On definition and construction of association measures

Ildar Z. Batyrshin*
Centro de Investigación en Computación (CIC), Instituto Politécnico Nacional, Mexico, Av. Juan de Dios Bátiz, Esq. Miguel Othón de Mendizábal S/N, Nueva Industrial Vallejo, Gustavo A. Madero, Ciudad de México, D.F., Mexico

Abstract. The definition and the general methods of construction of non-statistical association measures on different domains are discussed. An association measure is a function of two variables defined on a set X with involutive operation and satisfying the properties similar to the properties of the Pearson's correlation coefficient. Such measure can be used for analysis of the possible positive and negative relationships between variables. The methods of construction of association measures using similarity measures and pseudo-difference operations associated to t-conorms are discussed. The examples of association measures on different domains are considered.

Keywords: Association measure, Pearson's correlation coefficient, t-conorm, similarity measure, involution

1. Introduction

The association measures are widely used in data analysis. Different association and correlation measures have been introduced in statistics, data mining, fuzzy set theory etc. [1, 7, 12, 13, 17] for different types of data. The Pearson's correlation coefficient [12]

$$corr(x, y) = \frac{\sum_{i=1}^{n} (x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum_{i=1}^{n} (x_i - \bar{x})^2} \cdot \sqrt{\sum_{i=1}^{n} (y_i - \bar{y})^2}}.$$

(1)

is the most popular association measure used for analysis of possible relationships between variables. Many association measures similar to the correlation coefficient have been proposed but it is an interesting problem not only to introduce a new association measure for some type of data but to analyze a class of functions similar to the correlation coefficient and to propose the methods of their generation. In [1], it was proposed the measure of correlation between fuzzy membership functions satisfying to the set of properties similar to the properties of the Pearson's correlation coefficient. In [6], it was considered another set of properties similar to the properties of Pearson's correlation coefficient and defining the time series shape association measures. In [7], the general methods of construction of such association measures have been proposed and the sample Pearson's correlation coefficient was obtained as a particular case of the general approach. In [8], the methods proposed in [7] have been extended on the general case of functions $A : X \times X \rightarrow [-1, 1]$ defined on a set X with involutive operation N (called reflection) and satisfying the properties similar to the properties of the Pearson's correlation coefficient. The methods of construction of such measures [8, 9] use similarity measures and pseudo-difference operations associated with t-conorms [2, 15]. In [9], the problems appeared in the definition of the general class of functions similar to the Pearson's correlation coefficient have been discussed. These problems have the different reasons. First, the properties of the function (1): $corr(x,x) = 1$ and $corr(x,-x) = -1$, are, really, contradictive for the n-tuple

*Corresponding author. Ildar Z. Batyrshin, Centro de Investigación en Computación (CIC), Instituto Politécnico Nacional, Mexico, Av. Juan de Dios Bátiz, Esq. Miguel Othón de Mendizábal S/N, Nueva Industrial Vallejo, Gustavo A. Madero, 07738 Ciudad de México, D.F., Mexico. Tel.: +52 55 57296000; Ext. 56599; Fax: +52 55 57296000; Ext. 56504; E-mail: batyr1@gmail.com.

$x = (0, \ldots, 0)$ where it is fulfilled: $x = -x$. The similar problem appears, generally, for the fixed points of the reflection operation N used in the definition of the association measure A. Second, the function (1) does not defined for the constant n-tuples $x = (x_1, \ldots, x_n) = (s, \ldots, s)$ where s is some real value because the denominator of (1) equals to 0. Similarly, it is possible that an association measure cannot be defined on all set X. Such elements of X can be excluded from the domain of the association measure or this function should be additionally defined there. Third, depending on the domain X, additionally to the general properties of the association measures it is possible to consider other properties specific for this domain. See, for example, the definition of time series shape association measure [6, 7].

