International Journal of Computer Theory and Engineering, Vol. 1, No. 5, December, 2009
1793-8201

Fisher over Fuzzy Samples

B. Rajaei, H. Sadoghi Yazdi, S. Effati

***Abstract*— One of the main problems when handling the real world problems is the uncertainty degree of input data. Uncertainty factor can be a result of random variables existence, incomplete or inaccurate data, and approximations instead of measurements or incomparability of data (resulting from varying measurement or observation conditions). Interval and fuzzy numbers generally use for representation of real data. There are two main innovations in this paper: I) Classification of real data using fisher discriminator (FD), and II) Quadratic programming of FD problem with fuzzy parameters has led us to a quadratic fuzzy objective function and quadratic fuzzy constraints, that is solved for the first time in this paper. The proposed Fuzzy FD (FFD) obtain new version of classifier with two new points. I) Three region of decision are given include class 1, class 2, outlier class. II) We can classify real data with given uncertainty degree. Experimental results are performed over and Power of the FFD is seen nicely.**

***Index Terms*— Fisher discriminator; Fuzzy data; Fuzzy quadratic programming problems; Real data classification.**

I. Introduction

Pursuing of effective methods for real data analysis is considered with researchers recently. Progress in fuzzy mathematics encourage engineering and other scientist in application fields for presentation of real models. For example A. Bigand 2009 used interval-valued fuzzy sets for image filtering [1]. Interval-valued fuzzy sets make it possible to take into account the total uncertainty inherent to image processing, and particularly noise removal is considered. A. Salski 2007 [2] found ecological data as high uncertainty data and presented the extension and implementation of fuzzy data clustering method proposed by Yang and Liu 1999 [3]. R. Yang et al. [4] introduced a method for mapping 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 is simplified to that in one dimensional crisp data space.

In this paper we want to introduce a classifier based on fisher discriminator suited for working with real data. Fisher discriminator (FD) is a suitable approach in the field of pattern recognition for classification. FD is based on maximization of between class variance and minimization of within class scatter in linear transformation domain. FD includes quadratic programming with quadratic constraints. We study FD with fuzzy scatter matrix for within class and fuzzy covariance matrix for between class scatter. So, we must solve fuzzy quadratic programming with fuzzy quadratic constraint. Specifically, quadratic programming has been widely used in solving real problems and several efforts reported in literature developing efficient algorithms for solving this types of problems where crisp parameters are used [5,6]. In 2007 S.T. Liu and R.T. Wang [7] solved the problem of quadratic programming for interval parameters by formulating a pair of two level mathematical programs to calculate the upper bound and lower bound of the objective values of the interval quadratic program. After that, in 2009 S.T. Liu [8] solved the same problem by the same method, for dealing with fuzzy parameters. But the main limitation of Liu’s study is linearity assumption of constraints, something that is not confirmed in fisher problem. In the next sections we will represent our solution to quadratic objective functions and quadratic constraints with fuzzy parameters.

Organization of this paper includes as follows, Section 2 pay to preliminaries. Section 3 appropriates to explanation of FFD and solving fuzzy quadratic programming problems. Section 4 shows experimental results of the presented work. Finally, section 5 concludes the paper.

II. Preliminaries

A. *Fisher Discriminator Algorithm*

Fisher linear discriminator Finds linear transformation of predictor variables which provides a more accurate discrimination. Classes are separated nicely if we can find the direction to project data on so that (a) between classes variance is maximized (b) within class variance is minimized. Linear projection is $y = w^T x + w_0$. In the new space of y, between class and within class variance are calculated. Between class variance can be presented by $(\tilde{m}_2 - \tilde{m}_1)^2$, (where $\tilde{m}_1$, $\tilde{m}_2$ are means of classes w_1, w_2 respectively in the transform space) and within class variance for classes w_1, w_2 are shown to the form of $\tilde{s}_1^2 + \tilde{s}_2^2$, (where $\tilde{s}_1^2$, $\tilde{s}_2^2$ are covariance of two classes). So Fisher criteria can be formulated as

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

Between classes variance is calculated as follows:

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

Manuscript received August 1, 2009.

Boshra Rajaei (corresponding author) is with Computer Department of Ferdowsi University of Mashhad, Iran (phone: +985115028601; email: boshra.rajaei@stu-mail.um.ac.ir).

Hadi Sadoghi Yazdi is with Computer Department of Ferdowsi University of Mashhad, Iran (email: h-sadoghi@um.ac.ir, sadoghi@sttu.ac.ir).

