Interval-valued fuzzy clustering

M. Pagola[1] **A. Jurio**[1] **E. Barrenechea**[1] **J. Fernández**[1] **H. Bustince**[1]

[1]Universidad Pública de Navarra

Abstract

In this work we propose an objective function to obtain an interval-valued fuzzy clustering. After the process of optimization we obtain an interval-valued fuzzy partition in which the length of the intervals depends on the position of the points with respect of the clusters.

Keywords: Style files, LaTeX

1. Introduction

Clustering algorithms that are considered one of the methods in unsupervised machine learning are powerful tools for extracting knowledge from raw data. Their goal is to divide the data in to meaningful clusters. Basically clustering an unlabeled data set $X = \{x^{(1)}, \ldots, x^{(m)}\}$ is the partitioning of X into $C\{1 < C < m)$ subgroups such that each one represents a natural substructure in X [12].

One of the most widely used fuzzy clustering models is Fuzzy C-Means (FCM) [1]. The FCM algorithm assigns memberships to the elements which are inversely related to the relative distance of them to the point prototypes that are the cluster centres.

FCM algorithms have been shown to be closely related to Gaussian Mixture Models [8]. Gan et al. [9] showed that the GMMs can be translated to an additive fuzzy system. Ichihashi et al. [11] showed that the EM algorithm for the GMM can be derived from the FCM, when considering a regularized fuzzy objective function, for a proper selection of the distance metric.

Most of the existing clustering algorithms, including GMMs and FCMs,impose a probabilistic constraint on the utilized membership functions: the obtained cluster memberships of a data point must sum up to one over the derived clusters. Such constrained membership functions are not capable of distinguishing between data points which would be equally likely to belong to more than one cluster, also called as inliers, or data points that would be unlikely to belong to anyone of the known clusters, usually referred as outliers.

There exist different approaches to deal with this problem. Most common way to face it, is to remove the probabilistic constrain, such a way that the sum of the memberships of a point to the clusters must be between zero and one. The possibilistic c-means [13] and some of its improvements [12] are classical algorithms without the probabilistic constrain, or more recently we proposed the Fuzzy Mixture Models [16]. Another approach is to obtain an interval partition, instead of a fuzzy partition, such a way the length of the intervals show the uncertainty of the inliers and outliers.

Currently, there are three typical interval-valued fuzzy clustering algorithms, i.e. the Interval-valued Fuzzy C-means algorithm (IFCM) [5] (originlly called Interval type-2 fuzzy cluster means) and Interval-valued fuzzy Possibilistic C-Means algorithm (IPCM) [6], and the [5]. These algorithms are commonly viewed as the generalization of the standard FCM and PCM algorithms; However both algorithms are based on the application of the FCM or PCM several times, and from the different solutions obtained, then create an interval solution. These algorithms have been extended and used in several applications [17] and [18]. However the simple extension of the methodology and constrains of FCM to the interval case provokes errors. For example, some authors set a restriction to their interval partition, similar to the constrain of a fuzzy partition. This new constrain is such that the sum of the lower memberships of an element to every cluster must sum up to one, and the sum of the upper memberships of an element to every cluster must sum up to one also [18]. But it is only possible if the lower membership is equal to the upper membership, i.e. it is a point, not an interval.

In this work we propose an extension of the FCM algorithm that obtains an interval valued fuzzy partition. We propose an objective function to minimize from which the length of the intervals of the data points (inliers or outliers), that we know it is very difficult to assign those points to a cluster, is larger than the "normal" points. Such that after the clustering process we can identify those points. We propose to use a constrain in the interval fuzzy partition such that only the sum of the lower memberships must sum up to one.

The paper is organized as follows: First, some necessary concepts and definitions are explained. Next, in section 3, the problem of the probabilistic (or fuzzy) partition is explained in a detailed example. In Section 4 the method proposed is described, and in section 5 some numerical examples are shown and discussed. Finally, the conclusions and future research are detailed.

2. Preliminaries

In this section a set of definitions, theorems and algorithms are recalled to understand the rest of

the work.

