

# 模糊聚类光滑支持向量机

刘三阳, 吴 德<sup>†</sup>

(西安电子科技大学 计算机学院, 西安 710071)

**摘 要:** 为了提高光滑支持向量机的分类速度和精度, 构造一种模糊聚类光滑支持向量机 (FCSSVM). 运用模糊聚类将训练数据分解为若干子簇, 通过引入熵函数近似松弛向量的加函数, 并利用最优解处权重向量的表达式导出精确光滑模型; 定义测试样本的最近邻子空间, 以选择性集成策略组合若干近邻子空间中的分类决策函数. 数值实验表明, FCSSVM 的分类精度高, 迭代次数少, 鲁棒性好, 分类时间短.

**关键词:** 支持向量机; 模糊聚类; 熵函数; 光滑; 最近邻子空间; 选择性集成

**中图分类号:** TP273      **文献标志码:** A

## Fuzzy clustering smooth support vector machine

LIU San-yang, WU De<sup>†</sup>

(School of Computer Sciences, Xidian University, Xi'an 710071, China)

**Abstract:** To accelerate the training accuracy and speed of the smooth support vector machine, the fuzzy clustering smooth support vector machine (FCSSVM) is proposed. The fuzzy clustering means (FCM) is used to decompose the training data into several sub-clusters, the entropy function is introduced to approximate the plus function of the slack vector, and an exact smooth model is derived by using the expression of the weight vector at the optimal solution. The nearest neighbor subspaces of the test sample are defined, and the selective ensemble strategy is constructed to combine the classification results of several nearest neighbor subspaces. Numerical experiments demonstrate that FCSSVM has high accuracy, little iteration, good robustness and low training time.

**Keywords:** support vector machine; fuzzy clustering; entropy function; smooth; nearest neighbor subspaces; selective ensemble

## 0 引 言

光滑支持向量机 (SSVM) 的目标函数光滑可微, 便于引入高效的无约束优化算法来求解<sup>[1]</sup>, 从而吸引了许多学者进行研究, 如采用不同的近似函数构造多种光滑模型<sup>[2-4]</sup>, 或推广至回归、模糊建模等领域<sup>[5-6]</sup>. 这些研究均基于 Mangasarian 的光滑模型, 故存在 3 个问题: 1) 采用粗糙近似技术, 在核空间中以拉格朗日乘子向量的 2-范数近似代替权重向量的 2-范数; 2) 训练时间较长且内存占用较大, 不适用于大规模样本集; 3) 鲁棒性较差, 分类精度依赖于核参数的选择. 文献 [7] 提出一种精确的光滑模型, 证明了其与原模型的等价性, 但未解决训练时间长和内存占用大的问题, 且鲁棒性较差.

构造一种训练时间短、占用内存少且鲁棒性较好的光滑模型, 将会提高 SSVM 的效率并拓宽其在实

际问题中的应用. 分段识别<sup>[8-10]</sup> 提供了一种有效途径, 通过组合子空间的分段函数得到最终的分类决策, 该方法易于实现、分类精度高, 已被应用于图像识别领域<sup>[11]</sup>. 文献 [8] 采用分段识别的思想构造分片支撑向量机, 利用样本特征进行等距离分割, 导致某些子空间仅含一类样本, 分类精度较差. 此时需要缩减子空间的数目, 并重新分割训练数据集, 故而增加了训练时间. 文献 [9] 的分片模型分类精度较高, 但在子空间中训练双支持向量域分类器时, 求解约束凸二次规划的训练时间较长. 文献 [10] 运用分段函数近似支持向量机中松弛向量的加函数, 导出光滑可微的模型, 但这有悖分段识别思想的初衷.

本文构造一种模糊聚类光滑模型 (FCSSVM). 首先, 运用模糊  $C$  均值 (FCM) 将训练数据集分割为若干个子簇; 然后, 对子簇内样本采用熵函数近似松

收稿日期: 2015-12-10; 修回日期: 2016-05-12.

基金项目: 国家自然科学基金项目 (61373174, 61562001); 陕西省教育厅专项科研计划项目 (15JK1587).

作者简介: 刘三阳 (1959-), 男, 教授, 博士生导师, 从事最优化理论方法及应用、网络优化、数据挖掘等研究; 吴德 (1979-), 男, 高级工程师, 博士生, 从事网络优化、数据挖掘的研究.