Sohrab Effati is with Mathematics Department of Ferdowsi University of Mashhad (email: s-effati@um.ac.ir, effati@sttu.ac.ir).

where m_1, m_2 are means of class w_1, w_2 respectively in input space. If we define

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

Then numerator of fraction is $w^T S_B w$. For simplification of denominator, we have:

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

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

We define $S_w = S_1 + S_2$ so we can write criteria to following form,

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

For solving above fractional optimization problem we simplify it,

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

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

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

III. THE PROPOSED FUZZY FISHER DISCRIMINATOR

The well-known FD and other conventional classifiers are based on precise description of input data. But what about analyzing real world problems? As we know, in most situations, the real world is too complicated to obtaining precise descriptions and fuzziness must be introduced in reasonable and reality based models. In real world problems input data can have different interpretations. For example, in Fig. 1 we plotted graphically some fuzzy numbers representation. Here, triangular fuzzy numbers that are more applicable and more common in literature, used in simulations but formulations are general.

Intuitively, if input data in a classification or clustering problem is fuzzy, then S^B, S^W and the objective value should be fuzzy as well. So, the conventional FD objective function, introduced in the previous section, turn into fuzzy fisher discriminator (FFD) problem. Suppose we approximately known S^B and S^W coefficients and representing them in fuzzy sets format as $\tilde{S}^B$ and $\tilde{S}^W$ using $m_{\tilde{S}^B}$ and $m_{\tilde{S}^W}$ membership functions, respectively. We have

$$\begin{aligned}\tilde{S}_{ij}^B &= \{(s_{ij}^B, m_{\tilde{S}_{ij}^B}) \mid s_{ij}^B \in U(\tilde{S}_{ij}^B)\} \\ \tilde{S}_{ij}^W &= \{(s_{ij}^W, m_{\tilde{S}_{ij}^W}) \mid s_{ij}^W \in U(\tilde{S}_{ij}^W)\}\end{aligned} \tag{9}$$

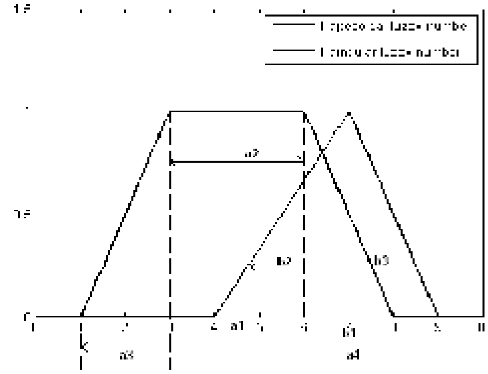


Figure 1 – Two samples of fuzzy numbers

where $U(\tilde{S}_{ij}^B)$ and $U(\tilde{S}_{ij}^W)$ are the supports of $\tilde{S}^B$ and $\tilde{S}^W$, respectively. Now the fuzzy fisher optimization problem can be formulated as follows:

$$\begin{aligned}&\max_W \quad \tilde{J}_F = \sum_{i=1}^n \sum_{j=1}^n s_{ij}^B w_i w_j \\ &s.t. \quad \sum_{i=1}^n \sum_{j=1}^n s_{ij}^W w_i w_j \le b\end{aligned} \tag{10}$$

or in minimization format

$$\begin{aligned}&\min_W \quad \tilde{J}_F = -\sum_{i=1}^n \sum_{j=1}^n s_{ij}^B w_i w_j \\ &s.t. \quad \sum_{i=1}^n \sum_{j=1}^n s_{ij}^W w_i w_j \le b\end{aligned} \tag{11}$$

Therefore fisher objective function is a quadratic programming problem with quadratic constraint. As mentioned before since we are only interested in finding direction of discriminator line, we can fix denominator of fisher fraction to any number. In constraint part we use the inequality format because it can simplify subsequent results and also don't reduce generality of fisher optimization function. Additionally, *b* can be any negative fuzzy number which adds fuzziness to the right-hand side of the above constraint and is defined using $\tilde{B}$ fuzzy set as follows

$$\tilde{B} = \{(b, m_{\tilde{B}}) \mid b \in U(\tilde{B})\} \tag{12}$$

Without loss of generality, $\tilde{S}^B$, $\tilde{S}^W$ and $\tilde{B}$ are assumed to be convex fuzzy numbers [1], as crisp values can be represented by degenerated membership functions which only have one value in their domains.