Definition 1 (Fuzzy set) *A fuzzy set A on a finite universe $U \neq \emptyset$ is a mapping $A \colon U \to [0,1]$.*

We denote by $\mathcal{FS}(U)$ the set of all fuzzy sets on U.

2.1. Interval valued fuzzy sets

Let us denote by $L([0,1])$ the set of all closed subintervals in $[0,1]$:

$$L([0,1]) = \{\mathbf{x} = [\underline{x}, \overline{x}] | 0 \leq \underline{x} \leq \overline{x} \leq 1\}.$$

Then $L([0,1])$ is a partially ordered set with respect to the relation $\leq_L$ defined in the following way: given $\mathbf{x}, \mathbf{y} \in L([0,1])$,

$$\mathbf{x} \leq_L \mathbf{y} \text{ if and only if } \underline{x} \leq \underline{y} \text{ and } \overline{x} \leq \overline{y} \,.$$

Moreover $(L([0,1]), \leq_L)$ is a complete lattice with smallest element $0_L = [0,0]$ and largest element $1_L = [1,1]$. Note that it is not a linear lattice, since there are elements which are not comparable.

The following definition can be found in [2].

Definition 2 (Interval-valued fuzzy set) *An interval-valued fuzzy set (IVFS) A on a finite universe $U \neq \emptyset$ is a mapping $A \colon U \to L([0,1])$.*

Let us denote by W the function that associates to a closed subinterval of $[0,1]$ its length, i.e. $W \colon L([0,1]) \to [0,1]$ with $W([\underline{x}, \overline{x}]) = \overline{x} - \underline{x}$. If $\mathbf{x} = [\underline{x}, \overline{x}] \in L([0,1])$, by interval arithmetics it is $1 - \mathbf{x} = [1,1] - [\underline{x}, \overline{x}] = [1 - \overline{x}, 1 - \underline{x}]$ and we have that $W(\mathbf{x}) = W(1 - \mathbf{x})$. Note that given an IVFS A the membership of each element $u_i \in U$ is represented by an interval value $A(u_i) = [\underline{\mu_A}(u_i), \overline{\mu_A}(u_i)]$ with length $W(\mu_A(u_i)) = \overline{\mu_A}(u_i) - \underline{\mu_A}(u_i)$. A study on the evolution of the IVFSs and several results on their representation, connective types and operations can be found in [2]. We denote by $\mathcal{IVFS}(U)$ the set of all IVFSs on U.

There are two different interpretations of IVFSs [15]:

1. The membership degree of an element to the set corresponds to a value in the considered membership interval. We can not say in a precise way what that value is, so we just provide bounds for that number are provided.
2. The membership degree of each element is the whole closed subinterval provided as membership, understood as a mathematical entity.

In this work we are going to use the first interpretation, in fact we are going to assign an interval membership to every point to each cluster. Such a way the larger the length, greater uncertainty we have on the membership of that point to the cluster.

2.2. Orders between intervals

In [3] the notion of admissible orders on $L([0,1])$ was introduced and investigated. The authors established that a binary relation $\preceq$ on $L([0,1])$ is an *admissible order* if it is a linear order on $L([0,1])$ refining $\leq_L$, i.e. if for all $[a,b], [c,d] \in L([0,1])$ such that $[a,b] \leq_L [c,d]$ then it is also $[a,b] \preceq [c,d]$. The use of admissible orders allow us to compare intervals using total orders between them. Next some examples of admissible orders are given.

