Teaching Mathematics and Statistics in Sciences - IPA HU-SRB/0901/221/088
Interesting Mathematical Problems in Sciences and Everyday Life - 2011

Selected mathematical and statistical methods in biology

Andreja Tepavčević

Department of Mathematics and Informatics, Faculty of Sciences, University of Novi Sad

Trg Dositeja Obradovića 4, 21000 Novi Sad, Serbia

andreja@dmi.uns.ac.rs

Some specific methods with applications in biology, ecology and related natural sciences will be presented. The stress will be on less known methods with a wide potential in applications. In the first part, correspondences between two sets will be considered as a tool for investigating various connections among their elements. One of the sets is called the set of objects and another is the set of attributes and the correspondence shows which object posseses which attribute. In this context, an algebraic discipline called formal concept analysis will be presented. Furthermore, binary relations illustrating internal connections within one set will be elaborated. Equivalence relations and ordering relations have a special role in applications. Equivalence relations correspond to partitions of the set into disjoint classes. Classification of organisms is based on partitions of sets and equivalence relations and every taxon represents an equivalence class. Ordering relations often appear in applications as natural orderings on a set of objects. Two elements can be comparable or uncomparable under an ordering. Finally, some new methods connected to fuzzy sets and fuzzy relations will be presented. These concepts allow possibility to erase strong borders between sets, and instead of strong belongingness of elements to sets, grades of belongingness are introduced. In this context fuzzy cluster analysis are explained and software packages in which this method is implemented are listed.

1. Correspondences and relations

Let A and B be non-empty sets. **A correspondence** R of sets A and B is a subset of a direct product of these sets, i.e., $R \subseteq A \times B$. If for elements $a \in A$ and $b \in B$ we have that $(a,b) \in R$, then a and b are in the correspondence and we can write it also by aRb. Therefore, a correspondence is a connection of some elements in set A with some elements in B and we define it by set of ordered couples of elements which are connected in this context.

Example 1. Let $A = \{p, r, q, s\}$ be a set of four species of fungi of genus *Amanita*, where by p we denote *Amanita phaloides*, by r *Amanita rubescens*, by q *Amanita pantherina* and by s *Amanita citrina*. Let $B = \{a, b, c\}$ be a set of some characteristics that may posses fungi of genus *Amanita*. Namely, with a we denote the characteristic "stem has a ring", with b - "a mushroom has a cap with striate margin" and c - "there is remaining of the veil on the cap". Let the correspondence R be defined as follows:

Species x is in the correspondence R with the characteristic y, if mushroom x has a characteristic y.

Then, we can obtain R as a subset of the set $A \times B$ in the following way:

$$R = \{(p,a),(r,a),(r,c),(q,a),(q,b),(q,c),(s,a),(s,c)\}.$$

Each of the ordered couples from the set R means that a species of mushrooms has a specific characteristic. In example: $(q,b) \in R$ means that the species *Amanita pantherina* has a cap with striate margin. Since $(r,b) \notin R$, this means that *Amanita rubescens* does not posses this characteristics.

A **binary relation** (relation) is a correspondence on a single set A only. Therefore, the relation ρ on a set A is a subset of the set $A \times A$ (denoted also by A^2). The relation is a connection between some elements of a set A. If elements a and b from A (the ordering of elements is here important) are in a relation ρ, we denote this by $(a,b) \in \rho$ or as $a \rho b$.

In the sequel, we start from a correspondence on sets A and B, and use it to define two particular relations on sets A and B.

Let R be a correspondence from a set A to a set B ($R \subseteq A \times B$). Let ρ be a relation defined on A using the correspondence R in the following way:

$$(x,y) \in \rho \text{ if and only if } \{z \in B \mid (x,z) \in R\} = \{z \in B \mid (y,z) \in R\}.$$

In other words, elements x and y are in the relation ρ if they are in the correspondence R with same elements from B.

Similarly, we can define a relation θ on B:

$$(x,y) \in \theta \text{ if and only if } \{z \in A \mid (z,x) \in R\} = \{z \in A \mid (z,y) \in R\}.$$

Example 2. Let $A = \{a_1, a_2, a_3, a_4, a_5, a_6, a_7, a_8, a_9, a_{10}\}$ be a set of following fungi growing on wood: a_1 - *Agrocybe aegerita*, a_2 - *Pleurotus ostreatus*, a_3 - *Polyporus sulphureus*, a_4 - *Fistulina*

hepatica, a_5 - *Auricularia auricula-judae*, a_6-*Polyporus squamosus*, a_7-*Meripilus giganteus*, a_8-*Flammulina velutipes*, a_9- *Panus tigrinus*, a_{10}-*Pholiota destruens*.

Let $B=\{b_1,b_2,b_3,b_4,b_5,b_6,b_7\}$ be a set of some characteristics that can be used for mushroom identification: b_1- "mushroom has gills underneath", b_2- "mushroom has pores or tubes underneath", b_3- "mushroom has a red cap", b_4- "mushroom has a yellow cap", b_5- "mushroom with decurrent gills", b_6 - "mushroom with cup having free gills or with adnexed to shortly adnate gills", b_7-"fungi has a tough, gelatinous, elastic texture".