Based on the extension principle, we have

[1] A fuzzy set A in R^n is convex if and only if

$$m_A[l x_1 + l x_2] \ge \min[m_A(x_1), m_A(x_2)]$$

for all $x_1, x_2 \in R^n$ and all $l \in [0,1]$.

$$
\begin{aligned}
& m_{\tilde{J}_F}(j_F) = \sup_{S^B, S^W, B} \min\{ m_{\tilde{S}^B_{ij}}(s^B_{ij}), m_{\tilde{S}^W_{ij}}(s^W_{ij}), m_{\tilde{B}}(b) \mid \\
& \qquad j_F = J_F(S^B, S^W, B)\} \\
& m_{\tilde{J}_F}(j_F) = a \leftrightarrow m_{\tilde{S}^B_{ij}}(s^B_{ij}) \geq a, \\
& \quad m_{\tilde{S}^W_{ij}}(s^W_{ij}) \geq a, \quad m_{\tilde{B}}(b) \geq a
\end{aligned} \tag{13}
$$

and at least one of these three membership functions must be equal to $a, \forall i, j$. Finding $m_{\tilde{J}_F}$ is equivalent to finding the upper bound and lower bound of objective function at each α-cut[2], named J^U_a and J^L_a, respectively. These bounds can be expressed as

$$
\begin{aligned}
& J^U_a = \max\{J_F(S^B, S^W, B) \mid (S^B_{ij})^L_a \leq s^B_{ij} \leq (S^B_{ij})^U_a, \\
& \qquad (S^W_{ij})^L_a \leq s^W_{ij} \leq (S^W_{ij})^U_a, (B)^L_a \leq b \leq (B)^U_a, \forall i, j\}, \\
& J^L_a = \min\{J_F(S^B, S^W, B) \mid (S^B_{ij})^L_a \leq s^B_{ij} \leq (S^B_{ij})^U_a, \\
& \qquad (S^W_{ij})^L_a \leq s^W_{ij} \leq (S^W_{ij})^U_a, (B)^L_a \leq b \leq (B)^U_a, \forall i, j\}.
\end{aligned} \tag{14}
$$

From the above relations, the largest and smallest values for J^U_a and J^L_a can be determined from the following two-level mathematical programming models:

$$
J^U_a = \max_{\substack{(S^B_{ij})^L_a \leq s^B_{ij} \leq (S^B_{ij})^U_a \\ (S^W_{ij})^L_a \leq s^W_{ij} \leq (S^W_{ij})^U_a \\ (B)^L_a \leq b \leq (B)^U_a \\ \forall i, j}} \left\{ \begin{array}{ll} \min\limits_{W} & -\sum_{i=1}^{n}\sum_{j=1}^{n} s^B_{ij} w_i w_j \\ s.t. & \sum_{i=1}^{n}\sum_{j=1}^{n} s^W_{ij} w_i w_j \leq b \end{array} \right. , \tag{15}
$$

$$
J^L_a = \min_{\substack{(S^B_{ij})^L_a \leq s^B_{ij} \leq (S^B_{ij})^U_a \\ (S^W_{ij})^L_a \leq s^W_{ij} \leq (S^W_{ij})^U_a \\ (B)^L_a \leq b \leq (B)^U_a \\ \forall i, j}} \left\{ \begin{array}{ll} \min\limits_{W} & -\sum_{i=1}^{n}\sum_{j=1}^{n} s^B_{ij} w_i w_j \\ s.t. & \sum_{i=1}^{n}\sum_{j=1}^{n} s^W_{ij} w_i w_j \leq b \end{array} \right. \tag{16}
$$

However these two level models are not solvable in their current form and must be transforming into conventional one level program.

The upper bound model is a two level mathematical program in different optimization directions, that is, one for outer level for maximization and inner level for minimization. The Lagrangian dual problem of inner level is

$$
L(W, l) = -\sum_{i=1}^{n}\sum_{j=1}^{n} s^B_{ij} w_i w_j + l\left(\sum_{i=1}^{n}\sum_{j=1}^{n} s^W_{ij} w_i w_j - b\right). \tag{17}
$$

By differentiating with respect to W and vanishing the result we have

$$
\frac{\partial L}{\partial W} = 0 \rightarrow -2\sum_{i=1}^{n} s^B_{ij} w_i + 2l \sum_{i=1}^{n} s^W_{ij} w_i = 0, \; j = 1, \ldots, n \tag{18}
$$

