Classification of Imprecise Data Using Interval Fisher Discriminator

Jafar Mansouri,[1,*] Hadi Sadoghi Yazdi,[2,†] Morteza Khademi[1,‡]
[1]*Electrical Engineering Department, Ferdowsi University of Mashhad, Mashhad, Islamic Republic of Iran*
[2]*Computer Engineering Department, Ferdowsi University of Mashhad, Mashhad, Islamic Republic of Iran*

In this paper, an imprecise data classification is considered using new version of Fisher discriminator, namely interval Fisher. In the conventional formulation of Fisher, elements of within-class scatter matrix (related to covariance matrix between clusters) and between-class scatter matrix (related to covariance matrix of centers of clusters) have single values; but in the interval Fisher, the elements of matrices are in the interval form and can vary in a range. The particle swarm optimization search method is used for solving a constrained optimization problem of the interval Fisher discriminator. Unlike conventional Fisher with one optimal hyperplane, interval Fisher gives two optimal hyperplanes thereupon three decision regions are obtained. Two classes with regard to imprecise scatter matrices are derived by decision making using these optimal hyperplanes. Also, fuzzy region lets us help in fuzzy decision over input test samples. Unlike a support vector classifier with two parallel hyperplanes, interval Fisher generally gives us two nonparallel hyperplanes. Experimental results show the suitability of this idea.

1. INTRODUCTION

In many theoretical and practical applications, there are imprecise, incomplete, and noisy data. These may be due to the lack of enough knowledge, low accuracy of measurement devices, noise, and so on. Therefore, it is necessary to deal with these data in their relevant applications.

A question is propounded in the field of classification: How imprecise data can be applied in the classification and clustering? Interval and fuzzy numbers are recently used for this purpose as Hiremath and Prabhakar[1] proposed symbolic kernel Fisher discriminant (KFD) analysis with radial basis function kernel. It is important to note the application of Hausdorff distance for conversion of interval data to a

*Author to whom all correspondence should be addressed: e-mail: afar.mansouri@gmail.com.
†e-mail: h-sadoghi@um.ac.ir.
‡e-mail: khademi@um.ac.ir.

INTERNATIONAL JOURNAL OF INTELLIGENT SYSTEMS, VOL. 26, 718–730 (2011)

View this article online at wileyonlinelibrary.com. • DOI 10.1002/int.20490

crisp value in the KFD and used classifier.[1] They adapted the symbolic KFD to extract interval-type nonlinear discriminating features. They applied a new RBF kernel function to map input data into highly nonlinear data in a high-dimensional space. Then, they extracted interval-type nonlinear discriminating features.

Yang et al.[2] presented a new model, which used a fuzzification of the Choquet integral for solving the classification problem, involving heterogeneous fuzzy data, including crisp data, fuzzy numbers, interval values, and linguistic variables. The proposed model acted as a projection tool, which can map a high-dimensional heterogeneous fuzzy data to a crisp virtual value on a real axis, so that the classification problem in high-dimensional heterogeneous fuzzy datum space was simplified to that in one-dimensional crisp data space.

Huhn and Hullermeie[3] introduced a fuzzy-rule-based classification method called fuzzy round robin, repeated incremental pruning to produce error reduction (RIPPER), (FR3). As the name suggests, FR3 builds upon the RIPPER algorithm. More specifically, in the context of polychotomous classification, it used a fuzzy extension of RIPPER as a base learner. A key feature of FR3, in comparison with its nonfuzzy counterpart (R3 learner that has recently been introduced in the literature), was its ability to represent different facets of uncertainty involved in a classification decision in a more faithful way. FR3 provides the basis for implementing "reliable classifiers" that might abstain from a decision when not being sure enough, or at least indicated that a classification was not fully supported by the empirical evidence at hand.

Zhao et al.[4] presented an interval set classification based on support vector machines (SVM). To make incomplete information patterns that could be classified correctly by trained SVM, they extended the input vectors of SVM to interval input vectors where each unmeasured attribute of input was represented by an interval that includes its possible value. Also, the operation in the classification function was extended to interval operation correspondingly. For the incomplete information input, the value of the classification function was the interval operation result. When the output of classification function satisfied the classification condition, the incomplete information input pattern could be classified correctly. Meanwhile, the attribute value prior to knowledge about interval representation could be utilized fully in the proposed algorithm. Lu et al.[5] also presented another algorithm using SVM. In Ref. 6, a new version of support vector machine with fuzzy data was presented. In that work, fuzzy input data were modeled on the form of fuzzy cost.

