Fuzzy Proximal Support Vector Classification Via Generalized Eigenvalues

Jayadeva[1], Reshma Khemchandani[2], and Suresh Chandra[2]

[1] Department of Electrical Engineering
[2] Department of Mathematics,
Indian Institute of Technology Delhi,
Hauz Khas, New Delhi 110 016, India
jayadeva@ee.iitd.ernet.in, reshmaiitd@gmail.com,
chandras@maths.iitd.ernet.in

Abstract. In this paper, we propose a fuzzy extension to proximal support vector classification via generalized eigenvalues. Here, a fuzzy membership value is assigned to each pattern, and points are classified by assigning them to the nearest of two non parallel planes that are close to their respective classes. The algorithm is simple as the solution requires solving a generalized eigenvalue problem as compared to SVMs, where the classifier is obtained by solving a quadratic programming problem. The approach can be used to obtain an improved classification when one has an estimate of the fuzziness of samples in either class.

Keywords: Support vector machines, fuzzy data classification, machine learning, generalized eigenvalue problem, proximal classifier.

1 Introduction

Recently, Mangsarian and Wild [1] proposed a non-parallel plane classifier for two category data classification problem. The classifier obtained from the non parallel planes comprises of the eigenvectors corresponding to smallest eigenvalue of two generalized eigenvalue problems which are close to data points of one class and is far from the data points of other class, respectively, whereas in the case of the proximal support vector machine classifier (PSVM) [2], two parallel planes are generated such that each plane is closest to one of the two data sets to be classified, and the planes are as far apart as possible.

Real data sets are usually corrupted with noise and are associated with a fuzziness in the membership of data points to a class. Data samples which are noisy are less informative. A classifier which is able to utilize information regarding this fuzziness can improve its performance and lessen the effect of outliers. Traditional SVM based classifiers lack such a mechanism [3].

In this paper, we propose a fuzzy extension of PSVM via generalized eigenvalues (GEPSVM), by incorporating a membership matrix S that indicates the fuzzy membership of data points (a number between 0 and 1) to the two

S.K. Pal et al. (Eds.): PReMI 2005, LNCS 3776, pp. 360–363, 2005.

classes. Experimental results show that the performance of Fuzzy GEPSVM (FGEPSVM) compares favourably with that of GEPSVM, PSVM, and SVM.

The paper is organized as follows: Section 2 briefly dwells on the theory of the multiplane linear kernel classifier for binary data classification. Section 3 introduces Fuzzy Proximal Support Vector Machines via Generalized Eigenvalues (FGEPSVM). Experimental results of the algorithm established in the paper are presented in Section 4. Section 5 contains concluding remarks.

2 The Multi-plane Linear Kernel Classifier

Let the data set consisting of m points in a n-dimensional space be represented by a $m \times n$ pattern matrix A. In this case data points belonging to classes 1 and -1 are represented by $A(+)$ and $A(-)$, respectively. The classifier yields two non parallel planes

$$x^T w_1 - \gamma_1 = 0 \text{ and } \quad x^T w_2 - \gamma_2 = 0, \tag{1}$$

the idea being to minimize the Euclidean distance of the planes from the data points of classes 1 and -1, respectively. This leads to the following optimization problem

$$\min_{(w,\gamma)\neq 0} \left[||A(+)w - e\gamma||^2 \Big/ \ ||\begin{bmatrix} w \\ \gamma \end{bmatrix}||^2 \right] \Big/ \left[||[A(-)w - e\gamma||^2 \Big/ \ ||\begin{bmatrix} w \\ \gamma \end{bmatrix}||^2 \right] \tag{2}$$

where e is a vector of ones. It is implicitly assumed that $(w, \gamma) \neq 0 \Rightarrow A(-)w - e\gamma \neq 0$. The optimization problem (2) can be regularized by introducing a Tikhonov regularization term [4] $\delta(> 0)$ to give the following Rayleigh Quotient

$$\min_{z \neq 0} \quad z^T G z \Big/ \ z^T H z, \tag{3}$$

where G and H are symmetric matrices in $\mathbf{R}^{(n+1)\times(n+1)}$ defined as: $G := [A(+) \quad -e]^T * [A(+) \quad -e] + \delta * I$; $H := [A(-) \quad -e]^T * [A(-) \quad -e]$; is positive semi-definite and z is an augmented vector defined as $z := [w \quad \gamma]^T$. Using well known properties of the Rayleigh Quotient [5], the solution of (3) is obtained by solving the generalized eigenvalue problem

$$Gz = \mu Hz, z \neq 0. \tag{4}$$

where μ is a vector of generalized eigenvalues.

If z_1 denotes the eigenvector corresponding to the smallest eigenvalue (μ_1) (4), then $z_1 = [w_1 \quad \gamma_1]^T$ determines the plane $x^T w_1 - \gamma_1 = 0$, which is closest to all data points of class 1. Using an entirely similar argument we define a minimization problem analogous to (2) by interchanging the roles of $A(+)$ and $A(-)$, for determining eigenvector z_2 corresponds to the smallest eigenvalue. The eigenvector z_2 will yield the plane $x^T w_2 - \gamma_2 = 0$, which is closest to all data points of class -1, respectively.

3 Fuzzy Proximal Support Vector Classification Via Generalized Eigenvalues

The FGEPSVMs is obtained by solving the optimization problem

$$\min_{(w,\gamma)\neq 0} \left[\left|\left|[S_1 * A(+)]w - e\gamma\right|\right|^2 / \ \left|\left|\begin{bmatrix} w \\ \gamma \end{bmatrix}\right|\right|^2 \right] \Big/ \left[\left|\left|[S_2 * A(-)]w - e\gamma\right|\right|^2 / \ \left|\left|\begin{bmatrix} w \\ \gamma \end{bmatrix}\right|\right|^2 \right] \tag{5}$$

where $||\cdot||$ denotes the Euclidean distance, and S_1 and S_2 are the matrices of membership values of two classes, respectively.

