A fuzzy clustering approach for fuzzy data based on a generalized distance

Ana Belén Ramos-Guajardo[1], Maria Brigida Ferraro[2]

[1]*Departamento de Estadística e I.O. y D.M., Universidad de Oviedo,*
Calle Federico García Lorca, 18 - 33007, Oviedo, Spain

[2]*Dipartimento di Scienze Statistiche, Sapienza Università di Roma,*
P.le Aldo Moro, 5 - 00185, Rome, Italy

Abstract

Most of the distances used in case of fuzzy data are based on the well-known Euclidean distance. In detail, a fuzzy number can be characterized by centers and spreads and the most common distances between fuzzy numbers are essentially defined as a weighted sum of the squared Euclidean distances between the centers and the spreads. In the multivariate case the Euclidean distance does not take into account the correlation structure between variables. For this reason, the Mahalanobis distance has been introduced which involves the corresponding covariance matrix between the variables. A generalization of that distance to the fuzzy framework is proposed. It is shown to be useful in different contexts and, in particular, in a clustering approach. As a result, non-spherical clusters, that generally are not recognized by means of Euclidean-type distances, can be recognized by means of the suggested distance. Clustering applications are reported in order to check the adequacy of the proposed approach.

Keywords: Fuzzy k-Means method, Gustafson-Kessel approach, Mahalanobis distance, Fuzzy data.

1. Introduction

Most of the statistical developments addressed in the literature refer to the analysis of real-valued data. Nevertheless, in the real world we can find different types of information regarding valuations, perceptions, ratings, imprecise descriptions of precise measurements or observations, etc., that lead to data which cannot be appropriately expressed by using real/vectorial values. This kind of imprecise data may be described by means of fuzzy numbers (see, for instance, [2, 11, 32, 33]).

On the other hand, fuzziness may also appear at any stage of the data analysis process [7]. In particular, concerning the cluster analysis of precise information some fuzzy methodologies have been proposed in the literature (see, for instance, [1, 4, 10, 12, 18, 20, 22, 24, 25]). For a deeper overview on fuzzy clustering in a real framework see, e.g., [17]. In the case of interval-valued data some fuzzy clustering approaches have been studied in [9, 14, 15], to mention a few. Besides, some fuzzy clustering methods have been also developed for dealing with fuzzy information (see, for instance, [8, 13, 16, 19, 28, 29], to name some of the most recent ones).

Those approaches for the fuzzy clustering of fuzzy numbers are extensions of the classical fuzzy k-means clustering procedure [4] and they are based on the renowned Euclidean distance. Here, the Euclidean distance between two fuzzy numbers is essentially defined as a weighted sum of the squared Euclidean distances among the so-called centers (or midpoints) and radii (or spreads) of the fuzzy sets. A main drawback of the Euclidean distance is that it does not take into account the existing interrelationships between the fuzzy numbers. Then, when a fuzzy k-means method based on the Euclidean distance is applied, it is only able to recognize clusters having spherical shape.

To overcome this drawback, Gustafson and Kessel [18] presented a more flexible fuzzy clustering algorithm in the real framework based on a version of the classical Mahalanobis distance, which involves the covariance matrix. This quadratic distance is defined by means of a positive definite symmetric matrix that must be inverted to make possible the updating of the distance in the clustering algorithm. Nonetheless, sometimes the matrix may be singular and its inversion is not possible. In this kind of situations an improved approach for estimating the covariance matrix proposed in [3] can be applied.

As an extension of the Gustafson-Kessel clustering algorithm to the field of symbolic data analysis, a partitioning fuzzy k-means clustering model for interval-valued data based on the Mahalanobis distance has been provided in [9]. Here, intervals are a particular kind of symbolic data (including sets of categories, interval histograms, etc.).

The aim of this work is to join the ideas provided in [8] about a fuzzy clustering methodology for fuzzy numbers using an Euclidean distance and the ideas in [18] about a fuzzy clustering methodology for crisp information using the Mahalanobis distance to develop a new fuzzy clustering approach for the case of fuzzy numbers.

The new approach is based on an extension of the Mahalanobis distance to the fuzzy framework with the intention of avoiding the drawbacks of the Euclidean distance described above. Thus, this extension attempts to detect non-spherical clusters having, for example, ellipsoidal shapes and that are not normally recognized by means of Euclidean-type distances. Then, the suggested approach is compared with the analogous Euclidean distance-based fuzzy k-means algorithm for fuzzy numbers developed by Coppi *et al.* in [8]. The behaviour of both approaches is observed and contrasted whenever either spherical or non-spherical shaped clusters are provided.

The rest of the paper is organized as follows. In Section 2 some preliminaries concerning the fuzzy numbers framework are presented as well as a brief summary of the Euclidean distance-based fuzzy k-means clustering for fuzzy data. Section 3 contains a motivating example. The generalization of the classical Mahalanobis framework is introduced in Section 4. Section 5 is devoted to the development of the fuzzy clustering algorithm for fuzzy data based on the suggested generalized distance. In Section 6 the empirical performance and the practical applicability of the algorithm is shown and compared with Euclidean distance-based fuzzy k-means clustering for fuzzy data. A real-life application is provided in Section 7. Finally, Section 8 includes some conclusions and future directions.

2. Preliminary concepts

In the following subsections, some preliminaries about the space of fuzzy numbers are briefly summarized. In addition, the Euclidean distance-based fuzzy k-means clustering procedure for fuzzy numbers is also recalled.

2.1. Fuzzy numbers

In many real-life situations there are measurements that may be imprecise and some observations which are vaguely defined. In such contexts it is appropriate to represent the information by means of either interval data or fuzzy data instead of considering crisp values.

The space of fuzzy numbers, denoted by $\mathcal{F}_c(\mathbb{R})$, is composed by the mappings $U : \mathbb{R} \to [0,1]$ such that for each $\alpha \in (0,1]$ the so-called α-level set (or α-cut) $U_\alpha = \{x \in \mathbb{R} | U(x) \geq \alpha\}$ belongs to the class of nonempty compact intervals in $\mathbb{R}$ (denoted by $\mathcal{K}_c(\mathbb{R})$). The 0-level, U_0, is the closure of the support of U.

To be more concrete, the so-called class of *LR-fuzzy numbers* is considered since it is the most common one in practice. An LR-fuzzy number U is determined by four parameters, $U = (c_1, c_2, r, l)$, so that c_1 and c_2 are the left and the right centers of the 1-level of U, and r and l are the right and left spreads of U (that is, the distances between the suprema and the infima of the 0-level and 1-level of U, respectively). The centers c_1 and c_2 are associated not only with the location of the fuzzy number but also with the imprecision of its 1-level whenever the size of the interval defined by $[c_1, c_2]$ is considered. On the other hand, the spreads r and l inform us about the imprecision of the fuzzy number too.

The membership degree of x to U is defined as

$$\mu_U(x) = \begin{cases} L\left(\frac{c_1 - x}{l}\right) & x \leq c_1 \\ 1 & c_1 \leq x \leq c_2 \\ R\left(\frac{x - c_2}{r}\right) & x \geq c_2 \end{cases}, \tag{1}$$

where $L : \mathbb{R} \to [0, 1]$ (and R) is a convex upper semi-continuous function so that $L(0) = 1$ and $L(x) = 0$, for all $x \in \mathbb{R} \setminus [0, 1]$ (see [33]). An example of an LR fuzzy number is shown in Figure 1.

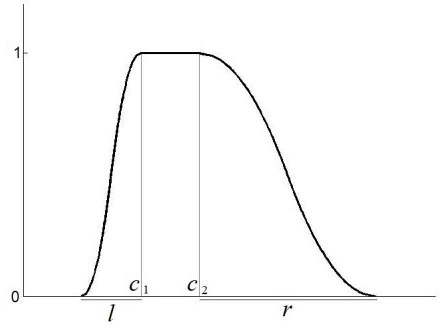

Figure 1: Example of an LR fuzzy number

2.2. Arithmetics and random fuzzy numbers

The usual arithmetic between fuzzy numbers is a level-wise extension of the standard arithmetic for intervals paying attention to the fuzzy meaning

([26, 32]). Given $U, V \in \mathcal{F}_c(\mathbb{R})$ and $\lambda \in \mathbb{R}$, the sum and the product by a scalar can be defined so that for each $\alpha \in [0,1]$ it is fulfilled that

$$(U + \lambda V)_\alpha = U_\alpha + \lambda V_\alpha = \{u + \lambda v : u \in U_\alpha, v \in V_\alpha\}\,. \tag{2}$$

It should be noticed that the arithmetic is non-linear due to the lack of symmetric element w.r.t. the Minkowski addition although $(\mathcal{F}_c(\mathbb{R}), +, \cdot)$ has a semilinear-conical structure since the addition extends level-wise the Minkowski sum of intervals. An example of the arithmetic is shown in Figure 2.

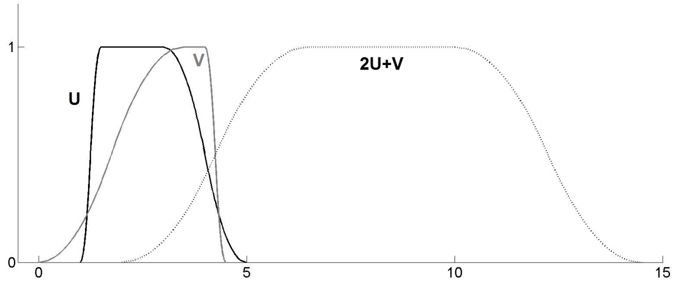

Figure 2: Example of addition and product by a scalar of fuzzy numbers

When the fuzzy numbers generation involves randomness, the sample is presumed to be obtained as a realization of a random fuzzy number. Let $(\Omega, \mathcal{A}, P)$ be a probability space. An $\mathcal{F}_c(\mathbb{R})$-valued *random fuzzy number* (RFN for short, also called fuzzy random number) in Puri and Ralescu's sense [27] is a mapping $\mathcal{X} : \Omega \to \mathcal{F}_c(\mathbb{R})$ so that the α-level mappings $\mathcal{X}_\alpha : \Omega \to \mathcal{K}_c(\mathbb{R})$ (s.t. $\mathcal{X}_\alpha(w) = (\mathcal{X}(w))_\alpha$) are random intervals for all $\alpha \in [0,1]$ (or, equivalently, $\inf \mathcal{X}_\alpha$ and $\sup \mathcal{X}_\alpha$ are real-valued random variables). In addition, it can be shown that a RFN is a Borel measurable mapping with respect to the metric presented below. One of the most important advantages of considering Borel measurable metric-space-valued mappings is that concepts such as induced distribution, independence, etc., can be stated as usual (see [5]).

Let's introduce now some notation which will be useful in the next developments. From now on, an LR fuzzy variable will be an RFN taking LR fuzzy numbers as outcomes. In the multidimensional framework, whenever p LR fuzzy variables are observed on a set of n objects, then a fuzzy data matrix can be defined as $\mathcal{X} = \{\mathcal{X}_{ij} = (c_{1ij}, c_{2ij}, l_{ij}, r_{ij})\}$ for $i = 1, \ldots, n$

and $j = 1, \dots, p$, where $\mathcal{X}_{ij}$ is the LR fuzzy variable j observed on the i-th unit with left center c_{1ij}, right center c_{2ij}, and left and right spreads l_{ij} and r_{ij}, respectively. The matrix $\mathcal{X}$ can be equivalently expressed as $\mathcal{X} = (\mathbf{C}_1, \mathbf{C}_2, \mathbf{L}, \mathbf{R})_{LR}$. In addition, $\mathcal{X}_i \equiv (\mathbf{c}_{1i}, \mathbf{c}_{2i}, \mathbf{l}_i, \mathbf{r}_i)_{LR}$ denotes the fuzzy vector of length p for object i, where $\mathcal{X}_i$, $\mathbf{c}_{1i}$, $\mathbf{c}_{2i}$, $\mathbf{l}_i$ and $\mathbf{r}_i$ are the i-th rows of $\mathcal{X}$, C_1, C_2, L and R, respectively.

