Similarity Measure Construction for Non-Convex Fuzzy Membership Function

Park, Hyun Jeong[1], Sungshin Kim[2] and Sang H. Lee[3]

1 Dept. of Mathematics Education, Ewha Womans University
2 Dept. of Electrical Engineering, Pusan National University
3 School of Mechatronics, Changwon National University

Abstract

The similarity measure is constructed for non-convex fuzzy membership function using well known Hamming distance measure. Comparison with convex fuzzy membership function is carried out, furthermore characteristic analysis for non-convex function are also illustrated. Proposed similarity measure is proved and the usefulness is verified through example. In example, usefulness of proposed similarity is pointed out.

Key Words: Fuzzy entropy, Similarity measure, Distance measure, Non-convex fuzzy set

1. Introduction

Research on the similarity measure between fuzzy sets has been applied to the pattern classification or many other fields. Similarity measure represents the degree of similarity between two or more informations. It also has been noticed as the complementary meaning of the distance measure, i.e, $s + d = 1$, where d and s are distance and similarity measure respectively. Until now the research of designing similarity measure has been made by numerous researchers [1-8]. For fuzzy set, there is an uncertainty knowledge in fuzzy set itself [9]. Fuzzy set uncertainty has been studied through analyzing and designing fuzzy entropy. Besides fuzzy entropy, similarity also represent the relatedness of two fuzzy set. Hence information of the data can be obtained from analyzing the fuzzy membership function. Thus most studies about similarity measure have been emphasized based on membership function.

In the previous results, similarity measures are obtained through fuzzy number[10-13]. Fuzzy number provide similarity measure easily. However considering similarity measures are restricted within triangular or trapezoidal membership functions[10-13]. In this paper we design similarity measure for general fuzzy membership which is based on distance measure. W also verify the usefulness of similarity via proving it. However two kinds of similarity measure that were mentioned before are all applied to convex fuzzy membership function. Here we have question about non-convex fuzzy mem-

bership functions. To apply similarity measure to non-convex fuzzy membership function, we analyze the similarity characteristics for fuzzy sets.

In the next chapter, the axiomatic definitions of entropy, distance measure and similarity measure of fuzzy sets are introduced and fuzzy entropy is constructed through distance measure. In Chapter 3, similarity measures are constructed and proved through the distance measure. Similarity application to non-convex fuzzy set is proposed by considering support average. In Chapter 4, definition and application of non-convex fuzzy set is illustrated, and another similarity is proposed and verified. Conclusions are followed in Chapter 5.

Notations: Through out this paper, $R^+ = [0, \infty)$, $F(X)$, and $P(X)$ represent the set of all fuzzy sets and crisp sets on the universal set X respectively. $\mu_A(x)$ is the membership function of $A \in F(X)$, and the fuzzy set A, we use A^C to express the complement of A, i.e. $\mu_{A^c}(x) = 1 - \mu_A(x)$, $\forall x \in X$. For fuzzy sets A and B, $A \cup B$, the union of A and B is defined as $\mu_{A \cup B}(x) = max(\mu_A(x), \mu_B(x))$, $A \cap B$, the intersection of A and B is defined as $\mu_{A \cap B}(x) = min(\mu_A(x), \mu_B(x))$. A fuzzy set A^* is called a sharpening of A, if $\mu_{A^*}(x) \geq \mu_A(x)$ when $\mu_A(x) \geq 1/2$ and $\mu_{A^*}(x) \leq \mu_A(x)$ when $\mu_A(x) < 1/2$. For any crisp sets D, A_{near} and A_{far} of fuzzy set A are defined as $\mu_{A_{near}}(x) = 1$ and 0 when $\mu_A(x) \geq 1/2$ and $\mu_A(x) < 1/2$, respectively. Furthermore $\mu_{A_{far}}(x) = 0$ and 1 when $\mu_A(x) \geq 1/2$ and $\mu_A(x) < 1/2$.

2. Preliminaries

We introduce some preliminary results about axiomatic definitions of distance measure and similarity measure. As the meaning of fuzziness, fuzzy entropy are introduced. By analyzing the fuzzy membership function pairs, we have investigated the similarity measure via fuzzy membership functions. It is also mentioned that common areas represent similarity measure.

Definition 2.1 [4] A real function d: $F^2 \rightarrow R^+$ is called a distance measure on $F(X)$, if d satisfies the following properties:

(D1) $d(A,B) = d(B,A), \forall A,B \in F(X)$

(D2) $d(A,A) = 0, \forall A \in F(X)$

(D3) $d(D,D^c) = max_{A,B \in F} d(A,B), \forall D \in P(X)$

(D4) $\forall A,B,C \in F(X)$, if $A \subset B \subset C$, then $d(A,B) \leq d(A,C)$ and $d(B,C) \leq d(A,C)$.