The current paper tries to avoid these problems by two ways. First, to consider explicitly the association measures defined on some subset V of X where these problems disappear. Second, to define the association measure on the set X and to correct some properties required from the association measure to avoid the possible contradiction between them.

The current paper also gives the proofs of some general results considered in the previous papers of the author without proofs. Some related details can be found also in [10].

The paper has the following structure. Section 2 discusses the definitions and the properties of association measures defined on the sets with involutive operation. For example, the simple association measure on the set of real values is introduced. Section 3 considers the basic definitions and the properties of operations of fuzzy logic used in the following sections. Section 4 considers the general methods of construction of association measures and gives the proofs of the related theoretical results. Section 5 considers an example of association measure constructed by proposed methods. The conclusions are given in the last section.

2. Association measures

Let X be a set and $|X| > 1$.

Definition 1. Let $N : X \to X$ be a function satisfying for all $x \in X$ the property:

$$N(N(x)) = x. \qquad \text{(involutivity)} \qquad (2)$$

N is called a *reflection* on X if it is not an identical function, i.e. for some $x \in X$ it is fulfilled $N(x) \neq x$. An element $x \in X$, such that

$$N(x) = x, \qquad (3)$$

is called a *fixed point* of N in X.

The fixed points will be denoted by x_{FP}, hence:

$$N(x_{FP}) = x_{FP}. \qquad (4)$$

Denote $FP(N, X)$ the set of all fixed points of N in X. This set can be empty.

Definition 2. Let X be a set with a reflection operation N on X, V be a subset of X, $|V| > 1$, from $x \in V$ it follows $N(x) \in V$ and the restriction of N on V is a reflection on V. A function $A : V \times V \to [-1, 1]$ satisfying for all $x, y \in V$ the properties:

$$A(x, y) = A(y, x), \qquad \text{(symmetry)} \qquad (5)$$

$$A(x, x) = 1, \qquad \text{(reflexivity)} \qquad (6)$$

$$A(x, N(y)) = -A(x, y), \qquad \text{(inverse relationship)} \qquad (7)$$

is called an *association measure* on V.

Proposition 1. *If A is an association measure on $V \subseteq X$ then $V \subseteq X \backslash FP(N, X)$.*

Proof. Suppose Proposition 1 does not true, i.e. A is a function satisfying on V the properties (5)-(7) and V contains some fixed point x_{FP} of the reflection N. Then from Equations (4) and (6) we obtain: $A(x_{FP}, N(x_{FP})) = A(x_{FP}, x_{FP}) = 1$, but from Equations (7) and (6) we have: $A(x_{FP}, N(x_{FP})) = -A(x_{FP}, x_{FP}) = -1$. The obtained contradiction proves the proposition. ∎

Consider the following properties of association measures:

$$A(N(x), N(y)) = A(x, y), \text{ (cancellation of reflections)} \qquad (8)$$

$$A(x, N(y)) = A(N(x), y). \text{ (permutation of reflections)} \qquad (9)$$

Proposition 2. *The association measure A on V satisfies for all $x, y \in V$ the properties (8), (9) and the following property*:

$$A(x, N(x)) = -1. \qquad (10)$$

Proof. (8) and (9) follow from Equations (7) and (5): $A(N(x), N(y)) = -A(N(x), y) = -A(y, N(x)) = A(y, x) = A(x, y)$. $A(N(x), y) = A(y, N(x)) = -A(y, x) = -A(x, y) = A(x, N(y))$. Equation (10)

follows from Equations (7) and (6): $A(x, N(x)) = -A(x, x) = -1$. ∎

Definition 3. Let X be a set with a reflection operation N on X and $FP(N, X)$ be a set of fixed points of N. A function $A : X \times X \to [-1, 1]$ will be called:

1) *an association measure of type* 1 on X if Equations (5) and (7) are fulfilled for all $x, y \in X$ and Equation (6) is fulfilled for all $x \notin FP(N, X)$;
2) *an association measure of type* 2 on X if Equations (5) and (6) are fulfilled for all $x, y \in X$ and Equation (7) is fulfilled for all $x \in X$ and all $y \notin FP(N, X)$.