Fernández et al.[7] considered the problem of classification with imbalanced data sets, that is, some data sets with a different class distribution among their patterns. This work improved the behavior of fuzzy rule based classification systems (FRBCSs) in the framework of imbalanced data sets by means of a tuning step. They adapted the 2-tuples-based genetic tuning approach to classification problems, showing the good synergy between this method and some FRBCSs. The 2-tuples-based genetic tuning increased the performance of FRBCSs in all types of imbalanced data.

Li et al.[8] proposed a probabilistic method for classification of noisy data. They presented two classifiers, probabilistic kernel Fisher and probabilistic Fisher, based on a probabilistic model proposed by Lawrence and Schölkopf[9] (in Ref. 9, the class noise was assumed to have been generated by a classification noise process. In this kind of noise process, the input feature distribution remained the same but their labels were independently and randomly reversed with some probabilities). The proposed methods in Ref. 8 were able to tolerate class noise and extended the earlier work of Lawrence and Schölkopf[9] in two ways. First, it presented a novel incorporation of their probabilistic noise model in the kernel Fisher discriminator; second, the distribution assumption previously made was relaxed in their work. The proposed method improved standard classifiers in noisy data sets and achieved larger performance gain in non-Gaussian data sets and small size data sets.

Yang et al.[10], Hung and Yang[11], and Yang et al.[12] presented more precise fuzzy C-Means clustering over symbolic and fuzzy data. Some other related works can be found in Refs. 13 and 14.

In this paper, Fisher discriminator over interval data (presentation of imprecise data) as a classifier is presented. Two optimal hyperplanes are obtained from the interval Fisher that led to three decision regions. This kind of decision results in fuzzy decision. This paper deals with modeling of imprecise data with interval data, solving interval quadratic problem using particle swarm optimization (PSO), and fuzzy decision and classification with utilization of interval Fisher discriminator (IFD).

The rest of the paper is organized as follows. Section 2 explains the Fisher linear discriminator, and Section 3 presents the proposed method. Experimental results are considered in Section 4. Finally, Section 5 concludes the paper and presents the future works.

2. FISHER LINEAR DISCRIMINATOR

Fisher linear discriminator uses linear transformation of predictor variables that provides a more accurate discrimination. Classes are separated satisfactorily if a direction can be found to project data on it when (a) between-class variance is maximized and (b) within-class variance is minimized. To simplify, suppose that there are two classes (and then the method can be generalized to multiple classes). Linear projection is $y = w^T x + w_0$. In the new space of y, between-class and within-class variances are calculated. Between-class variance can be presented by $(\tilde{m}_2 - \tilde{m}_1)^2$ (where $\tilde{m}_1$ and $\tilde{m}_2$ are means of classes c_1 and c_2, respectively, in the transform space) and within-class variance for classes c_1 and c_2 are shown to be in the form of $\tilde{s}_1^2 + \tilde{s}_2^2$ (where $\tilde{s}_1^2$ and $\tilde{s}_2^2$ are covariances of two classes). So, the Fisher criteria can be written in the following form:

$$\max J_F = \frac{(\tilde{m}_2 - \tilde{m}_1)^2}{\tilde{s}_1^2 + \tilde{s}_2^2} \tag{1}$$

Between-class variance is calculated as

$$(\tilde{m}_2 - \tilde{m}_1)^2 = w^T (m_2 - m_1)(m_2 - m_1)^T w \tag{2}$$

where m_1 and m_2 are means of class c_1 and c_2, respectively, in input space. Between-class scatter matrix is defined as

$$S_B = (m_2 - m_1)(m_2 - m_1)^T \tag{3}$$

So, the numerator of fraction is $w^T S_B w$. For simplification of denominator, it can be written as

$$\begin{aligned}\tilde{s}_1^2 &= \sum_{y_i : x_i \in R_1} (y_i - \tilde{m}_1)^2 = \sum_{x_i \in R_1} (w^T x_i - w^T m_1)^2 \\ &= w^T \left(\sum_{x_i \in R_1} (x_i - m_1)(x_i - m_1)^T \right) w = w^T S_1 w\end{aligned} \tag{4}$$

and

$$\begin{aligned}\tilde{s}_2^2 &= \sum_{y_i : x_i \in R_2} (y_i - \tilde{m}_2)^2 = \sum_{x_i \in R_2} (w^T x_i - w^T m_2)^2 \\ &= w^T \left(\sum_{x_i \in R_2} (x_i - m_2)(x_i - m_2)^T \right) w = w^T S_2 w\end{aligned} \tag{5}$$

where R_1 and R_2 are the sets of all data, which belong to Classes 1 and 2, respectively. Within-class scatter matrix is defined as $S_w = S_1 + S_2$. So, the criteria can be written in the following form:

$$\max \;\; J(w) = \frac{w^T S_B w}{w^T S_w w} \tag{6}$$

For solving the above fractional optimization problem, it is simplified as

$$\begin{aligned}&\max_w \;\; J(w) = w^T S_B w \\ &\text{s.t.} \;\; w^T S_w w = 1\end{aligned} \tag{7}$$