So the dual form of the inner level can be written as

$$
\begin{aligned}
& \max_{W, l} \; -\sum_{i=1}^{n}\sum_{j=1}^{n} s^B_{ij} w_i w_j + l\left(\sum_{i=1}^{n}\sum_{j=1}^{n} s^W_{ij} w_i w_j - b\right) \\
& s.t. \quad -\sum_{i=1}^{n} s^B_{ij} w_i + l \sum_{i=1}^{n} s^W_{ij} w_i = 0, \; j = 1, \ldots, n \; . \\
& \qquad l \geq 0
\end{aligned} \tag{19}
$$

Hence, previously two level optimization problems can be reformulated as

$$
J^U_a = \max_{\substack{(S^B_{ij})^L_a \leq s^B_{ij} \leq (S^B_{ij})^U_a \\ (S^W_{ij})^L_a \leq s^W_{ij} \leq (S^W_{ij})^U_a \\ (B)^L_a \leq b \leq (B)^U_a \\ \forall i, j}} \left\{ \begin{array}{l} \max\limits_{W, l} \; [-\sum_{i=1}^{n}\sum_{j=1}^{n} s^B_{ij} w_i w_j \\ \qquad + l\left(\sum_{i=1}^{n}\sum_{j=1}^{n} s^W_{ij} w_i w_j - b\right)] \\ s.t. \; -\sum_{i=1}^{n} s^B_{ij} w_i + l \sum_{i=1}^{n} s^W_{ij} w_i = 0 \\ \quad , j = 1, \ldots, n \\ \qquad l \geq 0 \end{array} \right. \tag{20}
$$

Now that the two levels is unidirectional the one level upper bound problem model can be formulated as

$$
\begin{aligned}
& J^U_a = \max_{W, l} \; -\sum_{i=1}^{n}\sum_{j=1}^{n} s^B_{ij} w_i w_j + l\left(\sum_{i=1}^{n}\sum_{j=1}^{n} s^W_{ij} w_i w_j - b\right) \\
& s.t. \qquad -\sum_{i=1}^{n} s^B_{ij} w_i + l \sum_{i=1}^{n} s^W_{ij} w_i = 0, \; j = 1, \ldots, n \\
& \qquad (S^B_{ij})^L_a \leq s^B_{ij} \leq (S^B_{ij})^U_a \\
& \qquad (S^W_{ij})^L_a \leq s^W_{ij} \leq (S^W_{ij})^U_a \\
& \qquad (B)^L_a \leq b \leq (B)^U_a \\
& \qquad l \geq 0, \quad \forall i, j
\end{aligned} \tag{21}
$$

Since for a specific W vector, we have $w_i w_j \geq 0$ (elements are negative or positive simultaneously) in more cases, we can simplified the above optimization function as follows

$$
\begin{aligned}
& J^U_a = \max_{W, l} \; -\sum_{i=1}^{n}\sum_{j=1}^{n} (S^B_{ij})^L_a w_i w_j + l\left(\sum_{i=1}^{n}\sum_{j=1}^{n} (S^W_{ij})^U_a w_i w_j - B^L\right) \\
& s.t. \qquad -\sum_{i=1}^{n} (S^B_{ij})^L_a w_i + l \sum_{i=1}^{n} (S^W_{ij})^U_a w_i = 0, \; j = 1, \ldots, n \\
& \qquad l \geq 0
\end{aligned} \tag{22}
$$

Obviously, this is a conventional quadratic programming problem with quadratic constraints and the optimal value of J^U_a will be the upper bound of the FFD objective value in the specified α level.

In the lower bound case, since inner and outer levels have the same optimization direction we can combine them and form a one level problem:

[2] An α-cut of a fuzzy set A is a crisp set A_a that is
$$A_a = \{x \in U \mid m_A(x) \geq a\}$$

$$\begin{aligned} J_a^L &= \min_W \; -\sum_{i=1}^{n}\sum_{j=1}^{n} s_{ij}^B w_i w_j \\ s.t. \quad & \sum_{i=1}^{n}\sum_{j=1}^{n} s_{ij}^W w_i w_j \le b \\ & (S_{ij}^B)_a^L \le s_{ij}^B \le (S_{ij}^B)_a^U \\ & (S_{ij}^W)_a^L \le s_{ij}^W \le (S_{ij}^W)_a^U \\ & (B)_a^L \le b \le (B)_a^U \\ & \forall i, j \end{aligned} \tag{23}$$

