A new fusion algorithm for fuzzy clustering

Ivan Vidović[1,*]   **Dražen Bajer**[1]   **and**   **Rudolf Scitovski**[2]

[1] *Faculty of Electrical Engineering Osijek, Josip Juraj Strossmayer University of Osijek*
Cara Hadrijana 10b, 31 000 Osijek, Croatia
E-mail: ⟨{*ividovi2, dbajer*}*@etfos.hr*⟩

[2] *Department of Mathematics, Josip Juraj Strossmayer University of Osijek*
Trg Lj. Gaja 6, 31 000 Osijek, Croatia
E-mail: ⟨*scitowsk@mathos.hr*⟩

Abstract. In this paper, we have considered the merging problem of two ellipsoidal clusters in order to construct a new fusion algorithm for fuzzy clustering. We have proposed a criterion for merging two ellipsoidal clusters π_1, π_2 with associated main Mahalanobis circles $E_j(c_j, \sigma_j)$, where c_j is the centroid and σ_j^2 is the Mahalanobis variance of cluster π_j. Based on the well-known Davies-Bouldin index, we have constructed a new fusion algorithm. The criterion has been tested on several data sets, and the performance of the fusion algorithm has been demonstrated on an illustrative example.

Key words: fusion algorithm, fuzzy clustering, Mahalanobis clustering, cluster merging, Davies-Bouldin index

1. Introduction

Given is a data points set $\mathcal{A} = \{a^i : i = 1, \ldots, m\} \subset [\alpha, \beta] \subset \mathbb{R}^n$, where $\alpha = (\alpha_1, \ldots, \alpha_n)^T$, $\beta = (\beta_1, \ldots, \beta_n)^T \in \mathbb{R}^n$ and $[\alpha, \beta] = \{x \in \mathbb{R}^n : \alpha_s \leq x_s \leq \beta_s\}$. A partition of the set $\mathcal{A}$ into k disjoint subsets $\pi_1, \ldots, \pi_k$, $1 \leq k \leq m$, such that

$$\bigcup_{j=1}^{k} \pi_j = \mathcal{A}, \qquad \pi_r \cap \pi_s = \emptyset, \quad r \neq s, \qquad |\pi_j| \geq 1, \quad j = 1, \ldots, k, \qquad (1)$$

will be denoted by $\Pi(\mathcal{A}) = \{\pi_1, \ldots, \pi_k\}$ and the set of all such partitions will be denoted by $\mathcal{P}(\mathcal{A}; k)$. The elements $\pi_1, \ldots, \pi_k$ of the partition Π are called *clusters*. Clustering or grouping a data set into conceptually meaningful clusters is a well-studied problem in recent literature, and it has practical importance in a wide variety of applications (see e.g., [3, 9, 11]).

If $d \colon \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}_+$, $\mathbb{R}_+ = [0, +\infty\rangle$ is some distance-like function (see e.g., [3, 8]), then to each cluster $\pi_j \in \Pi$ we can associate its center c_j defined by

$$c_j := \underset{x \in [\alpha, \beta]}{\operatorname{argmin}} \sum_{a \in \pi_j} d(x, a). \qquad (2)$$

---

*Corresponding author.

After that, by introducing the objective function $\mathcal{F}\colon \mathcal{P}(\mathcal{A};k) \to \mathbb{R}_+$, the quality of a partition can be defined, and searching for the *globally optimal k-partition* comes down to solving the following optimization problem

$$\operatorname*{argmin}_{\Pi \in \mathcal{P}(\mathcal{A};k)} \mathcal{F}(\Pi), \qquad \mathcal{F}(\Pi) = \sum_{j=1}^{k} \sum_{a \in \pi_j} d(c_j, a), \qquad c = (c_1, \dots, c_k). \tag{3}$$

Conversely, for a given set of centers $c_1, \dots, c_k \in [\alpha, \beta]$, by applying the minimal distance principle, the partition $\Pi = \{\pi(c_1), \dots, \pi(c_k)\}$ of the set $\mathcal{A}$ consisting of clusters:

$$\pi(c_j) = \{a \in \mathcal{A} : d(c_j, a) \leq d(c_s, a), \ \forall s = 1, \dots, k\}, \qquad j = 1, \dots, k,$$

can be defined. Therefore, the problem of finding an optimal partition of the set $\mathcal{A}$ can be reduced to the following *global optimization problem* (see e.g., [12, 3, 13])

$$\operatorname*{argmin}_{c \in [\alpha, \beta]^k} F(c), \qquad F(c) = \sum_{i=1}^{m} \min_{1 \leq j \leq k} d(c_j, a^i). \tag{4}$$

The solutions of (3) and (4) coincide [11, 12].