So $w = S_w^{-1}(m_2 - m_1)$ and optimum hyperplane is

$$\begin{aligned}y(x) &= w^T x - \frac{1}{2}(\tilde{m}_1 + \tilde{m}_2) = w^T x - \frac{1}{2} w^T (m_1 + m_2) \\ &= \left(S_w^{-1}(m_2 - m_1)\right)^T \left(x - \frac{1}{2}(m_1 + m_2) \right).\end{aligned} \tag{8}$$

The original problem of Equation 6 is also equivalent to

$$\begin{aligned} &\min_{w} \ Z(w) = w^T S_w w \\ &\text{s.t.} \ \ w^T S_B w = 1 \end{aligned} \tag{9}$$

where $Z(w)$ is the cost function that should be minimized and $w^T S_B w = 1$ is the constraint.

3. INTERVAL FISHER DISCRIMINATOR

The interval Fisher discriminator is similar to the Fisher linear discriminator with this distinction that the elements of matrices S_w and S_B are not single values, but they are in the interval form. It means that S_w and S_B can vary in a range. This case may occur because of noisy data, incomplete data, and any other reason that can lead to inaccurate and imprecise knowledge of parameters. These parameters are shown with S_w^I and S_B^I, where the superscript "I" is a representative of interval. It is worth mentioning that in fact each element of S_w^I (or S_B^I) is a random variable that can vary within its interval range. So, the problem that should be solved as

$$\begin{aligned} &\min_{w} \ Z(w) = w^T S_w^I w \\ &\text{s.t.} \ \ w^T S_B^I w = 1 \end{aligned} \tag{10}$$

Equivalently, Equation 10 can be written as

$$\begin{aligned} &\min_{w} \ Z(w) = \sum_{i=1}^{n} \sum_{j=1}^{n} s_{w,ij} w_i w_j \\ &\text{s.t.} \ \ \sum_{i=1}^{n} \sum_{j=1}^{n} s_{B,ij} w_i w_j = 1 \end{aligned} \tag{11}$$

where $s_{w,ij} \in s_{w,ij}^I$, $s_{B,ij} \in s_{B,ij}^I$ and the interval matrices $S_w^I = (s_{w,ij}^I)_{n\times n}$ and $S_B^I = (s_{B,ij}^I)_{n\times n}$ are symmetric positive semidefinite. Also, matrices are $n \times n$, in which n is the number of classes that should be separated. For two classes, $S_w^I = [\begin{smallmatrix} a & b \\ b & c \end{smallmatrix}]$ is a within-class scatter matrix, where a, b, and c are three numbers. For example, $a \in [4.6 \ \ 4.8]$, $b \in [0.1 \ \ 0.3]$, and "c" is a number in the interval of $[1.9 \ \ 2.1]$.

According to our knowledge, this kind of problem, optimization of an interval quadratic programming problem, has not been solved analytically so far and it is a very demanding problem. Li and Tian[15] and Liu and Wang[16] have presented numerical solution methods for interval quadratic programming. But their methods are applicable to optimization problems with linear interval constraints. Their methods cannot be used for the proposed problem in this paper because the proposed problem is confronted with quadratic interval constraint. It is useful to discuss more on Equation 10. For each known S_w and S_B, the Equation 10 is converted to the Equation 9. It means that for each $S_w \in S_w^I$ and $S_B \in S_B^I$, there is a specific w that minimizes the cost function with respect to related constraint.

Clearly, min $Z(w)$ is an interval. To find the interval that bounds the objective values, it suffices to find the lower bound and the upper bound of the objective values of Equation 10. The lower bound of min $Z(w)$ is indicated with Z^L and the upper bound with Z^U. Denote $S = \{(S_w, S_B) | S_w \in S_w^I, S_B \in S_B^I\}$. Therefore, the following formulas can be written:

$$
\begin{aligned}
Z^L &= \min_{(S_w, S_B \in S)} \min_w Z(w) = w^T S_w w \\
&\text{s.t.} w^T S_B w = 1
\end{aligned}
\tag{12}
$$

and

$$
\begin{aligned}
Z^U &= \max_{(S_w, S_B \in S)} \min_w Z(w) = w^T S_w w \\
&\text{s.t.} w^T S_B w = 1
\end{aligned}
\tag{13}
$$