Here in searching the minimum value of objective function, s_{ij}^B coefficients must set to their upper bounds. In addition to this the largest feasible region defined by the inequality constraint occurs when s_{ij}^W and b parameters reach their lower and upper bounds, respectively. So the lower bound objective function can be rewrite as

$$\begin{aligned} J_a^L &= \min_W \; -\sum_{i=1}^{n}\sum_{j=1}^{n} (S_{ij}^B)_a^U w_i w_j \\ s.t. \quad & \sum_{i=1}^{n}\sum_{j=1}^{n} (S_{ij}^W)_a^L w_i w_j \le B^U \end{aligned} \tag{24}$$

To have the same formulation format as upper bound optimization function, we use the dual form of the previous relation as follows

$$\begin{aligned} J_a^L &= \min_{W,l} \; -\sum_{i=1}^{n}\sum_{j=1}^{n} (S_{ij}^B)_a^U w_i w_j + l\Big(\sum_{i=1}^{n}\sum_{j=1}^{n} (S_{ij}^W)_a^L w_i w_j - B^U\Big) \\ s.t. \quad & -\sum_{i=1}^{n} (S_{ij}^B)_a^U w_i + l\sum_{i=1}^{n} (S_{ij}^W)_a^L w_i = 0, \; j = 1,\ldots,n \\ & l \ge 0 \end{aligned} \tag{25}$$

This is a conventional quadratic programming problem with quadratic constraints and similarly, the optimal value of J_a^L will be the lower bound of the FFD objective value in the specified α level.

IV. EXPERIMENTAL RESULTS

In this section, we utilize the proposed FFD algorithm using different samples and as comparison purposes the well-known fisher discriminator is used. Here, for ease of evaluation 2-dimensional data is used, but obviously the presented algorithm in the previous section is general and can be used in the case of higher dimensional data. As mentioned earlier we use symmetric triangular fuzzy numbers and in the case of fisher discriminator the input data fuzziness is assumed to be zero.

Example 1: As the first example, we aim to illustrating the solution method proposed in this paper for fuzzy quadratic objective functions with fuzzy quadratic constraints. Therefore for analysis simplification, completely separable data (shown in Fig. 2) was assumed. Here the train patterns are exactly separated and the linear fisher discriminator result is also depicted in this figure. In this case the normalized weight vector is: $w_1^* = 0.9998, w_2^* = 0.0214$.

Now we want to consider fuzziness of input data. We use triangular fuzzy numbers with unit support range from each side. Table 1 lists the upper and lower bounds of normalized weight vectors yield using eleven distinct α-cuts of objective value: 0.0, 0.1, … , 1.0. The α-cut of $\tilde{J}$ represents the possibility that the objective value will appear in the associated range.

The graphical representation of upper and lower bound discriminators in different α levels are provided in Fig. 3 for ease of investigation. Expectedly maximum between bounds variations (widest objective value interval) is occurred at $a = 0$ (fuzziness peak) and at $a = 1$ (no fuzziness) there is no difference between bounds and both of them are tends to the conventional fuzzy discriminator in Fig. 2. The membership function of this example $m_{\tilde{J}}$ is represented in Fig. 4. Same as input data it is a triangular like fuzzy number, too.

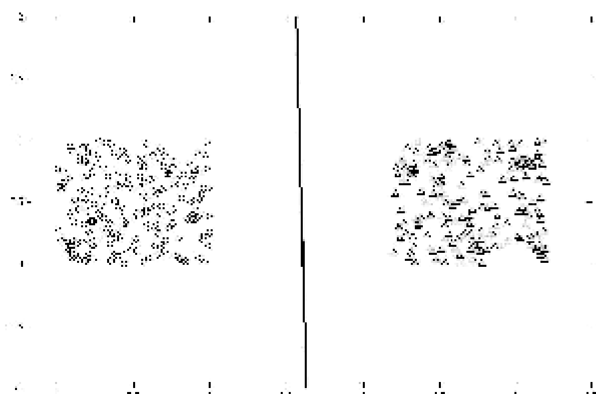
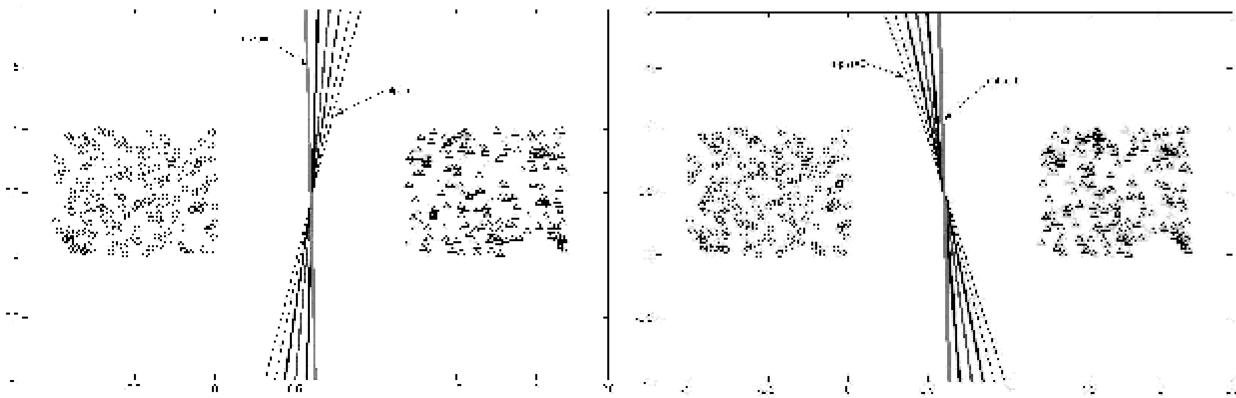