<sup>†</sup> 通讯作者. E-mail: wood1979@163.com

弛向量的加函数构造光滑支持向量机模型;最后,定义最近邻子空间,设置选择性集成策略,判断测试样本的类别指标.由于FCSSVM的每次训练仅针对单个子簇内样本进行,可以缩小单次训练的规模;利用高效无约束优化算法求解单个子簇内的光滑模型,能有效缩短训练时间;集成若干近邻子空间中的分类结果,可以保证较高的分类精度.数值实验表明了FCSSVM对比已有算法的优越性.

## 1 分片光滑支持向量机

给定二分类训练集  $T = \{X, Y\} = \{(x_i, y_i)\}_{i=1}^l$ . 其中:  $X = [x_1, x_2, \dots, x_l]^T, x_i \in R^n$  是第  $i$  个训练样本;  $Y = [y_1, y_2, \dots, y_l]^T, y_i \in \{1, -1\}$  是类别标签.

### 1.1 模糊 $C$ 均值

利用模糊  $C$  均值求解如下约束凸二次规划:

$$\begin{aligned} \min_{(U,V)} J(U, V; X) &= \sum_{i=1}^c \sum_{j=1}^l (u_{ij})^2 \|x_j - v_i\|^2; \\ \text{s.t.} \quad \sum_{i=1}^c u_{ij} &= 1, j = 1, 2, \dots, l. \end{aligned} \quad (1)$$

其中:  $J(U, V; X)$  是成本函数;  $U = (u_{ij})_{c \times l}$  是隶属度矩阵 ( $\sum_{i=1}^c u_{ij}^{(r)} = 1$ ),  $u_{ij} \in [0, 1]$  表示样本  $x_j$  归属聚类中心  $v_i$  的隶属度;  $V = (v_1, v_2, \dots, v_c)$  是聚类中心向量集;  $d_{ij} = \|x_j - v_i\|^2$  是  $v_i$  与  $x_j$  之间的欧氏距离.

FCM置成本函数关于各变量的偏导数为零,通过迭代得到最优分割  $(U^*, V^*)$ . 具体步骤如下.

**Step1:** 初始化. 输入阈值  $\varepsilon$ , 聚类数  $c (1 \leq c \leq l)$ , 初始隶属度矩阵  $U^{(r)} = (u_{ij}^{(r)})_{c \times l}$ . 记  $r = 0$ .

**Step2:** 计算聚类中心

$$v_i^{(r)} = \frac{\sum_{j=1}^l (u_{ij}^{(r)})^2 x_j}{\sum_{j=1}^l (u_{ij}^{(r)})^2}, i = 1, 2, \dots, c. \quad (2)$$

**Step3:** 修正模糊隶属度矩阵

$$u_{ij}^{(r+1)} = 1 / \sum_{k=1}^c \left( \frac{d_{ij}}{d_{kj}} \right)^2. \quad (3)$$

**Step4:** 估计成本函数. 如果两次迭代的成本函数的差小于阈值,即

$$\|J^r(U^r, V^r; X) - J^{r-1}(U^{r-1}, V^{r-1}; X)\| < \varepsilon, \quad (4)$$

则停止迭代;否则,记  $r = r + 1$ , 转 Step2.

### 1.2 SSVM 训练

经过 FCM 聚类,训练样本集  $X$  被分割为  $c$  个互不相交的子簇  $\{X_i\}_{i=1}^c$ . 记  $x_k(i) \in R^n$  为  $X_i$  中样本, 补充相应的类别指标  $y_k(i) \in \{1, -1\}$ , 并根据脚标  $k$

的值将样本升序排列,得到训练集

$$T_i = \{X_i, Y_i\} = \{(x_k(i), y_k(i))\}_{k=i_1}^{i_s}, \quad (5)$$

其中  $i_1 < i_2 < \dots < i_s$ .

子簇  $X_i$  中 SSVM 的训练即为

$$\begin{aligned} \min \quad & \frac{1}{2} w(i)^T w(i) + \frac{1}{2} P(i) \sum_{k=i_1}^{i_s} \xi_k^2(i). \\ \text{s.t.} \quad & y_k(i) \{w(i)^T \phi[x_k(i)] + b(i)\} \geq 1 - \xi_k(i); \\ & \xi_k(i) \geq 0, k = i_1, i_2, \dots, i_s. \end{aligned} \quad (6)$$

其中:  $w(i)$  和  $b(i)$  是最优分类超平面的权重和偏置;  $\phi: x_k(i) \rightarrow \phi[x_k(i)]$  是非线性映射;  $0 < P(i) \in R^1$  是惩罚因子;  $0 \leq \xi_k(i) \in R^1$  是样本  $x_k(i)$  的松弛.

