

# 反比例光滑支撑向量机

吴 振, 宇振盛\*

上海理工大学, 理学院, 上海  
Email: xv907177937@qq.com, \*zhsh-yu@163.com

收稿日期: 2020年10月26日; 录用日期: 2020年11月11日; 发布日期: 2020年11月18日

## 摘 要

为了提高光滑支撑向量机模型的分类性能, 文章给出了一种新的光滑函数反比例函数, 并利用光滑技术克服了支撑向量机模型的不可微性, 通过数学理论上严谨的论证和分析, 证明了所给出光滑函数的性质和基于此光滑函数所建立的反比例光滑支撑向量机模型(ISSVM)的收敛性。实验数值表明, 反比例光滑支撑向量机比多项式光滑支撑向量机在分类性能上更有优越性。

## 关键词

光滑支撑向量机, 反比例光滑函数, BFGS算法

# Inverse Proportional Smooth Support Vector Machine

Zhen Wu, Zhensheng Yu\*

University of Shanghai for Science and Technology, Shanghai  
Email: xv907177937@qq.com, \*zhsh-yu@163.com

Received: Oct. 26<sup>th</sup>, 2020; accepted: Nov. 11<sup>th</sup>, 2020; published: Nov. 18<sup>th</sup>, 2020

## Abstract

In order to improve the classified performance of smooth support vector machine model, a new smoothed function inverse proportional function is presented, and smoothing technique is used to overcome the non-differentiability of support vector machine model. Through rigorous demonstration and analysis in mathematical theory, the properties of the given smoothing function and the convergence of the inverse proportional smoothing support vector machine (ISSVM) model based on the smoothing function are proved. The experimental results show that the inverse proportional smoothing support vector machine is superior to the polynomial smoothing support vector machine in classified performance.

\*通讯作者。

## Keywords

Smooth Support Vector Machine, Inverse Proportional Smoothing Function, BFGS Algorithm

Copyright © 2020 by author(s) and Hans Publishers Inc.

This work is licensed under the Creative Commons Attribution International License (CC BY 4.0).

<http://creativecommons.org/licenses/by/4.0/>



Open Access

## 1. 引言

2001年 Lee 等人通过对支撑向量机的深入研究引入光滑的概念, 使用了 sigmoid 积分函数  $p(x, \alpha)$  对无约束的支撑向量机模型 SVM [1] 进行光滑化, 得出了分类性能较好的光滑支撑向量机 SSVM [2]。继而使用 Newton-Armijo 算法对 SSVM 进行求解, 结果展现出了 SSVM 比标准 SVM 具有更好的分类性能和较高的计算效率, 从此光滑函数作为 SSVM 模型的核心引起了人们的广泛关注, 并开辟了支撑向量机的一个新的研究方向。

2005年文献[3]提出了两个多项式光滑函数, 同时对无约束的支撑向量机模型进行光滑化处理, 引出了 PSSVM 模型。可以证明, PSSVM 比 SSVM 有更好的分类性能。之后文献[4]提出了三阶样条光滑函数, 从而引出了新的支撑向量机的 TSSVM 模型。通过分析 TSSVM 模型的收敛性, 并通过数据实验可以说明 TSSVM 分类器的效果要更优。2008年, 文献[5]提出了六阶光滑函数, 此光滑函数逼近正号函数的精度比三阶样条光滑函数更是高了一个数量级。2014年, 吴青等人在此基础上提出指数光滑支持向量分类机[6]的模型, 证明了其收敛性并通过数值实验表明了指数光滑支持向量机比多项式光滑支持向量机在分类性能上更有优势。

很明显可以得知, 光滑支撑向量机的分类性能随着光滑函数的逼近精度提高而改善, 那么, 能否寻求一种新的光滑函数使得光滑支撑向量机的分类性能获得提升, 一直是学者研究关于相关 SVM 问题的热点。本文通过借鉴指数光滑支持向量机模型中指数光滑函数的性质对反比例函数的研究提出了新的光滑函数, 并对支撑向量机模型光滑化处理得出了反比例光滑支撑向量机。

