Clustering with Interval-valued Fuzzy Sets

Ramón González del Campo[1] **Jose Luis González**[2]

[1]Complutense University of Madrid (e-mail: rgonzale@ucm.es)
[2]Complutense University of Madrid (e-mail: joselg04@ucm.es)

Abstract

Interval-Valued Fuzzy Sets handle uncertainty and vagueness effectively. These features are particularly useful for clustering. In this paper it is showed the utility of Interval-Valued Fuzzy Sets for clustering with no accurate information. An easy method for clustering is proposed by generating transitive closures under a pseudo-t-representable t-norm. Clusters are computed from transitive closures by generating alpha-cuts. It is found that some of these alpha-cuts are equivalence classes under the pseudo-t-representable min. It is also found that these transitive closures are closer to the original interval-valued fuzzy relation that the classical transitive closure under the t-norm $[\min, \min]$

Keywords: Clustering, Interval-valued fuzzy sets, Transitivity

1. Introduction

Cluster analysis is one of the major techniques in pattern recognition. The importance of clustering in various areas such as taxonomy, medicine, geology, business, engineering systems and image processing, etc., is well documented is well documented in [7, 8, 16]. The conventional (hard) clustering methods restrict that each point of the data set belongs to exactly one cluster. Fuzzy set theory proposed by Zadeh [20] in 1965 gave an idea of uncertainty of belonging which was described by a membership function. The use of fuzzy sets provides imprecise class membership information. Applications of fuzzy set theory in cluster analysis were early proposed in the work of Bellman, Kalaba and Zadeh [2] and Ruspini [18]. These papers open the door of research in fuzzy clustering. Now fuzzy clustering has been widely studied and applied in a variety of substantive areas. These methods become the important tools to cluster analysis.

Interval-valued fuzzy sets ($\mathcal{IVFS}s$) were introduced in the 60s by Grattan-Guinness [13], Jahn [15], Sambuc [19] and Zadeh [21]. They are extensions of classical fuzzy sets ($\mathcal{FS}s$) where the membership value between 0 and 1 is replaced by an interval in $[0, 1] \times [0, 1]$. They easily allow to model uncertainty and vagueness. Sometimes it is easier for experts to give a membership interval than a membership degree to describe a characteristic of objects on a universe. $\mathcal{IVFS}s$ are a special case of type-2 fuzzy sets that simplifies the calculations

while preserving their richness as well. Interval-valued fuzzy relations ($\mathcal{IVFR}s$) are fuzzy relations which experts express the relation degree between two objects by using interval numbers instead of numeric values.

Transitivity is a fundamental notion in decision theory. It is universally assumed in disciplines of decision theory and accepted in a principle of rationality for some kind of relations. A first task for decision science is the resolution of intransitivities when the transitive property is violated [17]. The transitive closure is a usual way to make transitive relation from an intransitive relation. The T-transitive closure of fuzzy relations has been studied for $\mathcal{FR}s$ by De Baets and De Meyer [1]. They showed that it always exists and it is unique. Gonzalez-del-Campo and Garmendia proposed an algorithm to compute the transitive closure for an $\mathcal{IVFR}s$ under a t-norm $\mathcal{T}$ [11].

In this paper, it is showed the utility of the Interval-Valued Fuzzy Relations to clustering.

This paper is organized as follows. In Section 2 some preliminaries are showed. In Section 3 it is exhibited how to use Interval-valued Fuzzy Relations for clustering. In Section 4 it is proposed an application example of this method in which data missing are processed. Finally, in Section 5 some conclusions are presented.

2. Preliminaries

Definition 2.1. *[4] Let $(L, \leq_L)$ be the lattice of intervals in [0,1] that satisfies:*

1. *$L = \{[x_1, x_2] \in [0, 1]^2 \text{ with } x_1 \leq x_2\}$.*
2. *$[x_1, x_2] \leq_L [y_1, y_2]$ if and only if $x_1 \leq y_1$ and $x_2 \leq y_2$*

Also by definition:

$[x_1, x_2] <_L [y_1, y_2] \Leftrightarrow x_1 < y_1, x_2 \leq y_2$ or
$x_1 \leq y_1, x_2 < y_2$
$[x_1, x_2] =_L [y_1, y_2] \Leftrightarrow x_1 = y_1, x_2 = y_2$.