Let R be a correspondence defined as in Example 1:

$(x,y)\in R$ if and only if the species x has a characteristic y.

The correspondence R is given in the following table, where elements of the set A are in the table heading left and elements of the set B are in the table heading up, and if $(x,y)\in R$ then in the place of intersection of the x row and y column a small circle is placed.

R	b_1	b_2	b_3	b_4	b_5	b_6	b_7
a_1	∘					∘	
a_2	∘				∘		
a_3		∘		∘			
a_4		∘	∘				
a_5							∘
a_6		∘					
a_7		∘					
a_8	∘		∘			∘	
a_9	∘				∘		
a_{10}	∘					∘	

Table **1**

Now, we can consider the relation ρ on A obtained from the correspondence R as follows: two fungi species are in the relation ρ if they have same characteristics. Obviously, we

have that every fungi is in relation with itself, moreover we have that $a_1 \rho a_{10}$ and $a_{10} \rho a_1$, $a_2 \rho a_9$ and $a_9 \rho a_2$, $a_6 \rho a_7$ and $a_7 \rho a_6$. Since we have that there are mushrooms which are in correspondence with other different mushrooms, this shows that using the characteristics from the set B only we can not guarantee the correct identification of mushrooms. In order to be sure that we are able to identify all mushrooms from the set A, we have to enlarge set B, i.e., we have to take into account some other characteristics and then to consider the correspondence from the set A to this larger set.

Further, we can consider also a relation θ on the set B determined by the correspondence R. Two characteristics are in a relation θ if sets of mushrooms having these properties coincide. In this example each characteristic is in relation θ only to itself. This means that there are no characteristics that are possessed by the same species of mushrooms. If this would happen (that two or more characteristics are possessed by the same species of mushrooms) this would mean that not all of the properties from the set B are necessary for the identification of mushrooms from the set A, i.e., we would be able to determine the mushrooms using the smaller number of characteristics.

Now, in order to be able to identify all mushrooms from A, we have to add some more characteristics into the set B. Now we taker a new set $B_1 = B \cup \{b_8, b_9, b_{10}\}$, where new characteristics are as follows: b_8 is the characteristic "cap, and especially its margin is covered with whitish irregular scales", b_9 - "cap is eccentric and asymmetric and stem is short and eccentrically attached to cap" and b_{10} - "flesh changes color when broken or bruised ".

The related correspondence R_1 on sets A and B_1 is given in the following table:

R_1	b_1	b_2	b_3	b_4	b_5	b_6	b_7	b_8	b_9	b_{10}
a_1	∘					∘				
a_2	∘				∘			∘		
a_3		∘		∘						
a_4		∘	∘							
a_5							∘			
a_6		∘								

a_7		∘						∘	
a_8	∘		∘			∘			
a_9	∘				∘				
a_{10}	∘					∘		∘	

Table 2

Now, there are no different species of mushrooms with the identical properties from the set B_1. Therefore, it is possible to identify fungi from the set A according to characteristics from the set B_1 only.

2. Formal concept analysis

Formal concept analysis is a technique that is used to identify similarities in data, starting from a correspondence. It is used in data analysis, knowledge processing, in various disciplines. It is developed at TU Darmstadt, Germany by R. Wille and B. Ganter. After publishing several research papers in which the theory is developed, B. Ganter and R. Wille in 1999 published a book Formal Concept Analysis - Mathematical Foundations, Springer, 1999. This book contains the detailed explanation of this theory and its applications.

A formal context $K=(G,M,I)$ consists of two sets G and M and a relation (correspondence) I between G and M. The elements of G are called objects and the elements of M attributes of the context. For a set $A\subseteq G$ of objects, we define:

$$A'=\{m\in M\mid (g,m)\in I \text{ for all } g\in A\}$$

(the set of attributes common to the objects in A), and for a set $B\subseteq M$ of attributes we define

$$B'=\{g\in G\mid (g,m)\in I \text{ for all } m\in B\}$$

(the set of objects common to all the attributes in B.)

A formal concept of the context $K=(G,M,I)$ is a pair (A,B), where $A\subseteq G$, $B\subseteq M$, satisfying $A'=B$ and $B'=A$.

A is then called the **extent** and B the **intent** of the concept (A,B).

Further, the set of all concepts of a formal context $K=(G,M,I)$ is considered and it is denoted by $B(G,M,I)$.

A natural order can be introduced in the set of all concepts, as follows:

If (A,B) and (C,D) are two concepts of the same context, then (A,B) is a **subconcept** of

(C,D) if $A \subseteq C$ or $D \subseteq B$.

Then, also (C,D) is a superconcept of (A,B), and we write $(A,B) \leq (C,D)$. The relation $\leq$ is called the hierarchical order of the concepts.

$B(G,M,I)$ can be drown as a diagram and posses a special structure (such diagram is called a lattice and it is out of scope of this article).

Example 3. Let the set of objects G consist of following mushrooms:
$G = \{\text{Boletus edulis, Agaricus xanthoderma, Morchella esculenta}\}$.
Further, let set of attributes be as follows:

$M = \{\text{saprophyte, edible, basidiomyceta}\}$.

I is the correspondence "mushroom species x has property y".

	saprophyte	edible	basidiomyceta
Boletus edulis		x	x
Agaricus xanthoderma	x		x
Morchella esculenta	x	x	

The correspondence I is presented in the Table 3.