## 2. 反比例光滑函数

根据光滑函数是支撑向量机研究的重点, 依据反比例函数的一些性质, 提出反比例光滑函数

$$i(x, k) = \begin{cases} x, & x > \frac{1}{10k} \\ \frac{1}{10k(2-10kx)}, & x \leq \frac{1}{10k} \end{cases}$$

**性质 1** 已知光滑函数  $i(x, k)$  并且  $x_+$  是正号函数, 那么

- 1)  $i(x, k)$  关于  $x$  一阶光滑。
- 2) 当  $x \rightarrow -\infty$  且  $k > 0$  时,  $i(x, k) \rightarrow 0$ 。

证明 对函数  $i(x, k)$  关于变量  $x$  求一阶导得分段函数

$$i'(x, k) = \begin{cases} 1, & x > \frac{1}{10k} \\ \frac{1}{(2-10kx)^2}, & x \leq \frac{1}{10k} \end{cases}$$

因为  $i'(x, k)$  在分段点处左右极限存在并且左极限

$$\lim_{x \rightarrow \frac{1}{10k}^-} \frac{1}{(2-10kx)^2} = 1$$

与右极限相等, 故为  $i'(x, k)$  在整个区间连续, 又因

$$i''(x, k) = \begin{cases} 0, & x > \frac{1}{10k} \\ \frac{20k}{(2-10kx)^3}, & x \leq \frac{1}{10k} \end{cases}$$

在分段点处左右极限皆存在, 但其右极限为 0 而左极限非 0, 所以  $i''(x, k)$  在断点处不连续。综上可知  $i(x, k)$  关于  $x$  一阶光滑。

当  $x \rightarrow -\infty$  并且  $k > 0$  时, 对  $i(x, k)$  求极限, 即得

$$\lim_{x \rightarrow -\infty} \frac{1}{10k(2-10kx)} = 0$$

**性质 2** 已知光滑函数  $i(x, k)$  并且  $x_+$  是正号函数, 那么

- 1)  $i(x, k) \geq x_+$ 。
- 2) 对于  $x \in (-\infty, +\infty)$ , 当  $k > 0$  时, 有

$$i^2(x, k) - x_+^2 \leq 0.00131105 \frac{1}{k^2}$$

证明 当  $x > \frac{1}{10k}$  时, 显然有  $i(x, k) \geq x_+$ , 而当  $0 \leq x \leq \frac{1}{10k}$  时, 构造函数

$$f_1(x, k) = i(x, k) - x_+ = \frac{1}{10k(2-10kx)} - x$$

对  $f_1(x, k)$  求导得

$$f_1'(x, k) = \frac{1}{(2-10kx)^2} - 1 < 0 \quad \left( 0 \leq x \leq \frac{1}{10k} \right)$$

故  $f_1(x, k)$  在  $\left( 0, \frac{1}{10k} \right)$  内单调递减, 又因  $f_1(x, k)$  在  $x = \frac{1}{10k}$  处取得最小值 0, 因此

$$f_1(x, k) = i(x, k) - x_+ \geq 0$$

也即  $i(x, k) \geq x_+$ 。另当  $x < 0$  时, 显然有

$$f_2(x, k) = i(x, k) - 0 = \frac{1}{10k(2-10kx)} > 0$$

因此也有  $i(x, k) \geq x_+$ 。综上可得结论(1)。

当  $x \geq \frac{1}{10k}$  时, 结论(2)显然成立。当  $0 < x < \frac{1}{10k}$  时,  $x_+ = x$ , 于是对