Then, $0_L =_L [0, 0]$ and $1_L =_L [1, 1]$ are the smallest and the greatest elements in L respectively.

For convenience, $x = [x_1, x_2]$ and $y = [y_1, y_2]$ are sometimes written as $x = [\underline{x}, \overline{x}]$ and $y = [\underline{y}, \overline{y}]$ respectively.

Definition 2.2. *[5] Let $\{[v_i, w_i]\}$ be a set of intervals on L. Then, the supremum (max) of $\{[v_i, w_i]\}$ is defined as follows:*

$$\max\{[v_i, w_i]\} \equiv [\sup\{v_i\}, \sup\{w_i\}]$$

Definition 2.3. *[4] An interval-valued fuzzy set A on a universe $X = \{e_1, \dots, e_n\}$ is a mapping $A : X \to L$:*

$$A = \{(e_i, [x_1, x_2]) \mid e_i \in X, [x_1, x_2] \in L\}$$

Definition 2.4. *[4] Let $X = \{e_1, \dots, e_n\}$ be a universe and A and B two interval-valued fuzzy sets. The equality between A and B is defined as: $A =_L B$ if and only if $A(e_i) =_L B(e_i) \ \forall e_i \in X$.*

Definition 2.5. *[4] Let $X = \{e_1, \dots, e_n\}$ be a universe and A and B two interval-valued fuzzy sets. The inclusion of A in to B is defined as: $A \subseteq_L B$ if and only if $A(e_i) \leq_L B(e_i) \ \forall e_i \in X$.*

Let $x =_L [x_1, x_2]$ and $y =_L [y_1, y_2]$ be two intervals on L.

Definition 2.6. *[4] A t-norm $\mathcal{T}$ on L is a monotone increasing, symmetric and associative operator, $\mathcal{T} : L^2 \to L$, that satisfies: $\mathcal{T}(1_L, [x_1, x_2]) =_L [x_1, x_2]$ for all $[x_1, x_2]$ in L.*

Definition 2.7. *[4] A t-norm $\mathcal{T}$ on L is t-representable in L if there are two t-norms: T_1 and T_2 (T_1, T_2, in ([0,1],$\leq$)) that satisfy:*

$$\mathcal{T}([x_1, x_2], [y_1, y_2]) =_L [T_1(x_1, y_1), T_2(x_2, y_2)]$$

where $T_1(v, w) \leq T_2(v, w) \ \forall v, w \in [0, 1]$.

Example 2.1. *The following product t-norm $\mathcal{T}$ on L is t-representable:*

$$\mathcal{T}([x_1, x_2], [y_1, y_2]) =_L [x_1 * y_1, x_2 * y_2]$$

Example 2.2. *Two generalizations of the Lukasiewicz t-norm [5] are the following:*

- $T_w([x_1, x_2], [y_1, y_2]) =_L$
  $[\max(0, x_1 + y_1 - 1), \max(0, x_2 + y_2 - 1)]$
- $T_W([x_1, x_2], [y_1, y_2]) =_L$
  $[\max(0, x_1 + y_1 - 1), \max(0, x_1 + y_2 - 1, x_2 + y_1 - 1)]$

Note that T_w is t-representable but T_W is not t-representable.

Example 2.3. *The t-representable t-norm minimum $T_{\min} : L^2 \to L$ is defined as follows:*

$$T_{\min}([x_1, x_2], [y_1, y_2]) =_L [\min(x_1, y_1), \min(x_2, y_2)]$$

Definition 2.8. *[6] A t-norm $\mathcal{T}$ on L is pseudo-t-representable if there exists a t-norm T in ([0,1],$\leq$) that satisfies:*

$$\mathcal{T}([x_1, x_2], [y_1, y_2]) =_L$$
$$[T(x_1, y_1), \max\{T(x_1, y_2), T(x_2, y_1)\}]$$

The t-norm T is called the representant of $\mathcal{T}$.