Let $U \in \{0,1\}^{m \times k}$ be a matrix such that

$$u_{ij} = \begin{cases} 1, & \text{if } a^i \in \pi_j \\ 0, & \text{if } a^i \notin \pi_j \end{cases}, \qquad i = 1, \dots m, \quad j = 1, \dots, k, \tag{5}$$

$$\sum_{j=1}^{k} u_{ij} = 1, \quad i = 1, \dots, m. \tag{6}$$

Then (3) can be written as [3, 13]

$$\operatorname*{argmin}_{c \in [\alpha, \beta]^k, \, U \in \{0,1\}^{m \times k}} F(c, U), \qquad F(c, U) = \sum_{i=1}^{m} \sum_{j=1}^{k} u_{ij} d(c_j, a^i). \tag{7}$$

In order to ensure all conditions from (1), the following should be added to conditions (5) and (6):

$$\sum_{i=1}^{m} u_{ij} \geq 1, \quad j = 1, \dots, k. \tag{8}$$

Assuming that elements $a^i \in \mathcal{A}$ can partially belong to different clusters, then, due to (6), it must be $u_{ij} \in [0, 1]$ (see e.g., [10]). According to [3, 13], the membership grade of a^i in cluster π_j is determined by u_{ij}^q, where parameter $q > 1$ is called the *fuzzifier*, and the objective function becomes

$$F(c, U) = \sum_{i=1}^{m} \sum_{j=1}^{k} u_{ij}^q(c) \, d(c_j, a^i). \tag{9}$$

The definition of membership functions $u_{ij} \colon \mathcal{A} \to [0,1]$ can be found in [6]. In applied research, the most common value is $q \in [1.5, 2.5]$.

A partition with the most appropriate number of clusters can be estimated by the following well-known validity indexes for fuzzy clustering (see e.g., [3, 7, 13, 14]): Xie-Beni index, Fuzzy Davies – Bouldin index, Fuzzy Hypervolume index, etc. In our paper, we specially use the fuzzy Davies – Bouldin index (DB) which is for the optimal partition $\Pi^\star(\pi_1^\star, \ldots, \pi_k^\star)$ defined by

$$\mathrm{DB}(k) = \frac{1}{k} \sum_{j=1}^{k} \max_{\substack{s \in \{1, \ldots, k\} \\ s \neq j}} \frac{V(\pi_j^\star) + V(\pi_s^\star)}{d(c_j^\star, c_s^\star)}, \qquad (10)$$

where $c_1^\star, \ldots, c_k^\star$ are centers of clusters and

$$V(\pi_j^\star) = \frac{1}{\sum\limits_{i=1}^{m} u_{ij}^q(c)} \sum_{i=1}^{m} u_{ij}^q(c)\, d(c_j, a^i).$$

The term $D_j := \max\limits_{s \in \{1,\ldots,k\},\, s \neq j} \frac{V(\pi_j^\star)+V(\pi_s^\star)}{d(c_j^\star,c_s^\star)}$ represents the worst case within–to–between cluster spread involving the cluster $\pi_j^\star$. Minimizing D_j for all clusters clearly minimizes the DB index. Therefore, more compact and better separated clusters in an optimal partition will result in a lower value of DB index (see [13, 14]).

It is well-known that by using the Least Squares (LS) distance-like function [3, 8] spherical clusters are obtained, and by using adaptive Mahalanobis clustering [5, 12, 13] ellipsoidal clusters are obtained. However, if different banana-form clusters should be recognized, then some fusion method can be applied [3, 4, 7, 15]. The idea is in the following: First, determine an optimal partition with relatively many clusters using a certain method and then keep merging close clusters for as long as the DB index value is decreasing. Such cluster-forms appear in different geographical, geological and medical research.

The paper is organized as follows. In the next section the merging problem of two spherical data sets is discussed. In Section 3, the problem is generalized for two ellipsoidal data sets and a corresponding merging algorithm is proposed. A new fusion algorithm is proposed in Section 4, and in Section 5 a numerical example is shown.

