第25卷 第4期
2015年4月
计算机技术与发展
COMPUTER TECHNOLOGY AND DEVELOPMENT
Vol. 25 No. 4
Apr. 2015

基于类内离散度的最小二乘支持向量机

薛 松,李 雷
(南京邮电大学 理学院,江苏 南京 210023)

摘 要:支持向量机是一种很流行的机器学习方法,在许多领域都有了广泛的使用。传统的支持向量机模型是寻求类之间的间隔最大化,而忽视了一个重要的信息—样本的类内结构,类内离散度。文中将 Fisher 判别分析里面的类内离散度引入到最小二乘支持向量机中,提出了基于类内离散度的最小二乘支持向量机模型。并通过核函数将样本映射到高维特征空间,在特征空间中进行样本分类。基于 UCI 数据库的数据集实验测试表明,基于类内离散度的最小二乘支持向量机提高了分类的准确度。

关键词:类内离散度;分类;最小二乘支持向量机;核方法
中图分类号:TP31 **文献标识码**:A **文章编号**:1673-629X(2015)04-0071-04
doi:10.3969/j.issn.1673-629X.2015.04.017

Least Squares Support Vector Machine Based on Within-class Scatter

XUE Song,LI Lei
(College of Science,Nanjing University of Posts and Telecommunications,Nanjing 210023,China)

Abstract:Support vector machine (SVM) is a popular machine learning technique,and it has been widely used in many real-world applications. Traditional SVMs aims at seeking the hyperplane that maximizes the margin and ignores an important prior knowledge,the within-class structure. It formulates a Least Square Support Vector Machine (LSSVM) based on within-class scatter for binary classification,which incorporates minimum within-class scatter in Fisher Discriminant Analysis (FDA) into LSSVM. The sample points are mapped to a high dimensional feature space where the samples are classified by using the kernel method. Experiments on four benchmarking datasets based on UCI show that the proposed WCS-LSSVM can improve the classification accuracy.

Key words:within-class scatter;classification;least squares support vector machine;kernel methods

0 引 言

支持向量机(SVM)[1]是 Vapnik 等于1995年提出的基于结构化风险的机器学习方法,它不仅有着统计学习理论的坚实基础,而且有直观的几何解释和完美的数学形式。通过将样本点映射到高维的特征空间,然后寻找一个最优分化超平面来使两类点之间的间隔最大化。因为支持向量机的全局最优、泛化能力好等特点,使得 SVM 已经应用到很多领域,比如:文本分类[2]、人脸识别[3-4]、生物信息学、语音识别、遥感图像分析以及时间序列预测等。

在过去的十几年内,研究者提出了一些新的支持向量机模型。比如 ν - SVM [5]、最小二乘支持向量机(LSSVM)[6]、模糊支持向量机(FSVM)[7]、拉格朗日支持向量机(LSVM)[8]、对支持向量机(TSVM)[9]等。在文献[10]中,受到 Fisher 判别分析法(FDA)的启发,A Tefas 提出了最小类内协方差支持向量机(minimum class variance support vector machine),他将 FDA 里面的类内离散度(within-class scatter)引入到支持向量机模型中并得到了具有鲁棒性的结果。An Wenjuan 和 Liang Mangui 在文献[11]中,通过将 FDA 里面的类内离散度与模糊支持向量机模型相结合,使得新的支持向量机模型提高了泛化性能和分类准确度。

文中将类内离散度与最小二乘支持向量机模型相结合,得到新的模型,称之为基于类内离散度的最小二乘支持向量机[12]。在 UCI 数据库里的四组数据集实验表明,新的算法能够提高分类准确度。

1 理论基础

1.1 最小二乘支持向量机

考虑一个具有 n 个训练点的训练集 $\{\boldsymbol{x}_i,y_i\}_{i=1}^{n}$,

收稿日期:2014-06-15 **修回日期**:2014-09-22 **网络出版时间**:2015-02-23
基金项目:国家自然科学基金资助项目(61373137)
作者简介:薛 松(1989-),男,硕士研究生,研究方向为机器学习方法;李 雷,教授,研究方向为智能信号处理、非线性分析与计算智能。
网络出版地址:http://www.cnki.net/kcms/detail/61.1450.TP.20150223.1252.053.html