Table 3.

Now we look at some formal concepts of the context $K = (G,M,I)$. Since
$\{\text{Boletus edulis}\}' = \{\text{edible, basidiomyceta}\}$ and

$\{\text{edible, basidiomyceta}\}' = \{\text{Boletus edulis}\}$, we have that $(\{\text{Boletus edulis}\},\{\text{edible, basidiomyceta}\})$ is one concept.
Also, since
$\{\text{Boletus edulis, Agaricus xanthoderma}\}' = \{\text{basidiomyceta}\}$ and
$\{\text{basidiomyceta}\}' = \{\text{Boletus edulis, Agaricus xanthoderma}\}$,
$(\{\text{Boletus edulis, Agaricus xanthoderma}\},\{\text{basidiomyceta}\})$ is another concept and so on.
There are 8 different concepts.

3. Equivalence and ordering relations

3.1 Equivalence relations

Let ρ be a relation on a set A. We define special properties of relations which are important in applications.

The relation ρ is **reflexive** relation on A if every element from A is in relation with itself, i.e., if the following is satisfied:

$$(\forall x \in A)(x,x) \in \rho.$$

The relation ρ is **symmetric** on A if from " x is in relation with y " it always follows that y is in relation with x, i.e.,

$$(\forall x, y \in A)\big((x,y) \in \rho \Rightarrow (y,x) \in \rho\big).$$

The relation ρ is **transitive on** A if from the facts that x is in relation with y and y is in relation with z, it follows that x is in relation with z, i.e.,

$$(\forall x, y, z \in A)((x,y) \in \rho \wedge (y,z) \in \rho \Rightarrow (x,z) \in \rho).$$

The relation ρ is **anti-symmetric** if (x,y) and (y,x) are not together in ρ unless $x = y$, i.e.,

$$(\forall x, y \in A)((x,y) \in \rho \wedge (y,x) \in \rho \Rightarrow x = y).$$

There are relations which are symmetric and antisymmetric at the same time: one example is ordinary equality relation.

There are two special types of relations well known according to good applicability in various fields: equivalence relations and ordering relations.

A relation ρ on a set A is an **equivalence relation** (equivalence) on this set if it is reflexive, symmetric and transitive.

Common examples of equivalence relations we can find in taxonomy. Wherever we divide a set into some classes, there is a connected equivalence relation, as in the following examples.

Example 4. Let A be a set of some animals on a habitat and ρ a relation "to be the same species with". This relation is an equivalence relation.

Example 5. Let B be a set of all species of mushrooms that were found in a wood and θ the relation "to be the same genus with". This is also an equivalence relation.

Let A be a set, ρ an equivalence relation on A and $x \in A$. **The equivalence class** C_x determined by an element x is the set of all elements from A that are in the relation ρ with x, i.e.,

$$C_x = \{y \in A \mid x \rho y\}.$$

For every element from the set A there is an equivalence class determined by this element. Every element belongs to the equivalence class determined by it, that follows by the

reflexivity and the definition of the class. Therefore, every element from A belongs to some of these equivalence classes. On the other hand, every equivalence class contains elements from the set A only. Hence, we can conclude that the union of all equivalence classes is equal to the set A. Some of these classes can be equal. To be more precise, two classes are either equal or disjoint (which means that their intersection is the empty set). Since the set A is a union of sets (the equivalence classes) which are mutually disjoint, these sets make a partition of the set A.

Let A be a set and ρ an equivalence relation on it. The set of all equivalence classes of the relation ρ, $\{C_x \mid x \in A\}$ is called **a quotient set** and it is denoted by A/ρ.

In the example 4, one quotient set (according to the definition) form all animals that are of the same species with some observed animal. These are all animals of the same species. Hence, each equivalence class is formed by all the animals of the same species. The quotient set is set of all animal species on this habitat.

In example 5, one equivalence class is formed by all the species belonging to the same genus and the quotient set is set of all mushroom genera found in this wood.

Every classification of objects to some classes is based on equivalence relations. Cluster analysis or clustering, which is an assignment of a set of observations into subsets (called clusters) so that observations in the same cluster are similar in some sense (and different from observations from different clusters) is based on equivalence relations.

3.2 Ordering relations

A relation ρ on a set A is an **ordering relation** (order) if it is reflexive, anti-symmetric and transitive. Usual relations $\leq$, $\geq$ on the set of real numbers R and $\leq$, $\geq$ and $|$ (divisibility relation) on the set of natural numbers N are ordering relations. The inclusion of a collection of all subsets of a set is also an ordering relation. Relations $<$ and $>$ on sets R and N are anti-symmetric and transitive, but not reflexive. Moreover, no element is in relation with itself. This last property is called irreflexivity. A relation which is irreflexive, antisymmetric and transitive is called **strong ordering relation**.

Example 6. Let A be a set of people making a wider family and ρ is the relation "to be descendant of". The relation ρ is strong ordering relation on the set A, because it is irreflexive, anti-symmetric and transitive.