Example 2.4. *The pseudo-t-representable t-norm minimum $T_{\sim\min} : L^2 \to L$ is defined as follows:*

$$T_{\sim\min}([x_1, x_2], [y_1, y_2]) =_L$$
$$[\min(x_1, y_1), \max\{\min(x_1, y_2), \min(x_2, y_1)\}]$$

Definition 2.9. *[3] Let X_1 and X_2 be two universes of discourse. An interval-valued fuzzy relation $R : X_1 \times X_2 \to L$ is a mapping:*

$$R = \{((a, b), [x, y]) \mid a \in X_1, b \in X_2, [x, y] \in L\}$$

where $x = \underline{R}(a, b)$ and $y = \overline{R}(a, b)$.

Let $X = \{e_1, \dots, e_n\}$ be a universe with n elements.

Definition 2.10. *[10] Let $\mathcal{T}$ be a t-norm on L and let R interval-valued fuzzy relation on X. R is $\mathcal{T}$-transitive if:*

$$\mathcal{T}(R(e_i, e_k), R(e_k, e_j)) \leq_L R(e_i, e_j) \forall e_i, e_j, e_k \in X$$

Definition 2.11. *[9] An interval-valued fuzzy relation $R : X^2 \to L$ is a generalized $\mathcal{T}$-indistinguishability if it is reflexive, symmetric and $\mathcal{T}$-transitive.*

Definition 2.12. *[9] Let $\mathcal{P}$ be a property of $\mathcal{IVFR}$s. Let $R : X^2 \to L$ be an interval-valued fuzzy relation on a finite universe X. The $\mathcal{P}$ closure of R is a $\mathcal{IVFR}$ $R^{\mathcal{P}} : X^2 \to L$ that satisfies:*

1. $R^{\mathcal{P}}$ satisfies P.
2. $R \subseteq_L R^{\mathcal{P}}$.
3. If $R \subseteq_L R'$ and R' satisfies P then $R^{\mathcal{P}} \subseteq_L R'$.

Lemma 2.1. *[9] Let R be an interval-valued fuzzy relation in a universe X and let $\mathcal{T}$ be an arbitrary t-norm on L. Then the $\mathcal{T}$-transitive closure of R always exists and it is unique.*

Let R be an interval-valued fuzzy relation on $X = \{e_1, \dots, e_n\}$. For convenience, $R(e_i, e_j)$ can be written $[\underline{R}(e_i, e_j), \overline{R}(e_i, e_j)]$ or $[\underline{R}, \overline{R}]$.

Proposition 2.1. *[9] If $\mathcal{T}$ is t-representable with T_1 and T_2 ($\mathcal{T} = [T_1, T_2]$) then an interval-valued relation $R : X^2 \to L$ is $\mathcal{T}$-transitive if and only if $\underline{R}$ is T_1-transitive and $\overline{R}$ is T_2-transitive.*

Proposition 2.2. *[9] Let R be an interval-valued relation on a finite universe X with cardinality n and let $\mathcal{T}$ be a t-norm on L. For simplicity $R(e_i, e_j)$ is written as $R(i, j)$. The $\mathcal{T}$-transitive closure of R, $R^{\mathcal{T}}$, can be computed using the following algorithm:*

```
for(k=1;k<n;k++)
   for(i=1;i<n;i++)
      for(j=1;j<n;j++)
         R(i,j)=Max(R(i,j),T(R(i,k),R(k,j)));
```

Definition 2.13. *[12] Let R be an interval-valued relation $R : X^2 \to L$. The α_1, α_2 cut of R, R_{α_1, α_2}, is a crisp relation defined for all α_1, α_2 in [0,1] as follows:*

$$R_{\alpha_1, \alpha_2}(e_i, e_j) = \begin{cases} 1 & R(e_i, e_j) \geq_L [\alpha_1, \alpha_2]; \\ 0, & otherwise. \end{cases}$$

Definition 2.14. *[14] Let A_X the set of interval-valued fuzzy sets on $X = \{e_1, \dots, e_n\}$. The Hamming distance d between M and N ($M, N \in A_X$) is defined by:*

$$d(M, N) = \sum \mid \overline{M}(e_i) - \overline{N}(e_i) \mid + \mid \underline{M}(e_i) - \underline{N}(e_i) \mid$$

for all e_i in X.