Proposition 3. *An association measure A of type 1 on X satisfies for all $x, y \in X$ the properties (8), (9) and*:

$$A(x, N(x)) = -1 \; if \; x \notin FP(N, X), \qquad (11)$$

$$A(x, x_{FP}) = A(x_{FP}, x) = 0, \text{for all } x_{FP} \in FP(N, X). \qquad (12)$$

Proof. For the proof of Equations (8), (9) and (11) see the proof of Proposition 2. Let us prove Equation (12). From Equations (4), (7) and (5) for all $x \in X$ and all $x_{FP} \in FP(N, X)$ it follows: $A(x, x_{FP}) = A(x, N(x_{FP})) = -A(x, x_{FP})$ hence $A(x, x_{FP}) = 0$ and $A(x_{FP}, x) = 0$. ∎

Note that from Proposition 3 it follows $A(x_{FP}, x_{FP}) = 0$. Although some papers require the fulfillment of (6) for all $x \in X$, in this paper the association measures of type 2 will be not considered. The property (12) of association measures of type 1 seems more reasonable. See [10].

2.1. Association measures on [0,1]

In [10], it was considered an association measure of type 1 on [0,1] related with the strong negation N.

Definition 4. A *strong negation* on $X = [0,1]$ is a continuous strictly decreasing function $N : [0, 1] \to [0, 1]$ satisfying for all $x \in [0, 1]$ the following properties:

$$N(N(x)) = x, \qquad (13)$$

$$N(0) = 1, N(1) = 0. \qquad (14)$$

A strong negation is a reflection operation on [0,1] with the unique fixed point denoted as c. In [10], it was considered the class of *c-separable* association

measures of type 1 satisfying for all $x, y \in [0, 1]$ the properties:

$$A(x, y) > 0 \; \text{if} \; x, y > c \; \text{ or } \; x, y < c, \qquad (15)$$

$$A(x, y) = 0 \; \text{if} \; x = c \; or \; y = c. \qquad (16)$$

$$A(x, y) < 0 \; \text{if} \; x < c < y \; \text{ or } \; y < c < x. \qquad (17)$$

Such association measures can be used for analysis of associations between truth or probability values of some plausible statements P and Q. For example, the association between them is negative when one statement has high plausibility value and another one has low plausibility value.

2.2. Association measures on the set of real values

Let X be a set of real values, $X = \mathbf{R}$, and $N(x) = -x$ for all x in $\mathbf{R}$. We have $X_{FP} = 0$ and $FP(N, X) = \{0\}$. The association measures of type 1 satisfy on R the properties:

$$A(x, y) = A(y, x), \qquad (18)$$

$$A(x, x) = 1 \text{ if } x \neq 0, \qquad (19)$$

$$A(x, -y) = -A(x, y). \qquad (20)$$

From Proposition 3 we obtain the following properties of the association measures on $\mathbf{R}$:

$$A(-x, -y) = A(x, y), \qquad (21)$$

$$A(x, -y) = A(-x, y), \qquad (22)$$

$$A(x, -x) = -1 \text{ if } x \neq 0, \qquad (23)$$

$$A(x, 0) = A(0, x) = 0. \qquad (24)$$

Similarly to the c-separable association measure on [0,1] introduce the following definition.

Definition 5. An association measure A on the set of real values $\mathbf{R}$ is called 0-*separable* (or simply "separable") if the following properties are fulfilled for all $x, y \in \mathbf{R}$:

$$A(x, y) > 0 \; \; if \; \; x \cdot y > 0, \qquad (25)$$

$$A(x, y) = 0 \; \text{if} \; x \cdot y = 0. \qquad (26)$$

$$A(x, y) < 0 \; \text{if} \; x \cdot y < 0, \qquad (27)$$

0-separable association measures have the simple interpretation: x and y are positively associated if they have the same sign and they are negatively associated if they have the opposite signs. Based on these considerations it can be proposed the following simplest association measure on the set of real values.