2. Motivation: Merging of two spherical data sets

Because of simplicity, we consider the merging problem of two spherical data sets in the plane. The problem in $\mathbb{R}^n$ is analogous.

Let d be the LS distance-like function, $c \in \mathbb{R}^2$ the point in the plane in the neighborhood of which m random points were generated by using Gaussian distributions with variance σ^2, then the circle $K(c, \sigma)$ contains about 68% data points [2]. The set of all data points constitutes the set $\mathcal{A}$. The circle $K(c, \sigma)$ is called the *main circle* of the set $\mathcal{A}$, and data points outside the circle $K(c, \sigma)$ are called *peripheral elements* of the set $\mathcal{A}$.

Let us suppose that two spherical data sets π_1, π_2 with their main circles $K_j(c_j, \sigma_j)$, $j = 1, 2$ are given, where

$$c_j = \frac{1}{|\pi_j|} \sum_{a \in \pi_j} a, \qquad \sigma_j^2 = \frac{1}{|\pi_j|} \sum_{a \in \pi_j} (c_j - a)^2. \qquad (11)$$

(a) $\|c_1 - c_2\| > \sigma_1 + \sigma_2$ (b) $\|c_1 - c_2\| = \sigma_1 + \sigma_2$

(c) $\|c_1 - c_2\| < \sigma_1 + \sigma_2 < 2\|c_1 - c_2\|$ (d) $\sigma_1 + \sigma_2 > 2\|c_1 - c_2\|$
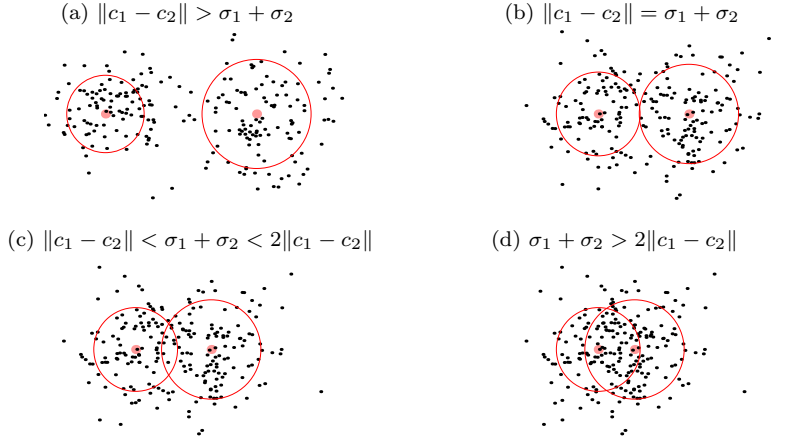
Figure 1: *Mutually different relationships between two spherical data sets*

Mutually different relationships between two spherical data sets with regard to the relation between their main circles are shown in Figure 1. As can be seen in Figure 1, merging of the sets can be thought of if the main circles K_1, K_2 intersect, i.e., if

$$\|c_1 - c_2\| \leq \sigma_1 + \sigma_2, \qquad (12)$$

what is also in accordance with the result from [1]: "a mixture of two normal distributions $\mathcal{N}(c_1, \sigma_1^2)$, $\mathcal{N}(c_2, \sigma_2^2)$ is unimodal if $\|c_1 - c_2\| \leq 2\min\{\sigma_1, \sigma_2\}$", since there holds: $2\min\{\sigma_1, \sigma_2\} \leq \sigma_1 + \sigma_2$. Therefore, the following coefficient can be used as an indicator of sets connectivity (see also [15])

$$\tau = \frac{\sigma_1 + \sigma_2}{\|c_1 - c_2\|}. \qquad (13)$$

Lemma 1. *Let $K_j(c_j, \sigma_j)$, $j = 1, 2$ be the main circles of sets π_1, π_2 and $K(c, \sigma)$ the minimal circle that includes both circles K_1, K_2. Then $\sigma = \frac{1}{2}(\|c_1 - c_2\| + \sigma_1 + \sigma_2)$, and $K_1 \cap K_2 \neq \emptyset$ if and only if*

$$(i) \quad \|c_1 - c_2\| \leq \sigma_1 + \sigma_2, \qquad or \qquad (14)$$

$$(ii) \quad \tau = \frac{\sigma_1 + \sigma_2}{\|c_1 - c_2\|} \geq 1, \qquad or \qquad (15)$$

$$(iii) \quad \partial K \leq \partial K_1 + \partial K_2, \qquad or \qquad (16)$$

$$(iv) \quad \sigma \leq \sigma_1 + \sigma_2. \qquad (17)$$

The proof of the lemma is trivial: (i) is obvious, (ii) immediately follows from (i), and (iii) and (iv) follow from (i) and

$$\partial K = 2\sigma\pi = (\|c_1 - c_2\| + \sigma_1 + \sigma_2)\pi \leq 2\sigma_1\pi + 2\sigma_2\pi = \partial K_1 + \partial K_2.$$