这里的 $\boldsymbol{x}_i \in R^m$ 是输入点, $y_i \in \{+1, -1\}$ 是对应的输入点的输出。最小二乘支持向量机是一个凸优化问题,它的模型如下:

$$\min_{w,b,\xi} \frac{1}{2}\boldsymbol{\omega}^{\mathrm{T}}\boldsymbol{\omega} + \frac{C}{2}\sum_{i=1}^{n}\xi_i^{\ 2} \tag{1}$$

$$\text{s. t. } y_i(\boldsymbol{\omega} \cdot \boldsymbol{\varphi}(x_i) + b) = 1 - \xi_i, i = 1,2,\cdots,n \tag{2}$$

其中, $\boldsymbol{\omega} \in R^m$; b 是一个偏置量; $\boldsymbol{\varphi}(x_i)$ 是非线性映射,它将训练点映射到高维特征空间;松弛变量 $\xi_i \geqslant 0$ 是用来软化限制条件 $y_i(\boldsymbol{\omega} \cdot \boldsymbol{\varphi}(x_i) + b) = 1$,即允许一些点不满足这个限制条件;参数 $C > 0$ 是一个惩罚因子,它对引入的松弛变量进行惩罚,使得松弛变量不能太大, C 的值体现了对最大化间隔和松弛变量之间的权衡。

最优化模型(1)和(2)的对偶问题为:

$$\max_{\alpha} \sum_{i=1}^{n}\alpha_i - \frac{1}{2}\sum_{i=1}^{n}\sum_{j=1}^{n}\alpha_i\alpha_j y_i y_j\left(K(x_i,x_j) + \frac{\delta_{ij}}{C}\right) \tag{3}$$

$$\text{s. t. } \sum_{i=1}^{n}\alpha_i y_i = 0 \tag{4}$$

引入记号

$$\delta_{ij} = \begin{cases} 1 & i = j \\ 0 & i \neq j \end{cases}$$

这里的 α_i 为拉格朗日乘子, $K(x_i,x_j)$ 为核函数。常用的核函数[13-14]有线性核函数、高斯核函数、Sigmoid 核函数。它们分别定义如下:

$$K(x_i,x_j) = ((x_i \cdot x_j) + b)^d$$

$$K(x_i,x_j) = \exp(-\|x_i - x_j\|^2/2\sigma^2)$$

$$K(x_i,x_j) = \tanh(\alpha(x_i,x_j) + \beta)$$

最优化问题(3)和(4)可以由二次规划(QP)求解。求得最优解 α 和 b 后,测试样本点 x 的决策值 y 可以由式(5)求得。

$$y(x) = \mathrm{sgn}\left[\sum_{i=1}^{n} y_i\alpha_i K(x,x_i) + b\right] \tag{5}$$

1.2 Fisher 判别分析

Fisher 判别分析法是模式识别中经常用到的算法,定义一个正类点样本集 $\chi_+ = \{x_1^1, x_2^1, \cdots, x_{n^+}^1\}$ 和一个负类点样本集 $\chi_- = \{x_1^2, x_2^2, \cdots, x_{n^-}^2\}$,它的主要思想是寻找一个合适的 $\boldsymbol{\omega}$ 使得类间离散度最大化和类内离散度最小化。它的数学表达如下:

$$\max J(\boldsymbol{\omega}) = \frac{\boldsymbol{\omega}\boldsymbol{S}_{\mathrm{b}}\boldsymbol{\omega}}{\boldsymbol{\omega}\boldsymbol{S}_{\mathrm{w}}\boldsymbol{\omega}} \tag{6}$$

这里的 $\boldsymbol{S}_{\mathrm{b}}$ 和 $\boldsymbol{S}_{\mathrm{w}}$ 分别为类间离散矩阵和类内离散矩阵。它们分别定义如下:

$$\boldsymbol{S}_{\mathrm{b}} = (m_+ - m_-)(m_+ - m_-)^{\mathrm{T}} \tag{7}$$

$$\boldsymbol{S}_{\mathrm{w}} = \frac{n_+}{n}\sum_{x\in\chi_1}(x - m_+)(x - m_+)^{\mathrm{T}} + \frac{n_-}{n}\sum_{x\in\chi_2}(x - m_-)(x - m_-)^{\mathrm{T}} \tag{8}$$