Example 1 *Let $[a,b], [c,d] \in L([0,1])$:*

- $[a,b] \preceq_{L1} [c,d] \Leftrightarrow a < c$ *or* $(a = c$ *and* $b \leq d)$;
- $[a,b] \preceq_{L2} [c,d] \Leftrightarrow b < d$ *or* $(b = d$ *and* $a \leq c)$;
- $[a,b] \preceq_{XY} [c,d] \Leftrightarrow a+b < c+d$ *or* $(a+b = c+d$ *and* $b - a \leq d - c)$ *(defined by Xu and Yager in [22]);*
- $[a,b] \preceq_{\alpha,\beta} [c,d] \Leftrightarrow K_\alpha(a,b) < K_\alpha(c,d)$ *or* $(K_\alpha(a,b) = K_\alpha(c,d)$ *and* $K_\beta(a,b) \leq K_\beta(c,d))$, *being $K_\alpha \colon [0,1]^2 \to [0,1]$ a mapping defined by $K_\alpha(a,b) = a + \alpha \cdot (b - a)$ for $\alpha, \beta \in [0,1]$ and $\alpha \neq \beta$.*

Definition 3 *An n-ary $(n \in \mathbb{N}, n \geq 2)$ aggregation function is a non-decreasing mapping in each argument, $M \colon [0,1]^n \to [0,1]$, verifying $M(0, \cdots, 0) = 0$ and $M(1, \cdots, 1) = 1$.*

Bustince *et al.* [3] described different construction methods of admissible orders by means of aggregation functions, as well as the notion of aggregation function of interval values.

2.3. Interval Type-2 Fuzzy Sets and IVFSs

From the notion given by Karnik and Mendel in 1998 of T2FSs, Interval Type 2 Fuzzy Sets were defined in 2006 (see [14]) as follows.

Definition 4 *When all $\mu_A(x, u) = 1$, then A is an IT2FS.*

Hence, according to [14], an IT2FS corresponds to

$$A(x) = \{(u, 1) \mid u \in J_x \subseteq [0,1]\} \qquad (1)$$

for every $x \in X$.

It has been proven in [4] that IT2FSs represent a generalization of the concept of IVFSs and that both concepts are not equivalent. Note that in the definition given in [14] $J_x \subseteq [0,1]$ is used. However, it is well known that with such mathematical expression J_x may be any subset of $[0,1]$, not just a closed subinterval.

In the literature, there has been a confusion about the IT2FSs. Mostly all of the works related with IT2FS supposes that the membership is just a closed subinterval, therefore, equivalent to IVFSs.

2.4. Interval valued fuzzy partition

The concept of I-fuzzy partition, that is a generalization of the fuzzy partition was proposed by V. Torra and Miyamoto [19]. They propose this generalization for the case of Intuitionistic fuzzy sets and Interval valued fuzzy sets. If we consider a set of IVFSs, it is not possible that all of the lower memberships sum up to one and all of the upper memberships sum up to one at the same time (unless $W = 0$). The generalization requires that only the lower memberships sum up to one.

Definition 5 *Let X be a referential set. Then a set of IVFSs $A = \{A_1, \cdots, A_C\}$ is an interval valued fuzzy partition if*

1. $\sum_{j=1}^{C} \underline{\mu_j}(x) = 1$ for all $x \in X$;
2. *there is at most one IVFS such that $\underline{\mu_j}(x) = \overline{\mu_j}(x)$ for all x.*

2.5. Fuzzy cluster means

Let an unlabeled data set $X = \{x^{(1)}, \ldots, x^{(m)}\}$ of m examples, and each example $x^{(i)} = (x_1^{(i)}, \ldots, x_n^{(i)})$ is a vector of n dimensions. The aim of FCM is to find the cluster centres (centroids) ν_c that minimize a dissimilarity function. The dissimilarity function measures de distance between a point $x^{(i)}$ and the cluster prototype ν_c:

$$d_{ic}^2 = ||x^{(i)} - \nu_c||_A^2 \qquad (2)$$

being $||.||_A$ an induced norm on $\mathcal{R}^n$ with A a positive definite $(n \times n)$ weight matrix. The objective function $J(U, \nu)$ is the weighted sum of dissimilarities within each cluster:

$$J(U, \nu) = \sum_{i=1}^{m} \sum_{c=1}^{C} (\mu_c(x^{(i)}))^b d_{ic}^2 \qquad (3)$$

being U the fuzzy partition and b the weighting exponent called the degree of fuzziness. The processing of minimizing object function $J(U, \nu)$ depends on how centres find their ways to the best positions, as the fuzzy memberships $\mu_c(x^{(i)})$ and norm distance d_{ic}^2 would change along with the new centres' position. Approximate optimization of $J(U, \nu)$ is based on the iteration through the following necessary conditions for its local extrema.

