



Distance difference and linear programming nonparallel plane classifier

Qiaolin Ye^{a,*}, Chunxia Zhao^a, Haofeng Zhang^a, Ning Ye^b

^aSchool of Computer Science and Technology, Nanjing University of Science and Technology, Nanjing, China

^bSchool of Information Technology, Nanjing Forestry University, Nanjing, China

ARTICLE INFO

Keywords:

GEPSVM
LSTSVM
TWSVM
Standard eigenvalues
Feature selection
Input features
Kernel functions

ABSTRACT

We first propose Distance Difference GEPSVM (DGEPSVM), a binary classifier that obtains two nonparallel planes by solving two standard eigenvalue problems. Compared with GEPSVM, this algorithm does not need to care about the singularity occurring in GEPSVM, but with better classification correctness. This formulation is capable of dealing with XOR problems with different distribution for keeping the genuine geometrical interpretation of primal GEPSVM. Moreover, the proposed algorithm gives classification correctness comparable to that of LSTSVM and TWSVM, but with lesser unknown parameters. Then, the regularization techniques are incorporated to the TWSVM. With the help of the regularized formulation, a linear programming formation for TWSVM is proposed, called FETSVM, to improve TWSVM sparsity, thereby suppressing input features. This means FETSVM is capable of reducing the number of input features, for linear case. When a nonlinear classifier is used, this means few kernel functions determine the classifier. Lastly, this algorithm is compared on artificial and public datasets. To further illustrate the effectiveness of our proposed algorithms, we also apply these algorithms to USPS handwritten digits.

© 2011 Elsevier Ltd. All rights reserved.

1. Introduction

Eigenvalue based techniques are attractive for the classification of very large sparse datasets (Guarracino, Cifarelli, Seref, & Pardalos, 2007) such as generalized proximal SVM (GEPSVM for short) (Mangasarian & Wild, 2006). GEPSVM obtains each of the nonparallel planes by solving the eigenvector corresponding to a smallest eigenvalue of a generalized eigenvalue problem, such that each plane is as close as possible to the samples for its class and meantime as far as possible from the samples for the other classes (Mangasarian & Wild, 2006). The edges of two-class GEPSVM lie in its lower computational complexity and its better classification performance in terms of solving XOR problems with respect to standard SVM that find one plane that separates the two classes. In Mangasarian and Wild (2006), Mangasarian et al. presented a simple “cross planes” example that is a generalization of the XOR example, which indicated the effectiveness of GEPSVM over PSVM and SVM. Fig. 1 in Mangasarian and Wild (2006) demonstrates GEPSVM has classification correctness of 100% in XOR case. Recently, a lot of GEPSVM-based algorithms have been proposed. To improve the generalization of GEPSVM, Jayadeva et al. proposed Fuzzy GEPSVM (FGEPSVM) given its multi-category formulation. In 2007, Guarracino et al. (2007) introduced a new regularization technique to GEPSVM for reducing the time complexity of GEPSVM, but with two unknown parameters in linear case. These

algorithms obtain two planes by solving generalized eigenvalue problems as GEPSVM does. However, for the symmetric matrices occurring in these algorithms such as H and M in the formulation (5) and (6), if both are semi-positive, an ill-defined operation will be obtained. Moreover, these algorithms weaken the genuine geometrical interpretation of the nonparallel plane classifier due to the adoption of regularization term that improves their generalization. Recently, a twin SVM algorithm (TWSVM for short), proposed by Jayadeva et al., was published in TPAMI (Jayadeva & Chandra, 2007). This algorithm, which is in the spirit of GEPSVM, obtains two planes by solving two smaller quadratic programming problems (QPPs) than that of the standard SVM. Experimental results show the effectiveness of TWSVM over SVM and GEPSVM (Arun Kumar & Gopal, 2009; Jayadeva & Chandra, 2007). TWSVM takes $O(1/4m^3)$ operations which is 1/4 of standard SVM, whereas, GEPSVM takes $O(1/4n^3)$. Here, m is the number of training samples, n is the dimensionality and $m \gg n$ (Arun Kumar & Gopal, 2009; Jayadeva & Chandra, 2007). Obviously, GEPSVM is by far faster than TWSVM. To reduce the time complexity and keep the effectiveness of the twin SVM classifier, some scholars proposed its least squares version (LSTSVM for short) in 2009 (Arun Kumar & Gopal, 2009; Ghorai, Mukherjee, & Dutta, 2009). In fact, LSTSVM determines two nonparallel planes by solving two PSVM-type (Fung & Mangasarian, 2001) problems. Compared with TWSVM, LSTSVM has lesser computational time due to the fact that it only solves two systems of linear equations instead of two QPPs as for TWSVM. TWSVM and LSTSVM however, also lose the genuine geometrical interpretation of the nonparallel plane classifier. GEPSVM is

* Corresponding author.