其中, m_+、m_- 分别为正负样本点的均值。

Fisher 判别分析法直观的几何解释就是使不同类之间的间隔越大越好,相同类之间的点越凝聚越好。

和支持向量机一样,通过引入核函数,可以将样本映入到高维的特征空间,然后在特征空间上进行 Fisher 判别分析。核 Fisher 判别分析(KFDA)[15]是非线性统计分析中的一种很重要的算法。

2 基于类内离散度的最小二乘支持向量机

设一个具有 n 个训练点的训练集 $\{x_i, y_i\}_{i=1}^n$,这里的 $x_i \in R^m$ 是输入点, $y_i \in \{+1, -1\}$,正类点(即 $y_i = +1$)的个数为 n_+,负类点个数为 n_-,且 $n_+ + n_- = n$。

现将 Fisher 判别法里面的类内离散度 $\boldsymbol{\omega}^{\mathrm{T}}\boldsymbol{S}_{\mathrm{w}}\boldsymbol{\omega}$ 引入到最小二乘支持向量机中,得到新的模型,目标是要最大化类间的间隔且最小化类内离散度。它的数学模型如下:

$$\min \frac{1}{2}\boldsymbol{\omega}^{\mathrm{T}}\boldsymbol{\omega} + \frac{1}{2}\beta\boldsymbol{\omega}^{\mathrm{T}}\boldsymbol{S}_{\mathrm{w}}\boldsymbol{\omega} + \frac{C}{2}\sum_{i=1}^{n}\xi_i^2 \tag{9}$$

$$\text{s. t. } y_i(\boldsymbol{\omega}^{\mathrm{T}}\boldsymbol{x}_i + b) = 1 - \xi_i \tag{10}$$

其中 $\beta \geqslant 0$,是新加入项 $\boldsymbol{\omega}^{\mathrm{T}}\boldsymbol{S}_{\mathrm{w}}\boldsymbol{\omega}$ 的权重。

但在实际的应用中,接触的大多数数据都是非线性的,因此一个线性分类器处理这些问题很乏力。为了解决这一问题,引入核函数[16],通过一个非线性映射 $\varphi(x)$ 将样本映入新的空间 H,这里并不要知道明确的映射 $\varphi(x)$,而是最终通过核函数 $K(x_i,x_j) = (\varphi(x_i) \bullet \varphi(x_j))$ 来计算。在核空间的模型如下:

$$\min \frac{1}{2}\boldsymbol{\omega}^{\mathrm{T}}\boldsymbol{\omega} + \frac{1}{2}\beta\boldsymbol{\omega}^{\mathrm{T}}\boldsymbol{S}_{\mathrm{w}}^{\varphi}\boldsymbol{\omega} + \frac{C}{2}\sum_{i=1}^{n}\xi_i^2 \tag{11}$$

$$\text{s. t. } y_i(\boldsymbol{\omega}^{\mathrm{T}}\boldsymbol{\varphi}(x_i) + b) = 1 - \xi_i \tag{12}$$

由核再生空间理论(theory of reproducing kernels)所知,解 $\boldsymbol{\omega} \in H$ 一定在由样本所扩展的空间中,所以 $\boldsymbol{\omega}$ 可由所有样本点线性表示,因此 $\boldsymbol{\omega}$ 可由下式表示:

$$\boldsymbol{\omega} = \sum_{i=1}^{n}\alpha_i\varphi(x_i)$$

式(11)中的项 $\frac{1}{2}\boldsymbol{\omega}^{\mathrm{T}}\boldsymbol{\omega}$ 等价于:

$$\frac{1}{2}\boldsymbol{\omega}^{\mathrm{T}}\boldsymbol{\omega} = \frac{1}{2}\sum_{i=1}^{n}\sum_{j=1}^{n}\alpha_i\alpha_j K(x_i,x_j) = \frac{1}{2}\boldsymbol{\alpha}^{\mathrm{T}}\boldsymbol{G}\boldsymbol{\alpha}$$

这里 $\boldsymbol{G}$ 为 Gram 矩阵, $\boldsymbol{G} = \begin{bmatrix} K(x_1,x_1) & K(x_1,x_2) & \cdots & K(x_1,x_n) \\ K(x_2,x_1) & K(x_2,x_2) & \cdots & K(x_2,x_n) \\ \vdots & \vdots & \cdots & \vdots \\ K(x_n,x_1) & K(x_n,x_2) & \cdots & K(x_n,x_n) \end{bmatrix}$, $\boldsymbol{\alpha}$ 为列向量