3. Clustering with Interval-valued Fuzzy Relations

Let us consider a set of n objects $X = \{e_1, \dots, e_n\}$. These objects have characteristics and they can be classified according to their degree of similarity. Fuzzy Sets allows to process this information and generate a matrix of similarity. However, it is sometimes not possible to know the degree of similarity between two objects. Then, it is necessary to handle information with vagueness. Interval-valued Fuzzy Relations can do that effectively.

In this paper, it is proposed a method of clustering using Interval-valued Fuzzy Relations. This methods has the following steps:

1. Creation of an interval-valued fuzzy relation R that represents the similarity between each pair of objects.
2. Imposition of transitivity computing the transitive closure of this interval-valued fuzzy relation.
3. Generation of the clusters.

Creation of an interval-valued fuzzy relation R.

Traditionally, clustering processes use distance measures to compute the clusters. However, in many times it is not possible to know the distance or the similarity between two objects with absolute precision (for example: data missing). In such cases it is always possible to compute lower and higher bounds. According to the features of the objects, two functions are used to compute these bounds to perform the degree of similarity between each pair of objects. This way, it is possible to compute its similarity degree and the uncertainty of these measures. Let C_{min} and C_{max} be the lower and higher bounds respectively. Then $R(e_i, e_j) = [C_{min}(e_i, e_j), C_{max}(e_i, e_j)]$

Imposition of transitivity.

Transitivity is imposed as a principle of rationality. The similarity degree between two objects is used to determine if they belong to the same cluster or not. It not desirable two objects with no relation and they have relation with another element. If this happens it is necessary to modify the matrix of similarity matrix. Even more, it is necessary to change very little the relation R.

Cluster generation.

The clusters are obtained by computing the α_1, α_2 cuts of R. The obtained clusters must be a set of equivalence classes so it is not possible to use

any $\mathcal{T}$-norm. Traditionally, it is used the $\mathcal{T}$-norm minimum ($T = [\min, \min]$) to impose the transitivity. This t-norm is the only one that generate alpha-cuts witch are all equivalence classes. We think this probably is not the better to impose $T_{\min}$-transitivity due the fact $T_{\min}$ is the highest $\mathcal{T}$-norm [12] because this way R can be "overmodified".

In this paper it is propose another $\mathcal{T}$-norm ($T_{\sim\min}$) that allows to obtain Interval-Valued Fuzzy Relations closer to the the original interval-valued relation.

Lemma 3.1. *Let $T_{\sim\min}$ and $T_{\min}$ be two t-norms on L. Then:*

$$T_{\sim\min}([x_1, x_2], [y_1, y_2]) \leq_L T_{\min}([x_1, x_2], [y_1, y_2])$$

for all $[x_1, x_2], [y_1, y_2]$ on L

Proof. According Examples 2.3 and 2.4:

- $T_{\min}([x_1, x_2], [y_1, y_2]) =_L$
  $=_L [\min(x_1, y_1), \min(x_2, y_2)]$
- $T_{\sim\min}([x_1, x_2], [y_1, y_2]) =_L$
  $=_L [\min(x_1, y_1), \max\{\min(x_1, y_2), \min(x_2, y_1)\}]$

It is only necessary to prove:

$$\max\{\min(x_1, y_2), \min(x_2, y_1)\} \leq \min(x_2, y_2)$$

- $\min(x_1, y_2) \leq \min(x_2, y_2)$ due to $x_1 \leq x_2$.
- $\min(x_2, y_1) \leq \min(x_2, y_2)$ due to $y_1 \leq y_2$.

so $\max\{\min(x_1, y_2), \min(x_2, y_1)\} \leq \min(x_2, y_2)$ □

In general, for any $\mathcal{T}$-transitive interval-valued fuzzy relation its α_1, α_2-cuts are not equivalence classes. Next lemma shows it is possible to find some α_1, α_2-cuts for the t-norm $T_{\sim\min}$ witch generate equivalence classes.