Proposition 4. *The function*

$$A(x, y) = sign(x \cdot y) = \begin{cases} 1, & if\ x \cdot y > 0 \\ 0, & if\ x \cdot y = 0 \\ -1, & if\ x \cdot y < 0 \end{cases}, \quad (28)$$

is the 0-separable association measure of type 1 on the set of real values.

The proof is straightforward.

2.3. Association measures on the set of time series

Association measures on the set of time series are considered in [6, 7]. A time series of the length n, $(n > 1)$, is a sequence (n-tuple) of a real values $x = (x_1 \ldots, x_n)$. Consider the reflection operation $N(x) = -x = (-x_1, \ldots, -x_n)$ on the set X of all time series with the length n. Suppose p, q are real values and $p \neq 0$. Define $x + y = (x_1 + y_1, \ldots, x_n + y_n)$ and $py + q = (py_1 + q, \ldots, py_n + q)$. Denote $q_{(n)}$ a constant time series with the length n with all elements equal to q. The n-tuple $x_{FP} = 0_{(n)}$ is a unique fixed point of N. We write $x = const$ if $x = q_{(n)}$ for some q, and $x \neq const$ if $x_i \neq x_j$ for some $i \neq j$ from $\{1, \ldots, n\}$. Denote X_C a set of all constant time series from X.

Definition 6. Suppose V is a subset of X such that from $x \in V$ it follows $-x \in V$, and $x + q \in V$ for all real q. A function $A : V \times V \to [-1, 1]$ satisfying on V the properties Equations (5)-(7) and the property:

$$A(x + q, y) = A(x, y), \text{ for all real } q,$$
$$\text{(translation invariance)} \quad (29)$$

is called a *shape association measure* on V. If from $x \in V$ it is fulfilled $px \in V$ for all $p > 0$ and A satisfies on V the property:

$$A(px, y) = A(x, y), \text{ for all } p > 0, \text{ (scale invariance)} \quad (30)$$

then A is called a *scale invariant* association measure.

Proposition 5. *If A is an association measure on V then* $V \subseteq X \backslash X_C$.

Proof. Suppose Proposition 5 does not true, i.e. A is an association measure on V, and V contains constant time series $x = s_{(n)}$ where s is some real value. For $q = -2s$ we have $x + q = x - 2s = (s - 2s, \ldots, s - 2s) = (-s)_{(n)} = -(s)_{(n)} = -x$ and from Equations (29), (5)-(7) we obtain: $A(x, x) = A(x + q, x) = A(-x, x) = A(x, -x) = -A(x, x) = -1$, that contradicts to reflexivity of A. The obtained contradiction proves the Proposition ∎

In the next section, there are considered the basic properties of some operations of fuzzy logic that will be used further in construction of association measures.

3. Basic properties of operations of fuzzy logic

Consider the basic properties of the operations of fuzzy logic used in the following sections [2–5, 10, 11, 14–16, 18].

Definition 7. *t-conorm* is a function $S : [0, 1]^2 \to [0, 1]$ satisfying for all $x, y, z \in [0, 1]$ the following properties:

$$S(x, y) = S(y, x), \qquad (commutativity)$$
$$S(x, S(y, z)) = S(S(x, y), z), \qquad (associativity)$$
$$S(x, y) \leq S(x, z)\ if\ y \leq z, \qquad (monotonicity)$$
$$S(x, 0) = x. \qquad (boundary\ condition)$$

From the definition of t-conorm it follows for all $a \in [0, 1]$:

$$S(1, x) = S(x, 1) = 1, S(0, x) = S(x, 0) = x. \quad (31)$$

Definition 8. A t-conorm S is *nilpotent* if there exist $x, y \in]0, 1[$ such that $S(x, y) = 1$.