$$g_1(x, k) = i^2(x, k) - x^2 = \frac{1}{100k^2(2-10kx)^2} - x^2$$

关于  $x$  求得

$$g_1'(x, k) = \frac{1}{5k(2-10kx)^3} - 2x$$

用二分法[7]求得导函数的零点为

$$x_0 = 0.020312 \frac{1}{k}$$

可求得  $g_1(x, k)$  在  $x_0$  处取得最大值为

$$g_1(x_0, k) = 0.00131105 \frac{1}{k^2}$$

另当  $x < 0$  时,  $x_+ = 0$ , 于是

$$g_2(x, k) = i^2(x, k) = \frac{1}{100k^2(2-10kx)^2}$$

在  $(-\infty, 0]$  上单调递增, 且其最大值

$$g_2(0, k) = \frac{1}{400k^2} = 0.0025 \frac{1}{k^2}$$

综上可知结论(2)成立, 且

$$i^2(x, k) - x_+^2 \leq 0.00131105 \frac{1}{k^2} < 0.0188 \frac{1}{k^2}$$

性质 2 表明了反比例光滑函数对正号函数的逼近精度比六阶光滑函数提高了一个数量级。

光滑因子为  $k = 10$  时, 各个光滑函数对正号函数的逼近程度如图 1 所示, 从中可见, 反比例光滑函数的逼近效果更好。

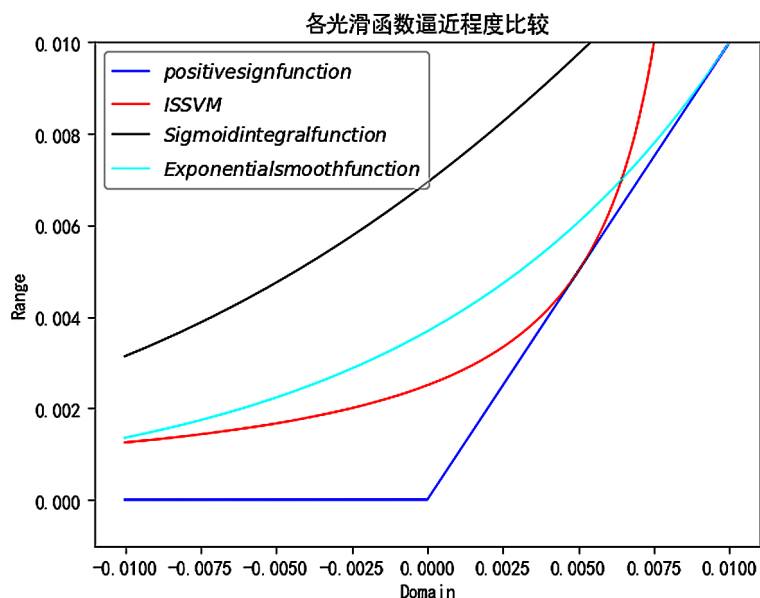


Figure 1. The approximating degree of smooth functions  
图 1. 不同光滑函数的逼近程度

### 3. ISSVM 模型及其收敛性

由软间隔支撑向量机的原始型和对偶形式我们可以对其进行求解得到  $\alpha_i^*$ , 于是分类函数

$$\begin{aligned} f(x) &= \text{sign} \left[ \sum_{i=1}^m y_i \alpha_i^* (\varphi(x_i), \varphi(x)) + b^* \right] \\ &= \text{sign} \left[ \sum_{i=1}^m y_i \alpha_i^* K(x_i, x) + b^* \right] \end{aligned} \quad (1)$$

其中

$$b^* = \frac{1}{m} \sum_{i=1}^m \left( y_i - \sum_{j=1}^m y_j \alpha_j^* (\varphi(x_j), \varphi(x_i)) \right) \quad (2)$$

$$\text{令 } M = (x_1, x_2, \dots, x_m)^T$$

$$x_i = (x_{i1}, x_{i2}, \dots, x_{in}) \quad (i = 1, 2, \dots, m),$$

$$e = \text{ones}(m, 1), \quad D = \text{diag}(y_1, y_2, \dots, y_m), \quad H = DK(M, M^T)D$$

则原始问题改写为

$$\begin{cases} \min_{\alpha, b} \frac{1}{2} \alpha^T H \alpha + C e^T \xi \\ \text{s.t. } D(K(M, M^T)D\alpha + eb) - e + \xi \geq 0, \xi \geq 0 \end{cases} \quad (3)$$