Figure 2 – Two separable data classes

TABLE 1 – LOWER AND UPPER BOUNDS OF OBJECTIVE FUNCTION VALUES AND NORMALIZED WEIGHTS USING DIFFERENT A VALUES.

A	0.0	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9	1.0
$\tilde{J}_a^U$	0.2336	0.2257	0.2177	0.2095	0.201	0.1923	0.1833	0.1739	0.1641	0.154	0.1434
$\tilde{J}_a^L$	0	0.0182	0.0354	0.0516	0.067	0.0815	0.0952	0.1082	0.1205	0.1322	0.1434
W_a^U	$\begin{bmatrix}-.98\\.19\end{bmatrix}$	$\begin{bmatrix}-.98\\.18\end{bmatrix}$	$\begin{bmatrix}.99\\-.16\end{bmatrix}$	$\begin{bmatrix}.99\\-.14\end{bmatrix}$	$\begin{bmatrix}.99\\-.12\end{bmatrix}$	$\begin{bmatrix}-1.0\\.09\end{bmatrix}$	$\begin{bmatrix}1.0\\-.07\end{bmatrix}$	$\begin{bmatrix}-1.0\\.05\end{bmatrix}$	$\begin{bmatrix}-1.0\\.03\end{bmatrix}$	$\begin{bmatrix}1.0\\-.00\end{bmatrix}$	$\begin{bmatrix}-1.0\\-.02\end{bmatrix}$
W_a^L	$\begin{bmatrix}-.97\\-.24\end{bmatrix}$	$\begin{bmatrix}.97\\.22\end{bmatrix}$	$\begin{bmatrix}.98\\.20\end{bmatrix}$	$\begin{bmatrix}-.98\\-.18\end{bmatrix}$	$\begin{bmatrix}.99\\.18\end{bmatrix}$	$\begin{bmatrix}.99\\.14\end{bmatrix}$	$\begin{bmatrix}.99\\.11\end{bmatrix}$	$\begin{bmatrix}1.0\\.09\end{bmatrix}$	$\begin{bmatrix}-1.0\\-.07\end{bmatrix}$	$\begin{bmatrix}-1.0\\-.04\end{bmatrix}$	$\begin{bmatrix}1.0\\.02\end{bmatrix}$

Note: as representation restrictions all numbers are rounded.



(a) (b)

Figure 3 – (a) upper bound of fuzzy discriminators (b) lower bound of fuzzy discriminator, using different confidence values and separable dataset.

TABLE 2 – ERROR RATE INTERVALS USING DIFFERENT LEVELS OF FUZZINESS IN INPUT DATA.