Remark 1. *Note that condition (ii) is in direct relationship with the DB index. Note also that merging of several sets inevitably leads to the merging problem of ellipsoidal sets. Namely, merging of two spherical sets forms an ellipsoidal set.*

3. Merging of two ellipsoidal data sets

Because of simplicity, we consider the merging problem of two ellipsoidal data sets in the plane. The problem in $\mathbb{R}^n$ is analogous.

Let $c \in \mathbb{R}^2$ be the point in the plane in the neighborhood of which the set $\mathcal{A} = \{a^i = (x_i, y_i) \in \mathbb{R}^2 : i = 1, \ldots, m\}$ of m random points was generated by using Gaussian distributions with covariance matrix S, and let $d_M : \mathbb{R}^2 \times \mathbb{R}^2 \to \mathbb{R}_+$ be the Mahalanobis distance-like function [3, 12, 13]

$$d_M(u, v; S) = \sqrt{\det S}(u - v)^T S^{-1}(u - v). \tag{18}$$

Note that then the ellipse

$$E(c, \sigma) = \left\{(x, y) \in \mathbb{R}^2 : d_M(c, (x, y); S) \leq \sigma^2\right\}, \quad \sigma^2 = \tfrac{1}{m} \sum_{i=1}^m d_M(c, a^i; S), \tag{19}$$

contains about 68% data points [2]. The ellipse $E(c, \sigma)$ will be called the *main Mahalanobis circle* (M-circle), and data points outside the ellipse $E(c, \sigma)$ will be called *peripheral elements* of the set $\mathcal{A}$.

Let us suppose that two ellipsoidal data sets π_1, π_2 with their main M-circles $E_j(c_j, \sigma_j)$, $j = 1, 2$ are given, where

$$c_j = \tfrac{1}{|\pi_j|} \sum_{a \in \pi_j} a, \qquad \sigma_j^2 = \tfrac{1}{|\pi_j|} \sum_{a \in \pi_j} d_M(c_j, a; S_j), \tag{20}$$

where S_j is the corresponding covariance matrix. Mutually different relationships between two ellipsoidal data sets with regard to the relation of their main M-circles are shown in Figure 2.

Of the criteria stated in Lemma 1, it seems that only criterion (iii) given by (16) could be adjusted, but unfortunately, there is no explicit formula for determining the minimal ellipse $E(c, \sigma)$ that includes both ellipses E_1, E_2. Assuming that the center c of the ellipse E coincides with the center of the set $\pi_1 \cup \pi_2$, then the M-radius σ of the ellipse E can be iteratively determined by Algorithm 1. According to this condition, sets π_1, π_2 with the main M-circles $E_j(c_j, \sigma_j)$, $j = 1, 2$, will be merged if the following criterion holds

Criterion 1: $\sigma < \sigma_1 + \sigma_2.$ $\hspace{5cm}$ (21)