对于任何核函数, 上式均是个凸二次规划问题, 不妨取  $H = I$ , 同时用  $b$  衡量分类间隔  $\left( \frac{2}{\|(\omega, b)\|_2} \right)$ ,

于是在对偶问题中  $b = -e^T D\alpha$ , 同时让松弛变量  $\xi^T \xi$  最小, 则(3)式转化为

$$\begin{cases} \min_{\alpha, b} \frac{1}{2} (\alpha^T \alpha + b^2) + \frac{C}{2} \xi^T \xi \\ \text{s.t. } D(K(M, M^T)D\alpha + eb) - e + \xi \geq 0, \xi \geq 0 \end{cases} \quad (4)$$

由约束条件知,

$$\xi = \left( e - D(K(M, M^T)D\alpha + eb) \right)_+$$

这里是非光滑的。将其代入问题(12), 得到一个强凸无约束优化问题

$$\min_{\alpha, b} \frac{1}{2} (\alpha^T \alpha + b^2) + \frac{C}{2} \left\| \left( e - D(K(M, M^T)D\alpha + eb) \right)_+ \right\|_2^2 \quad (5)$$

有唯一解。

我们用反比例函数对上述无约束支撑向量机模型 SVM 进行光滑化, 可以得到一个新的光滑支撑向量机模型

$$\min_{\alpha, b} \frac{1}{2} (\alpha^T \alpha + b^2) + \frac{C}{2} \left\| i \left( e - D(K(M, M^T)D\alpha + eb), k \right) \right\|_2^2 \quad (6)$$

称之为 ISSVM 模型。

记  $\omega = \alpha, \gamma = -b, A = K(M, M^T)D$ , 则上式可化为

$$\min_{\omega, \gamma} \frac{1}{2} (\omega^T \omega + \gamma^2) + \frac{C}{2} \left\| i \left( e - D(A\omega - e\gamma), k \right) \right\|_2^2$$

分析反比例光滑支撑向量机模型, 可以证明此模型收敛。该模型的最优解在  $k \rightarrow +\infty$  时无限逼近无约束 SVM 模型的最优解。

**定理 1** 设  $A \in \mathbb{R}^{m \times n}$ ,  $b \in \mathbb{R}^{m \times 1}$ , 定义实函数

$$f_+(x): \mathbb{R}^n \rightarrow \mathbb{R}$$

$$f_+(x) = \frac{1}{2} \|(Ax - b)_+\|_2^2 + \frac{1}{2} \|x\|_2^2$$

$$f_i(x, k): \mathbb{R}^n \times \mathbb{N} \rightarrow \mathbb{R}$$

$$f_i(x, k) = \frac{1}{2} \|i(Ax - b, k)\|_2^2 + \frac{1}{2} \|x\|_2^2$$

则有结论

- 1)  $f_+(x)$  和  $f_i(x, k)$  都是强凸函数。
- 2) 优化函数  $\min_x f_+(x)$  存在唯一解  $x^*$ , 优化函数  $\min_x f_i(x, k)$  存在唯一解  $(x^*)^k$ 。
- 3) 对于任意的  $k \geq 1$ , 有

$$\|(x^*)^k - x^*\|^2 \leq \frac{0.00131105m}{2k^2}$$

- 4)  $x^*$  和  $(x^*)^k$  满足

$$\lim_{k \rightarrow \infty} (x^*)^k = x^*$$

证明因为  $\|\cdot\|_2^2$  具有强凸性, 所以  $f_+(x)$  和  $f_i(x, k)$  也满足强凸性质。

由性质 1 可知, 水平集  $L_v(f_i(x, k))$  和水平集  $L_v(f_+(x))$  满足

$$L_v(f_i(x, k)) \subseteq L_v(f_+(x)) \subseteq \{x: \|x\|^2 \leq 2v\} (v \geq 0)$$

因此它们都是  $\mathbb{R}^n$  中的紧子集, 从而有  $\min_x f_+(x)$  和  $\min_x f_i(x, k)$  都有解存在, 而由  $\min_x f_+(x)$  和  $\min_x f_i(x, k)$  的强凸性可得解具有唯一性。