As a result, between all $S_w \in S_w^I$ and$S_B \in S_B^I$, two sets are found: One set is $S_{\min} = \{S_{w1}, S_{B1}\}$ that belongs to Z^L and $S_{\max} = \{S_{w2}, S_{B2}\}$ that belongs to Z^U. As it was pointed earlier, there is no analytical solution, like the Lagrange method, for this optimization problem. To overcome this problem, one of the popular metaheuristic algorithms is applied in this work. PSO is a powerful tool for solving this kind of problem, and the solution of this problem can be found easily with good approximation (obviously, other metaheuristic algorithms such as genetic algorithm can also be used). PSO is used two times in the program: one for finding $w1$ related to $S_{\min}$ and the other for finding $w2$ related to $S_{\max}$. Consequently, there is a range, that is $[w1, w2]$, that w can vary between $w1$ and $w2$. Of course, the answered range of w in Equation 10, namely $[w_{\min}, w_{\max}]$, maybe more than $[w1, w2]$, but since there is no analytical algorithm for finding the largest range of$w$, the later domain can also help in classification.

The code of program can be written as

- Finding $w1$

1. Initialize K randomly $S_w \in S_w^I$ and $S_B \in S_B^I$, and finding the related w and cost(k) $= Z_k(w)$ using an analytical method (e.g., the Lagrange method), where K is the population size for PSO.
2. PSO receives $S_w \in S_w^I$ and $S_B \in S_B^I$, and cost(k)and generates new K of $S_w \in S_w^I$ and $S_B \in S_B^I$.
3. For new $S_w \in S_w^I$ and$S_B \in S_B^I$, find the related w and cost(k) $= Z_k(w)$ using an analytical method.

4. If stopping criteria are satisfied go to 5, else go to 2.
5. The output of PSO is the$S_{\min}$, and the related w is$w1$.

- Finding $w2$

1. Initialize K randomly $S_w \in S_w^I$ and $S_B \in S_B^I$, and finding the related w and cost(k) $= -Z_k(w)$ using an analytical method, where K is the population size for PSO.
2. PSO receives $S_w \in S_w^I$ and $S_B \in S_B^I$, and cost(k) and generates new K of $S_w \in S_w^I$ and $S_B \in S_B^I$.
3. For new $S_w \in S_w^I$ and$S_B \in S_B^I$, find the related w and cost(k) $= -Z_k(w)$ using an analytical method.
4. If stopping criteria are satisfied go to 5, else go to 2.
5. The output of PSO is the$S_{\max}$, and the related w is $w2$.

It should be noted that "stopping criteria" are the parameters in the PSO algorithm that lead to the ending of the PSO algorithm. For example, maximum number of runs for the algorithm can be a stopping criterion.

It is worth mentioning in detail the concept of interval Fisher classifier. First, suppose that there is a conventional Fisher classifier, that is, Equation 9. Solving this optimization problem yields us one w that gives us one line or hyperplane, for n classes, (namely Equation 8) that it can linearly separate classes. Now consider the interval case. Solving Equations 12 and 13 gives us two hyperplanes. To simplify the discussion, suppose there are two classes. With these hyperplanes, for any data, its degree of belonging to any of the classes can be determined. This subject is more clarified in the Experimental results section. It should be noticed that the IFD is just applicable to classes that are linearly separable.

4. EXPERIMENTAL RESULTS

The proposed classifier, IFD, is the first interval type of linear discriminator analysis (LDA). The proposed method and its properties are illustrated with some examples. It is worth noting that the data of classes presented in Sections 4.1 and 4.2 are generated by the "rand" function of MATLAB in different manners.

4.1. IFD over Low-Imprecision Training Samples

At first, two classes of data are generated. S_w^I and S_B^I have 20% tolerance; that is, the lower bound of S_w^I is 20% smaller than the center of S_w^I interval and the upper bound of S_w^I is 20% larger than the center of S_w^I interval (this case is for S_B^I, too). Then, using PSO $w1$ and$w2$ are found, and after that the resultant lines $y_1 = w1^T x + w01$ and $y_2 = w2^T x + w02$ are plotted. Figure 1 shows two classes of data, one class with (.)symbols and the other class with (+) symbols.