Definition 9. An element $x \in]0, 1[$ is a *nilpotent element* of t-conorm S if there exists $y \in]0, 1[$ such that $S(x, y) = 1$.

It is clear that t-conorm S has no nilpotent elements if and only if for all $x, y \in [0, 1]$ it is fulfilled:

$$S(x, y) = 1 \text{ implies } x = 1 \text{ or } y = 1. \quad (32)$$

Consider the simplest, basic, t-conorms:

$$S_M(x, y) = \max\{x, y\}, \qquad \text{(maximum)}$$

$$S_P(x, y) = x + y - x \cdot y, \qquad \text{(probabilistic sum)}$$

$$S_L(x, y) = \min\{x + y, 1\}. \quad \text{(Lukasiewicz t - conorm)}$$

Maximum and probabilistic sum have no nilpotent elements but Lukasiewicz t-conorm has.

Definition 10. Let S be a t-conorm.

(i) The $S - difference \overset{s}{-}$ is defined for all $a, b \in [0, 1]$ as follows:

$$a \overset{s}{-} b = inf \{c \in [0, 1] \mid S(b, c) \geq a\}. \qquad (33)$$

(ii) The *pseudo-difference* $\ominus_s$ associated to S is defined for all $a, b \in [0, 1]$ as follows:

$$a \ominus_s b = \begin{cases} a \overset{s}{-} b & if\ a > b \\ -\left(b \overset{s}{-} a\right) & if\ a < b \\ 0 & if\ a = b \end{cases} \qquad (34)$$

Pseudo-difference $\ominus_s$ associated to S has the following properties (see [10, 15] for details):

1) $1 \ominus_s b = 1$, if $b = 0$ or

if $b < 1$ and t - conorm S has no nilpotent elements. (35)

2) For any $a, b \in [0, 1]$ it is fulfilled:

$$a \ominus_s b = -(b \ominus_s a). \qquad (36)$$

3) If t-conorm S is continuous at the point 0 in both arguments then the following is fulfilled for all $a, b \in [0, 1]$:

$$a \ominus_s b > 0 \ if\ a > b. \qquad (37)$$

The following pseudo-differences are associated to the basic t-conorms S_M, S_P and S_L:

$$a \ominus_M b = \begin{cases} a & if\ a > b \\ -b & if\ a < b \\ 0 & if\ a = b \end{cases}, \qquad (38)$$

$$a \ominus_P b = \begin{cases} \frac{a-b}{1-\min(a,b)}, & if\ a \neq b \\ 0, & if\ a = b \end{cases}, \qquad (39)$$

$$a \ominus_L b = a - b. \qquad (40)$$

Definition 11. An automorphism of the interval $[0,1]$ is a continuous, strictly increasing function φ :

$[0, 1] \rightarrow [0, 1]$ satisfying boundary conditions $\varphi(0) = 0$, $\varphi(1) = 1$.

Theorem 1. [3, 18]. *A function $N : [0, 1] \rightarrow [0, 1]$ is a strong negation if and only if there exists an automorphism φ of the unit interval such that*

$$N(x) = \varphi^{-1}(1 - \varphi(x)). \qquad (41)$$

The function φ in Equation (41) is called a generator of N.

Example 1. The standard negation:

$$N(x) = 1 - x, \qquad (42)$$

has the generator $\varphi(x) = x$ and the fixed point $c = 0.5$.

Example 2. Yager negation:

$$N_p(x) = \sqrt[p]{1 - x^p}, p > 0, \qquad (43)$$

has the generator $\varphi(x) = x^p$ and the fixed point $c = \sqrt[p]{0.5}$.