根据最优解处,权重向量的表达式为

$$w(i) = \sum_{k=i_1}^{i_s} y_k(i) \alpha_k(i) \phi[x_k(i)], \quad (7)$$

可以导出下列两式:

$$\begin{aligned} y_k(i) w(i)^T \phi[x_k(i)] &= \\ \sum_{p=i_1}^{i_s} y_k(i) y_p(i) \alpha_p(i) \phi[x_p(i)]^T \phi[x_k(i)] &= \\ \sum_{p=i_1}^{i_s} Q_{pk}(i) \alpha_p(i) &= [Q(i) \alpha(i)]_k, \quad (8) \\ w(i)^T w(i) &= \\ \sum_{p,q=i_1}^{i_s} y_p(i) \alpha_p(i) \phi[x_p(i)]^T y_q(i) \alpha_q(i) \phi[x_q(i)] &= \\ \sum_{p=i_1}^{i_s} \alpha_p(i) [Q(i) \alpha(i)]_p &= \alpha(i)^T Q(i) \alpha(i). \quad (9) \end{aligned}$$

其中:  $A(i) = [x_{i_1}(i), x_{i_2}(i), \dots, x_{i_s}(i)]^T$  为训练集矩阵,  $K[x_s(i), x_k(i)] = \phi[x_s(i)]^T \phi[x_k(i)]$  为核函数,  $D(i)$  是以  $y_k(i) (k = i_1, i_2, \dots, i_s)$  为对角线元素的对角矩阵,  $Q(i) = D(i) K[A(i), A(i)^T] D(i)$ .

代入式(8)和(9),将规划(6)转化为

$$\begin{aligned} \min \quad & \alpha(i)^T Q(i) \alpha(i) + \frac{1}{2} P(i) \|\xi(i)\|^2; \\ \text{s.t.} \quad & Q(i) \alpha(i) + b(i) y(i) \geq e(i) - \xi(i), \\ & \xi(i) \geq 0. \end{aligned} \quad (10)$$

其中:  $\xi(i) = \{\xi_{i_1}(i), \xi_{i_2}(i), \dots, \xi_{i_s}(i)\}$  是松弛向量;  $y(i) = \{y_{i_1}(i), y_{i_2}(i), \dots, y_{i_s}(i)\}$  是类别指标向量;  $e(i)$  是维数与  $\xi(i)$  相同、分量均为1的列向量.

令  $r(i) = e(i) - [Q(i) \alpha(i) + b(i) y(i)]$ , 将规划(10)中的松弛向量改写为

$$\xi(i) = \max\{0, r(i)\} = \{r(i)\}_+. \quad (11)$$

其中

$$\{r(i)\}_+ = \{[r_1(i)]_+, [r_2(i)]_+, \dots, [r_s(i)]_+\}; \quad (12)$$

$$[r_k(i)]_+ = \begin{cases} r_k(i), & r_k(i) \geq 0; \\ 0, & r_k(i) < 0. \end{cases} \quad (13)$$

于是规划(10)被转换为

$$\min \frac{1}{2} \alpha(i)^T Q(i) \alpha(i) + \frac{1}{2} P(i) \cdot \|\{e(i) - [Q(i)\alpha(i) + b(i)y(i)]\}_+\|^2. \quad (14)$$

记  $\beta > 0$  为光滑参数, 文献[8]中引理2保证可用熵函数

$$P_\beta(r) = r + \frac{\ln[1 + \exp(-\beta r)]}{\beta} \quad (15)$$

近似加函数  $r_+ = \max(0, r)$ , 从而将规划(14)转化为

$$\min \frac{1}{2} \alpha(i)^T Q(i) \alpha(i) + \frac{1}{2} P(i) \cdot \|P_\beta\{e(i) - [Q(i)\alpha(i) + b(i)y(i)]\}\|^2. \quad (16)$$