Ordering and strong ordering relations can be presented graphically, by a special type of diagram. In the diagram, elements of the sets are presented by small circles and the relation by a line connecting these circles. Two elements x and y are connected by a line if and only if they are in a direct connection (i.e., if x and y are in this relation, and there is no z different from x and y such that x is in the relation with z, and z is in the relation with y. If x is in the relation with y, then we represent this in a diagram by an arrow from x to y. In mathematics there is an agreement that the arrow is dropped and that x is on the line below y in the diagram.

Example 7. At a family meeting there are 14 people: three brothers A, B and C; their wives E, F and G respectively, their mother X and father Y, and also mother of the mother Z. There are also children: P and Q are

children of A and E; R and S are children of B and F, and T is a child of C and G. A diagram of strong ordering relation "to be descendant of" on this set of people is given in Figure 1.

From the diagram we can deduce the relation as follows: x is in the relation with y if and only if x is below y in the diagram and there is a sequence of lines in the same direction connecting them. In this way, from the diagram in Figure 1 we can read that T is a descendant of Z, and that G is not a descendant of Z since they are not connected by a line, etc.

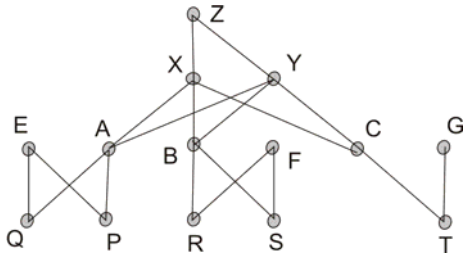


Figure 1.

Elements a and b are **comparable** according to an ordering or a strong ordering relation ρ if either $(a,b)\in\rho$ or $(b,a)\in\rho$. If a and b are not comparable, we say that they are **incomparable**.

Now, we return back to the relations defined in Section 1. Let R be a correspondence from a set A to a set B and ρ and θ relations defined on A and B, respectively, using R as follows:

$$(x,y)\in\rho \text{ if and only if } \{z\in B\,|\,(x,z)\in R\}=\{z\in B\,|\,(y,z)\in R\}, \text{ and}$$

$$(x,y)\in\theta \text{ if and only if } \{z\in A\,|\,(z,x)\in R\}=\{z\in A\,|\,(z,y)\in R\}.$$

The relation ρ is an equivalence relation on A and the relation θ is an equivalence relation on B. The relation ρ unifies elements which are similar taking into account the starting correspondence R. Now, we can consider equivalence classes of ρ and the set of all equivalence classes (quotient set) A/ρ. On A/ρ we can define the following relation:

Let C_x and C_y be two equivalence classes from A/ρ and let $\leq$ be a relation defined on A/ρ as follows: $C_x\leq C_y$ if and only if $x\rho y$. We can prove that this relation is well defined, since if $x\rho y$, then every element from C_x is in relation with every element from C_y. After checking reflexivity, anti-symmetry and transitivity, we can conclude that $\leq$ is an ordering relation on the set of all equivalence classes A/ρ. Analogue situation we have when we consider the relation θ on B. Since this is also an equivalence relation, we can take the quotient set, and the ordering relation on this quotient set can be defined similarly.

Now, we return back to the set all concepts of a formal context $B(G,M,I)$ defined in Section 2. The relation "to be subconcept of" defined on $B(G,M,I)$ as follows:
if (A,B) and (C,D) are two concepts then (A,B) is a subconcept of (C,D) if $A\subseteq C$ or

$D \subseteq B$ is an ordering relation.

4. Theory of fuzzy sets and fuzzy relations

The set is one of the basic notions in mathematics. The set usually contains some elements.The main relationship between sets and elements is membership. In classical set theory, an element either belongs to a set or it does not belong to the set. There is no other options. However, in applications there are situations in which it is not so easy to single out all elements satisfying some property from a set, in case the property is not precisely defined. In example, if we would like to consider a set of all small bird species, it would not be so simple. The hummingbird would obviously belong to the set of small birds, and also the wren, and maybe the canary, but what about a dove or a duck? For this set is difficult, or impossible that it is determined in a way that all people would accept this choice of elements. Situations in which for an object it is difficult to make a proper classification are very common. This type of problems can be overcome mathematically by introducing a new type of sets: fuzzy sets. These are sets with un-clear borders in which a grade of belongingness or grade of membership of an element to a set is introduced. In this context, grade of membership is a real number between 0 and 1. For an element that fully belongs to a set the grade of membership is set to be 1 and if an element totally does not belong to a set, the grade of membership is set to be 0. For elements that partially belong to the set (as e.g. a duck to a set of small bird species), the grade of membership would be a number between 0 and 1. An element that according to an estimation belongs more to the set then the other element, would have a higher grade of membership. For example, the grade of membership of a hummingbird and wren to a set of small birds might be 1, for a dove the grade of membership might be 0.4, for a duck it might be 0.2. Obviously, for a king vulture, the grade of membership to the set of small birds would be 0.

The formal definition of a fuzzy set (sometimes called a fuzzy subset) would be introduced in the sequel.

A set X we are starting with is called a **universe**. **A fuzzy subset** A of a set X is determined by a function $\mu_A : X \to [0,1]$, where $[0,1]$ is the interval of real numbers. This function is called a **membership function.** For every $x \in X$, the value $\mu_A(x)$ is called a **grade of membership** of an element x to the set A. Hence, elements belong to a fuzzy set with higher or lower grade of membership. It is common to identify a fuzzy set with its membership function, and this notation will be also used here.