Theorem 1 *FCM [1] If $d_{ic} = ||x^{(i)} - \nu_c|| > 0$ for all i, and $c, b > 1$, and X contains at least c distinct points, then (U, ν) may minimize $J(U, \nu)$ only if*

$$\mu_c(x^{(i)}) = \left[\sum_{j=1}^{C} \left(\frac{d_{ic}^2}{d_{ij}^2} \right)^{2/(b-1)} \right]^{-1} \qquad (4)$$

$$\nu_c = \frac{\sum_{i=1}^{m} \mu_c(x^{(i)})^b x^{(i)}}{\sum_{i=1}^{m} \mu_c(x^{(i)})^b} \qquad (5)$$

	Data		FCM	
Point	x	y	$\mu_1(x^{(i)})$	$\mu_2(x^{(i)})$
1	-5,	0	0.9583	0.0417
2	-3.34	1.67	0.9755	0.0245
3	-3.34	0	0.9756	0.0244
4	-3.34	-1.67	0.9318	0.0682
5	-1.67	0	0.9541	0.0459
6	1.67	0	0.0456	0.9544
7	3.34	1.67	0.0245	0.9755
8	3.34	0	0.0245	0.9755
9	3.34	-1.67	0.0683	0.9317
10	5	0	0.0419	0.9581
11	0	0	0.5002	0.4998
12	0	10	0.5000	0.5000

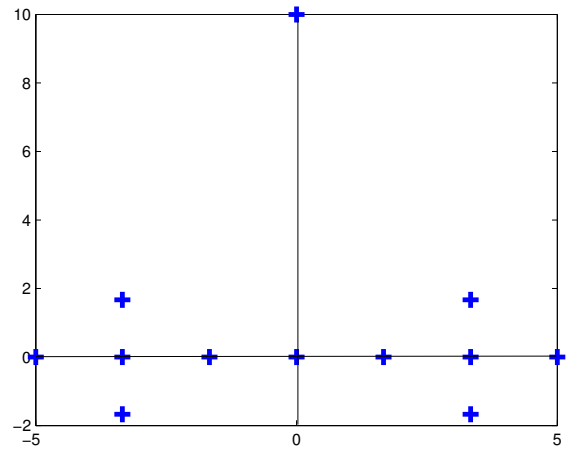Table 1: Data set and its memberships obtained from FCM

Figure 1: Points representing the data set of Table 1

3. The problem of Fuzzy partitions

Many authors have described the problem of fuzzy partitions. In this section we show an example that was described in [12] by Pal et. al.. We have the data set $\{x^{(1)}, x^{(2)}, \cdots .., x^{(11)}, x^{(12)}\}$ that are shown in Table 1 and Figure 1.

We can see that the data set has two clusters with five points each, with $x^{(11)}$ and $x^{(12)}$ as two points difficult to assing to each cluster. Usually $x^{(11)}$ is called an inlier (bridge) and $x^{(12)}$ is called an outlier or noise. These inliers and outliers can affect the partitions found by FCM. In [12] it was shown how the terminal centroids obtained by FCM vary their position on the presence of inliers and outliers. In fact, taking the data set of points $\{x^{(1)}, x^{(2)}, \cdots .., x^{(10)}\}$, an applying FCM, the centroids obtained are different from the ones obtained, with the same parameters and initial conditions, with the data set $\{x^{(1)}, x^{(2)}, \cdots .., x^{(11)}, x^{(12)}\}$. However the memberships of points $x^{(1)}, x^{(2)}, \cdots .., x^{(10)}$ do not change very much at all. Points $x^{(11)}$ and $x^{(12)}$ have

membership values around 0.5 in each cluster because both are equidistant from the centroids, even though $x^{(12)}$ is further away from the clusters than $x^{(11)}$. The problem is caused by the definition of the partition, that is, the summation constraint $\sum_{c=1}^{C} \mu_c(x^{(i)}) = 1$. Therefore, to obtain a partition in which we can identify inliers and outliers is an interesting challenge.