Example 3. The negation, introduced by Batyrshin in [10]:

$$N(x) = \begin{cases} 1 - \frac{1-c}{c}x & if\ x \leq c \\ \frac{c}{1-c}(1 - x) & otherwise \end{cases}, \qquad (44)$$

It has the generator:

$$\varphi(x) = \begin{cases} \frac{x}{2c} & if\ x \leq c \\ \frac{x+1-2c}{2(1-c)} & otherwise \end{cases}. \qquad (45)$$

This simple strong negation connects by line segments the fixed point (c, c) with the points $(0,1)$ and $(1,0)$. It can be used for construction of strong negations with any fixed point $c \in]0, 1[$.

4. Constructing association measures

Consider the methods of construction of association measures using a similarity measure and pseudo-difference operation associated with some t-conorm and prove the related results [8, 9].

Definition 12. A function $SIM : X \times X \rightarrow [0, 1]$ is a *similarity measure* on X if it satisfies for all $x, y \in X$ the properties:

$$SIM(x, y) = SIM(y, x), \qquad \text{(symmetry)} \quad (46)$$

$$SIM(x, x) = 1. \qquad \text{(reflexivity)} \qquad (47)$$

Definition 13. A similarity measure *SIM* on *X* is *strict reflexive* if

$$SIM(x, y) < SIM(x, x) \quad \text{for all } x \neq y \text{ in } X. \quad (48)$$

For strict reflexive similarity measure *SIM* on *X* with reflection *N* it is fulfilled the property of *weak similarity of reflections*:

$$SIM(x, N(x)) < 1 \quad \text{for all } x \in X \backslash FP(N, X). \quad (49)$$

Proposition 6. *The similarity measure SIM satisfies for all $x, y \in X$ the cancellation of reflections property:*

$$SIM(N(x), N(y)) = SIM(x, y), \qquad (50)$$

if and only if it satisfies the permutation of reflections property:

$$SIM(x, N(y)) = SIM(N(x), y). \qquad (51)$$

Proof. Equation (50) follows from Equation (51) and involutivity of $N : SIM(N(x), N(y)) = SIM(N(N(x), y)) = SIM(x, y)$. Equation (51) follows from Equation (50) and involutivity of $N : SIM(x, N(y)) = SIM(N(x), N(N(y))) = SIM(N(x), y)$. ∎

Theorem 2. *Suppose X is a set with a reflection N, $V \subseteq X \backslash FP(N, X)$, $|V| > 1$, V is closed under N which is a reflection on V, S is a t-conorm and SIM is a similarity measure on X satisfying the permutation of reflections property then the function $A_{SIM,S} : V \times V \to [-1, 1]$ defined for all $x, y \in V$ by*

$$A_{SIM,S}(x, x) = 1 \qquad (52)$$

$$A_{SIM,S}(x, N(x)) = -1 \qquad (53)$$

$$\begin{aligned} A_{SIM,S}(x, y) = SIM(x, y) \ominus_S SIM(x, N(y)) \\ \text{if } y \neq x, y \neq N(x), \end{aligned} \qquad (54)$$

is an association measure on V.

Proof. We need to prove only Equations (5) and (7).

Let us prove Equation (5). For $y = x$ Equation (5) is fulfilled trivially.

For $y = N(x)$ from involutivity of N we have $x = N(y)$, and from Equation (53) we obtain: $A_{SIM,S}(x, y) = A_{SIM,S}(x, N(x)) = -1$ and $A_{SIM,S}(y, x) = A_{SIM,S}(y, N(y)) = -1$, hence $A_{SIM,S}(x, y) = A_{SIM,S}(y, x)$.