In this figure, the bottom line is related to line of $w1$ and the upper line is related to$w2$. For instance in this example, $w1 = [0.4821 - 0.4300]^T$, $w01 = 0.0708$; and

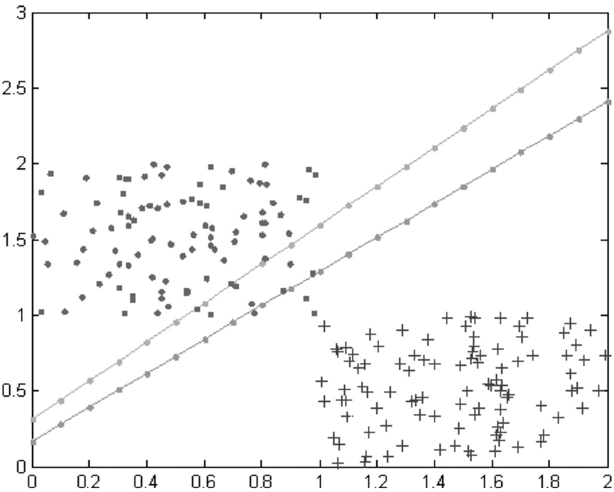

Figure 1. An example of two classes and two classifier lines.

the other vector is $w2 = [0.5455 - 0.4347]^T$, $w02 = 0.1075$. As it is observed, these two lines separate two classes with good approximation and also give us a margin (the region between two lines). This margin indicates that data in this region have fuzzy belonging and for any test data, its degree of belonging to any of classes can be determined. For any data below the lower line, the data belong to Class 1 and its degree of belonging is 100%. Also for any data above the upper line, the data belong to Class 2 and its degree of belonging is 100%. For the data inside the margin, it has fuzzy belonging. In fact, the result of classification is fuzzy. For more explanation of classification, Figures 2–4 show three cases.

In Figure 2, the data belong to Class 1. First of all, for the data,x, calculate $y_1 = w1^T x + w01$ and $y_2 = w2^T x + w02$, (for instance in the experiment, $y_1 = -0.4030$ and $y_2 = 0.3444$). Also for other x in this region, y_1 is negative and y_2 is positive. All data in this region have this characteristic.

Now consider Figure 3, where the data belong to Class 2. For data x, calculate $y_1 = w1^T x + w01$ and $y_2 = w2^T x + w02$, (for instance in the experiment, $y_1 = 0.6365$ and $y_2 = -0.9352$). Also for the other x in this region, y_1 is positive and y_2 is negative. All data in this region have this property.

Now consider Figure 4, where data are placed in the fuzzy region. y_1 and y_2 have the same sign in this region, (for instance in this example $y_1 = -0.556$ and $y_2 = -0.1440$). For other data in this region, this property exists, that is, if y_1 and y_2 have negative sign, x is in the fuzzy region. For data in this region, its degree of belonging to any of the classes can be calculated. For this case, the Euclidean distance of data with each of the lines is determined. Suppose d_1 is distance of x from the lower line and d_2 is distance of x from the upper line (in this example $d1 = 0.0995$ and $d2 = 0.1607$). So $d1\big/(d1 + d2)$ is degree of belonging x to the Class 1, and $d2\big/(d1 + d2)$ is degree of belonging x to the Class 2.

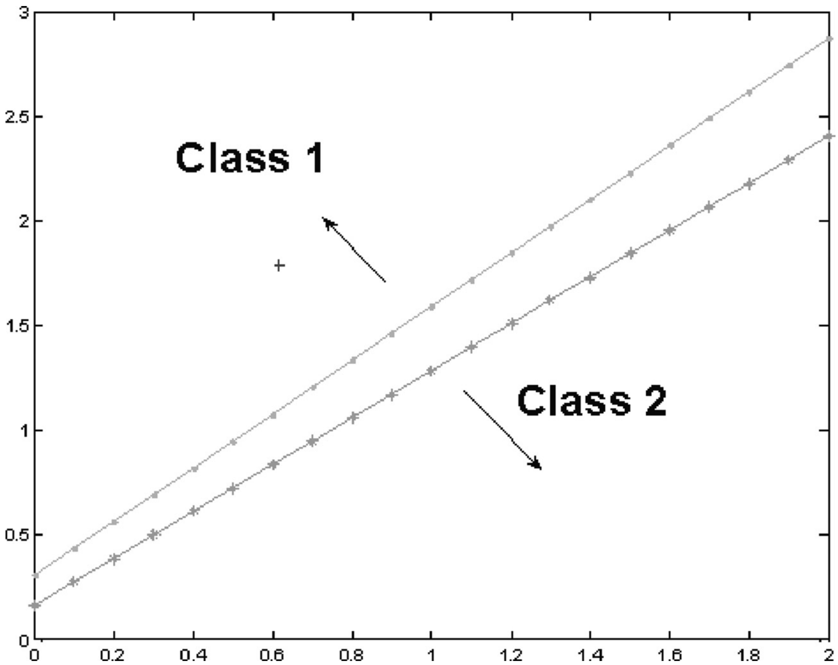

Figure 2. Test sample in the region of Class 1.