α	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9	1.0
e (%)	[6.5 22.75]	[6.5 20.25]	[6.5 15.5]	[6.5 10.25]	[6.5 7.75]	[6.5 6.75]	[6.5 6.75]	[6.5 6.5]	[6.5 6.5]	[6.5 6.5]

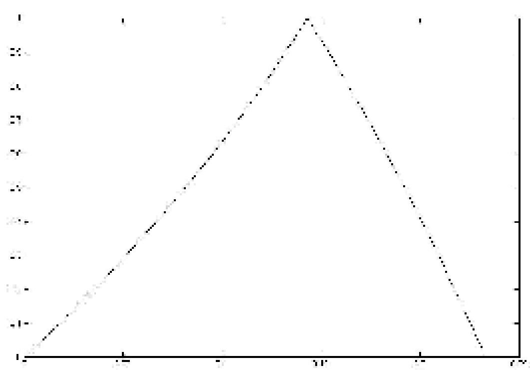


Figure 4 – Resulted fuzzy objective function

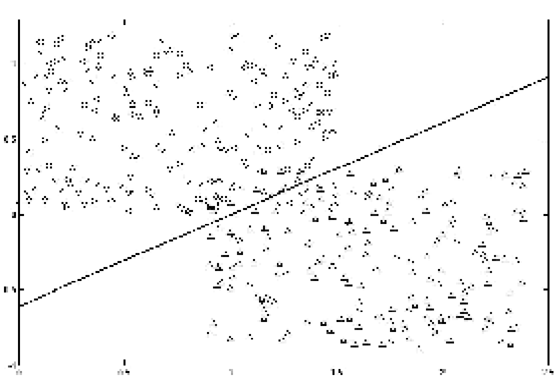


Figure 5 – Non-separable Training dataset distribution.

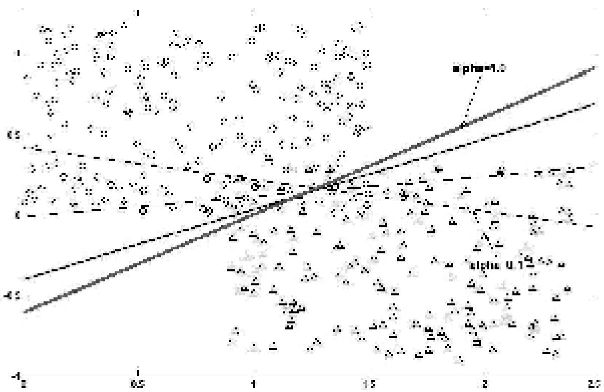
Example2: As our second example, we want to study and compare efficiency of conventional fisher algorithm and our FFDA as encountering fuzzy input data classes with different certainty levels in train and test phases. We will use samples as shown in Fig. 5 with 200 patterns in each class (triangle and circle) and overlapping area between them. Similar to previous example, using represented data as train set, discriminator lines with distinct α values: 0.1, 0.4, 0.7 and 1.0 are presented in Fig. 6.

In this case, as a suitable validation factor, classification error rate interval using different levels of fuzziness is defined below, and is listed in Table 2. Lower and upper levels of each α-cut interval are obtained using W_a^L and W_a^U vectors (Fig. 6). We must mention that error rates using crisp and non-fuzzy input data is 6.5%.

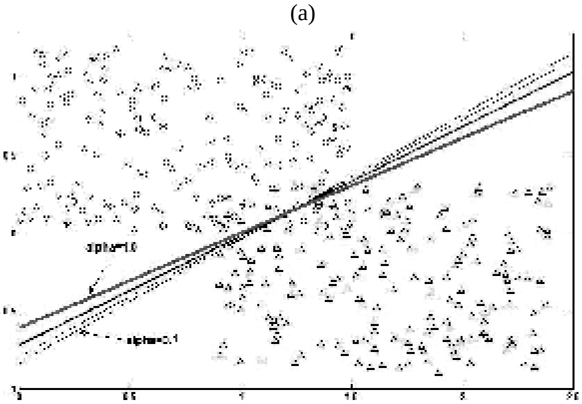
$$e = \frac{\text{num of missclassified samples}}{\text{total number of samples}} \times 100 \tag{26}$$

Expectedly, according to Table 2, as fuzziness decreases (or α value increases) we have smaller error interval and in addition to this property in each case the lower bound of error is approximately equivalent to the non-fuzziness situations.

But, different factors, such as environment noise, measurement errors and etc., can cause different uncertainty levels in real data. For example, in the previous dataset, the real value of a test data that is reported as $\mathbf{x} = [1.4\ 0.2]$ with 40% confidence value, can range from $\mathbf{x}_l = [0.8\ \text{-}\ 0.4]$ to $\mathbf{x}_h = [2.0\ 0.8]$ and misclassified as triangle class using conventional fuzzy classifier. Whereas in our FFDA method a reasonable margin with respect to input space confidence value is created and samples in test phase is classified as triangle, circle and unclassifiable classes (spaces belongs to these triple classes for previous example and 40% certainty level is shown as Fig. 7). Here we will call unclassifiable samples (samples lies in margin spaces) as *outlier*. It seems to be a good decision that when the uncertainty factor in input space is high and samples can range in a wide interval from their reported position, the algorithm don't confide on boundary samples and work on the basis of samples that their dependency to one class is certain. Now, one can design another classifier with appropriate parameters to classify outlier samples into one of triangle or circle classes.