2.3. Distance for fuzzy data

In order to develop clustering methods for LR fuzzy numbers a suitable dissimilarity measure is required. In the literature, several metrics for fuzzy data have been proposed. For instance, an L_2 type metric for fuzzy sets in terms of centers and spreads is provided in [31]. In this case, a weighted dissimilarity measure for fuzzy data firstly introduced in [8] is adopted. Thus, the distance is defined as a weighted sum of the (squared) Euclidean distances between centers and spreads. In detail, given two LR fuzzy observations, $\widetilde{\mathbf{x}}_i$ and $\widetilde{\mathbf{x}}_{i'}$, the distance between them is defined as

$$\begin{aligned} d_w^2(\widetilde{\mathbf{x}}_i, \widetilde{\mathbf{x}}_{i'}) = \; & w_C^2[d^2(\mathbf{c}_{1i}, \mathbf{c}_{1i'}) + d^2(\mathbf{c}_{2i}, \mathbf{c}_{2i'})] \\ + \; & w_S^2[d^2(\mathbf{l}_i, \mathbf{l}_{i'}) + d^2(\mathbf{r}_i, \mathbf{r}_{i'})], \end{aligned} \tag{3}$$

where $d(\cdot, \cdot)$ is the standard Euclidean distance (for non-fuzzy data), and w_C and w_S are weights for the center component and the spread component, respectively, depending on the importance given to the centers and to the spreads of the corresponding fuzzy numbers. By means of $d_w^2(\widetilde{\mathbf{x}}_i, \widetilde{\mathbf{x}}_{i'})$ we compute the dissimilarity between two LR fuzzy observations. In general, the weights given to the center component are larger to the ones corresponding to the spread one, since the membership function of a fuzzy number takes the maximum value in the centers. Therefore, the coherence condition $w_C \geq w_S \geq 0$ is adopted. In addition, the normalization condition $w_C + w_S = 1$ is taken into account, so following both conditions we have that $0.5 \leq w_C \leq 1$.

2.4. Euclidean distance-based fuzzy k-means for fuzzy data

The formalization of the fuzzy k-means clustering model for fuzzy data (FkM-F, for short) proposed by Coppi *et al.* in [8] is briefly summarized below. Thus, in order to cluster n observations described by p LR fuzzy variables, the FkM-F optimization problem can be written as

$$\begin{aligned}
\min_{\mathbf{U},\widetilde{\mathbf{H}},w} J_{\mathrm{F}k\mathrm{M\text{-}F}} &= \sum_{i=1}^{n}\sum_{g=1}^{k} u_{ig}^{m} d_{w}^{2}\left(\widetilde{\mathbf{x}}_i, \widetilde{\mathbf{h}}_g\right), \\
\text{s.t.} \quad & u_{ig} \geq 0, \quad i=1,\ldots,n, \quad g=1,\ldots,k, \\
& \sum_{g=1}^{k} u_{ig} = 1, \quad i=1,\ldots,n, \\
& w \in [0.5, 1],
\end{aligned} \tag{4}$$

where u_{ig} is the membership degree of observation i to cluster g, stored in the matrix $\mathbf{U}$ of order $(n \times k)$, and

$$\widetilde{\mathbf{H}} = \left\{ \widetilde{H}_{gj} \equiv \left(h_{gj}^{C_1}, h_{gj}^{C_2}, h_{gj}^{L}, h_{gj}^{R}\right)_{LR}, g=1,...,k, j=1,...,p \right\}. \tag{5}$$

In (5), $\widetilde{H}_{gj} \equiv \left(h_{gj}^{C_1}, h_{gj}^{C_2}, h_{gj}^{L}, h_{gj}^{R}\right)_{LR}$ represents the j-th LR fuzzy variable of the g-th centroid with left center $h_{gj}^{C_1}$, right center $h_{gj}^{C_2}$, left spread h_{gj}^{L} and right spread h_{gj}^{R}. In order to facilitate the understanding of the adopted notation, we can define the centroid matrices of the left centers ($\mathbf{H}^{C_1}$), of the right centers ($\mathbf{H}^{C_2}$), of the left spreads ($\mathbf{H}^{L}$) and of the right spreads ($\mathbf{H}^{R}$) of order $(k \times p)$ with generic elements $h_{gj}^{C_1}$, $h_{gj}^{C_2}$, h_{gj}^{L} and h_{gj}^{R}, respectively. Therefore, $\widetilde{\mathbf{h}}_g \equiv (\mathbf{h}_g^{C_1}, \mathbf{h}_g^{C_2}, \mathbf{h}_g^{L}, \mathbf{h}_g^{R})_{LR}$ is the fuzzy vector of length p for centroid g, where $\widetilde{\mathbf{h}}_g$, $\mathbf{h}_g^{C_1}$, $\mathbf{h}_g^{C_2}$, $\mathbf{h}_g^{L}$ and $\mathbf{h}_g^{R}$ are the g-th rows of $\widetilde{\mathbf{H}}$, $\mathbf{H}^{C_1}$, $\mathbf{H}^{C_2}$, $\mathbf{H}^{L}$ and $\mathbf{H}^{R}$, respectively. Thus, the centroids are assumed to have a complex structure inherited from the observed data. In other words, the imprecision of the observed data is propagated to the centroids that are of fuzzy nature. Moreover, the squared dissimilarity $d_w^2\left(\widetilde{\mathbf{x}}_i, \widetilde{\mathbf{h}}_g\right)$ recalled in (3) is used for comparing the observation i with centroid g. Finally, $m > 1$ is the fuzziness parameter. The membership degrees of the observations to the clusters are such that they are inversely related to the *relative* dissimilarities between the observations and the centroids. For this reason, the membership degrees can be interpreted as degrees of sharing (of the observations to the clusters).

The iterative solution is obtained by solving the constrained quadratic minimization problem in (4) through the Lagrangian multiplier method with respect to u_{ig} and by setting the first derivatives of $J_{\mathrm{F}k\mathrm{M\text{-}F}}$ with respect to $\mathbf{h}_g^{C_1}$, $\mathbf{h}_g^{C_2}$, $\mathbf{h}_g^{L}$, $\mathbf{h}_g^{R}$ and w_C equal to zero [8].

3. Motivating example

As we have noticed in the introductory section, since the FkM-F involves Euclidean distances, then it is only able to detect clusters having spherical shapes. This drawback is shown in the example below.

Example 1. Suppose that a fuzzy 2-dimensional vector of LR fuzzy numbers, $\mathcal{X}_i \equiv (\mathbf{c}_{1i}, \mathbf{c}_{2i}, \mathbf{l}_i, \mathbf{r}_i)_{LR}$ for $i \in \{1, \ldots, 100\}$, is generated by considering two clusters so that:

- for $i \in \{1, \ldots, 50\}$, $c_{1i1} \equiv U(15, 20)$, $c_{1i2} \equiv c_{1i1} \cdot 4 - 60 + \mathcal{N}(2, 1)$, $\mathbf{c}_{2i} \equiv \mathbf{c}_{1i} + [U(2, 3), U(2, 3)]$, $\mathbf{l}_i \equiv [U(0, 1), U(0, 1)]$ and $\mathbf{r}_i \equiv [U(0, 1), U(0, 1)]$;
- for $i \in \{51, \ldots, 100\}$, $c_{1i1} \equiv U(25, 31)$, $c_{1i2} \equiv c_{1i1} \cdot 4 - 100 + \mathcal{N}(2, 1)$, $\mathbf{c}_{2i} \equiv \mathbf{c}_{1i} + [U(2, 3), U(2, 3)]$, $\mathbf{l}_i \equiv [U(0, 1), U(0, 1)]$ and $\mathbf{r}_i \equiv [U(0, 1), U(0, 1)]$.

A representation of the two clusters is gathered in the left part of Figure 3. For the sake of the graphical visualization of the sample, each fuzzy vector $\mathcal{X}_i = (\mathcal{X}_{i1}, \mathcal{X}_{i2})$ is represented as a rectangle whose base and height correspond to the 0-levels of $\mathcal{X}_{i1}$ and $\mathcal{X}_{i2}$, respectively. Besides, as it is shown in the left part of Figure 3, the shape of both clusters is not spherical but ellipsoidal.

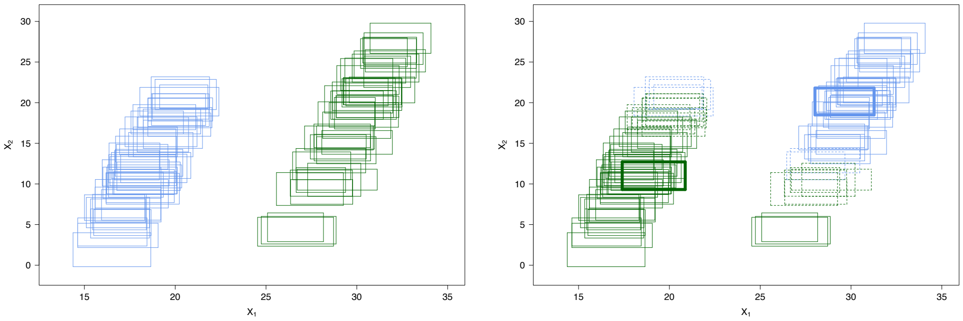

Figure 3: Example of two non-spherical clusters (on the left) and the corresponding partition when applying the FkM-F algorithm (on the right)

The FkM-F algorithm has been applied to the sample of LR fuzzy numbers generated above. In the right part of Figure 3 it can be observed that the application of the algorithm distinguishes objects having a degree of assignment to the first cluster greater than or equal to 0.5 (solid line) from those having a membership degree lower than 0.5 (dashed line). The centroids of each cluster are also highlighted by using a thicker line. It can be concluded that there are some situations in which the FkM-F algorithm is not able to detect clusters having a non-spherical shape.

4. Generalized distance for fuzzy data

The employment of the Euclidean distance in the fuzzy clustering framework (as, for example, in the fuzzy k-means procedure) leads to the good detection of the clusters when they are spherical or well separated. So, clusters of other different geometrical shapes are not well recognized [18] and, to overcome this problem, it is suitable to consider an adaptive distance norm. Gustafson and Kessel [18] proposed the use of the Mahalanobis distance, which includes the covariance matrices corresponding to each one of the clusters.

Thus, we propose to generalize the distance introduced in (3) by taking into account the Mahalanobis distance. If $\widetilde{\mathbf{x}}_i$ and $\widetilde{\mathbf{x}}_{i'}$ are two fuzzy numbers, the distance between them is given by

$$\begin{aligned} d^2_{M,w}(\widetilde{\mathbf{x}}_i, \widetilde{\mathbf{x}}_{i'}) &= w_C^2[d^2_M\left(\mathbf{c}_{1i}, \mathbf{c}_{1i'}\right) + d^2_M\left(\mathbf{c}_{2i}, \mathbf{c}_{2i'}\right)] \\ &+ w_S^2[d^2_M\left(\mathbf{l}_i, \mathbf{l}_{i'}\right) + d^2_M\left(\mathbf{r}_i, \mathbf{r}_{i'}\right)], \end{aligned} \tag{6}$$

where $d_M(x,y) = \sqrt{(x-y)^T M(x-y)}$ is the usual Mahalanobis distance, and M is a symmetric and definite positive matrix so that M^{-1} is the covariance matrix of x and y.