规划(16)的梯度向量和海塞矩阵分别为

$$\begin{aligned} \text{Grad} = & \begin{bmatrix} Q(i)\alpha(i) \\ 0 \end{bmatrix} + P(i) \begin{bmatrix} -Q(i)^T \\ -Y(i)^T \end{bmatrix} \cdot \\ \text{diag}\{P_\beta[r(i)]\} \text{diag}[e(i) + v(i)]^{-1} e(i), & \quad (17) \end{aligned}$$

$$\begin{aligned} \text{Hessian} = & \begin{bmatrix} Q(i) & 0 \\ 0 & 0 \end{bmatrix} + P(i) \begin{bmatrix} -Q(i)^T \\ -Y(i)^T \end{bmatrix} \cdot \\ \text{diag}[e(i) + v(i)]^{-2} \{I(i) + \beta \text{diag}[v(i)]\} \cdot & \\ \text{diag}[P_\beta(r(i))]\} [-Q(i) & -Y(i)]. & \quad (18) \end{aligned}$$

其中:  $v(i) = \exp[-\beta r(i)]$ ,  $\text{diag}(\gamma)$  是以  $\gamma$  的各个分量为对角线元素的对角矩阵.

采用无约束优化算法中简便易行的牛顿法求得规划(16)的最优解  $[\alpha(i)^T, b(i)]$ , 构造分类决策函数

$$\begin{aligned} y_x(i) = \text{sgn}\{w(i)^T \phi(x) + b(i)\} = & \\ \text{sgn}\left\{\alpha(i)^T \sum_{k=i_1}^{i_s} y_k(i) k[x_k(i), x] + b(i)\right\}. & \quad (19) \end{aligned}$$

### 1.3 选择性集成策略

选择性集成策略根据变尺度距离检测测试样本的  $k$  近邻子空间, 并综合相应的分类决策函数.

**定义1**(平均聚类半径) 给定子簇  $X_i = \{x_1, x_2, \dots, x_m\}$  及聚类中心  $v_i$ , 平均聚类半径即簇内样本在核空间中到聚类中心  $v_i$  的距离平均值, 有

$$\begin{aligned} r_i = \sum_{x \in X_i} \frac{\|\phi(x) - \phi(v_i)\|}{m} = & \\ \sum_{x \in X_i} \frac{\sqrt{k(x, x) - 2k(x, v_i) + k(v_i, v_i)}}{m}. & \quad (20) \end{aligned}$$

**定义2**(变尺度距离) 给定测试样本  $x$ 、子簇  $X_i$  及聚类中心  $v_i$ , 变尺度距离定义为核空间中  $x$  到  $v_i$  的距离与平均聚类半径  $r_i$  之比, 即

$$d(x, v_i) = \frac{\sqrt{k(x, x) - 2k(x, v_i) + k(v_i, v_i)}}{r_i}. \quad (21)$$

**定义3**(最近邻子空间) 记  $\{X_i\}_{i=1}^c$  为FCM分割样本集所得子簇,  $\{v_i\}_{i=1}^c$  为相应的聚类中心集(记为  $V$ ). 测试样本  $x$  的最近邻子空间是包含  $v_j$  的子簇  $X_j$ , 这里  $v_j$  是以聚类中心集  $V$  作为参考对象时  $x$  的最近邻, 满足

$$d(x, v_j) \leq d(x, v_i), \quad \forall v_i \in V, i \neq j. \quad (22)$$

**定义4**( $k$  近邻子空间) 记  $\{X_i\}_{i=1}^c$  为FCM分割样本集所得子簇,  $\{v_i\}_{i=1}^c$  为相应的聚类中心集(记为  $V$ ). 测试样本  $x$  的  $k$  近邻子空间是包含  $\{v_j\}_{j=p_1}^{p_k}$  的子空间  $\{X_j\}_{j=p_1}^{p_k}$ , 这里  $\{v_j\}_{j=p_1}^{p_k}$  是以聚类中心集  $V$  作为参考对象时  $x$  的  $k$  近邻, 满足

$$\begin{aligned} d(x, v_{p_1}) \leq d(x, v_i), \quad \forall v_i \in V, i \neq p_1; & \\ d(x, v_{p_2}) \leq d(x, v_i), \quad \forall v_i \in V, i \neq p_1, i \neq p_2; & \\ \vdots & \\ d(x, v_{p_k}) \leq d(x, v_i), \quad \forall v_i \in V, i \neq p_1, & \\ i \neq p_2, \dots, i \neq p_{k-1}. & \quad (23) \end{aligned}$$

选择性集成策略将测试样本的变尺度距离升序排列, 根据前  $k$  个最小值相应的近邻子空间中的分类决策函数, 构造如下最终分类判别函数:

$$\begin{aligned} y = \text{sgn} \sum_{i=p_1}^{p_k} y_x(i) = & \\ \text{sgn}\left\{\sum_{i=p_1}^{p_k} \text{sgn}[w(i)^T \phi(x) + b(i)]\right\}. & \quad (24) \end{aligned}$$

其中:  $p_1 = \arg(\min d_i), p_2 = \arg(\min_{i \neq p_1} d_i), \dots, p_k = \arg(\min_{i \neq p_1, p_2, \dots, p_{k-1}} d_i)$ .