4. Interval valued fuzzy clustering proposal

In our proposal we want to obtain an interval valued fuzzy partition satisfying Definition 5 after the clustering process. We demand two conditions:

1. The length of the intervals of the points that are inside of a cluster must as short.
2. The length of the intervals of the points that are far from every cluster must be large.

Moreover the partition obtained should represent the inherent structure of the data. Taking into account all those considerations we propose a new objective function (Equation 6) in which the memberships of the elements to the clusters are intervals.

$$J(U, \nu) = \sum_{i=1}^{m} \sum_{c=1}^{C} \underline{\mu_c}(x^{(i)})^b d_{ic}^2 + \qquad (6)$$

$$+ \sum_{i=1}^{m} \sum_{c=1}^{C} K \frac{\overline{\mu_c}(x^{(i)}) - \underline{\mu_c}(x^{(i)})}{min_{1 \leq c \leq C}\{d_{ic}^2\}}$$

The first term of the objective is the same as the classical FCM, using the lower membership value of the intervals, and it deals with the objective of getting a partition that represents the structure of the data. The second term deals with the length of the intervals regarding the position of the element with respect to the clusters. If the distance to every cluster is short, then the length of the interval must be short, due to it is a minimization process. On the contrary, if the distance of the point is large to every cluster, then the length of its interval should be large, as we have demanded.

There are several constrains. The first ones are related to the interval valued fuzzy partition in which the sum of all the lower membership degrees must sum up to one. There is a constraint for every element.

$$\sum_{c=1}^{C} \underline{\mu_c}(x^{(i)}) = 1 \qquad (7)$$

Furthermore the intervals membership must be intervals., i.e. the lower membership must be less than or equal the upper memberships, and both of them are bounded by zero and one. There is a constraint for all of the cluster memberships of every element.

$$0 \leq \underline{\mu_c}(x^{(i)}) \leq \overline{\mu_c}(x^{(i)}) \leq 1 \qquad (8)$$

The centroids of the clusters are calculated using the lower membership of the interval:

$$\nu_c = \frac{\sum_{i=1}^{m} \underline{\mu_c}(x^{(i)})^b x^{(i)}}{\sum_{i=1}^{m} \underline{\mu_c}(x^{(i)})^b} \qquad (9)$$

the distance d_{ic}^2 between the ith element and the cluster c is calculated by the following expression (similar to FCM):

$$d_{ic}^2 = ||x^{(i)} - \nu_c||_A^2 \qquad (10)$$

There is also a parameter K in the second term of the objective function (Equation 6). This parameter is used to control the length of the intervals in the final partition. If the value of K is high, as it is a minimization process, the second term will have a lot of weight and then the length of the intervals will be short. On the contrary if K is low then we will obtain a partition with large length of the intervals.

Solving this optimization problem is a non-trivial task and may become computationally expensive. The number of constrains is really high, $(m*c+m)$ and increase linearly with the number of examples and clusters to be identified. We have applied the solver `fmincon` function implemented in the optimization toolbox of Matlab. This function provides a method for constrained non-linear optimization based on interior point methods.

It is left for future research to derive the objective function and obtain and approximate solution by means of a set of iterative constrains, similar to FCM and others.

5. Illustrative Examples

Next, we show the results of our proposal in two different data sets. First we apply our method to the data of Table (1) (section 3). We use the euclidean distance, the parameters $b = 1$, $K = 1$, $C = 2$ and the initial partition is random. In Table 2 are the interval valued fuzzy memberships obtained. We can observe that $\{x^{(1)}, x^{(2)}, \cdots .., x^{(10)}\}$ have high memberships with very short length of the intervals to the clusters which they belong and very low memberships with also short interval length to the other cluster. Points $x^{(11)}$ and $x^{(12)}$ have lower membership values around 0.5 in each cluster and the length of their intervals are larger than the other elements. In Figure 2 the size of the dots is proportional to the length of the intervals of the memberships to cluster 2 (fourth and fifth columns of Table 2). The color of the dots is chosen depending on the cluster assigned. To calculate which interval is greater we have used the following admissible order: $[a, b] \preceq_{L1} [c, d] \Leftrightarrow a < c$ or ($a = c$ and $b \leq d$).