$\boldsymbol{\alpha} = (\alpha_1, \alpha_2, \cdots, \alpha_n)^{\mathrm{T}}$。

式(11)中的项可以写为:

$$\frac{1}{2}\boldsymbol{\omega}^{\mathrm{T}}\boldsymbol{S}_{\mathrm{w}}^{\varphi}\boldsymbol{\omega} = \frac{1}{2}\boldsymbol{\omega}^{\mathrm{T}}[\frac{n_+}{n}\sum_{x \in \chi_+}(\varphi(x) - m_+^{\varphi})(\varphi(x) - m_+^{\varphi})^{\mathrm{T}} + \frac{n_-}{n}\sum_{x \in \chi_-}(\varphi(x) - m_-^{\varphi})(\varphi(x) - m_-^{\varphi})^{\mathrm{T}}]\boldsymbol{\omega}$$

现在将其中一项展开:

$$\boldsymbol{\omega}^{\mathrm{T}}\sum_{x \in \chi_+}(\varphi(x) - m_+^{\varphi})(\varphi(x) - m_+^{\varphi})^{\mathrm{T}}\boldsymbol{\omega} = \boldsymbol{\omega}^{\mathrm{T}}\sum_{x \in \chi_+}(\varphi(x)\varphi(x) - m_+^{\varphi} m_+^{\varphi\,\mathrm{T}})\boldsymbol{\omega} =$$

$$\sum_{x \in \chi_+}\sum_{i=1}^{n}\alpha_i K(x_i, x)\sum_{j=1}^{n}\alpha_j K(x, x_j) - \frac{1}{n_+}\sum_{s=1}^{n}\alpha_s\sum_{t=1}^{n}\alpha_t\sum_{p=1}^{n_+}K(x_s, x_p^+)\sum_{q=1}^{n_+}K(x_q^+, x_t) = (1 - \frac{1}{n_+})\boldsymbol{\alpha}^{\mathrm{T}}\boldsymbol{K}_1\boldsymbol{K}_1^{\mathrm{T}}\boldsymbol{\alpha}$$

这里的 $\boldsymbol{K}_1$ 为一个 $n \times n_+$ 的矩阵: $\boldsymbol{K}_1 = \begin{bmatrix} K(x_1, x_1^+) & K(x_1, x_2^+) & \cdots & K(x_1, x_{n_+}^+) \\ K(x_2, x_1^+) & K(x_2, x_2^+) & \cdots & K(x_2, x_{n_+}^+) \\ \vdots & \vdots & \cdots & \vdots \\ K(x_n, x_1^+) & K(x_n, x_2^+) & \cdots & K(x_n, x_{n_+}^+) \end{bmatrix}$。

类似的可以得到:

$$\boldsymbol{\omega}^{\mathrm{T}}\sum_{x \in \chi_-}(\varphi(x) - m_-^{\varphi})(\varphi(x) - m_-^{\varphi})^{\mathrm{T}}\boldsymbol{\omega} = (1 - \frac{1}{n_-})\boldsymbol{\alpha}^{\mathrm{T}}\boldsymbol{K}_2\boldsymbol{K}_2^{\mathrm{T}}\boldsymbol{\alpha}$$

因此

$$\frac{1}{2}\boldsymbol{\omega}^{\mathrm{T}}\boldsymbol{S}_{\mathrm{w}}^{\varphi}\boldsymbol{\omega} = \frac{1}{2}[\frac{n_+ - 1}{n}\boldsymbol{\alpha}^{\mathrm{T}}\boldsymbol{K}_1\boldsymbol{K}_1^{\mathrm{T}}\boldsymbol{\alpha} + \frac{n_- - 1}{n}\boldsymbol{\alpha}^{\mathrm{T}}\boldsymbol{K}_2\boldsymbol{K}_2^{\mathrm{T}}\boldsymbol{\alpha}]$$

令 $\boldsymbol{Q} = \boldsymbol{G} + \beta(\frac{n_+ - 1}{n}\boldsymbol{K}_1\boldsymbol{K}_1^{\mathrm{T}} + \frac{n_- - 1}{n}\boldsymbol{K}_2\boldsymbol{K}_2^{\mathrm{T}})$,再将 $\boldsymbol{Q}$ 代入到式(11)中,得到:

$$\min \frac{1}{2}\boldsymbol{\alpha}^{\mathrm{T}}\boldsymbol{Q}\boldsymbol{\alpha} + \frac{C}{2}\sum_{i=1}^{n}\xi_i^2 \quad \text{s.t. } y_i(\sum_{j=1}^{n}\alpha_j K(x_j, x_i) + b) = 1 - \xi_i \tag{13}$$