## 2 数值实验

程序采用 Matlab 7.01 编制, 实验在 P4 CPU, 3.06 GHz, 1 GB 内存的计算机上进行. 采用性能良好的径向基核函数,  $k(x, y) = \exp\left(-\frac{\|x - y\|^2}{\sigma^2}\right)$ , 以梯度向量和海塞矩阵的极限值 ( $\beta \rightarrow \infty$  时) 代入运算.

### 2.1 分类表现随聚类和近邻子空间数目的变化

采用 Diabetics 和 Waveform 数据集, 训练集、测试集的规模及样本的类别和属性见表1.

表1 分类数据集特征

数据集	训练规模	测试规模	类别	属性
Waveform	400	4600	2	20
Diabetics	400	368	2	8

设定FCSSVM的惩罚参数 $P = 1$ ,核宽参数 $\sigma = 0.5$ ,10次随机抽取实验的结果列于表2和表3.其中:“/”表示近邻子空间数 $k$ 超过了聚类数 $c$ 而无需计算; $c = 1$ 意味着训练之前没有进行“FCM”聚类.

表2 FCSSVM在Waveform数据集上的分类表现

聚类数 $c$	聚类时间	性能指标	近邻子空间数 $k$		
			1	3	5
1	0	分类精度 / %	85.93	/	/
		分类时间 / s	62.30	/	/
3	4.57	分类精度 / %	90.59	92.12	/
		分类时间 / s	32.79	44.72	/
5	5.29	分类精度 / %	<b>92.45</b>	<b>93.59</b>	<b>91.87</b>
		分类时间 / s	19.17	28.23	34.67
7	6.03	分类精度 / %	79.59	83.11	81.23
		分类时间 / s	12.83	17.37	21.65
9	8.41	分类精度 / %	73.35	79.30	77.19
		分类时间 / s	<b>9.26</b>	<b>12.29</b>	<b>15.09</b>

表3 FCSSVM在Diabetics数据集上的分类表现

聚类数 $c$	聚类时间	性能指标	近邻子空间数 $k$		
			1	3	5
1	0	分类精度 / %	71.33	/	/
		分类时间 / s	21.76	/	/
3	1.43	分类精度 / %	<b>76.63</b>	<b>77.01</b>	/
		分类时间 / s	8.09	10.07	/
5	1.67	分类精度 / %	71.47	75.32	<b>75.32</b>
		分类时间 / s	5.11	6.82	12.18
7	1.99	分类精度 / %	67.66	69.55	70.16
		分类时间 / s	3.28	4.54	9.07
9	2.38	分类精度 / %	64.15	65.26	65.11
		分类时间 / s	<b>1.94</b>	<b>2.25</b>	<b>4.93</b>

观察表2和表3中数据,得到以下结论:

1) 随着聚类数的增加,FCSSVM的分类精度先增加后减少,而训练时间则逐渐减少.

不失一般性,取 $k = 1$ 为例.当聚类数从 $c = 1$ 增至 $c = 3$ 再增至 $c = 9$ 时,FCSSVM在Diabetics数据集上的分类精度先从71.33%增至76.63%再降至64.15%;当聚类数从 $c = 1$ 增至 $c = 9$ 时,FCSSVM的训练时间直接从21.76s降至1.94s,后者仅为前者的8.91%.

2) 随着近邻子空间数的增加,FCSSVM的分类精度变化微小,而分类时间的变化则较为显著.

以 $c = 5$ 为例.当近邻数从 $k = 1$ 增加至 $k = 3$ 再增加至 $k = 5$ 时,FCSSVM在Waveform数据集上的分类精度先从92.45%增加至93.59%再降至91.87%;而其分类时间从19.17s增至28.23s再增加至34.67s. FCSSVM分类精度的变化不超过1.72%,而

分类时间变化了1.81倍.换言之,FCSSVM的分类精度随着 $k$ 的增加产生微小变化,而分类时间则产生显著变化.