Lemma 3.2. *Let $T_{\sim\min}$ be pseudo-t-representable t-norm minimum $T_{\sim\min} : L^2 \to L$ defined in Example 2.4. Let R be a $T_{\sim\min}$-transitive interval-valued fuzzy relation on $X^2 \to L$. Then $R_{\alpha,\alpha}$ is a equivalence relation for all $[\alpha, \alpha]$ in L.*

Proof. It is necessary to prove that $R_{\alpha,\alpha}(e_i, e_k) = 1$ and $R_{\alpha,\alpha}(e_k, e_j) = 1$ implies $R_{\alpha,\alpha}(e_i, e_j) = 1$.

For convenience, $R(e_i, e_j)$ is written as R_{ij}. Then, the previous expression can be rewritten as follows:

if $R_{ik} \geq_L [\alpha, \alpha]$ and $R_{kj} \geq_L [\alpha, \alpha]$ then $R_{ij} \geq_L [\alpha, \alpha]$ or equivalently:

- if $\underline{R}_{ik} \geq \alpha$ and $\underline{R}_{kj}$ then $\underline{R}_{ij} \geq \alpha$ and
- if $\overline{R}_{ik} \geq \alpha$ and $\overline{R}_{kj}$ then $\overline{R}_{ij} \geq \alpha$

Due to the fact of R is $T_{\sim\min}$-transitive it is verified:

- $\min\{\underline{R}_{ik}, \underline{R}_{kj}\} < \underline{R}_{ij}$ (Condition 1) and
- $\max\{\min\{\overline{R}_{ik}, \underline{R}_{kj}\}, \min\{\underline{R}_{ik}, \overline{R}_{kj}\}\} < \overline{R}_{ij}$ (Condition 2)

Then:

- If $\underline{R}_{ik} \geq \alpha$ and $\underline{R}_{kj} \geq \alpha$ then $\min\{\underline{R}_{ik}, \underline{R}_{kj}\} \geq \alpha$ and $\underline{R}_{ij} \geq \alpha$ by Condition 1.
- If $\overline{R}_{ik} \geq \alpha$ and $\underline{R}_{kj} \geq \alpha$ then $\min\{\overline{R}_{ik}, \underline{R}_{kj}\} \geq \alpha$.
  In similar way it can be proved $\min\{\underline{R}_{ik}, \overline{R}_{kj}\} \geq \alpha$.
  So $\max\{\min\{\overline{R}_{ik}, \underline{R}_{kj}\}, \min\{\underline{R}_{ik}, \overline{R}_{kj}\}\} \geq \alpha$ and $\overline{R}_{ij} \geq \alpha$ by Condition 2.

Conclusion:
$\underline{R}_{ij} \geq \alpha$ and $\overline{R}_{ij} \geq \alpha$ so $R_{\alpha,\alpha}(e_i, e_j) = 1$

$\square$

This way, if the interval-valued fuzzy relation that represents the similarity degree between objects is not transitive, then it is better to impose the transitivity under $T_{\sim\min}$ instead of $T_{\min}$ because the final interval-valued fuzzy relation will be closer to the original relation.

In order to computing the distance between Interval-Valued Fuzzy Relations a measure of distance based on the Hamming distance is defined.

Definition 3.1. *Let R_X the set of interval-valued fuzzy relations on $X = \{e_1, \ldots, e_n\}$. The distance d between R and S ($R, S \in R_X$) is defined by:*

$$d(R, S) = \sum_{\forall i,j} | \overline{R}(e_i, e_j) - \overline{S}(e_i, e_j) | + | \underline{R}(e_i, e_j) - \underline{S}(e_i, e_j) |$$

4. An illustrative example

Let $q = \{q_1, \ldots, q_m\}$ be a set of boolean questions about a topic. Each question is composed of s items. Let q_a and q_b be the answers from two people. It is possible to compute the distance between q_a and q_b using next expression:

$$d(q_a, q_b) = \frac{dr}{s}$$

where dr is the number of different items.