It is clear that $d^2_{M,w}$ is indeed a distance since it is a linear combination of distances. As for the distance in (3), the role played by the distance for the centers is supposed to be more relevant than that played by the distance for the spreads. It is the usual choice in a fuzzy context.

5. Fuzzy *k*-Means for fuzzy data based on a generalized distance

In this section a new clustering approach for fuzzy data based on the generalized distance introduced in Section 4 is proposed. It will be denoted by FkMgk-F method and it can be formalized as follows:

$$\min_{\mathbf{U},\widetilde{\mathbf{H}},\mathbf{M}_1,\cdots,\mathbf{M}_k,w} J_{\text{F}k\text{Mgk-F}} = \sum_{i=1}^{n}\sum_{g=1}^{k} u_{ig}^m d^2_{M,w}\left(\widetilde{\mathbf{x}}_i, \widetilde{\mathbf{h}}_g\right), \tag{7}$$

$$
\begin{aligned}
\text{subject to}\quad & u_{ig} \geq 0, \quad i = 1, \ldots, n, \quad g = 1, \ldots, k, \\
& \sum_{g=1}^{k} u_{ig} = 1, \quad i = 1, \ldots, n, \\
& |\mathbf{M}_g^{C_1}| = \rho_g^{C_1} > 0, |\mathbf{M}_g^{C_2}| = \rho_g^{C_2} > 0, \quad g = 1, \ldots, k, \\
& |\mathbf{M}_g^{L}| = \rho_g^{L} > 0, |\mathbf{M}_g^{R}| = \rho_g^{R} > 0, \quad g = 1, \ldots, k, \\
& w \in [0.5, 1],
\end{aligned}
\tag{8}
$$

where $\widetilde{\mathbf{H}} \equiv \left(\mathbf{H}^{C_1}, \mathbf{H}^{C_2}, \mathbf{H}^{L}, \mathbf{H}^{R}\right)_{LR}$ is the prototype matrix, $d_{M,w}$ is the generalized distance in (6), $\mathbf{M}_g^{C_1}$, $\mathbf{M}_g^{C_2}$, $\mathbf{M}_g^{L}$, $\mathbf{M}_g^{R}$ are symmetric and definite positive matrices and $\rho_g^{C_1}$, $\rho_g^{C_2}$, ρ_g^{L}, ρ_g^{R} are the volume parameters (usually equal to 1).

The optimal solution of FkMgk-F can be found by minimizing the constrained optimization problem in (7) with respect to every group of parameters. For the sake of clarity, the computation of the parameter updates are reported in the Appendix.

At every update the loss function to minimize decreases and all the parameter entities are also renewed. Then, the updates are repeated until the value of the loss function decreases less than a specified threshold (as, for instance, 10^{-5}) from the previous function value. The steps of the algorithm are provided below.

Algorithm FkMgk-F ($\widetilde{\mathbf{X}}$, m, k)

Inizialization. Generate randomly a feasible membership degree matrix $\mathbf{U}^{(0)}$ subject to (8).

Step 1. Compute the centroid matrix $\widetilde{\mathbf{H}}^{(t)}$ according to (14)-(15) using $\mathbf{U}^{(t-1)}$.

Step 2. Update the matrix $\mathbf{M}_g^{(t)}$ according to (16)-(19), keeping fixed $\mathbf{U}^{(t-1)}$ and $\widetilde{\mathbf{H}}^{(t)}$.

Step 3. Update the weight $w_C^{(t)}$ according to (20), keeping fixed $\mathbf{U}^{(t-1)}$, $\widetilde{\mathbf{H}}^{(t)}$ and $\mathbf{M}_g^{(t)}$. If $w_C^{(t)} < 0.5$, then $w_C^{(t)} = 0.5$ ($w_S^{(t)}$=1-$w_C^{(t)}$).

Step 4. Update the fuzzy membership degree matrix $\mathbf{U}^{(t)}$ according to (13), keeping fixed $\widetilde{\mathbf{H}}^{(t)}$, $\mathbf{M}_g^{(t)}$, $w_C^{(t)}$ and $w_S^{(t)}$.

Step 5. Check convergence. If the convergence condition is not satisfied, go to *Step 1.*

Remark 1. *As in case of non-fuzzy data, the cluster covariance matrices may be singular, and hence the inverse matrices cannot be calculated. This may occur when the number of objects in a cluster is small or when the data within a cluster are linearly correlated. In these situations, a proper estimation of the fuzzy covariance matrices can be obtained by adopting the approach proposed in [3]. It consists in fixing the ratio between the maximal and minimal eigenvalue of each matrix. Furthermore, in order to avoid local optima due to a poor initialization, different random starts can be considered.*

6. Simulation study

In this section the results of a simulation study are reported in order to evaluate the performance of the FkMgk-F procedure in comparison with the closest competitor, FkM-F, proposed by Coppi et *al.* in [8]. Different scenarios have been considered. In detail, samples of LR fuzzy numbers have been randomly generated from two-dimensional LR fuzzy variables. In case of 2 clusters, the sample size n is in $\{60, 120, 180, 240\}$, whilst for $k = 3$ $n \in \{90, 180, 270, 360\}$. Clusters of equal sizes have been generated in all the cases. In order to take into account the shape of the clusters, two different situations have been accounted for: spherical shape (*s-shape*) or elongated shape (*e-shape*). Finally, we have distinguished two levels of overlapping of the clusters, depending on the shape of the clusters. For spherical clusters the levels are overlapped clusters (*o-clusters*) and partially overlapped clusters (*po-clusters*). In case of spherical shape separated clusters are well recognized by both the algorithms. For elongated clusters the considered levels are separated clusters (*s-clusters*) and overlapped clusters (*o-clusters*). In case of 3 clusters there are two types of overlapping: two separated clusters and one ovelapped with respect the other ones (*2s1o-clusters*) and three overlapped clusters (*o-clusters*).

6.1. Case $k = 2$ clusters

Table 1 summarizes the details about the random generation process for the centers and the spreads of the LR random variables under the different conditions proposed above for $k = 2$ clusters.

Table 1: Set-up of the simulation study (2 clusters, 2 dimensions, $\mathcal{X} = (\mathbf{C}_1, \mathbf{C}_2, \mathbf{L}, \mathbf{R})$)

Case	Cluster	Dim	$\mathbf{C}_1$	$\mathbf{C}_2$	$\mathbf{L}$	$\mathbf{R}$
s-shape	1	1	U(0,1)	U(0,1)+1	U(0,1)	U(0,1)
po-clusters		2	U(0,1)	U(0,1)+1	U(0,1)	U(0,1)
	2	1	U(0,1)+.5	U(0,1)+1.5	U(0,1)	U(0,1)
		2	U(0,1)+.5	U(0,1)+1.5	U(0,1)	U(0,1)
s-shape	1	1	U(0,1)	U(0,1)+1	U(0,1)	U(0,1)
o-clusters		2	U(0,1)	U(0,1)+1	U(0,1)	U(0,1)
	2	1	U(0,1)	U(0,1)+1	U(0,1)	U(0,1)
		2	U(0,1)	U(0,1)+1	U(0,1)	U(0,1)
e-shape	1	1	U(1,6)	c_{111}+U(2,3)	U(0,1)	U(0,1)
s-clusters		2	4c_{111}-$\mathcal{N}(5,1)$	c_{112}+U(2,3)	U(0,1)	U(0,1)
	2	1	U(7,12)	c_{121}+U(2,3)	U(0,1)	U(0,1)
		2	4c_{121}-$\mathcal{N}(30,1)$	c_{122}+U(2,3)	U(0,1)	U(0,1)
e-shape	1	1	U(1,6)	c_{111}+U(2,3)	U(0,1)	U(0,1)
o-clusters		2	4c_{111}-$\mathcal{N}(15,1)$	c_{112}+U(2,3)	U(0,1)	U(0,1)
	2	1	U(3,8)	c_{121}+U(2,3)	U(0,1)	U(0,1)
		2	$-4c_{121}$+$\mathcal{N}(15,1)$	c_{122}+U(2,3)	U(0,1)	U(0,1)

The value of the parameter of fuzziness m has been chosen to be equal to 1.5 (it is the most used value in most of the fuzzy clustering procedures). Finally, for every level of every design variable (number of statistical units, cluster shape, cluster separation), 1000 random samples have been generated.

To evaluate the performance of the clustering methods, we observe their ability to recover the true centroids by using the REC measure:

$$\text{REC} = \sum_{g=1}^{k} \left[d^2\left(\mathbf{h}_g^{C_1}, \mathbf{h}_g^{C_1 *}\right) + d^2\left(\mathbf{h}_g^{C_2}, \mathbf{h}_g^{C_2 *}\right) + d^2\left(\mathbf{h}_g^{L}, \mathbf{h}_g^{L*}\right) + d^2\left(\mathbf{h}_g^{R}, \mathbf{h}_g^{R*}\right) \right], \quad (9)$$

where the superscript '$*$' refers to the matrices of the centers and the spreads of the true centroids. In addition, we take into account the percentage of the objects properly assigned to the clusters (POPA) and their membership degrees (MD), and the adjusted rand index (ARI) (see [21]).

6.1.1. Case s-shape po-clusters

We start with the simplest case: spherical and partially overlapped clusters. An example of this case as well as the application of the FkM-F and FkMgk-F algorithms are shown in Figure 4.

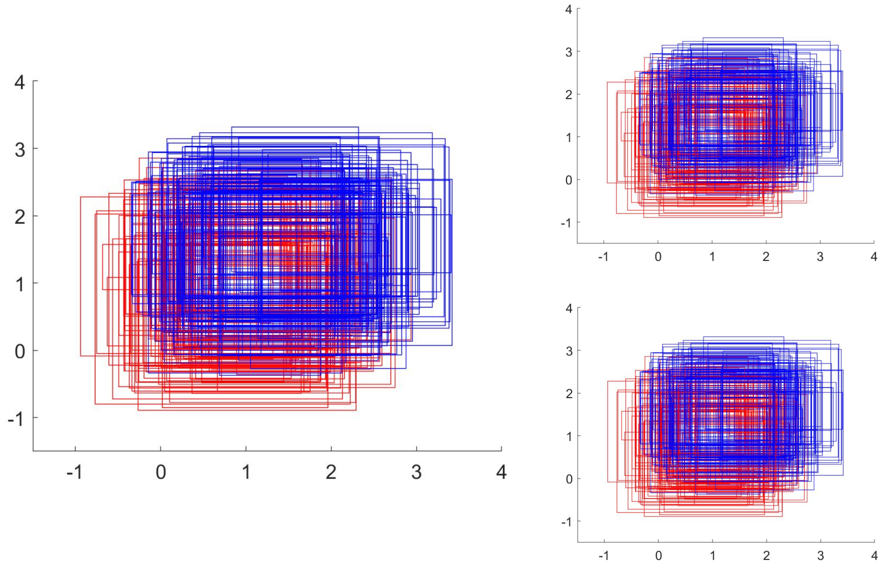

Figure 4: Example of two spherical shaped and partially overlapped clusters (left part) and application of the the FkM-F and FkMgk-F algorithms (upper and lower right parts)

The mean and median values of REC, POPA, MD of the objects properly assigned to the clusters and ARI, for both FkM-F and FkMgk-F, are reported in Table 2. As we may expect in case of spherical and partially overlapped clusters, both FkM-F and FkMgk-F work well in practice, in terms of recovery (values close to 0), proper assignment of the objects to the the clusters (high values of percentage and corresponding membership degree) and validity index, for the different sample sizes.