## 2.2 分类表现随核宽参数的变化

Image数据集包含2310个18维数据.随机选取1300个样本作为训练集,其余样本作为测试集.选取SVM、SSVM、文献[8]中精确模型SDWNSVM和文献[10]中分片支持向量域模型CPDSVDC作为比较对象,设置惩罚参数 $P = 1$ ,FCSSVM的聚类数和近邻数为 $c = 3$ 和 $k = 3$ ,CPDSVDC的分片数 $c = 3$ .图1给出了分类精度随径向基核参数的变化趋势.

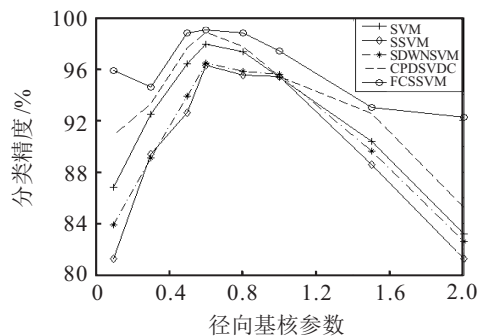


图1 分类精度随径向基核参数的变化

观察图1,可以得出如下结论:

1) FCSSVM具有最高的分类精度;CPDSVDC的分类精度略低;SVM、SSVM和SDWNSVM的分类精度相当,均低于FCSSVM和CPDSVDC的分类精度.

2) FCSSVM的鲁棒性最好,分类精度随径向基核参数变化产生微小变化. CPDSVDC、SVM、SSVM和SDWNSVM的鲁棒性较差,分类精度随径向基核参数变化产生较为明显的波动.

聚类数取 $c = 3$ 时,本例的聚类时间仅为5.26s.考虑到聚类时间不随径向基核参数变化,图2展示了不同算法的分类时间随径向基核参数的变化趋势.

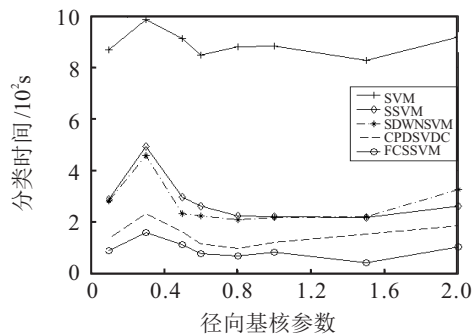


图2 分类时间随径向基核参数的变化

显然,FCSSVM的分类时间最短;CPDSVDC的分类时间略高;SSVM和SDWNSVM的分类时间远低于SVM的分类时间,却明显高于FCSSVM和CPDSVDC的分类时间;SVM的分类时间最长.

### 2.3 大数据集的分类表现

将区间  $[0,1] \times [0,1]$  分割为 10 000 个点,生成棋盘格数据集;抽取规模递增的样本作为训练集,其余作为测试集.以 SVM、SSVM、SDWNSVM 和 CPDSVDC 作为比较对象.设定惩罚参数  $P = 1$ ,径向基核宽参数  $\sigma = 0.3$ ,FCSSVM 和 CPDSVDC 的聚类数为  $c = 5$ ,FCSSVM 的近邻子空间数  $k = 3$ ;10 次随机抽取实验的平均结果列于表 4,其中时间是聚类时间和分类时间的总和,黑体数字表示最优结果.

表 4 不同规模数据集上的分类表现

样本规模	算法	精度 / %	时间 / s	迭代次数
1 000	SVM	99.30	12.02	/
	SSVM	98.16	4.68	2.2
	SDWNSVM	99.48	4.23	2.0
	CPDSVDC	99.53	2.67	/
	FCSSVM	<b>99.57</b>	<b>1.25</b>	<b>1.1</b>
2 000	SVM	99.33	30.09	/
	SSVM	98.57	11.29	4.1
	SDWNSVM	99.54	10.59	3.9
	CPDSVDC	99.65	8.65	/
	FCSSVM	<b>99.77</b>	<b>6.12</b>	<b>1.5</b>
5 000	SVM	99.69	292.58	/
	SSVM	99.09	87.27	8.3
	SDWNSVM	99.31	81.10	8.3
	CPDSVDC	99.89	56.81	/
	FCSSVM	<b>99.99</b>	<b>21.74</b>	<b>2.9</b>
10 000	SVM	/	/	/
	SSVM	/	/	/
	SDWNSVM	/	/	/
	CPDSVDC	98.08	462.70	/
	FCSSVM	<b>99.69</b>	<b>131.86</b>	<b>16.1</b>