So, in this data set we have achieved our objective. We can identify inliers and outliers based on their interval memberships. If the lower membership is near 0.5 we know that this point is really difficult to assign to a cluster. Furthermore, if the

| | Cluster 1 | | Cluster 2 | |
Pt.	$\underline{\mu_1}(x^{(i)})$	$\overline{\mu_1}(x^{(i)})$	$\underline{\mu_2}(x^{(i)})$	$\overline{\mu_2}(x^{(i)})$
1	0.9427	0.9690	0.0573	0.2600
2	0.9377	0.9640	0.0623	0.1719
3	0.9377	0.9641	0.0623	0.1717
4	0.9245	0.9581	0.0755	0.2756
5	0.9083	0.9431	0.0917	0.1910
6	0.0917	0.1266	0.9083	1.0
7	0.0623	0.0887	0.9377	1.0
8	0.0623	0.0886	0.9377	1.0
9	0.0755	0.1091	0.9245	1.0
10	0.0573	0.0836	0.9427	1.0
11	0.5000	0.5997	0.5000	0.6939
12	0.5000	0.6311	0.5000	0.8762

Table 2: Interval memberships obtained from IV-FCM
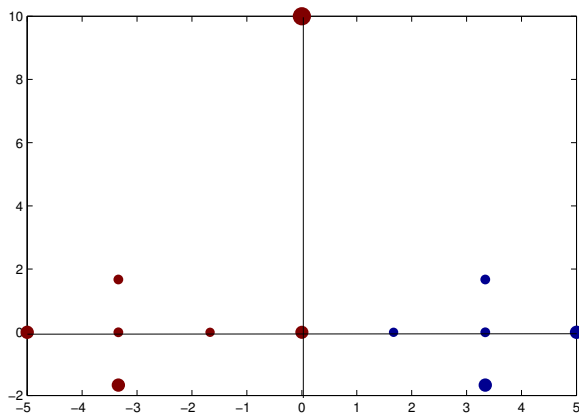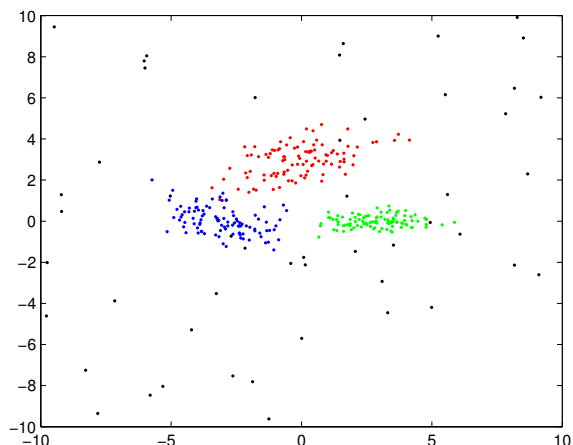
Figure 3: Synthetic data

length of the interval is greater than the majority of the the elements, it should be an outlier, and if the length is similar, then it should be an inlier.

Due to we are using an approximate minimization function, there are some residuals that makes that the memberships generated are not absolutely symmetric. This makes that the detection of the outliers must be done within a membership function. The lengths of the intervals are incomparable between different cluster memberships. In this example the length of the interval of the inlier is between the values of the normal points and the length of the outlier.