- 5) 我们不妨假设  $x^*$  和  $(x^*)^k$  分别是  $\min_x f_+(x)$  和  $\min_x f_i(x, k)$  的唯一解, 则有

$$f_+((x^*)^k) - f_+(x^*) \geq \nabla f_+(x^*) \left( (x^*)^k - x^* \right) + \frac{1}{2} \|(x^*)^k - x^*\|_2^2 = \frac{1}{2} \|(x^*)^k - x^*\|_2^2$$

$$f_i(x^*, k) - f_i((x^*)^k, k) \geq \nabla f_i((x^*)^k, k) \left( x^* - (x^*)^k \right) + \frac{1}{2} \|(x^*)^k - x^*\|_2^2 = \frac{1}{2} \|(x^*)^k - x^*\|_2^2$$

两式相加得

$$\|(x^*)^k - x^*\|_2^2 \leq (f_i(x^*, k) - f_+(x^*)) - (f_i((x^*)^k, k) - f_+((x^*)^k))$$

$$\leq f_i(x^*, k) - f_+(x^*) = \frac{1}{2} \|i(Ax^* - b, k)\|_2^2 - \frac{1}{2} \|(Ax^* - b)_+\|_2^2$$

又根据性质 2 可知

$$\|(x^*)^k - x^*\|^2 \leq \frac{0.00131105m}{2k^2}$$

根据上式可知

$$\lim_{k \rightarrow \infty} \left\| (x^*)^k - x^* \right\|^2 \leq \lim_{k \rightarrow \infty} \frac{0.00131105m}{2k^2} = 0$$

因此

$$\lim_{k \rightarrow \infty} (x^*)^k = x^*$$

#### 4. 求解 ISSVM 模型的 BFGS 算法

由性质 1 可知光滑函数  $i(x, k)$  是一阶光滑函数, 我们可选用 BFGS 算法[8] [9]来对上述 ISSVM 模型进行优化。其具体算法步骤如下:

步骤 1 初始化

$$H^0 = I, \left( (\omega^0), \gamma^0 \right) = p^0 \in \mathbb{R}^{n+1}, \varepsilon = 10^{-8}, \alpha^0 = I, i = 0,$$

其中  $I$  为单位矩阵。

步骤 2 计算

$$F^i = F(p^i, k), g^i = \nabla F(p^i, k),$$

其中  $F(p^i, k)$  指光滑函数。

步骤 3 如果

$\|g^i\|_2 \leq \varepsilon$  或迭代次数达到最大, 那么迭代停止, 并取  $\left( (\omega^i)^T, \gamma^i \right) = p^i$  为 ISSVM 模型的最优参数解;

否则计算梯度方向  $d^i = -H^i g^i$ 。

步骤 4 沿着方向  $d^i$  采用线搜索计算步长  $\alpha^i$ , 于是有

$$s^i = -\alpha^i H^i g^i$$

再计算出

$$F^{i+1} = F(p^{i+1}, k), g^{i+1} = \nabla F(p^{i+1}, k), y^i = g^{i+1} - g^i$$

步骤 5 计算

$$H^{i+1} = \begin{pmatrix} I - \frac{s^i (y^i)^T}{(s^i)^T y^i} \\ (s^i)^T y^i \end{pmatrix} H^i \begin{pmatrix} I - \frac{y^i (s^i)^T}{(s^i)^T y^i} \\ (s^i)^T y^i \end{pmatrix} + \frac{s^i (s^i)^T}{(s^i)^T y^i}$$

步骤 6 令  $i = i + 1$ , 转步骤 2。

#### 5. 数值实验

对于 SSVM, FSSVM, TSSVM, ESSVM 和 ISSVM5 种模型采用 BFGS 算法求解无约束优化模型, 算法采用的最大的迭代次数为 1000, 且取  $\varepsilon = 10^{-8}$ 。使用 matlab2018a 作为运行环境, 实验结果记录 CPU 耗时、样本分类的训练正确率和测试正确率, 根据这 3 个指标对这五种模型的性能进行比较和分析。