It is also possible to obtain an expression to measure the similarity degree:

$$S(q_a, q_b) = 1 - d(q_a, q_b) = 1 - \frac{dr}{s}$$

However, in many cases people do not answer all questions. This way our data always have missing data and the previous expressions are not applicable. Then, it is necessary to compute an upper bound and a lower bound to express the similarity degree between q_a and q_b.

To compute $C_{min}(q_a, q_b)$ it is supposed the missing data in q_a does not have the same value than in q_b and viceversa. To compute $C_{max}(q_a, q_b)$ it is supposed the missing data in q_a have the same value than in q_b.

Let $q_a = \{a_1, \ldots, a_s\}$ and $q_b = \{b_1, \ldots, b_s\}$ be two answers with s items. The mapping md let us know the number of missing data. It is defined as follows:

$$md(a_i, b_i) = \begin{cases} 1, & \text{if } a_i \text{ or } b_i \text{ is a missing data;} \\ 0, & \text{othercase.} \end{cases}$$

then, the number of missing data is defined as follows:

$$Nmd(q_a, q_b) = \sum_{k=1}^{s} md(a_k, b_k)$$

Let C_{min} and C_{max} be these bounds defined as following:

$$C_{min}(q_a, q_b) = 1 - \frac{dr + Nmd(q_a, q_b)}{s}$$

$$C_{max}(q_a, q_b) = 1 - \frac{dr}{s}$$

Example 4.1. *Let q_a and q_b be two answers with 8 items:*

$$q_a = (yes, yes, yes, -, no, yes, -, no)$$
$$q_b = (no, -, yes, -, yes, no, -, no)$$

Then $dr(q_a, q_b) = 3$ and $Nmd(q_a, q_b) = 3$ so $C_{min}(q_a, q_b) = 1 - 6/8 = 0.25$ and $C_{max}(q_a, q_b) = 1 - 3/8 = 0.625$. The similarity degree is $[0.25, 0.625]$.

Let $q = \{q_1, \ldots, q_6\}$ be a set of answers to one question with six items. The character '-' means that the person did not answer for the corresponding item.

q_1	'y'	'y'	'y'	'-'	'n'	'y'
q_2	'n'	'n'	'n'	'n'	'y'	'n'
q_3	'y'	'y'	'y'	'-'	'n'	'y'
q_4	'y'	'y'	'y'	'n'	'y'	'n'
q_5	'y'	'-'	'n'	'-'	'y'	'n'
q_6	'n'	'y'	'-'	'-'	'n'	'-'

4.1. Creation of an interval-valued fuzzy relation R.

It is possible to apply the previous bounds (C_{min} and C_{max}) to compute the similarity interval between two responses. Let R be the matrix of similarity intervals, then:

$$R$$

$$\begin{pmatrix}
(1,1) & (0,0.4) & (0.8,1) & (0.5,0.8) & (0.2,0.6) & (0.3,0.9) \\
(0,0.4) & (1,1) & (0,0.4) & (0.5,0.6) & (0.5,0.9) & (0.2,0.8) \\
(0.8,1) & (0.0,0.4) & (1,1) & (0.5,0.8) & (0.2,0.6) & (0.3,0.9) \\
(0.5,0.8) & (0.5,0.6) & (0.5,0.8) & (1,1) & (0.5,0.9) & (0.2,0.8) \\
(0.2,0.6) & (0.5,0.9) & (0.2,0.6) & (0.5,0.9) & (1,1) & (0,0.8) \\
(0.3,0.9) & (0.2,0.8) & (0.3,0.9) & (0.2,0.8) & (0,0.8) & (1,1)
\end{pmatrix}$$