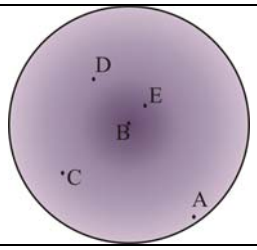
Example 8. Let X be a set of bird species $X = \{A, B, C, D, E\}$, such that A is a hoopoe, B is a hummingbird, C is a duck, D is a swallow and E is a wren. If we consider "a set of small birds", A would hardly belong to that set (i.e. the corresponding membirship degree would be near 0) , B would belong to that set (membership grade would be 1 and C, D and E would belong to that fuzzy set with the membership grade between 0 and 1, taking into account that if bird is smaller its membership grade is higher.

One fuzzy set in this context

is defined by following membership function:

$$\mu : \begin{pmatrix} A & B & C & D & E \\ 0.01 & 1 & 0.2 & 0.5 & 0.9 \end{pmatrix},$$

i.e., $\mu(D) = 0.5,$ etc.



Fuzzy sets have been introduced in 1965 by Lotfi Zadeh. Since then, they are applied in numerous fields. Fuzzy logic (a type of many valued logic based on fuzzy sets) is a basic tool for fuzzy controllers, fuzzy expert systems, fuzzy databases, etc. Fuzzy controllers are based on fuzzy relations and fuzzy correspondences.

One of the most important notions in this framework is the notion of a cut-set of a fuzzy set. Let $\mu : X \to [0,1]$ be a fuzzy set of a set X and let $p \in [0,1]$. Then a p-cut set is a classical subset of X usually denoted by μ_p and defined by

$$\mu_p = \{x \in X \mid \mu(x) \geq p\}.$$

Fuzzy correspondences (relations) are defined as mappings from direct products of sets to a [0,1] real interval. They are in fact fuzzy sets on a direct product of sets. As for classical relations, we will mostly use binary fuzzy relations.

Let A and B be sets, and $[0,1]$ a real interval. A mapping $R : A \times B \to [0,1]$ is a **fuzzy correspondence**. In a natural way, to every fuzzy correspondence there corresponds a matrix having elements from $[0,1]$, with rows indexed by elements from A and columns indexed by elements from B.

For $a \in A$ and $b \in B$, a value $R(a,b)$ is determined by the grade of relationship between elements a and b (which is the grade of membership of the ordered couple (a,b) to the correspondence R).

If $R : A \times B \to [0,1]$ is a fuzzy correspondence, then for every $p \in [0,1]$, a p-cut correspondence is a classical correspondence R_p on A and B defined by:

$$R_p = \{(a,b) \in A \times B \mid R(a,b) \geq p\}.$$

Futher important notions are x-row fuzzy sets and y-column fuzzy sets, defined as follows.

If $R : A \times B \to [0,1]$ is a fuzzy correspondence, then a fuzzy set $R^x : B \to [0,1]$ defined by

$R^x(y) = R(x,y)$ is an x-row fuzzy set of the correspondence R, and a fuzzy set $R^y : A \to [0,1]$ defined by $R^y(x) = R(y,x)$ is an y-column fuzzy set.

If A and B are finite sets, $A = \{x_1, ..., x_m\}$ and $B = \{y_1, ..., y_n\}$, we can present a fuzzy correspondence in a matrix:

$$R=\begin{bmatrix} R(x_1,y_1) & R(x_1,y_2) & . & . & . & R(x_1,y_n) \\ R(x_2,y_1) & R(x_2,y_2) & . & . & . & R(x_2,y_n) \\ . & . & . & . & . & . \\ . & . & . & . & . & . \\ . & . & . & . & . & . \\ R(x_m,y_1) & R(x_m,y_2) & . & . & . & R(x_m,y_n) \end{bmatrix}.$$

x -row fuzzy sets are then presented in rows of this matrix and similarly y-column fuzzy sets are in columns of this matrix.

If we have a relation on a set A, then we obtain a square matrix.

There are many problems in artificial intelligence that are based on fuzzy relations and correspondences, like fuzzy control theory and fuzzy databases. In real applications we often have a situation when output values are determined in advance by input values, and the problem is to find a fuzzy relation which performs such a transition. A fuzzy controller is rule based and fuzzy IF-THEN rules appear in this context. Rules are implications of the following type:

IF WASHING IS HEAVILY SOILED THEN ADD MORE DETERGENT

(in the control system of wash machine)

In this approach, each rule is of the type:

IF U is B THEN V is D.

This can be translated into a form

the pair (B,D) of (U,V) takes the value in R,

where R is a fuzzy relation.

Consequently, if U takes fuzzy input values from a set A, then we have that (U,V) is in relation G, where $G = A \circ R$.

Considering G and R as fuzzy relations on sets X and A as a fuzzy set on the same universe, we obtain the following equality:

$$G(x,y) = A(x) \wedge R(x,y).$$

The Mamdani approach to fuzzy controllers starts from a fuzzy relation R which is deduced from actual control process, and which from input values creates output values using the following compositional rule of inference:

$$\mu_{A \circ R}(y) = \sup_{x \in X}(\mu(x) \wedge R(x,y))$$

Here output values are determined in advance by input values so the problem is to find a fuzzy relation that satisfies the equation above.