**实验 1** 为了说明 ISSVM 具有解决大规模数据集的能力, 采用的数据集是 NDC 数据集[10], 数据集的样本数数量级至少在万以上。在采用 BFGS 算法对模型进行求解时, 为求结果的准确性, 对训练数据使用 10 折交叉验证方法[11]。结果如表 1 所示。

**Table 1.** Experiment of NDC data set  
**表 1.** NDC 数据集实验

Data	Models	time/s	Training error rate (%)	Test error rate (%)
<b>10000 × 10</b>	SSVM	0.0532	96.7700	96.7700
	FSSVM	0.0642	96.8744	96.8700
	TSSVM	0.0652	96.9722	96.9700
	ESSVM	0.0567	97.4744	97.4800
	ISSVM	0.0557	97.5237	97.5300
<b>10000 × 20</b>	SSVM	0.0743	97.3211	97.2800
	FSSVM	0.0811	97.4233	97.3900
	TSSVM	0.1047	97.5089	97.4800
	ESSVM	0.0795	98.1022	98.1200
	ISSVM	0.0790	98.1024	98.1200
<b>20000 × 10</b>	SSVM	0.1118	97.4394	97.4350
	FSSVM	0.1319	97.5461	97.5550
	TSSVM	0.1730	97.6061	97.6200
	ESSVM	0.1120	97.7833	97.7850
	ISSVM	0.1086	97.8124	97.8200
<b>20000 × 20</b>	SSVM	0.2004	97.8078	97.8000
	FSSVM	0.1920	97.9656	97.9650
	TSSVM	0.2508	98.0317	98.0400
	ESSVM	0.1971	98.7033	98.6900
	ISSVM	0.1902	98.8422	98.8500
<b>30000 × 10</b>	SSVM	0.1145	95.7356	95.7300
	FSSVM	0.1365	95.8507	95.8467
	TSSVM	0.1722	95.9496	95.9300
	ESSVM	0.1133	96.8707	96.8700
	ISSVM	0.1002	96.9852	96.9900
<b>30000 × 20</b>	SSVM	0.1510	96.3856	96.3933
	FSSVM	0.1718	96.5107	96.5100
	TSSVM	0.2083	96.5726	96.5700
	ESSVM	0.1496	97.0181	97.0067
	ISSVM	0.1298	97.2532	97.2500

由表 1 可知, 新提出的反比例光滑支撑向量机 ISSVM 模型具有解决大规模数据问题的能力, 并且在分类正确率和测试正确率上都有着较好的表现。在解决大规模数据问题时, ISSVM 在 CPU 耗时和正确率上有着一定的优势。

数值实验数据来自用 python 编写的随机线性不可分数据集, 统共样本数据有 1200 样本, 测试样本有 288 个, 每个样本有 29 个特征, 采用了高斯核来将其映射到高维特征空间上得到分类结果如图 2~5。



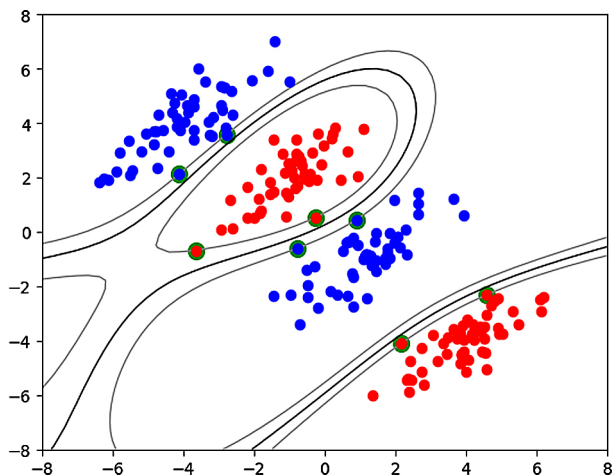


Figure 2. The result (1) of classification  
图 2. 分类结果(1)