Suppose $y \neq x$, $y \neq N(x)$ then from the involutivity of N it follows $N(y) \neq x$. From Equations (54), symmetry and permutation of reflections properties of *SIM* we obtain:

$$\begin{aligned} A_{SIM,S}(y, x) \\ = SIM(y, x) \ominus_S SIM(y, N(x)) \\ = SIM(x, y) \ominus_S SIM(N(x), y) \\ = SIM(x, y) \ominus_S SIM(x, N(y)) \\ = A_{SIM,S}(x, y). \qquad \blacksquare \end{aligned}$$

Let us prove Equation (7). For $y = x$ Equation (7) follows from Equations (53) and (52): $A_{SIM,S}(x, N(y)) = A_{SIM,S}(x, N(x)) = -1 = -A_{SIM,S}(x, x) = -A_{SIM,S}(x, y)$.

For $y = N(x)$ Equation (7) follows from the involutivity of N, Equations (52) and (53): $A_{SIM,S}(x, N(y)) = A_{SIM,S}(x, N(N(x))) = A_{SIM,S}(x, x) = 1 = -A_{SIM,S}(x, N(x)) = -A_{SIM,S}(x, y)$.

If $y \neq x$, $y = N(x)$ then Equation (7) follows from Equation (54), involutivity of N and Equation (36):

$$\begin{aligned} A_{SIM,S}(x, N(y)) \\ = SIM(x, N(y)) \ominus_S SIM(x, N(N(y))) \\ = SIM(x, N(y)) \ominus_S SIM(x, y) \\ = -(SIM(x, y) \ominus_S SIM(x, N(y))) \\ = -A_{SIM,S}(x, y) \qquad \blacksquare \end{aligned}$$

Theorem 3. *If in the conditions of the Theorem 2 SIM is a similarity measure on X satisfying the properties of permutation of reflections and weak similarity of reflections then the function $A_{SIM,S} : V \times V \to [0, 1]$ defined for all $x, y \in X$ by:*

$$A_{SIM,S}(x, y) = SIM(x, y) \ominus_S SIM(x, N(y)), \quad (55)$$

is an association measure on *V* if one of the following is fulfilled:

1) $SIM(x, N(x)) = 0$, for all $x \in V$,

    (non - similarity of reflections)   (56)

2) the t - conorm *S* has no nilpotent elements. (57)

Proof. The symmetry of $A_{SIM,S}$ follows from the symmetry and the permutation of reflections properties of *SIM*:

$$A_{SIM,S}(x, y)$$
$$= SIM(x, y) \ominus_S SIM(x, N(y))$$
$$= SIM(y, x) \ominus_S SIM(N(y), x)$$
$$= SIM(y, x) \ominus_S SIM(y, N(x))$$
$$= A_{SIM,S}(y, x).$$

The reflexivity of $A_{SIM,S}$ follows from the reflexivity and weak similarity of reflections properties of *SIM* and from (35) that requires the fulfilment of (56) or (57):

$$A_{SIM,S}(x, x) = SIM(x, x) \ominus_S SIM(x, N(x))$$
$$= 1 \ominus_S SIM(x, N(x)) = 1.$$

The inverse relationship of $A_{SIM,S}$ follows from the involutivity of *N* and Equations (36) and (55):

$$A_{SIM,S}(x, N(y))$$
$$= SIM(x, N(y)) \ominus_S SIM(x, N(N(y)))$$
$$= SIM(x, N(y)) \ominus_S SIM(x, y)$$
$$= -(SIM(x, y) \ominus_S SIM(x, N(y)))$$
$$= -A_{SIM,S}(x, y) \qquad \blacksquare$$

5. Examples of association measures

The examples of association measures on different domains constructed by the methods discussed in the previous section can be found in [7–10]. The similarity measures satisfying the conditions of Theorems 2 and 3 can be obtained from the distance measures used together with some data transformation [7], from generators of strong negations [3, 5, 9, 10] etc. For example, suppose $\phi, \psi : [0, 1] \to [0, 1]$ are automorphisms of [0,1] and ϕ defines by (41) a strong negation *N* on [0,1]. Then the function

$$SIM(x, y) = 1 - \psi(|\varphi(x) - \varphi(y)|), \qquad (58)$$

is a similarity measure on [0,1] that can be used for constructing association measure on [0,1] related with strong negations (42)-(44) (see [10] for details). Below is an example of the simplest association measure on [0,1] related with the standard negation (42) [10]:

$$A_{SIM,S_M}(x, y) =$$
$$\begin{cases} 1 - |x - y| & if \ x, y > 0.5 \ or \ x, y < 0.5 \\ 0, & if \ x = 0.5 \ or \ y = 0.5 \\ |x + y - 1| - 1 & otherwise \end{cases}$$