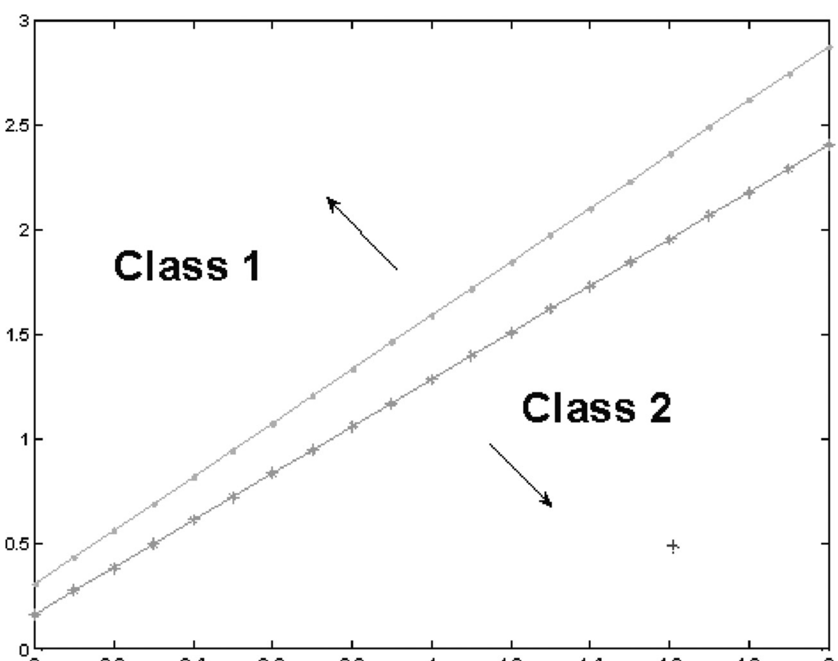

Figure 3. Test sample in the region of Class 2.

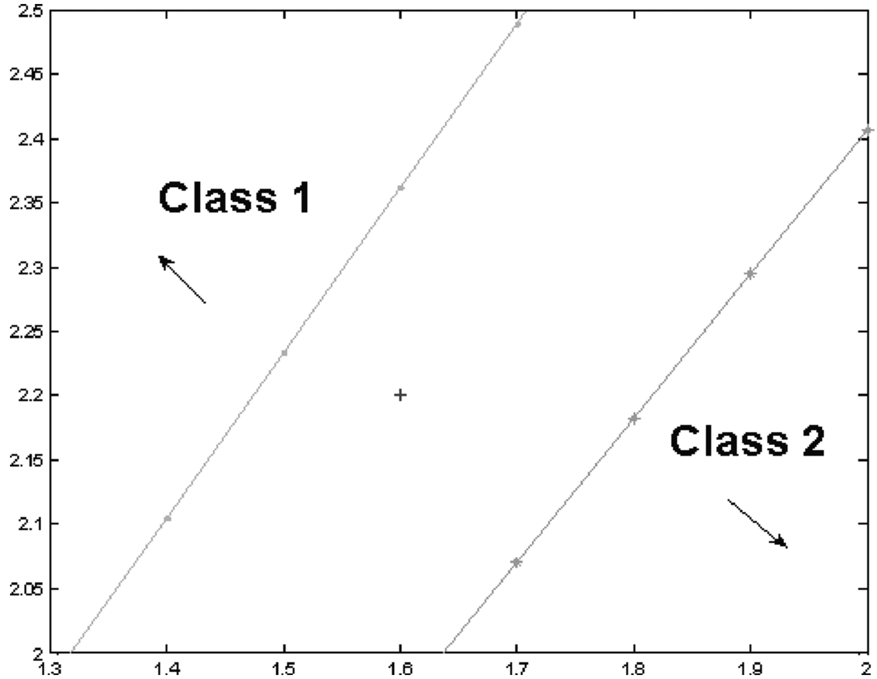

Figure 4. Test sample in the fuzzy region.

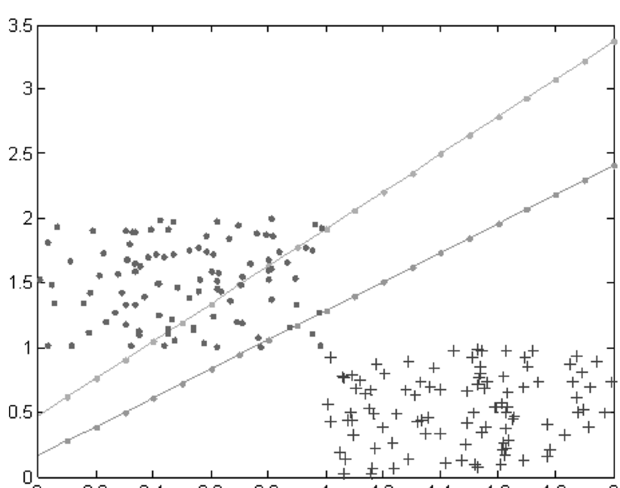

Figure 5. An example of classifier lines with increased range of S_w^I and S_B^I.