We can present these problems schematically, as a usual matrix equation, taking into account that instead of usual number operations $+$ and $\cdot$, we have here operations $\sup$ and $\inf$ on $[0,1]$ interval (that sometimes we also denote, respectively by $\vee$ and $\wedge$).

Therefore, we have a problem of the type: Knowing two vectors that contain input and output values (two sequences of values of finite fuzzy sets) $[x_1,...,x_n]$ and $[y_1,...,y_m]$, problem is to find a fuzzy relation satisfying:

$$[x_1,...,x_n] \wedge R = [y_1,...,y_m].$$

If first fuzzy set is a fuzzy set on a set A, and the second fuzzy set is a fuzzy set on a set B, then R is a correspondence on A and B (i.e. a subset of $A \times B$).

Sometimes this relational equation does not have a solution, and sometimes it has more than one solution. In applications, it is an important problem to find the greatest solution of the equation.

In the sequel we give an example of such a problem and a solution.

Example 9. Let $\mu = [0.8, 0.9, 0.4, 0.1]$ be a fuzzy set on a set $A = \{a,b,c,d\}$. We consider this fuzzy set as a function from A to $[0,1]$ and values are defined respectively by the order elements are listed:

$$\mu(a) = 0.8, \mu(b) = 0.9, \mu(c) = 0.4, \mu(d) = 0.1.$$

Further, let $\nu = [0.9, 0.6, 0.2]$ be a fuzzy set on a set $B = \{p,q,r\}$, defined by $\nu(p) = 0.9, \nu(q) = 0.6, \nu(r) = 0.2$. The problem is to find a fuzzy correspondence that from input values defined by μ creates the output values defined by ν. We have to solve the following relational equation:

$$[0.8, 0.9, 0.4, 0.1] \wedge \begin{bmatrix} R(a,p) & R(a,q) & R(a,r) \\ R(b,p) & R(b,q) & R(b,r) \\ R(c,p) & R(c,q) & R(c,r) \\ R(d,p) & R(d,q) & R(d,r) \end{bmatrix} = [0.9, 0.6, 0.2]$$

This matrix equation is equivalent with the following system of equations:

$$(0.8 \wedge R(a,p)) \vee (0.9 \wedge R(b,p)) \vee (0.4 \wedge R(c,p)) \vee (0.1 \wedge R(d,p)) = 0{,}9$$

$$(0.8 \wedge R(a,q)) \vee (0.9 \wedge R(b,q)) \vee (0.4 \wedge R(c,q)) \vee (0.1 \wedge R(d,q)) = 0{,}6$$

$$(0.8 \wedge R(a,r)) \vee (0.9 \wedge R(b,r)) \vee (0.4 \wedge R(c,r)) \vee (0.1 \wedge R(d,r)) = 0{,}2$$

,

where we take into account that $\vee$ is the supremum and $\wedge$ the infimum on $[0,1]$ real interval.

Analysing the first equation, by the fact that all members (except the second one) are less than $0{,}9$, we can conclude that $R(b,p)$ must be greater or equal to $0{,}9$. Analysing the second equation, we have that either $R(a,q)$ or $R(b,q)$ must be equal to $0{,}6$ (and no one can be greater than 0,6). Further, analysing the third equation, we have that some of the elements $R(a,r)$, $R(b,r)$ and $R(c,r)$ must be equal to $0{,}2$ (and no one is greater than 0,2). Therefore, one of the solution of this relational equation is e.g.

$$R = \begin{bmatrix} 0.5 & 0.6 & 0.2 \\ 0.9 & 0.6 & 0.2 \\ 0.7 & 0.8 & 0.2 \\ 0.8 & 0.5 & 0.9 \end{bmatrix}.$$

This solution satisfies the conditions mentioned above, so we just enter the determined elements, others are taken randomly. This is not a unique solution of the equation, there are infinitely many of them.

A further application of fuzzy relations and correspondences that will be presented here is in connection with methods described in Section 3 for ordinary relations. First, a special type of fuzzy relations will be defined and used in the framework of relations determined by fuzzy correspondences.

Let $R: A\times A\to[0,1]$ be a fuzzy relation. Then the fuzzy relation R is:
reflexive if $R(x,x)=1$ for all $x\in A$;
symmetric if $R(x,y)=R(y,x)$, for all $x,y\in A$;
transitive if $R(x,y)\wedge R(y,z)\le R(x,z)$, for all $x,y,z\in A$;
antisymmetric if $R(x,y)\wedge R(y,x)=0$, for all $x,y\in A, x\neq y$.

The fact is that cut-relations of a fuzzy relation satisfying special properties above are ordinary relations satisfying analogous properties.

The fuzzy relation R is

-**fuzzy similarity** (equivalence) relation if it is reflexive, symmetric and transitive;
-**fuzzy quasi-ordering** relation if it is reflexive and transitive;
-**fuzzy ordering relation** if it is reflexive, antisymmetric and transitive.

Cut relations of fuzzy similarity relations are ordinary equivalence relations.
Cut relations of fuzzy quasi-ordering and ordering relations are, ordinary quasi ordering and ordering relations (respectively).