利用拉格朗日乘子法可以解决式(13)的最优化问题,引入拉格朗日乘子 h_i:

$$L(\alpha, b, \xi, h) = \frac{1}{2}\boldsymbol{\alpha}^{\mathrm{T}}\boldsymbol{Q}\boldsymbol{\alpha} + \frac{C}{2}\sum_{i=1}^{n}\xi_i^2 - \sum_{i=1}^{n}h_i[y_i(\sum_{j=1}^{n}\alpha_j K(x_j, x_i) + b) - 1 + \xi_i] = \frac{1}{2}\boldsymbol{\alpha}^{\mathrm{T}}\boldsymbol{Q}\boldsymbol{\alpha} + \frac{C}{2}\boldsymbol{\xi}^{\mathrm{T}}\boldsymbol{\xi} - \boldsymbol{h}^{\mathrm{T}}(\boldsymbol{Y}\boldsymbol{G}^{\mathrm{T}}\boldsymbol{\alpha} + b\boldsymbol{y} - \boldsymbol{e} + \boldsymbol{\xi}) \tag{14}$$

分别对式(14)中的 α, b, ξ_i, h_i 求偏导,得到:

$$\nabla_{\alpha}L = \boldsymbol{Q}\boldsymbol{\alpha} - \sum_{i=1}^{n}h_i y_i\sum_{j=1}^{n}K(x_j, x_i) = 0 \Rightarrow \boldsymbol{\alpha} = \boldsymbol{Q}^{-1}\boldsymbol{G}\boldsymbol{Y}\boldsymbol{h} \tag{15}$$

$$\nabla_{\xi_i}L = \boldsymbol{C}\xi_i - h_i = 0 \Rightarrow \boldsymbol{C}\boldsymbol{\xi} = \boldsymbol{h} \tag{16}$$

$$\nabla_b L = \sum_{i=1}^{n}h_i y_i = 0 \Rightarrow \boldsymbol{y}^{\mathrm{T}}\boldsymbol{h} = 0 \tag{17}$$

$$\nabla_{h_i}L = y_i(\sum_{j=1}^{n}\alpha_j K(x_j, x_i) + b) - 1 + \xi_i = 0 \Rightarrow \boldsymbol{Y}\boldsymbol{G}^{\mathrm{T}}\boldsymbol{\alpha} + b\boldsymbol{y} = \boldsymbol{e} - \boldsymbol{\xi} \tag{18}$$

(这里的 $\boldsymbol{\xi} = (\xi_1, \xi_2, \cdots, \xi_n)^{\mathrm{T}}$,$\boldsymbol{h} = (h_1, h_2, \cdots, h_n)^{\mathrm{T}}$,$\boldsymbol{Y} = \mathrm{diag}(y_1, y_2, \cdots, y_n)$,$\boldsymbol{y} = (y_1, y_2, \cdots, y_n)\mathrm{T}$,$\boldsymbol{e}$ 是一个 $n \times 1$ 的全 1 向量)。

将 $\boldsymbol{\alpha}$ 用 $\boldsymbol{\alpha} = \boldsymbol{Q}^{-1}\boldsymbol{G}\boldsymbol{Y}\boldsymbol{h}$ 代入式(13)中,得到一个关于以 $\boldsymbol{h}$ 为变量的最优化问题:

$$\min \frac{1}{2}\boldsymbol{h}^{\mathrm{T}}[\boldsymbol{Y}\boldsymbol{G}(\boldsymbol{Q}^{-1})^{\mathrm{T}}\boldsymbol{G}\boldsymbol{Y} + \frac{\boldsymbol{I}}{C}]\boldsymbol{h} - \boldsymbol{e}^{\mathrm{T}}\boldsymbol{h} \quad \text{s.t. } \boldsymbol{y}^{\mathrm{T}}\boldsymbol{h} = 0 \tag{19}$$

式中,$\boldsymbol{Y}\boldsymbol{G}(\boldsymbol{Q}^{-1})^{\mathrm{T}}\boldsymbol{G}\boldsymbol{Y} + \frac{\boldsymbol{I}}{C}$ 为对称半正定矩阵。