We observe from (5) that the term in the numerator (denominator) is the weighted euclidean distance in the input space of points in the class 1 (class -1) to the plane $x^T w - \gamma = 0$. On simplification, and working on the lines of [1] we obtain the following generalized eigenvalue problem.

$$Kz = \mu Lz, z \neq 0. \tag{6}$$

where the symmetric matrices K and L in $\mathbf{R}^{(n+1)\times(n+1)}$ are defined as $K := [S_1 * A(+) \quad -e]^T * [S_1 * A(+) \quad -e] + \delta * I; \quad L := [S_2 * A(-) \quad -e]^T * [S_2 * A(-) \quad -e]$; $z = [w \quad \gamma]^T$. By an entirely similar argument we define a minimization problem analogous to (5) for determining z_2 and hence (w_2, γ_2).

$$\min_{(w,\gamma)\neq 0} \left[\left|\left|[S_2 * A(-)]w - e\gamma\right|\right|^2 / \ \left|\left|\begin{bmatrix} w \\ \gamma \end{bmatrix}\right|\right|^2 \right] \Big/ \left[\left|\left|[S_1 * A(+)]w - e\gamma\right|\right|^2 / \ \left|\left|\begin{bmatrix} w \\ \gamma \end{bmatrix}\right|\right|^2 \right] \tag{7}$$

Solving (7) is equivalent to solving the following generalized eigenvalue problem

$$Pz = \mu Qz, z \neq 0. \tag{8}$$

where the symmetric matrices P and Q in $\mathbf{R}^{(n+1)\times(n+1)}$ are defined as $P := [S_2 * A(-) \quad -e]^T * [S_2 * A(-) \quad -e] + \delta * I; \quad Q := [S_1 * A(+) \quad -e]^T * [S_1 * A(+) \quad -e]$; $z = [w \quad \gamma]^T$.

Hence the plane (1) gets characterized with the eigenvectors z_1, and z_2, corresponding to minimum eigenvalues of (6) and (8), respectively. A new data sample $x \in \mathbf{R}^n$ is assigned to a class depending on its proximity to the non-parallel planes corresponding to z_1 and z_2 i.e,

$$x^T w_r - \gamma_r = \min_{l=1,2} |x^T w_l - \gamma_l|. \tag{9}$$

4 Experimental Results

To demonstrate the performance of our approach, we report results on publicly available datasets from the UCI Repository [7]. The classification methods were implemented using MATLAB 6.5 [6] running on a PC with an Intel P4 processor (2.4 MHz), and 512 MB RAM. These results are based on ten fold cross-validation approach. In our case, the optimal value of δ is obtained by using a tuning set comprising of 5% of the dataset.

Table 1. Test Set Classification Accuracy(%) using Linear kernel

Dataset	FGEPSVM (s_1, s_2)	GEPSVM	PSVM	SVM
Heart-c	85.13±04.99 (1,0.8)	84.80±05.21	84.48±05.11	82.87±06.64
Heart-statlog	84.44±05.93 (1,0.9)	83.70±06.67	82.59±07.04	81.48±10.64
Heart-h	81.61±10.26 (1,0.8)	78.85±15.83	84.36±05.94	83.39 ± 08.61
Pima-indians	73.04±03.50 (1,0.9)	72.13±03.95	56.56±02.97	71.36±04.83
Sonar	78.93±08.46 (1,0.9)	75.12±11.14	62.93±09.74	75.52±10.15

In particular, FGEPSVMs may prove valuable when one is willing to obtain an improved classification performance for one class even at the expense of the other. In order to understand the importance of the membership function values, we took a simple illustration in which all samples belonging to class 1 are assigned a membership value s_1, while all samples belonging to class -1 are assigned a membership value s_2 [8]. In general, there may be many ways to calculate fuzzy membership matrix S [8]. The experimental results in Table 1 demonstrate that FGEPSVM for data classification provides improved generalization ability in comparison with that of GEPSVM, PSVM and SVM on the UCI datasets.

5 Conclusion

In this paper, we have proposed a fuzzy extension of Proximal Support Vector Machines via Generalized Eigenvalues, which uses knowledge of the uncertainty associated with the membership of a data sample to a given class, to improve generalization.

References

1. Mangasarian, O.L., and Wild,E.W.: Multisurface Proximal Support Vector Classification via Generalized Eigenvalues. Data Mining Institute Technical Report. 04-03 June 2004.
2. Fung, G., and Mangasarian, O.L.: Proximal Support Vector Machines. Proc. KDD-2001 August 26-29, San Francisco (2001) 77-86.
3. Burges, C.: A Tutorial on Support Vector Machines for Pattern Recognition. Data Mining and Knowledge Discovery, **2** (1998).
4. Tikhonov, A.N., and Arsenin,V.Y.: Solution of Ill Posed Problems, John Wiley and Sons, New York, 1977.
5. Parlett,B.N.: The symmetric Eigen value Problem, SIAM, Philadelphia, PA, 1998.
6. http://www.mathworks.com
7. Blake, C.L, and Merz,C.J.: UCI Repository for Machine Learning database, Irvine, CA: University of California, Department of Information and Computer Sciences. On-line at http://www.ics.uci.edu/ mlearn/MLRepository.html
8. Jayadeva, Khemchandani, R. and Chandra. S.: Fast and Robust Learning Through Fuzzy Linear Proximal Support Vector Machines. Journal of Neurocomputing. **61**(2004) 401-411.