4.2. Imposition of transitivity.

Transitivity is imposed by generating transitive closures under t-norms. First, the transitive closure under T_{min} is showed:

$$R^{T_{min}}$$

$$\begin{pmatrix}
(1,1) & (0.5,0.8) & (0.8,1) & (0.5,0.8) & (0.5,0.8) & (0.3,0.9) \\
(0.5,0.8) & (1,1) & (0.5,0.8) & (0.5,0.9) & (0.5,0.9) & (0.3,0.8) \\
(0.8,1) & (0.5,0.8) & (1,1) & (0.5,0.8) & (0.5,0.8) & (0.3,0.9) \\
(0.5,0.8) & (0.5,0.9) & (0.5,0.8) & (1,1) & (0.5,0.9) & (0.3,0.8) \\
(0.5,0.8) & (0.5,0.9) & (0.5,0.8) & (0.5,0.9) & (1,1) & (0.3,0.8) \\
(0.3,0.9) & (0.3,0.8) & (0.3,0.9) & (0.3,0.8) & (0.3,0.8) & (1,1)
\end{pmatrix}$$

Finally, the transitive closure under $T_{\sim min}$ is showed:

$$R^{T_{\sim min}}$$

$$\begin{pmatrix}
(1,1) & (0.5,0.5) & (0.8,1) & (0.5,0.8) & (0.5,0.6) & (0.3,0.9) \\
(0.5,0.5) & (1,1) & (0.5,0.5) & (0.5,0.6) & (0.5,0.9) & (0.3,0.8) \\
(0.8,1) & (0.5,0.5) & (1,1) & (0.5,0.8) & (0.5,0.6) & (0.3,0.9) \\
(0.5,0.8) & (0.5,0.6) & (0.5,0.6) & (1,1) & (0.5,0.9) & (0.3,0.8) \\
(0.5,0.6) & (0.5,0.9) & (0.5,0.6) & (0.5,0.9) & (1,1) & (0.3,0.8) \\
(0.3,0.9) & (0.3,0.8) & (0.3,0.9) & (0.3,0.8) & (0.3,0.8) & (1,1)
\end{pmatrix}$$

It is possible to compute the distances between interval-valued fuzzy relations using the distance given in Definition 3.1:

$$d(R, R^{T_{\sim min}}) = 6.3 < d(R, R^{T_{min}}) = 8.3$$

so $R^{T_{\sim min}}$ is closer than $R^{T_{min}}$ to R.

4.3. Cluster generation.

Next, two alpha-cuts for R are showed with $\alpha = 0.5$ and $\alpha = 0.8$

$$R^{T_{\sim min}}_{0.5,0.5} = \begin{pmatrix}
1 & 1 & 1 & 1 & 1 & 0 \\
1 & 1 & 1 & 1 & 1 & 0 \\
1 & 1 & 1 & 1 & 1 & 0 \\
1 & 1 & 1 & 1 & 1 & 0 \\
1 & 1 & 1 & 1 & 1 & 0 \\
0 & 0 & 0 & 0 & 0 & 1
\end{pmatrix}$$

There exist two clusters:

- $C_1 = \{q_1, q_2, q_3, q_4, q_5\}$
- $C_2 = \{q_6\}$

$$R^{T_{\sim min}}_{0.8,0.8} = \begin{pmatrix}
1 & 0 & 1 & 0 & 0 & 0 \\
0 & 1 & 1 & 0 & 0 & 0 \\
1 & 0 & 1 & 0 & 0 & 0 \\
0 & 0 & 0 & 1 & 0 & 0 \\
0 & 0 & 0 & 0 & 1 & 0 \\
0 & 0 & 0 & 0 & 0 & 1
\end{pmatrix}$$

There exist five clusters:

- $C_1 = \{q_1, q_3\}$
- $C_2 = \{q_2\}$
- $C_3 = \{q_4\}$
- $C_4 = \{q_5\}$
- $C_5 = \{q_6\}$

5. Conclusions

It is shown that $\mathcal{IVFRs}$ are useful to clustering with no accurate information.