证明:因为 $\boldsymbol{G}$ 是 Gram 矩阵,所以 $\boldsymbol{G}$ 是对称半正定的。

$\boldsymbol{Q} = \boldsymbol{G} + \beta(\frac{n_+ - 1}{n}\boldsymbol{K}_1\boldsymbol{K}_1^{\mathrm{T}} + \frac{n_- - 1}{n}\boldsymbol{K}_2\boldsymbol{K}_2^{\mathrm{T}})$,$\boldsymbol{Q}$ 的转置 $\boldsymbol{Q}^{\mathrm{T}} = \boldsymbol{Q}$,所以 $\boldsymbol{Q}$ 是对称的,对于任意一个非零向量 $\boldsymbol{a} = (a_1, a_2, \cdots, a_n)$,有

$$\boldsymbol{a}^{\mathrm{T}}\boldsymbol{Q}\boldsymbol{a} = \boldsymbol{a}^{\mathrm{T}}\boldsymbol{G}\boldsymbol{a} + \beta\boldsymbol{a}^{\mathrm{T}}(\frac{n_+ - 1}{n}\boldsymbol{K}_1\boldsymbol{K}_1^{\mathrm{T}} + \frac{n_- - 1}{n}\boldsymbol{K}_2\boldsymbol{K}_2^{\mathrm{T}})\boldsymbol{a} = \boldsymbol{a}^{\mathrm{T}}\boldsymbol{G}\boldsymbol{a} + \beta\boldsymbol{b}^{\mathrm{T}}\boldsymbol{S}_{\mathrm{w}}^{\varphi}b$$

式中,$\boldsymbol{b} = \sum_{i=1}^{n}a_i\boldsymbol{\varphi}(x_i)$。因为 $\boldsymbol{S}_{\mathrm{w}}^{\varphi}$ 是类内离散矩阵,所以 $\boldsymbol{S}_{\mathrm{w}}^{\varphi}$ 是半正定的,因此 $\boldsymbol{a}^{\mathrm{T}}\boldsymbol{Q}\boldsymbol{a} \geqslant 0$,即 $\boldsymbol{Q}$ 是半正定的。所以 $\boldsymbol{Q}$ 是对称半正定的,又因为 $\boldsymbol{Y}$ 为对角矩阵,所以 $\boldsymbol{Y}\boldsymbol{G}(\boldsymbol{Q}^{-1})^{\mathrm{T}}\boldsymbol{G}\boldsymbol{Y}$ 是对称半正定的。$\boldsymbol{I}$ 是 $n \times n$ 单位矩阵,综上 $\boldsymbol{Y}\boldsymbol{G}(\boldsymbol{Q}^{-1})^{\mathrm{T}}\boldsymbol{G}\boldsymbol{Y} + \frac{\boldsymbol{I}}{C}$ 是一个对称半正定矩阵。

问题(19)为凸二次优化问题。由此求解 $\boldsymbol{h}^*$,然后 $\boldsymbol{\alpha}$ 由 $\boldsymbol{\alpha}^* = \boldsymbol{Q}^{-1}\boldsymbol{G}\boldsymbol{Y}\boldsymbol{h}^*$ 求得。偏置 b 由(18)求解:

$$b^* = \frac{1}{n}\sum_{i=1}^{n}[(1 - \frac{h_i^*}{C})y_i - \sum_{j=1}^{n}\alpha_j^* K(x_j, x_i)]$$

最终对于一个输入 x,它的决策函数为:

$$y(x) = \mathrm{sgn}[\sum_{i=1}^{n}\alpha_i^* K(x_i, x) + b^*]$$

3　实　验

为了验证所提出的新模型的分类效果,选用了 Ripley 数据集和 UCI 数据库上的三组数据集,Statlog(Australian Credit Approval),Statlog(German Credit Data),Statlog(Heart),作为实验数据。从中随机选取一部分点作为训练样本点,余下的点作为测试样本点。所有实验都是在个人电脑上的 Matlab2009 上进行的。

实验得到的分类的准确度如表 1 所示。

表 1　分类准确度

数据集	SVM 准确度/%	LSSVM 准确度/%	WCS-LSSVM 准确度/%
Ripley	90.80	90.70	91.40
Australian	85.50	85.00	86.52
German	72.71	75.26	75.26
heart	83.89	84.44	85.00