6. Conclusion

The paper gives the definitions of the association measures generalizing the Pearson's correlation coefficient and proposes the general methods of construction of such measures. The proofs of the main results are provided. The simple association measure on the set of real numbers is introduced. The considered methods of generation of association measures can be used for construction of association measures on different domains.

Acknowledgments

References

[1] C.A. Murthy, S.K. Pal and D. Dutta Majumder, Correlation between two fuzzy membership functions, *Fuzzy Sets and Systems* **17** (1985), 23–38.

[2] E.P. Klement, R. Mesiar and E. Pap, *Triangular Norms*, Kluwer, Dordrecht, 2000.

[3] E. Trillas, Sobre funciones de negacion en la teoria de conjuntos difusos, *Stochastica* **111** (1979), 47–59.

[4] G. Beliakov, A. Pradera and T. Calvo, *Aggregation functions: A guide for practitioners, Springer Publishing Company*, Incorporated, 2008.

[5] I. Batyrshin, On the structure of involutive, contracting and expanding negations, *Fuzzy Sets and Systems* **139** (2003), 661–672.

[6] I. Batyrshin, L. Sheremetov and J.X. Velasco-Hernandez, On axiomatic definition of time series shape association measures, in *ORADM 2012, Workshop on Operations Research and Data Mining*, Cancun, 2012, pp. 117–127.

[7] I. Batyrshin, Constructing time series shape association measures: Minkowski distance and data standardization, in: BRICS CCI 2013, Brasil, Porto de Galhinas, 2013. http://arxiv.org/pdf/1311.1958v3

[8] I. Batyrshin, Association measures and aggregation functions, in: Advances in Soft Computing and Its Applications, *Lecture Notes in Computer Science*, Springer, vol. 8266, 2013, pp. 194–203.

[9] I. Batyrshin, Association measures on sets with involution and similarity measure, *in Proc 4th World Conference on Soft*

Computing, Berkeley, California, 2014. Will be printed also by Springer, 2015.

[10] I.Z. Batyrshin, Association measures on [0,1], *Journal of Intelligent and Fuzzy Systems* **29** (2015), 1011–1020.

[11] J. Fodor and M. Roubens, *Fuzzy Preference Modelling and Multi-Criteria Decision Support*, Kluwer, Dordrecht, 1994.

[12] J.F. Hair Jr, E.R. Anderson and R.L. Tatham, Multivariate data analysis with readings, Macmillan Publishing Co., Inc., 1986.

[13] J. Han and M. Kamber, *Data mining: Concepts and techniques*, 2nd ed., Morgan Kaufmann, Amsterdam, 2006.

[14] L.A. Zadeh, Fuzzy sets, *Information and Control* **8** (1965), 338–353.

[15] M. Grabisch, J.-L. Marichal, R. Mesiar and E. Pap, *Aggregation Functions*, Cambridge University Press, Cambridge, 2009.

[16] R.R. Yager, On the measure of fuzziness and negation. II. Lattices, *Information and Control* **44** (1980), 236–260.

[17] S. Brin, R. Motwani and C. Silverstein, Beyond market baskets: Generalizing association rules to correlations, *in ACM SIGMOD Record*, 1997, pp. 265–276.

[18] S. Ovchinnikov and M. Roubens, On strict preference relations, *Fuzzy Sets and Systems* **43** (1991), 319–326.