In Figure 2, several typical situations are shown. The numerical complexity of determining the parameter σ can be mentioned as the lack of Criterion 1, but also the fact that clusters in Figure 2b would not be merged according to this criterion, although that would be expected.
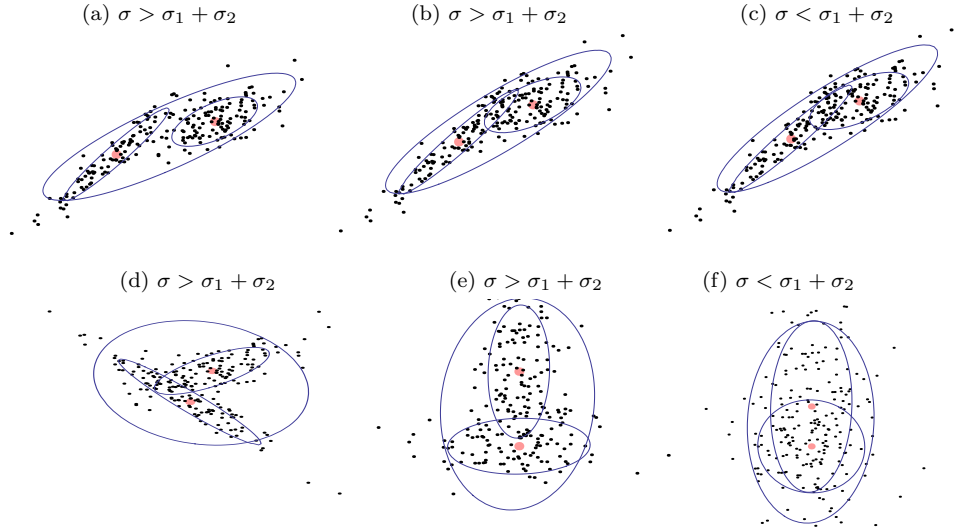
(a) $\sigma > \sigma_1 + \sigma_2$ (b) $\sigma > \sigma_1 + \sigma_2$ (c) $\sigma < \sigma_1 + \sigma_2$

(d) $\sigma > \sigma_1 + \sigma_2$ (e) $\sigma > \sigma_1 + \sigma_2$ (f) $\sigma < \sigma_1 + \sigma_2$

Figure 2: *Mutually different relationships between two ellipsoidal data sets*

Algorithm 1. *(Merging of two ellipsoidal sets)*

Step 1: Input π_1, π_2, and set: $\mathcal{A} = \pi_1 \cup \pi_2$, $m_1 = |\pi_1|$, $m_2 = |\pi_2|$, $m = m_1 + m_2$, $\alpha = 1$, $\epsilon > 0$;

Step 2: For each $j = 1, 2$, determine:

$$c_j = \frac{1}{m_j} \sum_{a \in \pi_j} a, \qquad\qquad S_j = \frac{1}{m_j} \sum_{a \in \pi_j} (c_j - a)(c_j - a)^T,$$

$$\sigma_j^2 = \frac{1}{m_j} \sum_{a \in \pi_j} d_M(c_j, a; S_j), \qquad E_j = \{a \in \pi_j : d_M(c_j, a; S_j) \leq \sigma_j^2\},$$

$$c = \frac{1}{m} \sum_{a \in \mathcal{A}} a, \quad S = \frac{1}{m} \sum_{a \in \mathcal{A}} (c - a)(c - a)^T, \quad \sigma_0^2 = \frac{1}{m} \sum_{a \in \mathcal{A}} d_M(c, a; S);$$

Step 3: Set: $\sigma = \alpha \sigma_0$;
  While $[\{a \in E_1 \cup E_2 : d_M(c, a; S) \geq \sigma^2\} \neq \emptyset, \alpha := \alpha + \epsilon; \sigma = \alpha \sigma_0]$.

Therefore, we try to define new conditions which will simplify the calculation, but also to refine the conditions for merging. Instead of Criterion 1, we will define a criterion which is analogous to condition (i) given by (14), i.e., which is analogous to condition (ii) given by (15)

Criterion 2:

$$\Delta(c_1, c_2) := \min\{\sqrt{d_M(c_1, c_2; S_1)}, \sqrt{d_M(c_1, c_2; S_2)}\} < \sigma_1 + \sigma_2. \tag{22}$$

According to this criterion, the sets in Figure 2b would be merged, but sets in Figure 2d and Figure 2e would also be merged, which should definitely be avoided. Namely, sets should not be merged if the angle between the main directions of their M-circles is greater than a limit value (e.g., 65°). Therefore, let us introduce the following additional criterion:

Criterion 3: $sp_{12} := |\langle v_{max}^1, v_{max}^2 \rangle| > \cos 65^o \approx 0.42,$ (23)

where v_{max}^1, v_{max}^2 are unit eigenvectors corresponding to the largest eigenvalues of covariance matrices S_1, S_2. However, a new problem arises here: This criterion would not allow for the merging of sets in Figure 2f! Namely, the merging of sets that are not essentially ellipsoidal should be allowed for, although the angle between the main directions of their M-circles is greater than the given limit value. We will assume that two sets π_1 and π_2 are not essentially ellipsoidal if the relative difference of their eigenvalues is less than 0.5. Criterion 3 will be refined as follows:

Criterion 3′:

$$\delta_1 = \frac{\lambda_1^{(1)} - \lambda_2^{(1)}}{\lambda_2^{(1)}} < 0.5 \ \vee \ \delta_2 = \frac{\lambda_1^{(2)} - \lambda_2^{(2)}}{\lambda_2^{(2)}} < 0.5 \ \vee \ sp_{12} = |\langle v_{max}^1, v_{max}^2 \rangle| > \cos 65^o,$$

$$(24)$$

where $\lambda_1^{(1)} > \lambda_2^{(1)} > 0$ and $\lambda_1^{(2)} > \lambda_2^{(2)} > 0$ are eigenvalues of covariance matrices S_1, S_2. Note also that, since there holds

$$\neg(A \vee B \vee C) = (\neg A) \wedge (\neg B) \wedge (\neg C),$$

Criterion	Fig. 2a	Fig. 2b	Fig. 2c	Fig. 2d	Fig. 2e	Fig. 2f
Criterion 1	-	-	+	-	-	+
Criterion 2	-	+	+	+	+	+
Criterion 3	+	+	+	-	-	-
Criterion 3′	+	+	+	-	-	+
Criterion 4	-	+	+	-	-	+

Table 1: *Analysis of the impact of the criteria on situations given in Figure 2*

according to Criterion 3′, the sets π_1, π_2 will be merged if at least one of them is almost spherical or if the angle between their main directions is less than 65°, and that they will not be merged if both are essentially ellipsoidal or if the angle between their main directions is greater than 65°. In this way, Criterion 3′ will not allow for the merging of sets shown in Figures 2d and 2e, but it will allow for the merging of sets shown in Figure 2f. The analysis of the impact of all criteria is given in Table 1. Thereby, marks "+" and "-" mean that the sets should be merged and not merged, respectively. As can be seen in Figure 2, the criterion composed of Criterion 2 and Criterion 3′ can be acceptable. This criterion will be denoted as Criterion 4.

4. A new fusion algorithm

Given is a data points set $\mathcal{A} = \{a^i : i = 1, \ldots, m\} \subset [\alpha, \beta] \subset \mathbb{R}^n$. Suppose that by some method (see [5, 3]) an optimal fuzzy partition $\Pi = \{\pi_1, \ldots, \pi_k\}$ with membership matrix $U \in \mathbb{R}^{m \times k}$ and fuzzifier $q \in [1.5, 2.5]$ is determined. Based on the considerations from Section 3, we propose a new fusion algorithm.

As an input, the algorithm takes the set $\mathcal{A}$ and an optimal fuzzy k-partition with membership matrix $U \in \mathbb{R}^{m \times k}$ and fuzzifier $q \in [1.5, 2.5]$ (Step 1). First, for each

cluster the algorithm determines their center, covariance matrix with their largest and smallest eigenvalue and the corresponding eigenvectors, and the radius of the main M-circle (Step 2). In Step 3, the DB-index of the initial partition is calculated.

The next step is to see if there are possible candidates for merging in the given k-partition (Step 4). If there are such candidates, among all pairs of clusters for which condition (22) is satisfied up to an $\epsilon > 0$, Algorithm 2 looks for that pair π_r, π_s which coincides most (Step 4) and defines a new partition with $(k-1)$ clusters $(\{\pi_1, \ldots, \pi_k\} \setminus \{\pi_r, \pi_s\}) \cup \hat{\pi}$, where $\hat{\pi}$ is a cluster resulting form merging of two clusters π_r, π_s (Step 5). The procedure is repeated until there are candidates for merging and as long as the value of the DB index decreases.

By using Algorithm 2, a new partition with the most appropriate number of clusters is obtained.