6.1.2. Case s-shape o-clusters

We consider now spherical shaped and overlapped clusters. An example of this case and the application of the FkM-F and FkMgk-F algorithms are shown in Figure 5. The values of the evaluating measures are reported in Table 3. As one may expect, the mean and median values of the REC are very close to 0 for both the methods in case of spherical clusters, and are lower as the sample size increases. Since the clusters are overlapped, the mean and median values of the POPA are, for both FkM-F and FkMgk-F, a little bit greater than 50% and the MDs are close to 0.5. In this case is very difficult to distinguish the two clusters. Finally, the values of ARI (in mean and median) are all close to 0, due to the overlapping.

Table 2: Evaluating measures of clustering results corresponding to the case of two spherical shaped and partially overlapped clusters ($k = 2$)

Measure	n	Mean FkM-F	Mean FkMgk-F	Median FkM-F	Median FkMgk-F
REC	60	.0061	.0104	.0055	.0087
	120	.0042	.0089	.0038	.0083
	180	.0035	.0082	.0034	.0072
	240	.0031	.0081	.0028	.0074
POPA	60	.953	.9412	.95	.95
	120	.9567	.9533	.9583	.9583
	180	.9544	.9501	.9556	.95
	240	.9578	.956	.9583	.9542
MD	60	.8231	.8048	.8235	.8098
	120	.8204	.7948	.8207	.7952
	180	.8182	.7905	.8188	.7908
	240	.818	.7885	.8181	.7865
ARI	60	.8208	.7795	.8068	.8067
	120	.8347	.8228	.8389	.8389
	180	.826	.8102	.8292	.8089
	240	.8384	.8315	.8396	.8243

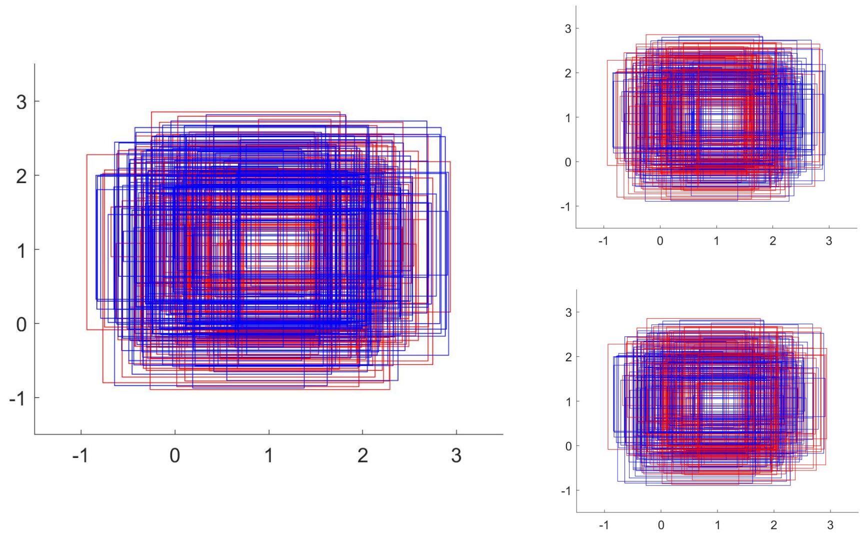

Figure 5: Example of two spherical shaped and overlapped clusters (left part) and application of the the FkM-F and FkMgk-F algorithms (upper and lower right parts)

Table 3: Evaluating measures of clustering results corresponding to the case of two spherical shaped and overlapped clusters ($k = 2$)

Measure	n	Mean FkM-F	Mean FkMgk-F	Median FkM-F	Median FkMgk-F
REC	60	.0557	.0759	.0529	.0729
	120	.0198	.0318	.0160	.0281
	180	.0091	.0128	.0079	.01
	240	.0062	.0074	.0061	.0064
POPA	60	.5595	.5668	.55	.5667
	120	.5417	.5456	.5333	.5417
	180	.5351	.5415	.5278	.5444
	240	.5334	.5312	.5292	.525
MD	60	.52	.5257	.5155	.5203
	120	.5056	.5087	.5047	.5069
	180	.5016	.5039	.5002	.5025
	240	.5003	.501	.5001	.5001
ARI	60	.0008	.001	-.0125	-.007
	120	.0001	-.002	-.0059	-.0059
	180	.0003	.0002	-.0025	-.0036
	240	.0009	-.0009	-.0017	-.0025

6.1.3. Case e-shape s-clusters

We take into account now the case of elongated clusters and we start with the well-separated ones. An example as well as the application of the FkM-F and FkMgk-F algorithms are shown in Figure 6.

Besides, by looking at the values in Table 4, we can observe that FkM-F fails in terms of REC (very high differences between estimated and true prototypes), the corresponding values of POPA and their MDs are low and the mean and median values of the ARI, for all the sample sizes, are close to 0. On the other hand, in this case, the median values corresponding to FkMgk-F are the optimal ones for all the evaluating measures. The mean values are a little bit worse because there are some anomalous values on 1000 random replications.

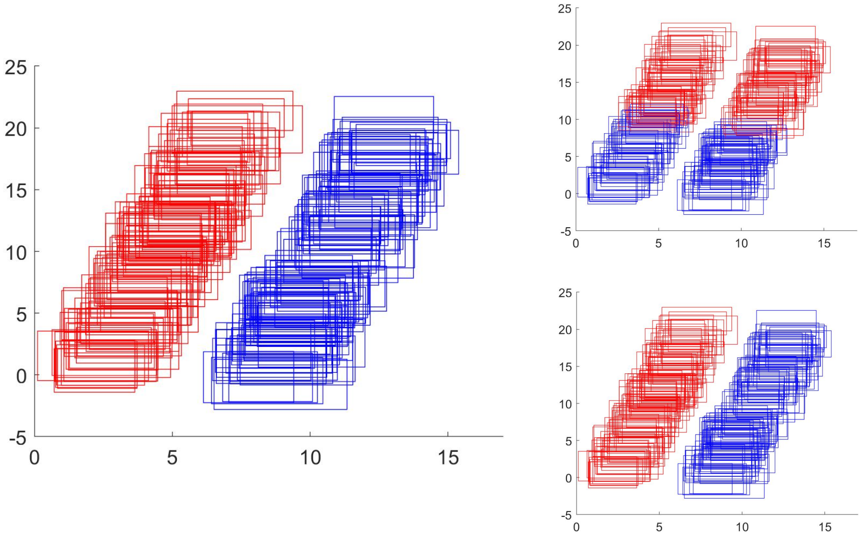

Figure 6: Example of two elongated shaped and separated clusters (left part) and application of the the FkM-F and FkMgk-F algorithms (upper and lower right parts)

Table 4: Evaluating measures of clustering results corresponding to the case of two elongated shaped and separated clusters ($k = 2$)

Measure	n	Mean FkM-F	Mean FkMgk-F	Median FkM-F	Median FkMgk-F
REC	60	117.7322	1.0144	119.0116	0
	120	125.2554	11.4096	125.6293	0
	180	131.6623	0	132.5226	0
	240	130.1286	5.3727	131.3544	0
POPA	60	.5843	.9963	.575	1
	120	.5635	.9468	.5583	1
	180	.5454	1	.5389	1
	240	.5447	.9818	.5417	1
MD	60	.5796	.9965	.57	1
	120	.5601	.9465	.5554	1
	180	.5434	1	.5386	1
	240	.5415	.9815	.5381	1
ARI	60	.0153	.9906	-.0066	1
	120	.0072	.8736	-.0011	1
	180	.0021	1	-.0011	1
	240	.0038	.96	-.0008	1

6.1.4. Case e-shape o-clusters

We consider now the most complicated case for $k = 2$ clusters: elongated and overlapped clusters. Figure 7 shows an example of this scenario and the application of the FkM-F and FkMgk-F algorithms.

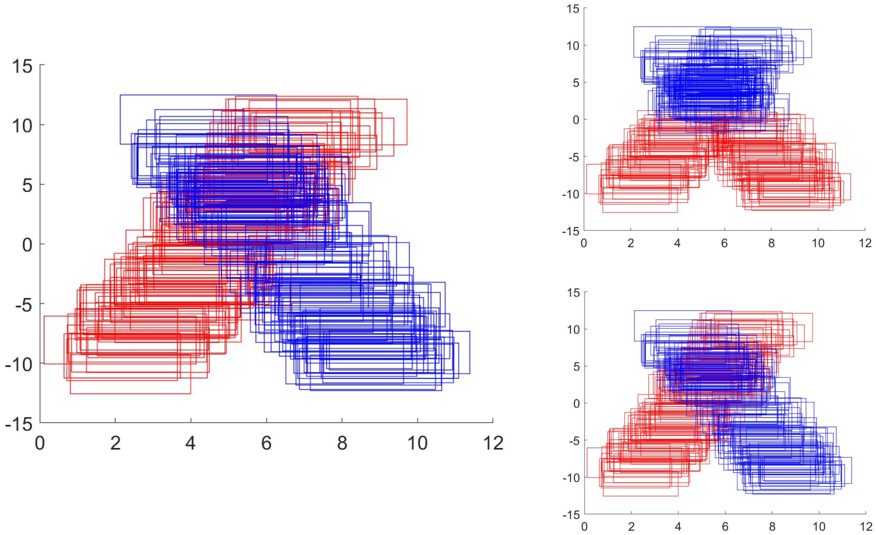

Figure 7: Example of two elongated shaped and overlapped clusters (left part) and application of the the FkM-F and FkMgk-F algorithms (upper and lower right parts)

As we expected, FkM-F doesn't work properly in this case. This is shown by the mean and median values reported in Table 5: the REC values are very high, the POPA is between 0.55 and 0.59, the MDs of objects properly assigned to the clusters is close to 0.5 and the ARI values is close to 0. Conversely, our proposal works very well in this case in terms of REC, assignment and ARI, even if the values are a little bit worse than those obtained in case of well-separated clusters, since the overlapped objects are difficult to be assigned properly.

Table 5: Evaluating measures of clustering results corresponding to the case of two elongated shaped and overlapped clusters ($k = 2$)

Measure	n	Mean FkM-F	Mean FkMgk-F	Median FkM-F	Median FkMgk-F
REC	60	90.1828	1.5819	90.2644	.2113
	120	89.5314	.327	88.9886	.194
	180	91.408	.2397	88.984	.1593
	240	88.5092	1.3067	88.8014	.1806
POPA	60	.5818	.9478	.5833	.95
	120	.5582	.9523	.55	.9583
	180	.544	.9548	.5472	.9556
	240	.5457	.9505	.5458	.95
MD	60	.5771	.9362	.5678	.9416
	120	.5555	.9409	.5527	.9425
	180	.5423	.9426	.5444	.9428
	240	.5438	.9395	.5419	.943
ARI	60	-.001	.8062	-.0118	.8068
	120	.0029	.818	-.0015	.8389
	180	.0003	.8275	-.0025	.8292
	240	.0031	.8145	-.0002	.8092

6.2. Case $k = 3$ clusters

Table 6 summarizes the details about the random generation process for the centers and the spreads of the LR random variables under the different conditions proposed for $k = 3$ clusters. The evaluating measures for clustering results are the same used in case of $k = 2$ (with the appropriate modifications).