If ρ is a fuzzy correspondence from a set A to a set B, $\rho: A\times B\to[0,1]$, then a quasi-ordering relation θ on B can be defined similarly as in construction on ordinary relations, and then a fuzzy equivalence can be defined analogously as in ordinary case.

5. Many valued contexts

Another application of fuzzy correspondences are many-valued contexts. In section 2 we presented ordinary contexts, which are based on ordinary correspondences and relations. In classical formal concept analysis either an object has an attribute or it does not have the attribute. In the case of many valued contexts, an object can have an attribute to some extent.

A **many valued context** is an ordered quadruple (G,M,W,I), where G, M and W are sets and $I\subseteq G\times M\times W$ is a ternary relation. Here, as above, G is a set of objects and M is a set of attributes. What makes a difference here is a set W, that contains attribute values. I is a ternary relation between G, M and W satisfying the following condition:

$$\text{If } (g,m,w)\in I \text{ and } (g,m,v)\in I \text{ then } w=v.$$

This means that an attribute for an object should have a unique attribute value.
Therefore, it is possible to consider a special mapping $J: G\times M\to W$ instead of I. If W is an ordered set of attribute values, this represents a fuzzy correspondence.

This is the way how we can consider many valued contexts as fuzzy correspondences.

In the following, we continue Example 3.

Example 10. Let the set of objects G consist of following mushrooms:

$G = \{\text{Boletus edulis, Agaricus xanthoderma, Morchella esculenta}\}$

and the set of attributes contains following characteristics, as in Example 3:

$M = \{\text{saprophyte, edible, basidiomyceta}\}$.

Let $W = \{0, s, m, g, 1\}$ be a set of attribute values, with the following meaning:

0 - an object does not have an attribute at all

s - an object has an attribute to a small extent

m - an object has an attribute to a middle extent

g - an object has an attribute to a large extent

1 - an object has an attribute (completely).

The attribute values are ordered in the following way:

$$0<s<m<g<1.$$

In the following table a mapping $J : G \times M \to W$ is defined, taking into account mushrooms and their characteristics and attribute values:

	saprophyte	edible	basidiomyceta
Boletus edulis	s	1	1
Agaricus xanthoderma	1	s	1
Morchella esculenta	1	m	0

Table 4

Now, we can read from the table that e.g. Morchella esculenta is edible to a middle extent (this practically means that it is edible when cooked).
Now, we can consider a cut-correspondence, e.g. for m-cut, and then we obtain a usual context, identical to the one from Example 3.

6. Cluster analysis

Cluster analysis (clustering) is a method for assigning a set of different objects into groups (clusters), taking care objects in each group to be as similar as possible, and to be as different as possible from objects in other groups. Frases "as similar as possible" and "as different as possible" are related to a selected measure of similarity or distance. According to these criteria, some groups, called clusters are produced . Therefore, the **cluster** is a group of objects, which is separated as an entity because of the relatively high similarity (low diversity) within the group and relatively low similarities (big difference) with members of other groups. In most methods of cluster analysis the final result is a partition of the set of considered objects. Then each cluster can be divided further into smaller clusters using the same methods, and so on.

Clustering is a method of statistical data analysis with wide applications, e.g. machine learning, data mining, pattern recognition, image analysis, information retrieval, bioinformatics, etc.

Methods of cluster analysis are of the two main types:

1. agglomerative methods (methods in which objects are grouped into clusters according to characteristics considered.

2. divisive (the methods of dividing a set of objects into several groups (clusters)).

Final results is an equivalence relation (partition of a set of objects we started with).

There are several different types of methods of cluster analysis. Here we present two:

1. Method of k-means clustering
2. Method of fuzzy k-means clustering

Usualy we start from a set of objects we would like to classify into different groups and we have a set of characteristics of these objects. Usualy these characteristics are numerical, although there are methods that allow classification by qualitative characteristics.

K- means clustering

This technique is used in general when we already have some hypotheses concerning the number of clusters. So, we chose a number k in advance. Then this method will produce exactly k clusters that are as distinct as possible.
This method starts with k random clusters, and then move objects between obtained clusters with the goal to minimize variability within clusters and to maximize variability between clusters. Clusters are defined by their centers (so firstly we chose k random centers). Then, k-means algorithm assigns each point to the cluster whose center is nearest. Then, centers are calculated again, so that the center is the average of all the points in the cluster (the arithmetic mean for each characteristic separately).

So, we start from k random cluster centers, which determine k clusters. With $c^i, i = 1,...,k$ we denote the cluster centers and with $K(c_i)$ clusters determined by the center c_i. Now we consider all objects and we calculate the distance (using some of the similarity or distance measures) of each object from each of the k centers. For each object we determine the center with the minimum distance. Then, we join the object to the nearest center. In case there are more centers with the same distance we randomly chose one of them. Now we have all objects divided between k clusters. Further we calculate new center coordinates for each cluster (that are average coordinates of all objects in the cluster). Then we continue the procedure in the same way: again we consider all objects and calculate distances from the new centers. And then again assign objects to the closest clusters. In case clusters are identical to those in the previous step, we stop the procedure and adopt the obtained division. If clusters are different then in the previous step, we continue the procedure until some convergence criterion is met (usually that the clusters have not changed).