Algorithm 2. *(Fusion algorithm)*

Step 1: Input: $\mathcal{A} \subset \mathbb{R}^n$, $U \in \mathbb{R}^{m \times k}$ and set $\epsilon > 0$;

Step 2: For each $j = 1, \ldots, k$ determine $c_j = \frac{1}{\sum_{i=1}^m u_{ij}^q} \sum_{i=1}^m u_{ij}^q a^i$,

$$S_j = \frac{1}{\sum_{i=1}^m u_{ij}^q} \sum_{i=1}^m u_{ij}^q (c_j - a^i)(c_j - a^i)^T, \quad \sigma_j^2 = \frac{1}{\sum_{i=1}^m u_{ij}^q} \sum_{i=1}^m u_{ij}^q d_M(c_j, a^i; S_j),$$

and the largest and smallest eigenvalue $(\lambda_{max}^j, \lambda_{min}^j)$ and the corresponding eigenvectors (v_{max}^j, v_{min}^j) of covariance matrix S_j;

Step 3: Determine DB-index $DB(k) = \frac{1}{k} \sum_{r=1}^k \max_{\substack{s \in \{1,\ldots,k\} \\ s \neq r}} \frac{\sigma_r^2 + \sigma_s^2}{\max\{d_M(c_r,c_s;S_r), d_M(c_r,c_s;S_s)\}}$;

Step 4: Determine the set

$$\mathcal{J}_0 = \{(r,s) \in J \times J : r \neq s, \kappa_{rs} + \epsilon > 0 \wedge (\delta_r < 0.5 \vee \delta_s < 0.5 \vee sp_{rs} > \cos 65^o)\},$$

where $\delta_r, \delta_s, sp_{rs}$ are given by (24), and

$$\kappa_{rs} = \sigma_r + \sigma_s - \min\{\sqrt{d_M(c_r, c_s; S_r)}, \sqrt{d_M(c_r, c_s; S_s)}\};$$

If $\mathcal{J}_0 = \emptyset$, Print[There are no candidates for merging]; STOP;
Else, $\{r_0, s_0\} = \underset{(r,s) \in \mathcal{J}_0}{\operatorname{argmax}} \kappa_{rs}$;

Step 5: Define a new partition $(\{\pi_1, \ldots, \pi_k\} \setminus \{\pi_{r_0}, \pi_{s_0}\}) \cup \hat{\pi}$, where $\hat{\pi}$ is defined according to Step 2 of Algorithm 1;
Determine a new DB-index $DB(k-1)$;

Step 6: If $DB(k-1) < DB(k)$, put $k := k-1$ and GoTo Step 4;
Else, STOP.

5. Numerical examples

We will demonstrate our fusion algorithm by means of the following illustrative example.

Example 1. $k = 11$ *points C_j:* $(5,5), (3,3), (2.5, 4.2), (2.3, 6), (2.5, 7.3), (3, 9), (6, 1.5),$ $(7, 2),$ $(7.5, 3.5), (7.8, 5.2), (7.5, 6.5)$ *were chosen in the square* $[0, 10]^2$ *and k uniform distributed random numbers $m_1, \ldots, m_k$ were chosen in the interval $[70, 80]$. In the neighborhood of each point C_j, m_j random points were generated by using Gaussian normal distributions. In this way, we obtain a data set $\mathcal{A}$ (see Figure 3a). By using Algorithm 2, a partition with the most appropriate number of clusters should be determined.*
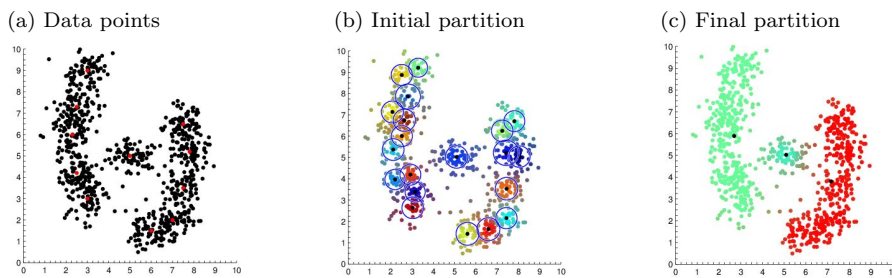
(a) Data points   (b) Initial partition   (c) Final partition

Figure 3: *Searching for a partition with the most appropriate number of clusters*

First, by using the fuzzy c-means algorithm [3], we determine a locally optimal fuzzy partition with 20 clusters (see Figure 3b) and after that, by using Algorithm 2, we obtain the final optimal partition with three clusters (see Figure 3c).