Table 6: Set-up of the simulation study (3 clusters, 2 dimensions, $\mathcal{X} = (\mathbf{C}_1, \mathbf{C}_2, \mathbf{L}, \mathbf{R})$)

Case	Cluster	Dim	$\mathbf{C}_1$	$\mathbf{C}_2$	$\mathbf{L}$	$\mathbf{R}$
s-shape po-clusters	1	1	U(0,1)+2	U(0,1)+3	U(0,1)	U(0,1)
		2	U(0,1)+2	U(0,1)+3	U(0,1)	U(0,1)
	2	1	U(0,1)+3	U(0,1)+4	U(0,1)	U(0,1)
		2	U(0,1)+3	U(0,1)+4	U(0,1)	U(0,1)
	3	1	U(0,1)+3	U(0,1)+4	U(0,1)	U(0,1)
		2	U(0,1)+1.5	U(0,1)+2.5	U(0,1)	U(0,1)
s-shape o-clusters	1	1	U(0,1)+1	U(0,1)+2	U(0,1)	U(0,1)
		2	U(0,1)+1	U(0,1)+2	U(0,1)	U(0,1)
	2	1	U(0,1)+1.1	U(0,1)+2.1	U(0,1)	U(0,1)
		2	U(0,1)+1.1	U(0,1)+2.1	U(0,1)	U(0,1)
	3	1	U(0,1)+0.9	U(0,1)+1.9	U(0,1)	U(0,1)
		2	U(0,1)+0.9	U(0,1)+1.9	U(0,1)	U(0,1)
e-shape s-clusters	1	1	U(1,6)	c_{111}+U(2,3)	U(0,1)	U(0,1)
		2	$4c_{111}$-$\mathcal{N}(5,1)$	c_{112}+U(2,3)	U(0,1)	U(0,1)
	2	1	U(7,12)	c_{121}+U(2,3)	U(0,1)	U(0,1)
		2	$4c_{121}$-$\mathcal{N}(30,1)$	c_{122}+U(2,3)	U(0,1)	U(0,1)
	3	1	U(0,1)	U(0,1)+1	U(0,1)	U(0,1)
		2	U(0,1)	U(0,1)+1	U(0,1)	U(0,1)
e-shape 2s1o-clusters	1	1	U(2,7)	c_{111}+U(2,3)	U(0,1)	U(0,1)
		2	$4c_{111}$-$\mathcal{N}(20,1)$	c_{112}+U(2,3)	U(0,1)	U(0,1)
	2	1	U(12,17)	c_{121}+U(2,3)	U(0,1)	U(0,1)
		2	$4c_{121}$-$\mathcal{N}(60,1)$	c_{122}+U(2,3)	U(0,1)	U(0,1)
	3	1	U(0,25)	c_{131}+U(2,3)	U(0,1)	U(0,1)
		2	$\mathcal{N}(0,1)$	c_{132}+U(2,3)	U(0,1)	U(0,1)
e-shape o-clusters	1	1	U(2,7)	c_{111}+U(2,3)	U(0,1)	U(0,1)
		2	$4c_{111}$-$\mathcal{N}(20,1)$	c_{112}+U(2,3)	U(0,1)	U(0,1)
	2	1	U(3,8)	c_{121}+U(2,3)	U(0,1)	U(0,1)
		2	$-4c_{121}$+$\mathcal{N}(25,1)$	c_{122}+U(2,3)	U(0,1)	U(0,1)
	3	1	U(0,13)	c_{131}+U(2,3)	U(0,1)	U(0,1)
		2	$\mathcal{N}(0,1)$	c_{132}+U(2,3)	U(0,1)	U(0,1)

6.2.1. Case s-shape po-clusters

In case of $k = 3$ we start with the simplest scenario: spherical shaped and partially-overlapped clusters. An example and the application of the FkM-F and FkMgk-F algorithms in this case are shown in Figure 8.

As one may expect, the mean and the median values of the evaluating measures, gathered in Table 7, analogously to the case of $k = 2$, confirm that both the methods, FkM-F and FkMgk-F, work well with spherical shaped and partially-overlapped clusters.

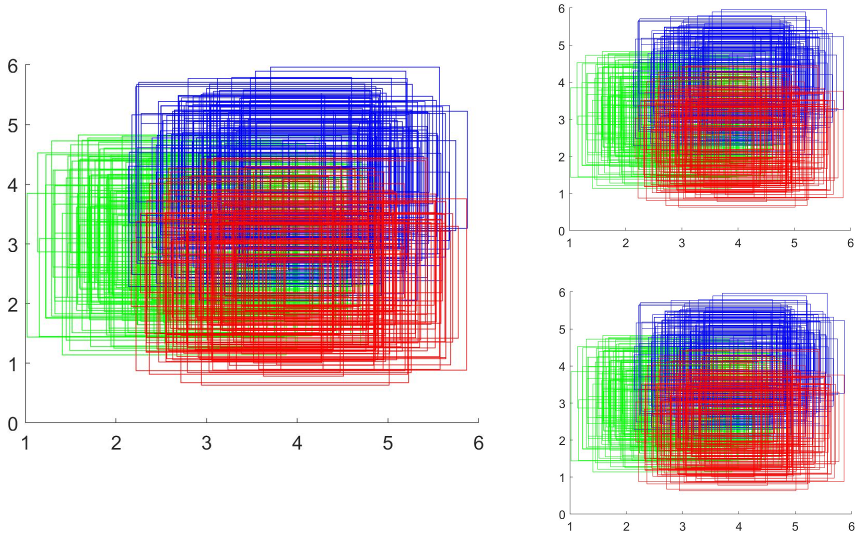

Figure 8: Example of three spherical shaped and partially-overlapped clusters (left part) and application of the the FkM-F and FkMgk-F algorithms (upper and lower right parts)

Table 7: Evaluating measures of clustering results corresponding to the case of three spherical shaped and partially-overlapped clusters ($k = 3$)

Measure	n	Mean FkM-F	Mean FkMgk-F	Median FkM-F	Median FkMgk-F
REC	90	.0009	.0012	.0008	.0009
	180	.0006	.0007	.0005	.0006
	270	.0004	.0006	.0004	.0005
	360	.0004	.0005	.0004	.0004
POPA	90	.9993	.9986	1	1
	180	.9995	.9989	1	1
	270	.9991	.9989	1	1
	360	.9994	.9992	1	1
MD	90	.9364	.9367	.9363	.9365
	180	.935	.9339	.935	.9341
	270	.9337	.9312	.9339	.9312
	360	.9337	.9314	.9331	.9308
ARI	90	.9987	.9967	1	1
	180	.9987	.9973	1	1
	270	.9974	.9968	1	1
	360	.9987	.9981	1	1

6.2.2. Case s-shape o-clusters

The second scenario for $k = 3$ corresponds to spherical shaped and overlapped clusters. An example of this case and the application of the FkM-F and FkMgk-F algorithms are shown in Figure 9.

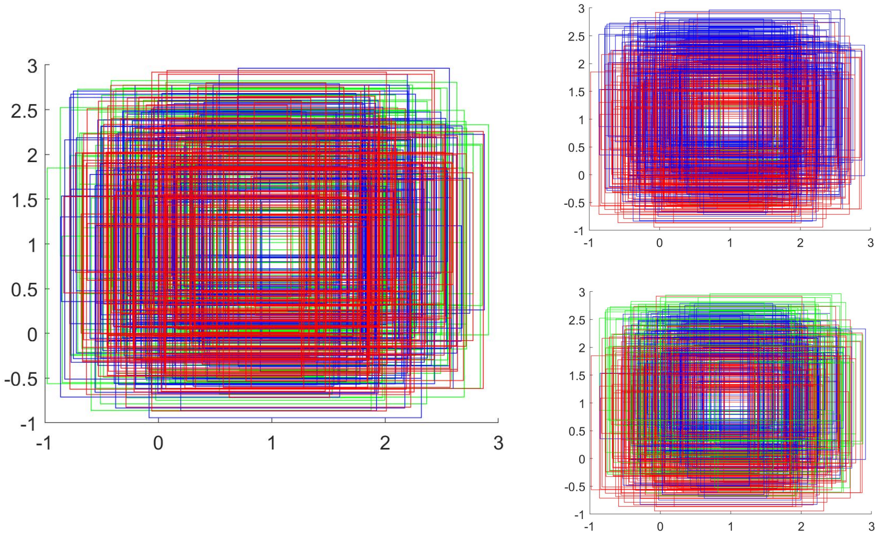

Figure 9: Example of three spherical shaped and overlapped clusters (left part) and application of the the FkM-F and FkMgk-F algorithms (upper and lower right parts)

The recovery values for both FkM-F and FkMgk-F are close to 0 for all the considered sample sizes (see Table 8). The POPA values are higher as the sample size increases. The median values for both methods are close to 1 for $n \geq 270$. Since the clusters are overlapped, the MDs are close to 0.33 and the ARI values are close to 0.

6.2.3. Case e-shape s-clusters

We take into account now the case of three elongated shaped clusters. In particular, we start with the separated ones. An example as well as the application of the FkM-F and FkMgk-F algorithms are shown in Figure 10.

As for $k = 2$, also for three elongated shaped and separated clusters, as reported in Table 9, the median values corresponding to FkMgk-F are the optimal ones. On the contrary, FkM-F works worse, even if the values of the POPA and of the ARI are higher than those obtained for $k = 2$.

Table 8: Evaluating measures of clustering results corresponding to the case of three spherical shaped and overlapped clusters ($k = 3$)

Measure	n	Mean FkM-F	Mean FkMgk-F	Median FkM-F	Median FkMgk-F
REC	90	.1050	.1465	.1051	.1428
	180	.0756	.0806	.0747	.0779
	270	.0822	.0886	.0832	.0901
	360	.0832	.0887	.0841	.0872
POPA	90	.6131	.5471	.6	.5444
	180	.8094	.7098	.7444	.6639
	270	.914	.9114	1	.9963
	360	.9563	.9882	1	1
MD	90	.3699	.3817	.3704	.3808
	180	.3468	.3491	.3383	.3467
	270	.3384	.3367	.334	.334
	360	.3363	.3342	.3338	.3337
ARI	90	.0103	.0161	.0032	.0096
	180	.0016	.0028	0	0
	270	.0002	.0001	0	0
	360	0	0	0	0

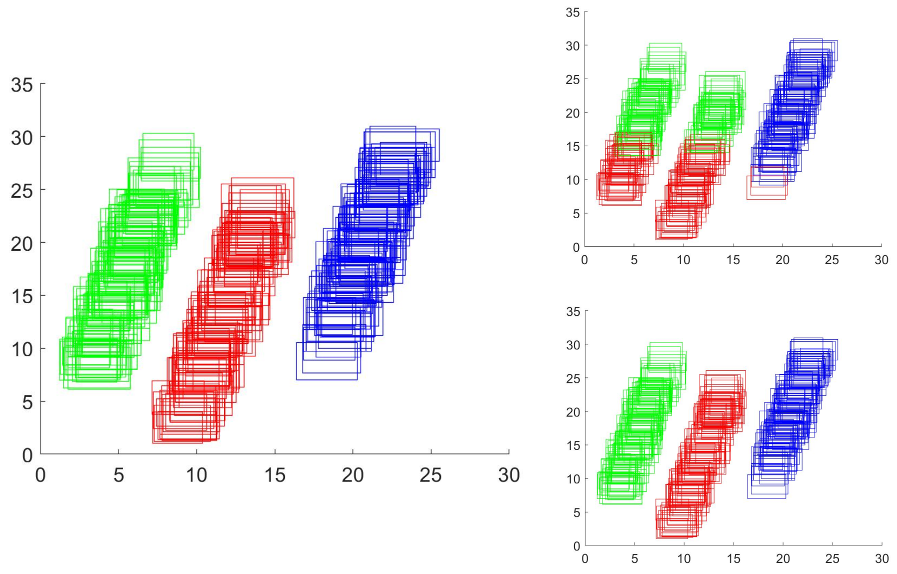

Figure 10: Example of three elongated shaped and separated clusters (left part) and application of the the FkM-F and FkMgk-F algorithms (upper and lower right parts)