Next, we are going to use a simulated data (similar to the one used in [7]) to illustrate the results of our methodology in a set with a large number of elements. The synthetic data is generated by means of 100 points of each of these three bivariate Gaussian distribution functions: $f_a = \mathcal{N}\left(\begin{bmatrix} 0 \\ 3 \end{bmatrix}, \begin{bmatrix} 2 & 0.5 \\ 0.5 & 0.5 \end{bmatrix}\right)$, $f_b = \mathcal{N}\left(\begin{bmatrix} 3 \\ 0 \end{bmatrix}, \begin{bmatrix} 1 & 0 \\ 0 & 0.1 \end{bmatrix}\right)$, $f_c = \mathcal{N}\left(\begin{bmatrix} -3 \\ 0 \end{bmatrix}, \begin{bmatrix} 2 & -0.5 \\ -0.5 & 0.5 \end{bmatrix}\right)$

The final 100 points, outliers and noise, are generated from a bivariate Gaussian distribution with each of its components in the interval $[-10, 10]$ (see Fig. 3).

We apply our method using the euclidean distance, the parameters $b = 1$, $K = 1$, $C = 3$ and the initial partition is random. In Figure (4) is similar to Figure 2, where the size of the dots is proportional to the length of the interval, and each point has a color that depends on the cluster that each point is assigned. The intervals are compared by means of the order $\preceq_{L1}$. The properties that we demand to the partition are also satisfied and the outliers have greater length then the points near the centroid of the cluster. There are also some inliers with large amplitude.

We have tested our algorithm in different datasets and we have obtained similar results.
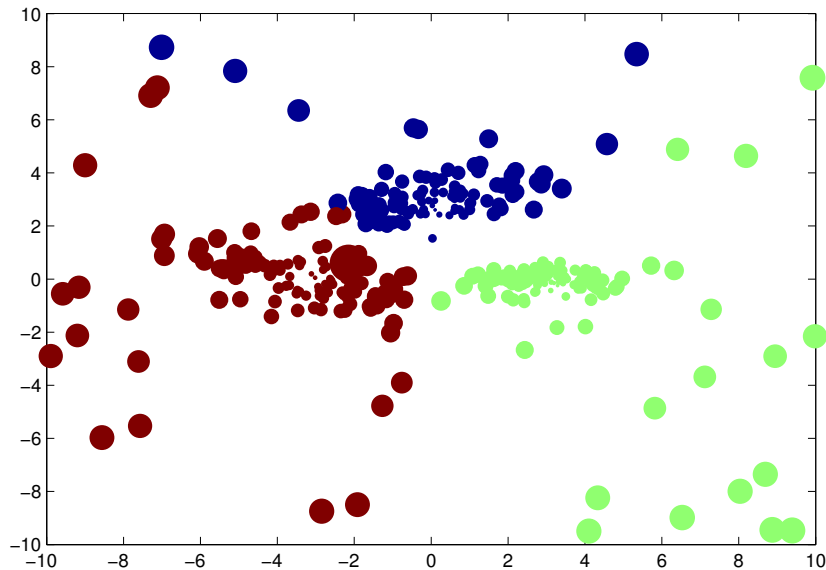

Figure 2: Picture of the length of the intervals in the final partition

Figure 4: Picture of the length of the intervals in the final partition of the synthetic data

6. Conclusions and future research

In this work we have proposed an objective function to obtain an interval-valued fuzzy partition.

References

[1] Bezdek, J.C.: Pattern Recognition With Fuzzy Objective Function Algorithms. New York: Plenum (1981).

[2] Bustince, H., Barrenechea, E., Pagola, M.: Generation of interval-valued fuzzy and Atanassov's intuitionistic fuzzy connectives from fuzzy connectives and from K_α operators: laws for conjunctions and disjunctions, amplitude, *International Journal of Intelligent Systems*, 23(6): 680–714, 2008.

[3] H. Bustince, H., Fernandez, J., Kolesárová, A., Mesiar,R.: Generation of linear orders for intervals by means of aggregation functions, *Fuzzy Sets and Systems*, 2012.

[4] Bustince, H., Fernandez, J., Hagras, H., Herrera, F., Pagola, M., Barrenechea, E.: Interval Type-2 Fuzzy Sets are generalization of IVFSs: Towards a Wider view on their relationship, in Press *IEEE Transactions on Fuzzy Systems* DOI: 10.1109/TFUZZ.2014.2362149.