实验结果表明,新提出的算法比支持向量机和最小二乘支持向量机的分类准确度要高。最小二乘支持向量机可以视为基于类内离散度的最小二乘支持向量机模型的一个特例(当 $\beta = 0$ 时)。

4　结束语

文中通过将类内离散度和最小二乘支持向量机结合得到了新的模型,并给出了它的数学模型。利用新的模型进行分类试验,结果表明分类准确度得到了提高。当然类内离散度矩阵有一个缺点,就是当样本个数少于样本维数时,类内离散度矩阵奇异,但是在绝大数情况下样本数量是要大于样本维数的,另外还可以进行降维处理。文中实验研究也仅局限在仿真和对比,在应用方面还需努力。

参考文献:

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

[2] 刘　文,吴　陈. 一种新的中文文本分类算法—One Class SVM-KNN 算法[J]. 计算机技术与发展,2012,22(5):83-86.

[3] 王　辉. 主成分分析及支持向量机在人脸识别中的应用[J]. 计算机技术与发展,2006,16(8):24-26.

[4] 雷　虎,樊泽明. 基于 SVM 和 D-S 理论的三维人脸识别[J]. 计算机技术与发展,2014,24(6):75-78.

[5] Scholkopf B,Smola A J,Williamson R C,et al. New support vector algorithms[J]. Neural Computation,2000,12(5):1207-1245.

[6] Suykens J A K,Vandewalle J. Least squares support vector machine classifiers[J]. Neural Processing Letters,1999,9(3):293-300.

[7] Lin Chunfu,Wang Shengde. Fuzzy support vector machines[J]. IEEE Transactions on Neural Network,2002,13(2):464-471.

[8] Mangasarian O L, Musicant D R. Lagrangian support vector machines[J]. Journal of Machine Learning Research,2001,1:161-177.

[9] 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.

[10] Zafeiriou S,Tefas A,Pitas I. Minimum class variance support vector machines[J]. IEEE Transactions on Image Processing,2007,16(10):2551-2564.

[11] An W J,Liang M G. Fuzzy support vector machine based on within-class scatter for classification problems with outliers or noises[J]. Neurocomputing,2013,110:101-110.

[12] 阎　辉,张学工,李衍达. 支持向量机与最小二乘法的关系研究[J]. 清华大学学报:自然科学版,2001,41(9):77-80.

[13] 汪廷华,陈峻婷. 核函数的选择研究综述[J]. 计算机工程与设计,2012,33(3):1181-1186.

[14] 郭丽娟,孙世宇,段修生. 支持向量机及核函数研究[J]. 科学技术与工程,2008,8(2):487-490.

[15] Mika S,Ratsch G,Weston J,et al. Fisher discriminant analysis with kernels[C]//Proceedings of the 1999 IEEE signal processing society workshop on neural networks for signal processing IX. Madison,WI:IEEE,1999:41-48.

[16] 邬　啸,魏　延,吴　瑕. 基于混合核函数的支持向量机[J]. 重庆理工大学学报:自然科学,2011,25(10):66-70.

(上接第 70 页)

models with testing coverage[J]. European Journal of Operational Research,2003,145(2):445-454.

[12] Huang C Y,Kuo S Y,Lyu M R. An assessment of testing-effort dependent software reliability growth models[J]. IEEE Transactions on Reliablility,2007,56(2):198-211.

[13] Mallick S,Kushwaha D S. LWSDM:layered Web service discovery mechanism[J]. Advanced in Information Sciences and Service Sciences,2010,2(3):25-31.

[14] Canfora G,Penta M D,Esposito R,et al. A framework for QoS-aware binding and re-binding of composite Web service[J]. Journal of Systems and Software,2008,81(10):1754-1769.

[15] 谢景燕,安金霞,朱纪洪. 考虑不完美排错情况的 NHPP 类软件可靠性增长模型[J]. 软件学报,2010,21(5):942-949.

[16] Dai Yuanshun,Yang Bo,Dongarra J. Cloud service reliability modeling and analysis[EB/OL]. 2009. http://www.netlib.org/utk/people/JackDongarra/PAPERS/Cloud-Shaun-Jack.pdf.

基于类内离散度的最小二乘支持向量机

作者：　薛松，　李雷，　XUE Song，　LI Lei
作者单位：　南京邮电大学 理学院,江苏 南京,210023
刊名：　计算机技术与发展ISTIC
英文刊名：　Computer Technology and Development
年，卷(期)：　2015(4)