Table 9: Evaluating measures of clustering results corresponding to the case of three elongated shaped and separated clusters ($k = 3$)

Measure	n	Mean FkM-F	Mean FkMgk-F	Median FkM-F	Median FkMgk-F
REC	90	74.0365	13.9067	68.9704	0
	180	73.4257	15.1809	70.3438	0
	270	76.4597	21.0297	73.9107	0
	360	75.7877	32.1718	73.4104	0
POPA	90	.7344	.963	.7333	1
	180	.7218	.9588	.7222	1
	270	.7138	.9321	.7148	1
	360	.7083	.9126	.7083	1
MD	90	.7259	.9625	.7277	1
	180	.7134	.9583	.7193	1
	270	.7061	.9314	.7085	1
	360	.7034	.9119	.7069	1
ARI	90	.4637	.9347	.4584	1
	180	.4578	.9305	.4685	1
	270	.4389	.8858	.439	1
	360	.4395	.8556	.4446	1

6.2.4. Case e-shape 2s1o-clusters

For $k = 3$ elongated shaped clusters, we consider two kinds of overlapping. The first one consists in two separated clusters and one overlapped with respect the other two. An example of this specific case and the application of the FkM-F and FkMgk-F algorithms are shown in Figure 11.

The mean and median values of the evaluating measures in Table 10 confirm what we expected in this case. The prototypes estimated by means of FkM-F are very different from the true prototypes (high values of REC), the POPA is more or less equal to 60% with a MD between 0.61 and 0.65 an the ARI values are low. Conversely, the values corresponding to FkMgk-F show that it works very well, even if this is a more complex scenario and the values are lower than those obtained for the separated clusters.

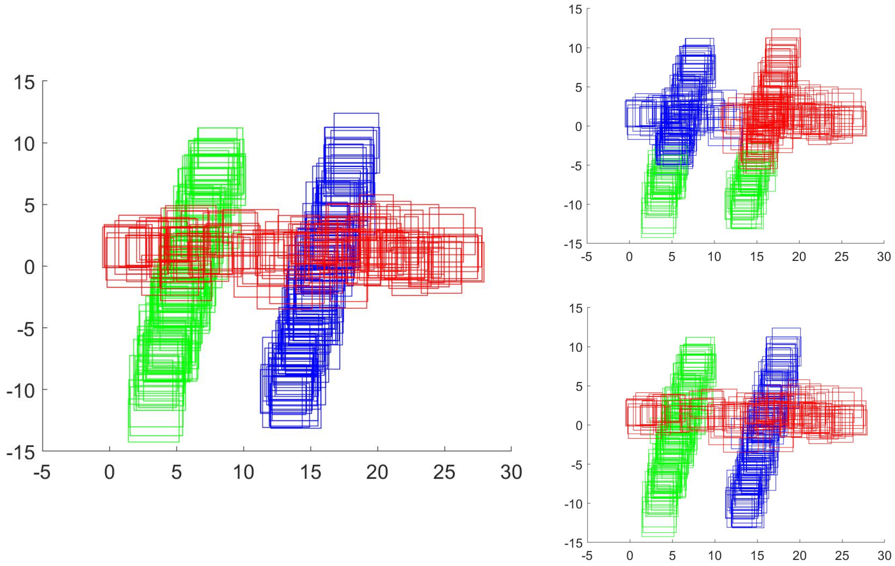

Figure 11: Example of three elongated shaped and separated clusters whose two are separated and one is overlapped (left part) and application of the the FkM-F and FkMgk-F algorithms (upper and lower right parts)

Table 10: Evaluating measures of clustering results corresponding to the case of three elongated shaped clusters whose two are separated and one is overlapped ($k = 3$)

Measure	n	Mean FkM-F	Mean FkMgk-F	Median FkM-F	Median FkMgk-F
REC	90	162.6707	15.7094	168.2827	2.2571
	180	179.036	8.4139	190.7022	1.5249
	270	198.3198	3.6163	207.6049	1.435
	360	196.7962	3.7305	201.5823	1.4458
POPA	90	.6541	.8993	.6556	.9333
	180	.6329	.9157	.6222	.9444
	270	.6211	.9385	.6222	.9444
	360	.6266	.9367	.6236	.9444
MD	90	.6389	.8928	.6429	.9317
	180	.6236	.9084	.6213	.9355
	270	.6152	.9302	.6142	.9368
	360	.6201	.9284	.6188	.9387
ARI	90	.3094	.7625	.2976	.8137
	180	.2922	.7941	.2784	.8419
	270	.2776	.8361	.2672	.8432
	360	.2865	.8333	.2781	.8434

6.2.5. Case e-shape o-clusters

We conclude with the more complex case: three elongated and overlapped clusters. Figure 12 shows an example of this scenario and the application of both algorithms.

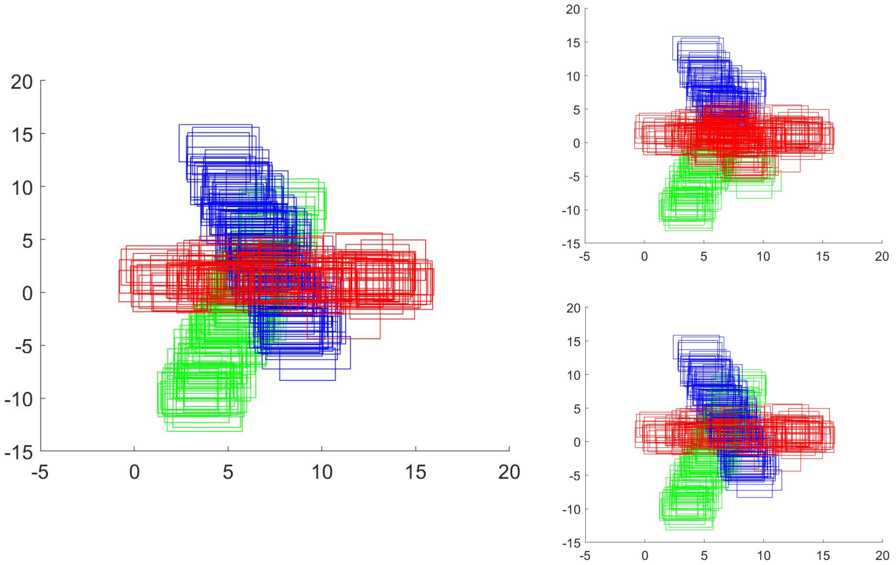

Figure 12: Example of three elongated shaped and overlapped clusters (left part) and application of the the FkM-F and FkMgk-F algorithms (upper and lower right parts)

Taking into account the values reported in Table 11, the conclusions are very similar to those of the previous case (three elongated shaped clusters so that two are separated and one is overlapped with respect the other two), but the values are a little bit worse.

7. Real-case study

In this section, a real-life application is provided in order to show the applicability of the procedure presented in this work. Thus, a questionnaire of 16 statements has been proposed to a group of students.

On one hand, some statements about the learning style of the students (independent or dependent style) have been drawn from the so called Grasha-Riechmann Learning Style Scales (GRLSS, [30]). Those questions are the following ones:

Table 11: Evaluating measures of clustering results corresponding to the case of three elongated shaped and overlapped clusters ($k = 3$)

Measure	n	Mean FkM-F	Mean FkMgk-F	Median FkM-F	Median FkMgk-F
REC	90	100.3127	14.4611	101.3849	1.0966
	180	96.7781	4.9723	93.672	.7752
	270	96.8149	3.5538	96.1229	.6491
	360	97.1758	2.8277	97.2105	.603
POPA	90	.6471	.8324	.65	.8889
	180	.6405	.8724	.6444	.8944
	270	.6411	.887	.6407	.9
	360	.6332	.8875	.6306	.8986
MD	90	.6159	.8228	.6193	.8789
	180	.6083	.8613	.6094	.8810
	270	.6074	.8742	.6086	.8869
	360	.6005	.8741	.5994	.8859
ARI	90	.2155	.6064	.2082	.6883
	180	.2113	.6750	.2092	.7109
	270	.2191	.7043	.2178	.7301
	360	.2142	.7046	.2118	.7186

I_1. “*I prefer to work on my own in the tasks of my subjects than to ask the teacher or my classmates for help.*”

D_1. “*I trust that the teachers will tell me what is important to study.*”

I_2. “*I learn many of the contents that are given in the classes on my own.*”

D_2. “*I complete assignments exactly the way teachers tell me to do it.*”

I_3. “*I feel confident with my ability to learn on my own.*”

D_3. “*The fact of having to decide what to study or how to perform tasks makes me feel uncomfortable.*

I_4. “*If I like a topic, I try to find more information about it on my own.*”

D_4. “*I prefer that the classes are very organized.*”

Questions I_j, for $j \in \{1, \dots, 4\}$, correspond to an independent learning style whereas D_j, for $j \in \{1, \dots, 4\}$, correspond to a dependent learning style.

On the other hand, some statements regarding students’ mathematics related beliefs have been proposed. Those are the following ones:

M_1. *“I think it’s interesting what I learn in math class.”*

M_2. *“I like to do math stuff.”*

M_3. *“I hope to do well on math assignments and math tests.”*

M_4. *“Compared with other colleagues I think I’m good at math.”*

M_5. *“I think I’m going to do well in math this year.”*

M_6. *“I understand everything we have done in math this year.”*

M_7. *“Doing the best I can in math I try to show my teacher that I am better than other classmates.”*

M_8. *“I work hard in math to show the teacher and my classmates how good I am.”*.

A group of 114 students attending the second course of the Degree in Primary Education of the University of Cantabria (Spain) are asked to reflect on these 16 statements. The respondents employed trapezoidal fuzzy sets in a scale ranging from 0 to 10 (where 0 represents totally disagree and 10 represents totally agree). The 0-level of each response is the set of values that the student considers compatible with his/her opinion at some extent (that is, the student considers that his/her opinion cannot be outside of this set). On the other hand, the 1-level of the trapezoidal fuzzy set is the set of values that the student considers completely compatible with his/her opinion. Finally, the corresponding limits of the 0-level and 1-level can be linearly interpolated in order to obtain a trapezoid. In practice we asked the students to mark only the four vertices of the trapezoid. An example of answer is shown in Figure 13.

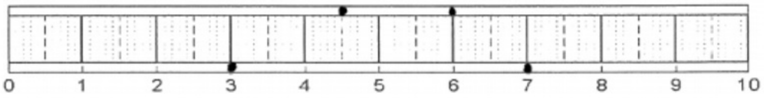

Figure 13: Example of answer of a question

We apply the FkM-F and the FkMgk-F procedures in two different scenarios: firstly, to the set of questions $Q_1 = \{I_1, D_1, I_2, D_2, I_3, D_3, I_4, D_4\}$, and later to the questions related to mathematics beliefs, i.e., $Q_2 = \{M_1, M_2, M_3, M_4, M_5, M_6, M_7, M_8\}$.

