙，提取模式类间隙多边形中轴，中轴线在模式空间的延展构成多类分类边界。

基于仿真数据集的实验表明，该方法所构造的分类边界在保证分类精度的同时，能够使分类空隙最大化，使得 SVM 能够实现对线性可分多类数据的最优分类。

本文提出的方法适用于模式类在二维判别域空间满足可分要求的多类分类问题。

参考文献

[1] 祁亨年. 支持向量机及其应用研究综述[J]. 计算机工程, 2004, 30(10): 6-9.

[2] 应自炉, 李景文, 张有为. 基于融合的多类支持向量机[J]. 计算机工程, 2009, 35(19): 187-188, 191.

[3] Theodoridis S, Koutroumbas K. Pattern Recognition[M]. 李晶皎, 王爱侠, 张广渊, 译. 北京: 机械工业出版社, 2006.

[4] 张 铃. 基于核函数的 SVM 机与三层前向神经网络的关系[J]. 计算机学报, 2002, 25(7): 1-5.

[5] 边肇祺, 张学工. 模式识别[M]. 北京: 清华大学出版社, 2004.

[6] 周文科. 一种简单多边形凸包的快速算法及程序设计[J]. 广州大学学报: 自然科学版, 2003, 2(6): 545-547.

[7] 魏许青. 计算多边形交集、并集面积的算法[J]. 计算机工程与科学, 2007, 29(12): 85-86.

[8] 刘建新, 卢新明, 岳 昊. 简单多边形快速 Delaunay 三角剖分算法[J]. 计算机技术与发展, 2006, 16(7): 126-128.

[9] 唐 英, 李应珍. 任意多边形的 Delaunay 三角剖分约束生长算法[J]. 计算机应用研究, 2009, 26(11): 134-135.

编辑 顾逸斐

(上接第 151 页)

参考文献

[1] 李立涛, 杨 旭, 李顺利. 针对非合作目标的中距离相对导航方法[J]. 吉林大学学报: 工学版, 2008, 38(4): 986-990.

[2] Ma Weihua, Luo Jianjun, Wang Mingming, et al. Space Relative Navigation Filter Based on Board Radar Observation[C]//Proc. of the 2nd International Congress on Image and Signal Processing. [S. l.]: IEEE Press, 2009: 1-5.

[3] Cao Zhigao, Gao Yuan, Cheng Hongwei. EKF Tracking of Low Earth Orbit Satellites Based on Bearings——Only Data of A Single Space-based Platform[J]. Journal of Spacecraft TT&C Technology, 2007, 26(6): 70-75.

[4] Li Qiang, Guo Fucheng, Li Jun, et al. Research of Satellite-to-Satellite Passive Tracking Using Bearings-only Measurements[C]// Proc. of International Conference on Radar. [S. l.]: IEEE Press, 2006.

[5] Yang Guosheng, Dou Lihu, Chen Jie, et al. Synergy Decision in the Multi-target Tracking Based on IRST and Intermittent-working Radar[J]. Information Fusion, 2001, 2(4): 243-250.

[6] Julier S J, Uhlmann J K. Unscented Filtering Nonlinear Estimation[J]. Proc. of the IEEE Aerospace and Electronic Systems, 2004, 92(3): 401-422.

[7] 郭晓松, 李奕芃, 郭君斌. 贝叶斯目标跟踪方法的研究[J]. 计算机工程, 2009, 35(12): 137-139.

编辑 顾逸斐