E-mail address: yeqiaolin65620868@163.com (Q. Ye).

proposed to solve the complex examples which are difficult classification cases for typical linear classifiers just as XOR example does (Mangasarian & Wild, 2006). Each of the planes obtained by GEPSVM is as close as possible to the samples for its class and meantime as far as possible from the samples for the other classes (Mangasarian & Wild, 2006). However, TWSVM requires each of planes obtained to be as close as possible to the samples for its class and meantime at a distance of at least 1 from the samples for the other classes (Jayadeva & Chandra, 2007). LSTSVM requires each of the planes to be as close as possible to the samples for its class and meantime at a distance of 1 from the samples for the other classes. In intuition, when handling XOR examples of different distribution, TWSVM and LSTSVM may yield poor classification performance due to the difference from the optimization criterion of GEPSVM, although they can obtain good classification performance on UCI datasets due to the use of the loss function. Moreover, another flaw of TWSVM and LSTSVM is that two penalty parameters are introduced to their objective functions instead of one regularization parameter as for GEPSVM. Undoubtedly, this will lead to the difficulty of parameter selection. In addition, when there are many noise variables, the 1-norm SVM (Zou, 2007; Zhou, Zhang, & Jiao, 2002) has advantages over the 2-norm SVM because the former is capable of generating sparse solutions that make the classifier easier to store and faster to compute. However, these GEPSVM-based algorithms, including GEPSVM, cannot generate very sparse solutions, even if we give their 1-norm formulations as in 1-norm SVM (Zou, 2007). This is so because the direction w_i and threshold r_i that determine the i^{th} separating planes combines with the input samples.

In this paper, we first propose a new but fast algorithm, termed as Difference GEPSVM (DGEPSVM). DGEPSVM need not consider the singularity occurring in GEPSVM due to the use of a similar formulation to the MMC (Jiang & Zhang, 2004). We show that the solution of DGEPSVM reduces to solving two simple eigenvalue problems. This property determines DGEPSVM is fast and at least comparable to GEPSVM. Moreover, DGEPSVM can deal with XOR examples with different distribution because it keeps the genuine geometrical interpretation of GEPSVM. Then, we further propose a feature selection algorithm for TWSVM, called FETSVM. This proposed algorithm can overcome such a flaw, that is, GEPSVM and other GEPSVM-based algorithms cannot generate the very sparse solutions. Lastly, the two algorithms are compared on artificial and UCI datasets. We also go onto illustrate their effectiveness for USPS handwritten digits application.

Given four facts: (1) DGEPSVM need not care about the singularity occurring in GEPSVM and performs better in classification correctness than GEPSVM; (2) DGEPSVM surpasses TWSVM and LSTSVM in terms of solving XOR examples with different distribution and gives comparable classification correctness on standard datasets; (3) DGEPSVM has lesser unknown parameters than TWSVM and LSTSVM; and (4) FETSVM performs faster than TWSVM and suppresses input features as well as giving comparable classification correctness.

2. Related work

2.1. Generalized Proximal Support Vector Machines (GEPSVM) (Mangasarian & Wild, 2006)

Given m training points in n dimension input space R^n , denoted by the $m_1 \times n$ matrix A belonging to class 1 and the $m_2 \times n$ matrix B belonging to class-1, where $m_2 + m_1 = m$.

The main purpose of GEPSVM is to find two nonparallel hyperplanes in n -dimension space, i.e.,

$$x^T w_1 - b_1 = 0, \quad x^T w_2 - b_2 = 0 \tag{1}$$

where $(w_i, b_i) \in (R^n \times R)$ ($i = 1, 2$). This algorithm requires each plane to be as close as possible to the samples for its class and as far as possible from the samples for the other classes at the same time.