In the first case, $k = 2$ and $m = 1.5$ has been considered to cluster the answers of the statements included in Q_1. The corresponding centers of the prototypes obtained by carrying out both procedures are the following ones:

$$\mathbf{H}^{C_1} = \begin{bmatrix} 6.3634 & 8.3314 & 6.6028 & 7.6 & 7.062 & 7.2614 & 6.9142 & 5.8514 \\ 4.3431 & 7.6274 & 5.3966 & 6.5613 & 3.6735 & 6.8554 & 5.8238 & 4.6524 \end{bmatrix}$$

$$\mathbf{H}^{C_2} = \begin{bmatrix} 6.6653 & 8.5596 & 7.1415 & 7.8426 & 7.4112 & 7.5203 & 7.1442 & 6.1401 \\ 4.7189 & 7.8416 & 5.8823 & 6.8028 & 3.917 & 7.0854 & 6.126 & 4.9589 \end{bmatrix}$$

$$\mathbf{H}^{C_1}_{GK} = \begin{bmatrix} 4.8368 & 9.09 & 5.6582 & 8.0035 & 4.9777 & 8.0185 & 5.3324 & 5.6371 \\ 5.6448 & 7.253 & 6.1975 & 6.5586 & 5.7267 & 6.4657 & 6.9844 & 5.0585 \end{bmatrix}$$

$$\mathbf{H}^{C_2}_{GK} = \begin{bmatrix} 5.1797 & 9.2476 & 6.8112 & 8.2918 & 5.2193 & 8.1644 & 5.6596 & 5.8678 \\ 5.9584 & 7.5139 & 6.3976 & 6.7788 & 6.0594 & 6.7731 & 7.1966 & 5.3878 \end{bmatrix}$$

In the FkMF case, the first cluster corresponds to high values for the eight statements whereas the second cluster includes low values for the eight statements. Instead, the FkMgk-F procedure provides two clusters so that the first one corresponds to higher values for the independent learning style statements and the second one considers higher values for the dependent learning style statements. In this scenario it has more sense that a student having a more dependent learning style gives higher values to the corresponding dependent style statements than to the independent style ones and vice versa. Therefore, the results obtained using the FkMgk-F procedure are more coherent and they adapt better to reality. The case k = 3 has not been contemplated in this first study since it is known in advance that the statements are associated to two groups of different learning styles.

The second study concerns a clustering of the answers of the statements included in Q_2. In this case, $k \in \{2, 3\}$ and $m = 1.5$ has been considered. The results are provided below.

<u>Case k=2:</u>

The corresponing centres of the prototypes are:

$$\mathbf{H}^{C_1} = \begin{bmatrix} 6.779 & 6.7554 & 8.5978 & 6.5743 & 7.3768 & 7.2523 & 4.8386 & 4.5758 \\ 4.6399 & 2.5791 & 7.5614 & 2.363 & 5.248 & 4.3625 & 2.0809 & 2.0074 \end{bmatrix}$$

$$\mathbf{H}^{C_2} = \begin{bmatrix} 6.9306 & 6.9868 & 8.7919 & 6.7872 & 7.629 & 7.4169 & 5.0289 & 4.6726 \\ 4.9549 & 2.7564 & 7.7823 & 2.6335 & 5.5515 & 4.6875 & 2.3018 & 2.1773 \end{bmatrix}$$

$$\mathbf{H}^{C_1}_{GK} = \begin{bmatrix} 4.9202 & 4.0956 & 8.8355 & 4.0448 & 6.1356 & 5.3711 & 3.6512 & 3.4002 \\ 6.4100 & 5.3410 & 7.4778 & 5.0011 & 6.5555 & 6.2914 & 3.3432 & 3.259 \end{bmatrix}$$

$$\mathbf{H}^{C_2}_{GK} = \begin{bmatrix} 5.1184 & 4.3150 & 8.9856 & 4.2845 & 6.4075 & 5.5874 & 3.8365 & 3.4899 \\ 6.6668 & 5.54 & 7.7337 & 5.2478 & 6.8341 & 6.5632 & 3.5614 & 3.4247 \end{bmatrix}$$

Again, in the FkM-F case the first cluster corresponds to high values for the eight statements whereas the second cluster includes low values for the eight statements. Thus, the FkM-F procedure does not distinguish groups of statements attending to different aspects concerning students' mathematics related beliefs but classifies individuals according to whether the assessments assigned to all statements have been high or low. Nonetheless, in the FkMgk-F case the first cluster includes lower values for statements M_1, M_2, M_4, M_5 and M_6 and the second one comprises lower values for statements M_3, M_7 and M_8. The first group of statements corresponds to math liking, math interest and selfconcept about mathematical abilities whereas the statements of the second group encompasses own expectations in math marks, competitiveness and insecurity when dealing with math. In other words, the FkMgk-F classifies on one hand individuals who like mathematics, who are interested in the subject, who have a more or less high selfconcept concerning their mathematics abilities and that, additionally, who are self-confident, not so much competitive and who don't have high expectations on their marks (since they are centered on learning and not only on their marks). On the other hand, we have individuals who don't like math so much, who don't show a lot of interest in the subject, and who have a low selfconcept about their abilities but, otherwise, who have high expectations on their marks and are insecure and competitive (which could mean that this kind of students are centered on passing the exams rather than on learning the contents of the subject).

Case k=3:
The corresponing centres of the prototypes are:

$$\mathbf{H}^{C_1} = \begin{bmatrix} 6.9264 & 7.3403 & 8.8379 & 7.3847 & 7.7795 & 7.7881 & 6.4017 & 5.8416 \\ 6.2601 & 5.3416 & 7.9498 & 4.8388 & 6.4442 & 6.0365 & 2.9041 & 2.8425 \\ 3.9104 & 1.5313 & 7.6441 & 1.5761 & 4.8296 & 3.7909 & 1.6783 & 1.7412 \end{bmatrix}$$

$$\mathbf{H}^{C_2} = \begin{bmatrix} 7.063 & 7.5421 & 9.0588 & 7.5738 & 7.9621 & 7.8879 & 6.5118 & 5.9099 \\ 6.519 & 5.6085 & 8.1327 & 5.1259 & 6.8242 & 6.3714 & 3.1761 & 3.0031 \\ 4.193 & 1.6707 & 7.8644 & 1.787 & 5.0571 & 4.0404 & 1.8723 & 1.8856 \end{bmatrix}$$

$$\mathbf{H}^{C_1}_{GK} = \begin{bmatrix} 6.5312 & 5.4932 & 8.1294 & 5.2632 & 6.7365 & 6.3884 & 3.3237 & 3.2858 \\ 4.3411 & 4.4503 & 8.7022 & 4.2154 & 5.9486 & 4.9992 & 3.7519 & 3.8194 \\ 6.0136 & 4.3089 & 7.659 & 4.1491 & 6.3368 & 6.0381 & 3.4945 & 3.0275 \end{bmatrix}$$

$$\mathbf{H}^{C_2}_{GK} = \begin{bmatrix} 6.8369 & 5.653 & 8.3029 & 5.5734 & 7.0608 & 6.7169 & 3.6204 & 3.4912 \\ 4.5507 & 4.6212 & 8.8925 & 4.3795 & 6.15 & 5.1658 & 3.8874 & 3.8794 \\ 6.1846 & 4.5966 & 7.9104 & 4.3855 & 6.6249 & 6.2635 & 3.6648 & 3.1435 \end{bmatrix}$$

In this case, by considering the FkM-F procedure, the first cluster corresponds to high values for the eight statements, the second one medium values and the third one low values. Following the same idea as before, the FkM-F procedure classifies individuals according to whether the assessments assigned to all statements have been high, medium or low but do not refer to aspects about students' mathematics related beliefs. On the other hand, in the FkMgk-F case we can outline three groups: the first one includes statements M_1, M_2, M_4, M_5 and M_6 (high values for the first cluster, medium values for the third cluster and low values for the second one), the second one comprises statements M_3 and M_8 (in general high values for the first cluster, medium values for the third cluster and low values for the second one, although statements M2 and M4 present similar results associated to clusters 2 and 3), and the third one comprehends statement M_7. The groups are the same as in the case $k = 2$ except for statement M_7, since it is different the intention when showing that a student is good at math than the intention when showing that a student is better than others. To summarize, the conclusions are similar to the case k=2 except for the statement M7 which highlights the competitiveness of the students in contrast to their insecurity and own expectations in math marks.

8. Concluding remarks

In this paper we have introduced a generalized distance for fuzzy data, taking into account the correlation structures among the variables, by means of the Mahalanobis distance. This is very useful in multivariate analyses, in particular, in a clustering context. The usual metrics for fuzzy data, based on the Euclidean one, have a limited use since they work properly only in case of spherical or well separated clusters. When the clusters have non spherical shape, and the variables are correlated, the classical fuzzy methods may fail. For this reason, we have proposed to consider the generalized adaptive distance in a fuzzy k-means like algorithm. The simulation studies in different scenarios have confirmed the adequacy of the proposal, in comparison with the classical fuzzy k-means. In addition, the usefulness of the proposal has been checked also by means of a very interesting real case study.

As a direction for a further research, we are working on developing robust versions of the current study in order to be able to detect anomalous data.

Acknowledgements

The research in this paper has been partially supported by the Spanish Government through MINECO-18-MTM2017-89632-P Grant and by the COST Action 1408. Their finantial support is greatfully acknowledged.

References

[1] Antoine, V., Quost, B., Masson, M.-H., Denoeux, T., 2012. CECM: Constrained Evidential C-Means algorithm. Computational Statistics and Data Analysis, 56 (4), 894–914.

[2] Blanco-Fernández, Á, Casals, M.R., Colubi, A., Corral, N., García-Bárzana, M., Gil, M.A., González-Rodríguez, G., López, M.T., Montenegro, M., Lubiano, M.A., Ramos-Guajardo, R.G., de la Rosa de Saa, S., Sinova, B., 2013. Random fuzzy sets: a mathematical tool to develop statistical fuzzy data analysis. Iranian Journal of Fuzzy Systems, 10(2), 1–28.

[3] Babuska, R., Van der Veen, P.J., Kaymak, U., 2002. Improved covariance estimation for Gustafson-Kessel clustering. In: Proc. 2002 IEEE International Conference on Fuzzy Systems, 2, Honolulu, HI, 1081–1085.

[4] Bezdek, J.C., 1981. Pattern recognition with fuzzy objective function algorithm. Plenum Press, New York.

[5] Billingsley, P., 1968. Convergence of probability measures. Wiley, New York.

[6] Colubi, A., González-Rodríguez, G., 2015. Fuzziness in data analysis: Towards accuracy and robustness. Fuzzy Sets and Systems, 281, 260–271.

[7] Coppi, R., Gil, M.A., Kiers, H., 2008. The fuzzy approach to statistical analysis. Computational Statistics and Data Analysis, 51, 196–214.

[8] Coppi, R., D'Urso, P., Giordani, P., 2012. Fuzzy and possibilistic clustering for fuzzy data. Computational Statistics and Data Analysis, 56, 915–927.

[9] De Carvalho, F. de A.T., Tenório, C.P., 2010. Fuzzy K-means clustering algorithms for interval-valued data based on adaptive quadratic distances. Fuzzy Sets and Systems, 161, 2978–2999.

[10] Doering, C., Lesot, M.J., Kruse, R., 2006. Data analysis with fuzzy clustering methods. Computational Statistics and Data Analysis, 51, 192–214.

[11] Dubois, D., Prade. H., 1980. Fuzzy Sets and Systems: Theory and Applications. Academic Press, New York.