[5] Hwang, c., Rhee F.C.: Uncertain fuzzy clustering: interval type-2 fuzzy approach to C-means, *IEEE Trans. Fuzzy Systems*, 15:107–120, 2007.

[6] Min, J.H., Shim, E.A., Rhee, F.C.H.: An interval type-2 fuzzy PCM algorithm for pattern recognition, *Proceedings of the 18th International Conference on Fuzzy Systems, FUZZ-IEEE'09*, 480–483, 2009.

[7] Chatzis, S.P., Tsechpenakis, G.:A possibilistic clustering approach toward generative mixture models, *Pattern Recognition* 45:1819–1825, 2012.

[8] Dempster, A. P., Laird, N. M., Rubin, D. B.: Maximum Likelihood from Incomplete Data via the EM Algorithm, *Journal of the Royal Statistical Society. Series B (Methodological)*, 39:1–38, 1977.

[9] Gan, M.T., Hanmandlu ,M., Tan, A.H.: From a Gaussian mixture model to additive fuzzy systems, *IEEE Transactions on Fuzzy Systems*, 13(3): 303–316, 2005.

[10] Hwang, C., Rhee, F.: Uncertain fuzzy clustering: interval type-2 fuzzy approach to C-means, *IEEE Transactions on Fuzzy Systems*, 15:107-120, 2007.

[11] Ichihashi, H., Honda, K., Tani, N.: Gaussian mixture pdf approximation and fuzzy c-means clustering with entropy regularization, In Proceedings of the *Fourth Asian Fuzzy System Symposium*, 217–221, 2000.

[12] Pal, N., Pal, K., Keller, J., Bezdek, J.: A possibilistic fuzzy c-means clustering algorithm, *IEEE Transactions on Fuzzy Systems*,13(4):517–530, 2005.

[13] Krishnapuram, R., Keller, J.M.: A Possibilistic Approach to Clustering, *IEEE Transactions on Fuzzy Systems*, 1(2):98-110, 1993.

[14] Mendel, J. M.; John, R. I., Liu, F.: Interval type-2 fuzzy logic systems made simple, *IEEE Transactions on Fuzzy Systems*, 14(6): 808–821, 2006.

[15] Montero, J., Gómez, D. Bustince, H.: On the relevance of some families of fuzzy sets, ,*Fuzzy Sets and Systems*, 158(22): 2429–2442, 2007.

[16] Pagola, M., Barrenechea, E., Aránzazu, J., Paternain, D., Bustince, H.: Clustering Based on a Mixture of Fuzzy Models Approach, *Information Processing and Management of Uncertainty in Knowledge-Based Systems Communications in Computer and Information Science*, 443:475-484, 2014.

[17] Rubio E., Castillo, O.: Designing Type-2 fuzzy systems using the interval type-2 fuzzy c-means algorithm, *Recent Advances on Hybrid approaches for designing Intelligent systems*, 37-50, 2014.

[18] Rubio E., Castillo, O.: Interval type-2 fuzzy clustering for membership function generation, *HIMA*, 13-18, 2013.

[19] Torra, V., Miyamoto, S.: A definition of I-fuzzy partitions, *Soft Computing*, 15:363–369, 2011.

[20] Lu, W., Zhang, L., Liu X., Yang J., Pedrycz W.: A Human-Computer Cooperation Fuzzy c-Means Clustering with Interval-Valued Weights, *International Journal of Intelligent Systems*,30:81–98, 2015.

[21] Ji Z., Xia Y., Sun Q., Cao G.: Interval-valued possibilistic fuzzy C-means clustering algorithm, *Fuzzy Sets and Systems*, 253(16):138–156, 2014.

[22] Xu, Z.S., Yager, R.R.: Some geometric aggregation operators based on intuitionistic fuzzy sets, *International Journal of General Systems*, 35: 417–433, 2006.