Suppose $(w_i, b_i) \neq 0$, binary GEPSVM classifiers can be written as the optimization formulation as follows:

$$(GEPSVM1) \quad \underset{(w_1, b_1) \neq 0}{\text{Min}} \quad \frac{(Aw_1 - e_1 b_1)^T (Aw_1 - e_1 b_1) + \delta \left(\begin{bmatrix} w_1 \\ b_1 \end{bmatrix} \right)^T \left(\begin{bmatrix} w_1 \\ b_1 \end{bmatrix} \right)}{(Bw_1 - e_2 b_1)^T (Bw_1 - e_2 b_1)} \tag{2}$$

$$(GEPSVM2) \quad \underset{(w_2, b_2) \neq 0}{\text{Min}} \quad \frac{(Bw_2 - e_2 b_2)^T (Bw_2 - e_2 b_2) + \delta \left(\begin{bmatrix} w_2 \\ b_2 \end{bmatrix} \right)^T \left(\begin{bmatrix} w_2 \\ b_2 \end{bmatrix} \right)}{(Aw_2 - e_1 b_2)^T (Aw_2 - e_1 b_2)} \tag{3}$$

where δ is a regularization constant. The formulation (2) enables GEPSVM to obtain the plane, which is closest to the points for set +1 and furthest from set -1, and (3) enables GEPSVM to obtain the plane which is closest to the points for set -1 and furthest from the points for set +1.

Let

$$\begin{aligned} G &= E^T E + \delta I, & H &= F^T F \\ L &= F^T F + \delta I, & M &= E^T E \end{aligned} \tag{4}$$

$$z_1 = \begin{bmatrix} w_1 \\ b_1 \end{bmatrix}, \quad z_2 = \begin{bmatrix} w_2 \\ b_2 \end{bmatrix}$$

where $E = [A - e_1]$, $F = [B - e_2]$, $E, F \in R^{n+1}$. The optimization problem (2) and (3) become:

$$(GEPSVM1) \quad \underset{z_1 \neq 0}{\text{Min}} \quad \frac{z_1^T G z_1}{z_1^T H z_1} \tag{5}$$

$$(GEPSVM2) \quad \underset{z_2 \neq 0}{\text{Min}} \quad \frac{z_2^T L z_2}{z_2^T M z_2} \tag{6}$$

Using the well-known properties of Rayleigh quotient, we can obtain the solutions of (5) and (6) by solving the following two generalized eigenvalue problems:

$$Gz_1 = \lambda_1 H z_1, \quad Lz_2 = \lambda_2 M z_2, \quad z_i \neq 0, \quad i = 1, 2. \tag{7}$$

It can be seen from (2) and (3), Tikhonov regularization (Tikhonov & Arsen, 1977) is applied to each GEPSVM pair. This can improve the generalization of GEPSVM due to the fact that the regularization term is used for penalty.

2.2. TWIN Support Vector Machines (TWSVM) (Jayadeva & Chandra, 2007)

In 2007, Jayadeva et al. proposed a stand-alone algorithm for binary classification, termed as Twin SVM (TWSVM) (Jayadeva & Chandra, 2007), which is in the spirit of GEPSVM. However, TWSVM has different formulation from GEPSVM. This algorithm obtains two nonparallel planes by solving two SVM-type problems. Experimental results show TWSVM outperforms GEPSVM and standard SVM, in terms of classification correctness. TWSVM can be written as follows:

$$(TWSVM1) \quad \begin{aligned} \text{Min} \quad & \frac{1}{2} (Aw_1 - e_1 b_1)^T (Aw_1 + e_1 b_1) + C_1 e_1^T \xi \\ \text{s.t.} \quad & - (Bw_1 - e_2 b_1) + \xi \geq e_2, \quad \xi \geq 0 \end{aligned} \tag{8}$$

$$(TWSVM2) \quad \begin{aligned} \text{Min} \quad & \frac{1}{2} (Bw_2 - e_2 b_2)^T (Bw_2 - e_2 b_2) + C_2 e_2^T \xi \\ \text{s.t.} \quad & (Aw_2 - e_1 b_2) + \xi \geq e_1, \quad \xi \geq 0 \end{aligned} \tag{9}$$

where C_1 and C_2 are two penalty coefficients. From (8) and (9), we find that only constraints of the other class appear. The objective function does not sum up error over patterns of both the classes. These features show TWSVM is effective on skewed or unbalanced datasets. This may be a reason results in a better classification why

متن کامل مقاله

دریافت فوری ←

ISIArticles

مرجع مقالات تخصصی ایران

- ✓ امکان دانلود نسخه تمام متن مقالات انگلیسی
- ✓ امکان دانلود نسخه ترجمه شده مقالات
- ✓ پذیرش سفارش ترجمه تخصصی
- ✓ امکان جستجو در آرشیو جامعی از صدها موضوع و هزاران مقاله
- ✓ امکان دانلود رایگان ۲ صفحه اول هر مقاله
- ✓ امکان پرداخت اینترنتی با کلیه کارت های عضو شتاب
- ✓ دانلود فوری مقاله پس از پرداخت آنلاین
- ✓ پشتیبانی کامل خرید با بهره مندی از سیستم هوشمند رهگیری سفارشات