4.2. IFD over High-Imprecision Training Samples

Now consider the case, where the range of S_w^I and S_B^I is larger than the previous case; S_w^I and S_B^I have 30% tolerance. The optimal lines are shown in Figure 5. As

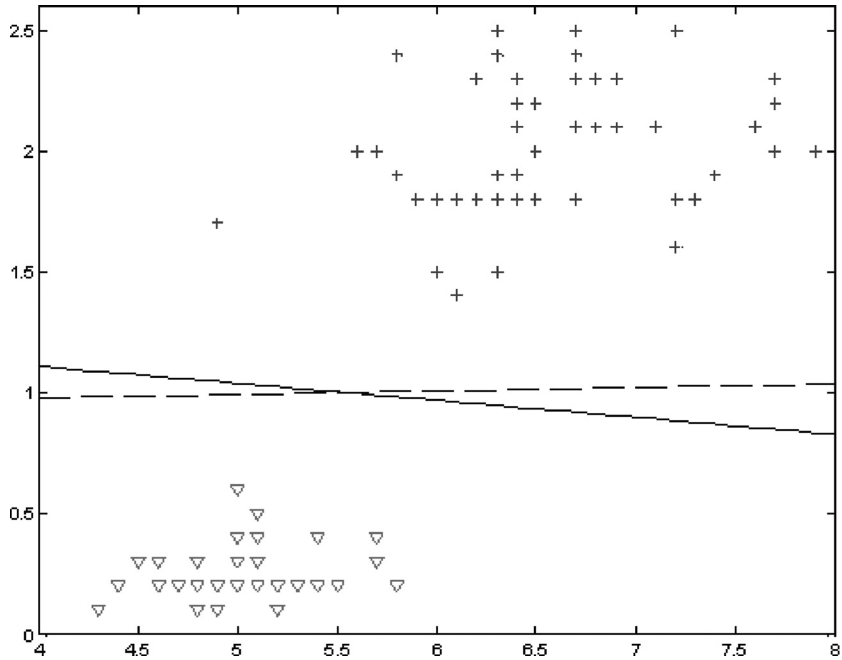

Figure 6. Classifier lines for Classes 1 and 3 in the iris data set.

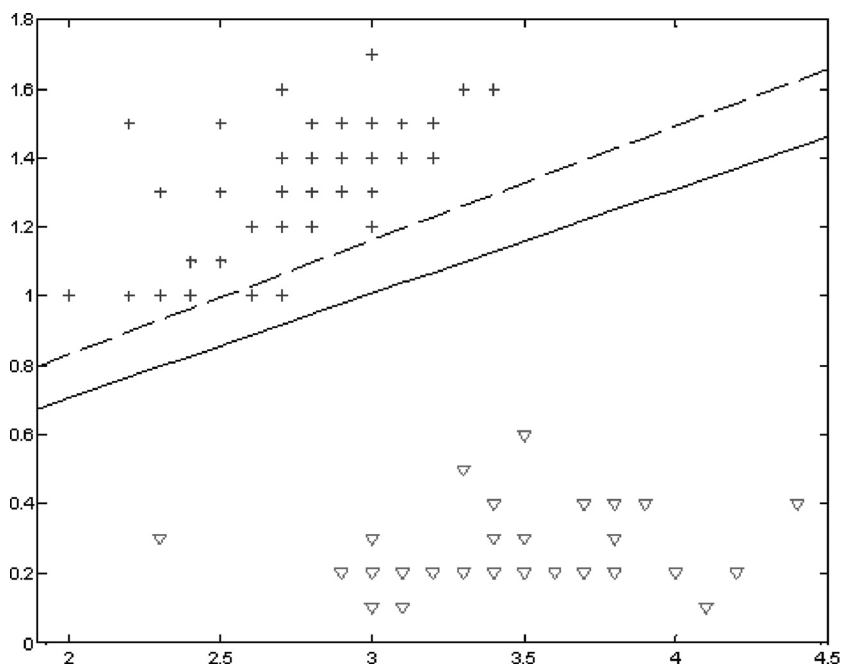

Figure 7. Classifier lines for Classes 2 and 4 in the iris data set.

it is observed, the line related to Z^L has not changed but the line of that Z^U has changed and the fuzzy region has been expanded. So with increasing the range of S_w^I and S_B^I, the fuzzy region becomes enlarged.

