Available online at www.sciencedirect.com
ScienceDirect

ELSEVIER

February 2016, 23(1): 68–72
www.sciencedirect.com/science/journal/10058885

The Journal of China Universities of Posts and Telecommunications

http://jcupt.bupt.edu.cn

# Smooth support vector machine based on circular tangent function

Zhang Jie (✉), Fan Xuhui, Ban Dengke

School of Electronics and Information, Northwestern Polytechnical University, Xi'an 710072, China

**Abstract**

Support vector machines (SVMs) have been intensively applied in the domains of speech recognition, text categorization, and faults detection. However, the practical application of SVMs is limited by the non-smooth feature of objective function. To overcome this problem, a novel smooth function based on the geometry of circle tangent is constructed. It smoothes the non-differentiable term of unconstrained SVM, and also proposes a circle tangent smooth SVM (CTSSVM). Compared with other smooth approaching functions, its smooth precision had an obvious improvement. Theoretical analysis proved the global convergence of CTSSVM. Numerical experiments and comparisons showed CTSSVM had better classification and learning efficiency than competitive baselines.

**Keywords** classifiers, SVM, circle tangent function, smooth technique

## 1 Introduction

As a novel machine learning method, SVM has been widely applied to various fields due to its generalization performance, such as speech recognition, text categorization, faults detection and financial application et al. [1–4]. So far, SVM has made great progress in classification.

In 2001, Lee et al. proposed the integral of sigmoid function to smooth the non-differentiable plus function, thus introduced smooth concept [5]. As the core of smooth SVM (SSVM) model, smooth function gained widespread attention of researchers and opened a new research direction of SVM. In Ref. [6], the author proposed two polynomial functions and got a fourth polynomial SSVM (FPSSVM) model and a quadratic polynomial SSVM (QPSSVM) model. Yuan et al. proposed a three-order spline function and obtained three-order spline SSVM (TSSVM) model [7]. A once continuously differentiable smooth function was used to approximate the plus function and obtained a corresponding SSVM [8]. In Refs. [9–10], Wu et al. introduced a novel SSVM model based on piecewise (PWESSVM), but recently, again used exponential function to replace the plus function and got exponential SSVM (ESSVM) model.

Above achievements are very significant to SSVM since many fast algorithms can be applied. Whether there are other smooth functions to improve the limited efficiency or precision of above schemes is a hot topic for researchers recently. The Brogde-Fletcher-Goldfarb-Shanno (BFGS) algorithm [11–12] is employed to train CTSSVM. Theoretical analysis and numerical experiments confirm that CTSSVM has higher efficiency than previous baselines. This paper, a novel function is proposed to serve as smooth function.

The paper is outlined as follows. In Sect. 2, we presented the derivation of the circle tangent function and described CTSSVM model. The approximation error of smooth functions to the plus function is analyzed in Sect. 3. And in Sect. 4, we introduced the algorithm of BFGS. The convergence performance of CTSSVM is stated in Sect. 5. Sect. 6 employed the BFGS algorithm to train CTSSVM and got the numerical results. In the last section, we make a brief conclusion.

## 2 A novel SSVM model

### 2.1 Derivation of circle tangent function

**Theorem 1** Let $L(x)=(x)_+=\max(0,x)$ be the plus

Received date: 23-06-2015
Corresponding author: Zhang Jie, E-mail: f769397603@163.com
DOI: 10.1016/S1005-8885(16)60010-9

function. The circle tangent function satisfies the first-order differentiation at $x=0$, and can be employed to approximate $L(x)$. The function form can be described as follows, where $\alpha>0$ is a real parameter.

$$\psi(x,\alpha)=\begin{cases} x; & x>\dfrac{\sqrt{2}}{2}\alpha \\ \left(1+\sqrt{2}\right)\alpha-\sqrt{\left(\left(1+\sqrt{2}\right)\alpha\right)^2-(x+\alpha)^2}; & -\alpha<x\leqslant\dfrac{\sqrt{2}}{2}\alpha \\ 0; & x\leqslant-\alpha \end{cases} \tag{1}$$

**Proof** Firstly, assume that the coordinate of point $P$ is $(-\alpha,0)$ and that point $O$ is the center of a circle whose tangent is $L(x)$ and point of tangency is point $P$ and $Q$ (see Fig. 1). According to the geometry, we can gain the coordinate of point Q $\left(\left(\sqrt{2}/2\right)\alpha,\left(\sqrt{2}/2\right)\alpha\right)$ and circle center $O\left(-\alpha,(1+\sqrt{2})\alpha\right)$. Moreover, we can further gain the radius of the circle $r=\left(1+\sqrt{2}\right)\alpha$ by geometric knowledge. A standard equation of a circle is presented in Eq. (2). Based on calculated radius and circle center.