As an illustration, Figure 4 shows the 10th iteration. Figure 4a shows the partition with 11 clusters in which two clusters for merging have been chosen, and Figure 4b shows the partition with 10 clusters where the previously mentioned clusters have been merged.
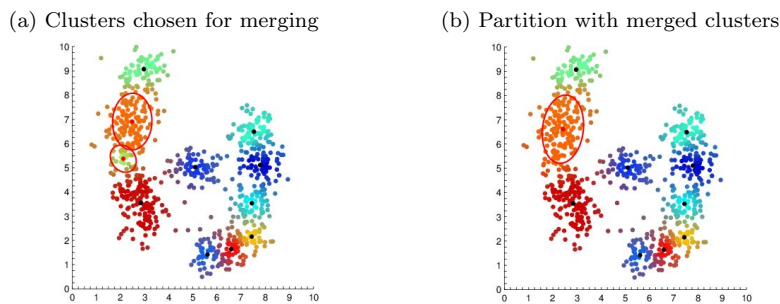
(a) Clusters chosen for merging   (b) Partition with merged clusters

Figure 4: *Selection of clusters for merging*

6. Conclusion

The paper presents a new fusion algorithm for fuzzy clustering. Criteria for merging two ellipsoidal clusters with associated Mahalanobis circles were defined. Starting with a larger number of clusters in every iteration the algorithm merges two candidate clusters as long as there are candidates for merging and the value of the adapted Davies-Bouldin index decreases. Numerical examples suggest that the proposed algorithm successfully finds various banana-form clusters. Future work could include the adaption of the proposed algorithm to higher dimensional problems as well as a comparison with fusion algorithms from literature on a larger number of data sets.

Acknowledgement

References

[1] Behboodian, J. (1979). On the modes of a mixture of two normal distribu-tions, Technometrics, 12, 131–139.

[2] Benšić, M. and Šuvak, N. (2012). Primijenjena statistika (Applied statistics), Odjel za matematiku (Department of Mathematics).

[3] Bezdek, J. C. Bezdek, Keller, J., Krisnapuram, R. and Pal, N. R. (2005). Fuzzy models and algorithms for pattern recognition and image processing, Springer.

[4] Frossyniotis, D., Pertselakis, M. and Stafylopatis, A. (2002). A multi-clustering fusion algorithm, In: Vlahavas, I. P. and Spyropoulos, C. D., editors, Methods and Applications of Artificial Intelligence, Second Hellenic Conference on AI, SETN 2002. Thessaloniki, Greece, 225–236.

[5] Gustafson, D. E. and Kessel, W. C. (1979). Fuzzy clustering with a fuzzy covariance matrix, In: Proc. IEEE Conf. Decision Control, San Diego, CA, 761–766.

[6] Höppner, F. and Klawonn, F. (2003). A contribution to convergence theory of fuzzy c-means and derivatives, IEEE Transactions on Fuzzy Systems, 11, 682–694.

[7] Kaymak, U. and Setnes, M. (2002). Fuzzy clustering with volume prototype and adaptive cluster merging, IEEE Transactions on Fuzzy Systems, 10, 705–712.

[8] Kogan, J. (2007). Introduction to Clustering Large and High-dimensional Data, Cambridge University Press.

[9] Sabo, K., Scitovski, R. and Taler, P. (2012). Uniform distribution of the number of voters per constituency on the basis of a mathematical model (in Croatian), Hrvatska i komparativna javna uprava (Croatian and Comparative Public Administration), 14, 229–249.

[10] Scitovski, R. and Sabo, K. (2014). Analysis of the k-means algorithm in the case of data points occurring on the border of two or more clusters, Knowledge-Based Systems, 57, 1–7.

[11] Scitovski, R. and Scitovski, S. (2013). A fast partitioning algorithm and its application to earthquake investigation, Computers & Geosciences, 59, 124–131.

[12] Späth, H. (1983). Cluster-Formation und Analyse, R. Oldenburg Verlag, München.

[13] Theodoridis, S. and Koutroumbas, K. (2009). Pattern Recognition, Academic Press, Burlington, 4th edition.

[14] Vendramin, L., Campello, R. J. G. B. and Hruschka, E.R. (2009). On the comparison of relative clustering validity criteria, In: Proceedings of the SIAM International Conference on Data Mining, SDM 2009, April 30 – May 2, 2009, Sparks, Nevada, USA, SIAM, 733–744.

[15] Xiong, X., Chan, K. and Tan, K. L. (2004). Similarity-driven cluster merging method for unsupervised fuzzy clustering, In: Proceedings of the Twentieth Conference on Uncertainty in Artificial Intelligence (UAI-04), AUAI Press, Arlington, Virginia, 611–618.