显然,FCSSVM 具有最高的精度,最少的迭代次数和最短的分类时间.以样本规模 5 000 为例,FCSSVM 的分类精度 99.99% 是最高的,而分类时间分别是 SVM、SSVM、SDWNSVM 和 CPDSVDC 的 7.43%、24.91%、26.80% 和 38.2%.当训练规模升至 10 000 时,FCSSVM 的分类精度 99.69% 优于 CPDSVDC 的分类精度 98.08%,且分类时间 131.86 s 仅为 CPDSVDC 分类时间 462.7 s 的 28.49%,而迭代次数仅为 16.1;SVM、SSVM 和 SDWNSVM 由于内存溢出而无法计算.

FCM 被用于对训练数据进行初始分割,尽管其鲁棒性较差且仅对球形分布数据有良好表现,但没有降低 FCSSVM 的分类精度;这一点可从 2.1~2.3 节的数值实验部分得到验证.这里由于 FCM 聚类的目的只是为了将原始数据分割为规模较小的子簇以方便训练,而子集上的分类模型才是决定算法性能优劣的关键;同时,选择性集成策略的设置可以减弱分类器对子簇的依赖性,保证较高的分类精度.

### 3 结论

FCSSVM 每次训练仅对子簇中样本进行,训练时间短且随聚类数的增加而减少;其测试时以选择性集成策略链接最近邻子空间的分类结果,具有较强的鲁棒性,分类精度几乎不随径向基核参数发生变化.数值实验验证了 FCSSVM 具有高的分类精度、短的分类时间和少的迭代次数,从而为大规模样本集提供了一种可行的手段.下一步将研究分割训练数据集的更有效的方法及构造支持向量机的快速训练算法.

### 参考文献(References)

- [1] Lee Yuh-Jye, Mangasarian O L. SSVM: A smooth support vector machine for classification[J]. Computational Optimization and Applications, 2001, 22(1): 5-21.
- [2] Yuan Yu-bo. Forecasting the movement direction of exchange rate with polynomial smooth support vector machine[J]. Mathematical and Computer Modelling, 2013, 57(3/4): 932-944.
- [3] Wu Qing, Zhang Leyou, Wang Wan. New family of piecewise smooth support vector machine[J]. J of System Engineering and Electronics, 2015, 26(3): 618-625.
- [4] Qin Chuan-dong, Liu San-yang. Fuzzy smooth support vector machine with different smooth functions[J]. J of Systems Engineering and Electronics, 2012, 23(3): 460-466.
- [5] Balasundaram S, Tanveer M. Smooth newton method for implicit lagrangian twin support vector regression[J]. Int J of Knowledge-Based and Intelligent Engineering Systems, 2013, 17(4): 267-278.
- [6] Ji Rui, Yang Yupu, Zhang Weidong. Incremental smooth support vector regression for Takagi-Sugeno fuzzy modeling[J]. Neurocomputing, 2014, 12(3): 281-291.
- [7] Liang Jinjin, Wu De. Smooth diagonal weighted newton support vector machine[J]. Mathematical Problems in Engineering, 2013, 2013(2): 1-10.
- [8] 任双桥,杨德贵,黎湘,等.分片支撑向量机[J].计算机学报, 2009, 32(1): 77-85.  
(Ren S Q, Yang D G, Li X, et al. Piecewise support vector machines[J]. Chinese J of Computers, 2009, 32(1): 77-85.)
- [9] 梁锦锦,吴德.聚类分片双支持向量域分类器[J].控制与决策, 2015, 30(7): 1298-1302.  
(Liang J J, Wu D. Clustering piecewise double support vector domain classifier[J]. Control and Decision, 2015, 30(7): 1298-1302.)
- [10] Wu Qing, Wang Wenqing. Piecewise smooth support vector machine for classification[J]. Mathematical Problems in Engineering, 2013, 2013(3): 301-312.
- [11] Ye Qixiang, Han Zhenjun, Jiao Jianbin, et al. Human detection in images via piecewise linear support vector machines[J]. IEEE Trans on Image Processing, 2013, 22(2): 778-789.