(a)



(b)

Figure 6 - (a) upper bound of fuzzy discriminators (b) lower bound of fuzzy discriminator, using different confidence values and non-separable samples.

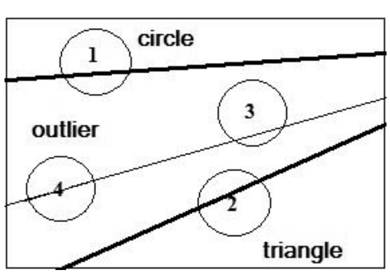


Figure 7 – FFDA lower and upper discriminator lines (bold) and FDA discriminator (normal). Obviously samples such as 1 and 2 can certainly classified as circle and triangle classes. But samples such as 3 and 4 are unclassifiable because of uncertainty factor reported values.

TABLE 3 – CLASSIFICATION RESULTS USING 400 TEST SAMPLES.

	Conventional FDA	FFDA
Classification rate	87.75%	87.75%
Misclassification rate	12.25%	3.5%
Outlier rate	----------	8.75%

Finally, according to previous observations, we examine 200 test samples from each class with 40% certainty level and the results are listed as Table 3. Lower misclassification in FFDA can be illustrated as attending fuzzy nature of real data.

V. CONCLUDING REMARKS

This paper solves quadratic programming problems with fuzzy parameters and quadratic fuzzy constraints and using this solution FFD classifier that is based on conventional FDA and suitable for real data that is usually imprecise, is introduced.

REFERENCES

[1] A. Bigand and O. Colot, "Fuzzy filter based on interval-valued fuzzy sets for image filtering," To be appeared in: *Fuzzy Sets and Systems* Accepted date: 25 March 2009.

[2] A. Salski, "Fuzzy clustering of fuzzy ecological data," *Ecological Informatics*, Vol. 2, pp. 262-269, 2007.

[3] M.S. Yang, H.H. Liu, "Fuzzy clustering procedures for conical fuzzy vector data". *Fuzzy Sets and Systems*, Vol. 106, pp. 189–200, 1999.

[4] R. Yang, Z. Wang P-A. Heng and K-S. Leung, "Classification of heterogeneous fuzzy data by Choquet integral with fuzzy-valued integrand," *IEEE Trans. Fuzzy Systems*, Vol. 15, No. 5, pp. 931– 942. Oct. 2007.

[5] Bazaraa MS, Sherali HD, Shetty CM. Nonlinear programming: theory and algorithm. second ed. New York: John Wiley & Sons; 1993.

[6] Reklatis GV, Ravindran A, Ragsdell KM. Engineering optimization: methods and applications. New York: John Wiley & Sons; 1983.

[7] S.T. Liu, R.T. Wang. "A numerical solution method to interval quadratic programming", *Applied Mathematics and Computation*, Vol. 189, pp. 1274–1281, 2007.

[8] S.T. Liu. "A revisit to quadratic programming with fuzzy parameters", To be appeared in: *Chaos, Solitons and Fractals*, Accepted date: 24 April 2008.

Boshra Rajaei received the B.S. degree in software engineering from Azzahra university of Tehran, Iran, in 2005 and the M.S. degree from Ferdowsi university of Mashhad, Iran, in 2007. She is currently pursuing Ph.D. degree at Ferdowsi university of Mashhad, Iran. Her research interests include image and multidimensional signal processing, multiscale geometric analysis, visual information representation and data compression.

Hadi Sadoghi Yazdi was born in Sabzevar, Iran, in 1971. He received the B.S. degree in electrical engineering from Ferdowsi University of Mashad in 1994, and then he received to the M.S. and PhD degrees in electrical engineering from Tarbiat Modarres University of Iran, Tehran, in 1996 and 2005 respectively. He works in Computer Department as an assistant professor at Ferdowsi University of Mashhad. His research interests include adaptive filtering, image and video processing, and optimization in signal processing. He has more than 140 journal and conference publications in subject of interesting area.

Sohrab Effati was born in Ghoochan, Iran. He received the PhD degree in Applied Mathematics. He works in Mathematics Department as an associated professor at Ferdowsi University of. His research interests include optimization, and operational research.