The main advantages of this algorithm are its simplicity and speed and thus it is possible to apply it on large datasets. Its disadvantage is that we do not obtain the same result with different initial random assignments. Another disadvantage is that we can use it only with numerical characteristics, since the mean has to be calculated.

Example 11. We start from a set of 8 objects with two defined characteristics and our aim is to apply k-means algorithm to obtain two clusters. This is same as we would like to define an equivalence relation on the set of objects with two equivalence classes.

Let $A = \{a_1, a_2, a_3, a_4, a_5, a_6, a_7, a_8\}$ be a set that contains eight species of mushrooms, and let $K = \{k_1, k_2\}$ be a set containing two numerical characteristics (e.g. let k_1 be an average size of the cap in centimeters and let k_2 be an average size of the spores in microns).

	k_1	k_2
a_1	3	4
a_2	2	1
a_3	10	9
a_4	4	4
a_5	5	5
a_6	9	4
a_7	8	9
a_8	3	2

Now we apply k-means algorithm to obtain two clusters. In

order to calculate distances we use Manhattan (or city-block) distance, where the distance between two objects is calculated simply as the sum of absolute differences on all coordinates. E.g. distance between objects with characteristics $(10,9)$ and (4.4) is calculated as $d = |10-4| + |9-4| = 11$.

Now, we start from two randomly chosen cluster centers, $c_1 = (3,3), \quad c_2 = (5,6)$.

In the first step we obtain the following clusters: $K_1 = K(c_1) = \{a_1, a_2, a_4, a_8\}$ and $K_2 = K(c_2) = \{a_3, a_5, a_6, a_7\}$.

Now we calculate new centers of the obtaining clusters as average characteristics.

$$c_1 = \frac{a_1 + a_2 + a_4 + a_8}{4} = (3, 2.75)$$

$$c_2 = \frac{a_3 + a_5 + a_6 + a_7}{4} = (8, 6.75).$$

Now, we have new cluster centres and we assign again the objects to the cluster to which it is closest. Now, we obtain slightly different clusters:

$K_1 = K(c_1) = \{a_1, a_2, a_4, a_5, a_8\}$, $K_2 = K(c_2) = \{a_3, a_6, a_7\}$.

Again we calculate centers of new clusters:

$$c_1 = \frac{a_1 + a_2 + a_4 + a_5 + a_8}{5} = (3.4, 3.2),$$

$$c_2 = \frac{a_3 + a_6 + a_7}{3} = (9, 7.33)$$

And again we assign objects to the clusters: $K_1 = K(c_1) = \{a_1, a_2, a_4, a_5, a_8\}$, $K_2 = K(c_2) = \{a_3, a_6, a_7\}$.

Since the obtained clusters are identical to those in the previous step, we stop the procedure and adopt the obtained division.

■ Fuzzy k-means clustering

The fuzzy k-means clustering technique (better known as fuzzy c-means) is an extension of the k-means clustering algorithm described above. The difference between two techniques is that the ordinary k-means make a partition of the starting set to the subsets, while the fuzzy c-means clustering produce a family of fuzzy sets instead of a family of ordinary sets. In other words, while in the usual k-means clustering an object belongs to only one cluster, in the fuzzy

k-means clustering a particular object can belong to more than one cluster with a certain grade of membership. Therefore, in the fuzzy clustering, each object has a degree of belonging to clusters, rather than belonging completely to only one cluster. Usually objects on the edge of a cluster belong to the cluster with a smaller degree than the objects in the center of the cluster. For each object a the degree of membership to the i-th cluster $k_i(a)$ is calculated. Usually, the sum of those membership degrees for any given a is defined to be 1.

The algorithm of the fuzzy k-means clustering is very similar to the one of k-means clusters:

First we choose a number of clusters.

Then we assign randomly for each object a grade of membership to every cluster.

Then we compute the center for each cluster (using the grades of membership).

Then we compute grades of membership of each element to every cluster taking into account previously computed centers of clusters.

Finally, we repeat last two steps until some convergence criterion is met (usually that the grades of membership to clusters have not changed or that the change of membership grades between two iterations is less than the given threshold) .

This procedure has advantages when compared to k-means algorithm, in some applications it is more convenient to have clusters with not clear border. The disadvantage is similar as for k-means procedure: that the final result depends on the initial choice of membership grades.

Ordinary k-means clustering is implemented in almost all statistical and mathematical software and fuzzy k-means clustering is also incorporated in some of them, like R-package, S-plus and Matlab.

References

[1] Bernhard Ganter, Rudolf Wille, Formal concept analysis, Springer, 1999.
[2] George J. Klir, Bo Juan, Fuzzy Sets and fuzzy logic, Theory and applications, Prentice Hall INC. 2002 (seventh printing).
[3] David N. Pegler, Mushrooms and toadstools, Mitchell Beazley Publishers Limited, 1983.
[4] Andreja Tepavčević, Zorana Lužanin, Mathematical methods in taxonomy, University of Novi Sad, 2006.
[5] Ronald R. Yager, Dimitar Filev, Essentials of fuzzy modeling and control, John Wiley & Sons, Inc. 1994.