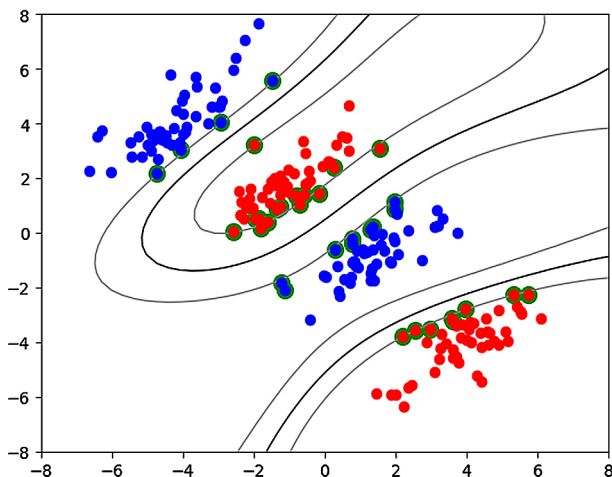


Figure 3. The result (2) of classification  
图 3. 分类结果(2)

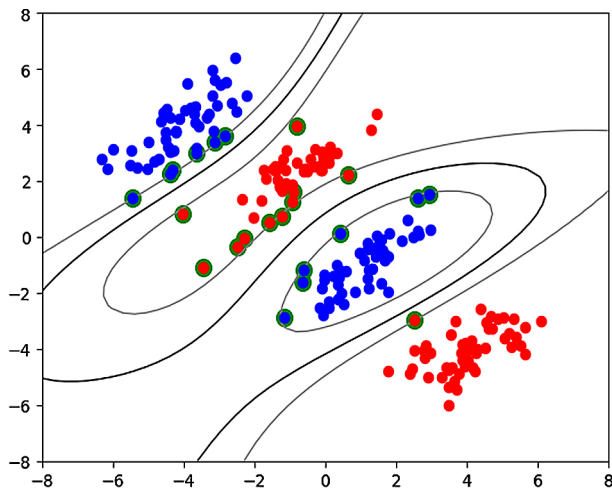
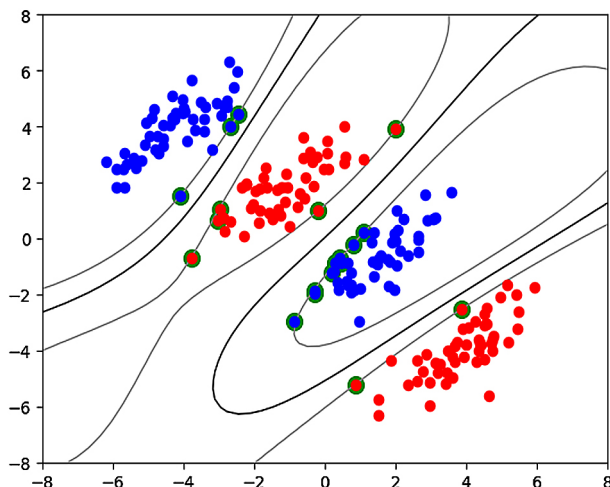


Figure 4. The result (3) of classification  
图 4. 分类结果(3)



**Figure 5.** The result (4) of classification  
**图 5.** 分类结果(4)

可以看出在小规模数据集下其分类的正确率都有着良好的性质。

除此之外, 我们还补充了对于非线性数据集 Checkerboard 数据集的实验结果。当采用的核函数为高斯核函数  $K(x, y) = \exp(-\mu\|x - y\|^2)$  时[12]结果如表 2 所示:

**Table 2.** Experiment of checkerboard data set  
**表 2.** Checkerboard 数据集实验

Data	Models	time/s	Training error rate(%)	Test error rate(%)
1000/9000	SSVM	2.170	93.900	93.430
	FSSVM	2.220	93.900	93.501
	TSSVM	2.102	93.900	93.370
	ESSVM	1.662	93.900	93.500
	ISSVM	1.420	93.890	93.920
2000/8000	SSVM	22.840	95.600	95.320
	FSSVM	23.920	95.902	95.331
	TSSVM	22.030	95.650	95.440
	ESSVM	20.702	95.900	95.664
	ISSVM	20.280	95.880	95.980
3000/7000	SSVM	103.480	97.430	97.311
	FSSVM	81.892	97.430	97.310
	TSSVM	80.925	97.534	97.361
	ESSVM	76.750	97.700	97.470
	ISSVM	74.830	97.938	98.004