4.3. IFD over Real Samples

In Figures 6 and 7, data of two classes of iris data (UCI Machine learning repository[17]) are shown. In these simulations, S_w and S_B have 10% tolerance. In these figures, the solid line is related to Z^U and the dashed line is related to Z^L. As it is seen in these figures, these lines separate the classes.

5. CONCLUSION AND FUTURE WORKS

In this paper, a new version of Fisher discriminator namely IFD was introduced. In IFD, the elements of the between-class scatter matrix (related to covariance matrix of center of clusters) and within-class scatter matrix (related to covariance matrix between clusters) are in the interval form. One of the important applications of IFD can be in classification of imprecise and incomplete data. Because there is no analytical solution for minimization of an interval quadratic programming, problem one of the metaheuristic methods, PSO, was used for solving the optimization problem of IFD. Unlike conventional Fisher classifier with one optimal hyperplane, IFD gave two optimal hyperplanes and three decision regions were obtained. Two classes with regard to imprecise scatter matrices were derived using these two optimal hyperplanes. Also, the fuzzy region let us help in fuzzy decision over input test samples. In fact, IFD is a kind of fuzzy classification. An important result that could be inferred is that with increasing of imprecision of parameters, the fuzzy region increased.Experimental results validated the ability of this proposed method. For the future work, we will pursue comparison of the proposed method with other methods such as fuzzy method over practical data sets, presentation of kernel IFD, and presentation of analytical approach for solving IFD and kernel-type of IFD.

References

1. Hiremath PS, Prabhakar CJ. A new kernel function to extract nonlinear interval type features using symbolic kernel Fisher discriminant method with application to face recognition. In: IEEE Int Symp on ISBAST; 2008. pp 1–7.
2. Yang R, Wang RZ, Heng PA, Leung KS. Classification of heterogeneous fuzzy data by Choquet integral with fuzzy-valued integrand. IEEE Trans Fuzzy Syst 2007;15:931–942.
3. Huhn JC, Hullermeier E. FR3: a fuzzy rule learner for inducing reliable classifiers. IEEE Trans Fuzzy Syst 2009;17:138–149.
4. Zhao Y, He Q, Chen Q. An interval set classification based on support vector machines. In: IEEE Int Conf ICAS/ICNS2005; 2005. pp 81–86.

5. Lu C, Li XW, Pan Hb. Application of SVM and fuzzy set theory for classifying with incomplete survey data. In: IEEE Int Conf on Service Systems and Service Management; 2007. pp 1–4.
6. Sadoghi Yazdi H., Royani T, Sadoghi Yazdi M, Effati S. Fuzzy cost support vector regression. Int J Intell Syst Technol 2008;3:238–243.
7. Fernández A, Del Jesus MJ, Herrera F. On the 2-tuples based genetic tuning performance for fuzzy rule based classification systems in imbalanced data-sets. Inform Sci 2010;180:1268–1291.
8. Li Y, Wessels LFA, De Ridder D, Reinders MJT. Classification in the presence of class noise using a probabilistic kernel Fisher method. Pattern Recognit 2007;40:3349–3357.
9. Lawrence ND, Schölkopf B. Estimating a kernel Fisher discriminant in the presence of label noise, In: Proc 18th Int Conf on Machine Learning. San Francisco, CA: Morgan Kaufmann; 2001. pp 306–313.
10. Yang MS, Hwang PY, Chen DH. Fuzzy clustering algorithm for mixed feature variables. Fuzzy Sets Syst 2004;141:301–317.
11. Hung WL, Yang MS. Fuzzy clustering on LR-type fuzzy numbers with an application in Taiwanese tea evaluation. Fuzzy Sets Syst 2005;150:561–577.
12. Yang MS, Hung WL, Cheng FC. Mixed-variable fuzzy clustering approach to part family and machine cell formation for GT applications. Int J Prod Econ 2006;103:185–198.
13. Ishibuchi H, Nojima Y. Comparison between fuzzy and interval partitions in evolutionary multiobjective design of rule-based classification systems, In: IEEE Int Conf on Fuzzy Systems; 2005. pp 430–435.
14. Lin CT, Yeh CM, Liang SF, Chung JF, Kumar N. Support-vector-based fuzzy neural networks for pattern classification. IEEE Trans Fuzzy Syst 2006;14:31–41.
15. Li W, Tian X. Numerical solution method for general interval quadratic programming. Appl Math Comput 2008;202:589–595.
16. Liu ST, Wang RT. A numerical solution method to interval quadratic programming. Appl Math Comput 2007;189:1274–1281.
17. UCI Machine learning repository, 2010, Available at http://archive.ics.uci.edu/ml/datasets/Low+Resolution+Spectrometer (donated by: MARSHALL%PLU@io.arc.nasa.gov).