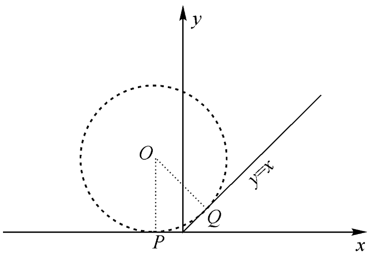

**Fig. 1** Circle tangent function

$$\left(y-\left(1+\sqrt{2}\right)\alpha\right)^2+(x+\alpha)^2=\left(\left(1+\sqrt{2}\right)\alpha\right)^2 \tag{2}$$

So we have $\psi(x,\alpha)=y=\left(1+\sqrt{2}\right)\alpha+\sqrt{\left(\left(1+\sqrt{2}\right)\alpha\right)^2-(x+\alpha)^2}$ for $-\alpha<x\leqslant\left(\sqrt{2}/2\right)\alpha$. Since $L(x)$ is the tangent of the arc $PQ$, $\psi(x,\alpha)=L(x)$ for $x>\left(\sqrt{2}/2\right)\alpha$ or $x\leqslant-\alpha$.

$$\psi'(x,\alpha)=\begin{cases} 1; & x>\dfrac{\sqrt{2}}{2}\alpha \\ \dfrac{x+\alpha}{\sqrt{\left(\left(1+\sqrt{2}\right)\alpha\right)^2-(x+\alpha)^2}}; & -\alpha<x\leqslant\dfrac{\sqrt{2}}{2}\alpha \\ 0; & x\leqslant-\alpha \end{cases} \tag{3}$$

Since the left limit and right limit of derivative $\psi'(x,\alpha)$ exist and equal at the piecewise point, $\psi'(x,\alpha)$ is continuous for $-\infty<x<\infty$.

$$\psi''(x,\alpha)=\begin{cases} 0; & x>\dfrac{\sqrt{2}}{2}\alpha \text{ or } x\leqslant-\alpha \\ \dfrac{\left(\left(1+\sqrt{2}\right)\alpha\right)^2}{\left[\left(\left(1+\sqrt{2}\right)\alpha\right)^2-(x+\alpha)^2\right]^{\frac{3}{2}}}; & -\alpha<x\leqslant\dfrac{\sqrt{2}}{2}\alpha \end{cases} \tag{4}$$

In the same way, the left limit and right limit of second derivative $\psi''(x,\alpha)$ exist at the piecewise point. But the left limit of $\psi''(x,\alpha)$ is 0 and its right limit is 1 at the piecewise point $x=-\alpha$. So $\psi''(x,\alpha)$ is discontinuous at the piecewise point. In summary, circle tangent function is first-order differentiable with respect to $x$.

## 2.2 CTSSVM

Now we consider the binary problem of classifying $m$ training points in $n$-dimensional real space $\mathbf{R}^n$. The set of each point $A_i$ is represented by $m\times n$ matrix $\boldsymbol{A}$ and its corresponding membership in the class 1 or $-1$ is represented by a given $m\times m$ matrix $\boldsymbol{D}$ with 1 or $-1$ along the diagonal. For this problem, the standard SVM is given by the following mathematical model with parameter $v$>0.

$$\left.\begin{aligned} &\min_{\omega,\gamma,y}\frac{1}{2}\|\omega\|_2^2+v\boldsymbol{e}^{\mathrm{T}}y \\ &\text{s.t. } \boldsymbol{D}(\boldsymbol{A}\omega-\boldsymbol{e}\gamma)+y-\boldsymbol{e}\geqslant0;\ y\geqslant0 \end{aligned}\right\} \tag{5}$$

Where $\boldsymbol{\omega}$ is the normal to the bounding plane and $\gamma$ is the distance of the bounding to the origin. The 1-norm of the slack variable $\boldsymbol{y}$ is minimized with weight $v$. The problem (5) can be converted into an unconstrained optimization problem (6) by replacing $\boldsymbol{e}^{\mathrm{T}}y$ with 2-norm vector $\boldsymbol{y}=\left(\boldsymbol{e}-\boldsymbol{D}(\boldsymbol{A}\boldsymbol{w}-\boldsymbol{e}\gamma)\right)_+$ and adding $\gamma^2/2$ to the objective function, which has little effect on the model.