由表 2 可知, ISSVM 在处理非线性数据时, 其耗时最短并且有着更高的分类正确率和测试正确率, 说明了 ISSVM 处理非线性数据集的性能更好。

## 6. 结语

本文给出了一种新的光滑函数反比例函数, 并基于光滑支撑向量机的过程应用此光滑反比例函数建立了 ISSVM 模型。与以往的光滑函数如多项式光滑函数、sigmoid 积分函数、样条光滑函数等相比较而言, 其逼近正号函数的精度提高了不同的数量级。实验数据表明, 该反比例光滑支撑向量机模型对于大数据集具有较好的分类性能, 相比于三阶样条光滑支撑向量机模型等而言, 反比例光滑支撑向量机模型对线性数据分类所花费的时间更短且正确率也有所提升, 整体比较来说对于非线性数据也能有较好的分类性能。反比例光滑函数相比其他光滑函数有着逼近程度更强的能力, 反比例光滑支撑向量机相比于其他光滑支撑向量机模型的性能也更优越。

## 参考文献

- [1] Vapnik, V.N. (1995) *The Nature of Statistical Learning Theory*. Springer Verlag, New York, 1-299. [https://doi.org/10.1007/978-1-4757-2440-0\\_1](https://doi.org/10.1007/978-1-4757-2440-0_1)
- [2] Lee, Y.J. and Mangasarian, O.L. (2001) SSVM: A Smooth Support Vector Machine for Classification. *Computational Optimization and Applications*, **20**, 5-22. <https://doi.org/10.1023/A:1011215321374>
- [3] 袁玉波, 严杰, 徐成贤. 多项式光滑的支撑向量机[J]. 计算机学报, 2005, 28(1): 9-17.
- [4] Yuan, Y.B., Fan, W.G. and Pu, D.M. (2007) Spline Function Smooth Support Vector for Classification. *Journal of Industrial and Management Optimization*, **3**, 529-542. <https://doi.org/10.3934/jimo.2007.3.529>
- [5] 黄仁泰, 熊金志. 一个支持向量机的新光滑函数及其性能分析[J]. 计算机工程与应用, 2008, 44(18): 64-65.
- [6] 吴青, 王婉, 王玲芝. 指数光滑支持向量分类机[J]. 西安邮电大学学报, 2014, 19(4): 9-14.
- [7] Song, J.L., Gan, X.B. and Song, D.X. (2004) Nonlinear Global Optimization Based on Bisection Method. *Journal of Information and Computational Science*, **1**, 229-233.
- [8] Yuan, Y.B. and Byrd, R. (1995) Non-Quasi-Newton Updates for Unconstrained Optimization. *Journal of Computing Mathematics*, **13**, 95-107.
- [9] Yuan, Y.B. (1991) A Modified BFGS Algorithm for Unconstrained Optimization. *IMA Journal of Numerical Analysis*, **11**, 325-332. <https://doi.org/10.1093/imanum/11.3.325>
- [10] 吴青, 赵雄. 一类新样条光滑支持向量机[J]. 西安邮电大学学报, 2013, 18(6): 68-74.
- [11] Stone, M. (1974) Cross-Validatory Choice and Assessment of Statistical Prediction. *Journal of the Royal Statistical Society*, **36**, 111-147. <https://doi.org/10.1111/j.2517-6161.1974.tb00994.x>
- [12] Mangasarian, O.L. and Musicant, D.R. (1999) Successive Over-Relaxation for Support Vector Machines. *IEEE Transactions on Neural Networks*, **10**, 1032-1037. <https://doi.org/10.1109/72.788643>