[12] Dunn, J.C., 1974. A fuzzy relative to the ISODATA process and its use in detecting compact, well-separated clusters. Journal of Cybernetics, 3, 32–57.

[13] D'Urso, P., Giordani, P., 2006. A weighted fuzzy c-means clustering model for fuzzy data. Computational Statistics and Data Analysis, 50, 1496–1523.

[14] D'Urso, P., De Giovanni, L., 2015. Trimmed fuzzy clustering for interval-valued data. Advances in Data Analysis and Classification, 9, 21–40.

[15] El-Sonbaty, Y., Ismail, M.A., 1998. Fuzzy clustering for symbolic data. IEEE Transactions on Fuzzy Systems, 6, 195–204.

[16] Ferraro, M.B., Giordani, P., 2013. On possibilistic clustering with repulsion constraints for imprecise data. Information Sciences, 245, 63–75.

[17] Ferraro, M.B., Giordani, P., 2015. A toolbox for fuzzy clustering using the R programming language. Fuzzy Sets and Systems, 279, 1–16.

[18] Gustafson, E., Kessel, W., 1979. Fuzzy clustering with a fuzzy covariance matrix. In: Proc. of IEEE CDC, San Diego, CA, USA, 761–766.

[19] Hariz, S.B., Elouedi, Z., Mellouli, K., 2006. Clustering approach using belief function theory, in: J. Euzenat, J. Domingue (Eds.), Artificial Intelligence: Methodology, Systems, and Applications, Springer, pp. 162–171.

[20] Havens, T., Bezdek, J., Leckie, C., Hall, L., Palaniswami, M., 2012. Fuzzy c-means algorithms for very large data. IEEE Trans. Fuzzy Syst., 20, 1130–1146.

[21] Hubert, L., Arabie, P., 1985. Comparing partitions. Journal of classification, 2, 19–218.

[22] Krishnapuram, R., Joshi, A., Yi, L., 1999. A fuzzy relative of the k-medoids algorithm with application to web document and snippet clustering In: Fuzzy Systems Conference Proceedings FUZZ-IEEE '99, 1281–1286.

[23] Maronna, R., Martin, D., Yohai, V., 2006. Robust Statistics: Theory and Methods, Wiley.

[24] Masson, M.-H., Denoeux, T., 2008. ECM: an evidential version of the fuzzy c-means algorithm. Pattern Recognition 41, 1384–1397.

[25] Masson, M.-H., Denoeux, T., 2009. RECM: relational evidential c-means algorithm. Pattern Recognition Letters 30, 1015–1026.

[26] Nguyen, H.T., 1978. A note on the extension principle for fuzzy sets. Journal of Mathematical Analysis and Applications 64, 369–380.

[27] Puri, M.L., Ralescu, D.A., 1986. Fuzzy random variables, Journal of Mathematical Analysis and Applications 114, 409–422.

[28] Quost, B., Denoeux, T., 2010. Clustering fuzzy data using the fuzzy EM algorithm, in: A. Deshpande, A. Hunter (Eds.), Proceedings of the 4th International Conference on Scalable Uncertainty Management, SUM'2010, Toulouse, France, Springer-Verlag, pp. 333–346.

[29] Quost, B., Denoeux, T., 2016. Clustering and classification of fuzzy data using the fuzzy EM algorithm, Fuzzy Sets and Systems, 286, 134–156.

[30] Riechmann, S.W., Grasha, A.F., 1974. A rational approach to developing and assessing the construct validity of a student learning style scales insstrument. The Journal of Psychology: Interdisciplinary and Applied, 87(2), 213–223.

[31] Trutschnig, W., González-Rodríguez, G., Colubi, A., Gil, M.A., 2009. A new family of metrics for compact, convex (fuzzy) sets based on a generalized concept of mid and spread. Information Sciences, 179, 3964–3972.

[32] Zadeh, L.A., 1965. Fuzzy sets. Information and Control 8, 338–353.

[33] Zimmermann, H.J., 1996. Fuzzy Set Theory and its Applications. Third edition. Ed. Kluwer Academic Publishers.

Appendix

In order to obtain the optimal fuzzy membership degree matrix $\mathbf{U}$ of the FkMgk-F method, we consider the following Lagrangian function defined as $L(\mathbf{U}, \widetilde{\mathbf{H}}, \mathbf{M}_1, \cdots, \mathbf{M}_k, w, \lambda, \beta^{C_1}, \beta^{C_2}, \beta^L, \beta^R)$ which is equal to

$$\begin{aligned}\sum_{i=1}^{n}\sum_{g=1}^{k} u_{ig}^{m} d_{M,w}^{2}\left(\widetilde{\mathbf{x}}_i, \widetilde{\mathbf{h}}_g\right) \quad & -\sum_{i=1}^{n} \lambda_i \left(\sum_{g=1}^{k} u_{ig} - 1\right) \\ & -\sum_{g=1}^{k} \beta_g^{C_1}\left(|\mathbf{M}_g^{C_1}| - \rho_g^{C_1}\right) \\ & -\sum_{g=1}^{k} \beta_g^{C_2}\left(|\mathbf{M}_g^{C_2}| - \rho_g^{C_2}\right) \qquad (10) \\ & -\sum_{g=1}^{k} \beta_g^{L}\left(|\mathbf{M}_g^{L}| - \rho_g^{L}\right) \\ & -\sum_{g=1}^{k} \beta_g^{R}\left(|\mathbf{M}_g^{R}| - \rho_g^{R}\right)\end{aligned}$$

To find the optimal value of $\mathbf{U}$ we compute the partial derivatives of (10) with respect to u_{ig} and λ_i and we set them equal to 0:

$$\frac{\vartheta L}{\vartheta u_{ig}} = 0 \Leftrightarrow m u_{ig}^{m-1} d_{M,w}^2 \left(\widetilde{\mathbf{x}}_i, \widetilde{\mathbf{h}}_g \right) - \lambda = 0, \tag{11}$$

$$\frac{\vartheta L}{\vartheta \lambda_i} = 0 \Leftrightarrow \sum_{g=1}^{k} u_{ig} - 1 = 0. \tag{12}$$

Then, by carrying out the usual calculations we get

$$u_{ig} = \frac{1}{\sum\limits_{g'=1}^{k} \left(\frac{d_{M,w}^2\left(\widetilde{\mathbf{x}}_i, \widetilde{\mathbf{h}}_g\right)}{d_{M,w}^2\left(\widetilde{\mathbf{x}}_i, \widetilde{\mathbf{h}}_{g'}\right)} \right)^{\frac{1}{m-1}}}, \quad i = 1, \ldots, n, \quad g = 1, \ldots, k. \tag{13}$$

On the other hand, by considering the partial derivatives of (10) with respect to $\mathbf{h}_g^{C_1}$, $\mathbf{h}_g^{C_2}$, $\mathbf{h}_g^{L}$ and $\mathbf{h}_g^{R}$, for $g \in \{1, \ldots, k\}$, and setting them equal to 0, the centroid matrix components are given by

$$\mathbf{h}_g^{C_1} = \frac{\sum\limits_{i=1}^{n} u_{ig}^m \mathbf{c}_{1i}}{\sum\limits_{i=1}^{n} u_{ig}^m}, \qquad \mathbf{h}_g^{C_2} = \frac{\sum\limits_{i=1}^{n} u_{ig}^m \mathbf{c}_{2i}}{\sum\limits_{i=1}^{n} u_{ig}^m}, \tag{14}$$

$$\mathbf{h}_g^{L} = \frac{\sum\limits_{i=1}^{n} u_{ig}^m \mathbf{l}_{i}}{\sum\limits_{i=1}^{n} u_{ig}^m}, \qquad \mathbf{h}_g^{R} = \frac{\sum\limits_{i=1}^{n} u_{ig}^m \mathbf{r}_{i}}{\sum\limits_{i=1}^{n} u_{ig}^m}. \tag{15}$$

In order to update $\mathbf{M}_g^{C_1}$, we consider the partial derivative of (10) with respect to $\mathbf{M}_g^{C_1}$, taking into account that, for a non-singular matrix $\mathbf{A}$ and any compatible vector $\mathbf{x}$, $\frac{\vartheta}{\vartheta \mathbf{A}} \mathbf{x}' \mathbf{A} \mathbf{x} = \mathbf{x}\mathbf{x}'$ and $\frac{\vartheta}{\vartheta \mathbf{A}} |\mathbf{A}| = |\mathbf{A}| \mathbf{A}^{-1}$. Thus, setting the partial derivative equal to 0 we get

$$(\mathbf{M}_g^{C_1})^{-1} = \frac{S_g^{C_1}}{\left(det(S_g^{C_1}) \right)^{1/p}} \tag{16}$$

where $S_g^{C_1} = \dfrac{\sum_{i=1}^{n} u_{ig}^m (\mathbf{c}_{1i} - \mathbf{h}_g^{C_1})(\mathbf{c}_{1i} - \mathbf{h}_g^{C_1})^T}{\sum_{i=1}^{n} u_{ig}^m}$ is the fuzzy covariance matrix of the left center for the g-th cluster. Analogously, it can be stated that

$$(\mathbf{M}_g^{C_2})^{-1} = \frac{S_g^{C_2}}{\left(det(S_g^{C_2})\right)^{1/p}} \text{ with } S_g^{C_2} = \frac{\sum_{i=1}^{n} u_{ig}^m (\mathbf{c}_{2i} - \mathbf{h}_g^{C_2})(\mathbf{c}_{2i} - \mathbf{h}_g^{C_2})^T}{\sum_{i=1}^{n} u_{ig}^m}, \quad (17)$$

$$(\mathbf{M}_g^{L})^{-1} = \frac{S_g^{L}}{\left(det(S_g^{L})\right)^{1/p}} \text{ with } S_g^{L} = \frac{\sum_{i=1}^{n} u_{ig}^m (\mathbf{l}_{i} - \mathbf{h}_g^{L})(\mathbf{l}_{i} - \mathbf{h}_g^{L})^T}{\sum_{i=1}^{n} u_{ig}^m}, \text{ and} \quad (18)$$

$$(\mathbf{M}_g^{R})^{-1} = \frac{S_g^{R}}{\left(det(S_g^{R})\right)^{1/p}} \text{ with } S_g^{R} = \frac{\sum_{i=1}^{n} u_{ig}^m (\mathbf{r}_{i} - \mathbf{h}_g^{R})(\mathbf{r}_{i} - \mathbf{h}_g^{R})^T}{\sum_{i=1}^{n} u_{ig}^m}. \quad (19)$$

Finally, to update the value of the weight, first we note that just the distance involved in (7) depends on w_C and, thus, the remaining terms can be ignored. Setting the partial derivative of (10) with respect to w_C equal to 0 and solving it, we obtain

$$w_C = \frac{\sum_{i=1}^{n} \sum_{g=1}^{k} u_{ig}^m [d_M^2\left(\mathbf{l}_i, \mathbf{h}_g^L\right) + d_M^2\left(\mathbf{r}_i, \mathbf{h}_g^R\right)]}{\sum_{i=1}^{n} \sum_{g=1}^{k} u_{ig}^m [d_M^2\left(\mathbf{c}_{1i}, \mathbf{h}_g^{C_1}\right) + d_M^2\left(\mathbf{c}_{2i}, \mathbf{h}_g^{C_2}\right) + d_M^2\left(\mathbf{l}_i, \mathbf{h}_g^L\right) + d_M^2\left(\mathbf{r}_i, \mathbf{h}_g^R\right)]}, \quad (20)$$

$(w_S = 1 - w_C)$. If $w_C < 0.5$, then we set $w_C = 0.5$.