A measure of distance between $\mathcal{IVFRs}$ is given.

It is found some alpha-cuts are equivalence relations for pseudo-t-representable t-norms.

It is studied how to compute the T-transitive closure using a t-norm $T_{\sim min}$. This way the T-transitive closure is closer to the original relation than the T-transitive closure using the t-norm T_{min}.

References

[1] B. De Baets and H. De Meyer. On the existence and construction of t-transitive closures. *Information Sciences*, 152(1):167–179, 2003.

[2] R. Bellman, R. Kalaba, and L. Zadeh. Abstraction and pattern classification. *Journal of Mathematical Analysis and Applications*, 13(1):1 – 7, 1966.

[3] H. Bustince and P. Burillo. Mathematical analysis of interval-valued fuzzy relations: application to approximate reasoning. *Fuzzy Sets Syst.*, 113(2):205–219, 2000.

[4] C. Cornelis, G. Deschrijver, and E. Kerre. Implication in intuitionistic fuzzy and interval-valued fuzzy set theory: construction, classification, application. *International Journal of Approximate Reasoning*, 35(1):55–95, 2004.

[5] C. Cornelis, G. Deschrijver, and E. Kerre. Advances and challenges in interval-valued fuzzy logic. *Fuzzy Sets and Systems*, 157(5):622–627, 2006.

[6] G. Deschrijver. Arithmetic operators in interval-valued fuzzy set theory. *Information Sciences*, 177(14):2906–2924, 2007.

[7] P.A. Devijver and J. Kitttler. *Pattern Recognition: A Statistical Approach.* Prentice-Hall, 1982.

[8] R.O. Duda and P.E. Hart. *Pattern Classification and Scene Analysis.* Wiley, 1973.

[9] R. González del Campo, L. Garmendia, and B. De Baets. Transitive closure of l-fuzzy relations and interval-valued fuzzy relations. In *WCCI 2010 IEEE World Congress on Computational Intelligence*, pages 3141–3148, 2010.

[10] R. González del Campo, L. Garmendia, and J. Recasens. Transitive closure of interval-valued relations. In *Proceedings IFSA-EUSFLAT'09*, pages 837–842, 2009.

[11] R. González del Campo, L. Garmendia, and J. Recasens. Transitive closure of interval-valued fuzzy relations. *International Journal of Computational Intelligence Systems*, 6(4):648–657, 2013.

[12] R. González del Campo, L. Garmendia, and R. Yager. Specificity for interval-valued fuzzy sets. *International Journal of Computational Intelligence Systems*, 5(3):452–459, 2012.

[13] I. Grattan-Guinness. Fuzzy membership mapped onto intervals and many-valued quantities. *Mathematical Logic Quarterly*, 22(1):149–160, 1976.

[14] P. Grzegorzewski. Distances between intuitionistic fuzzy sets and/or interval-valued fuzzy sets based on the hausdorff metric. *Fuzzy Sets and Systems*, 148(2):319 – 328, 2004.

[15] K. Jahn. Intervall-wertige Mengen. *Mathematische Nachrichten.*, 68:115–132, 1975.

[16] A.K. Jain and R.C. Dubes. *Algorithms for Clustering Data.* Prentice-Hall, 1988.

[17] A. Maas, T. Bezembinder, and P. Wakker. On solving intransitivities in repeated pairwise choices. *Mathematical Social Sciences*, 29(2):83–101, 1995.

[18] E. H. Ruspini. A new approach to clustering. *Information and Control*, 15(1):22–32, July 1969.

[19] E. Sanchez and R. Sambuc. Fuzzy relationships. phi -fuzzy functions. application to diagnostic aid in thyroid pathology. *Proceedings of an International Symposium on Medical Data Processing*, pages 513–524, 1976.

[20] L.A. Zadeh. Fuzzy sets. *Information And Control*, 8:338–353, 1965.

[21] L.A. Zadeh. The concept of a linguistic variable and its application to approximate reasoning I. *Information Sciences*, 8:199–249, 1975.