Fuzzy normal entropy on F is obtained by the division of $max_{C,D \in F} s(C,D)$. Liu also pointed out that there is an one-to-one relation between all distance measures and all similarity measures, $d + s = 1$. In this paper, among distance measures, Hamming distance is commonly used as distance measure between fuzzy sets A and B,

$$d(A,B) = \frac{1}{n}\sum_{i=1}^{n}|\mu_A(x_i) - \mu_B(x_i)|$$

where $X = \{x_1, x_2, \cdots x_n\}$, $|k|$ is the absolute value of k.

With this Hamming distance measure we had proposed fuzzy entropy induced by distance measure which is different from Theorem 3.1 of Fan, Ma and Xie [7].

Fuzzy entropy 1. If distance d satisfies $d(A,B) = d(A^C, B^C)$, $A,B \in F(X)$, then

$$e(A) = 2d((A \cap A_{near}), [1]) + 2d((A \cup A_{near}), [0]) - 2$$

is fuzzy entropy.

Proposed entropy needs only A_{near} crisp set, and it has the advantage in computation of entropy. Furthermore we considered another entropy, which considers only A_{far}, and it has more compact form than another one.

Fuzzy entropy 2 If distance d satisfies $d(A,B) = d(A^C, B^C)$, $A,B \in F(X)$, then

$$e(A) = 2d((A \cap A_{far}), [0]) + 2d((A \cup A_{far}), [1])$$

is also fuzzy entropy.

Proofs of Fuzzy entropy 1 and 2 are found in [9].

Proposed entropies have some advantages to the Liu's, they don't need extra assumptions of Liu. Furthermore they use only one crisp sets A_{near} and A_{far}, respectively.

For the similarity analysis we introduce definition of similarity measure, and concept based on fuzzy membership functions.

Definition 2.2 [4] A real function s: $F^2 \rightarrow R^+$ is called a similarity measure, if s has the following properties:

(S1) $s(A,B) = s(B,A), \forall A,B \in F(X)$

(S2) $s(A,A^c) = 0, \forall A \in F(X)$

(S3) $s(D,D) = max_{A,B \in F} s(A,B), \forall A,B \in F(X)$

(S4) $\forall A,B,C \in F(X)$, if $A \subset B \subset C$, then $s(A,B) \geq s(A,C)$ and $s(B,C) \geq s(A,C)$.

Fuzzy normal similarity measure on F is also obtained by the division of $max_{C,D \in F} s(C,D)$. With the definition of similarity measure we derive the similarity between fuzzy sets. Consider the two gaussian type membership functions as in Fig. 1.
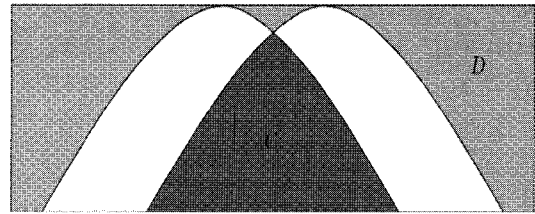

Fig.1 Gaussian type two membership functions

From Fig. 1, we conjecture that the shaded area can be considered as the component of similarity measure. At first we conjecture that the area C only be the similarity between two membership functions. However area D is also common area between two membership function. By proposing those areas as equations, next two equations

$$1 - d((A \cap B), [1]) \text{ and } 1 - d((A \cup B), [0])$$

will be denoted as C and D. Part C is common area of membership function, whereas D also represents the common information of the two membership functions. As the area C goes to 1, area D satisfies 0. If D satisfies 1, then two membership functions are the exact same membership function. Hence proper similarity measure is obtained by combining two values.

Those common areas are constructed via distance measure. In the following chapter we derived similarity measure and proved. In this paper we do not consider similarity measure using fuzzy number. Similarity measures based on fuzzy number are found in references [10-13].

3. Similarity measure derivation

We have described that the common areas of two fuz-

zy membership functions satisfy similarity between two fuzzy membership functions. By the Fig. 1, it is obvious summarizing two common areas mean similarity measure for corresponding two fuzzy membership functions.

Area C represents $1 - d((A \cap B), [1])$, and D also satisfies $1 - d((A \cup B), [0])$. Hence similarity for two fuzzy set A and B is obtained as follows.

$$s(A, B) = 2 - d((A \cap B), [1]) - d((A \cup B), [0])$$

With the following theorem we represent similarity measure using distance measure.

Theorem 3.1 For fuzzy set $A \in F(X)$, if d satisfies distance measure, then

$$s(A, B) = 2 - d((A \cap B), [1]) - d((A \cup B), [0]) \quad (1)$$

is the similarity measure between fuzzy set A and B.

proof. We verify that (1) satisfies similarity measure with proving the similarity definition. (S1) means the commutativity of set A and B, hence it is clear from (1) itself. From (S2), $s(A, A^c) = 0$ is shown by

$$\begin{aligned} s(A, A^c) &= 2 - d((A \cap A^c), [1]) - d((A \cup A^c), [0]) \\ &= 2 - d([0], [1]) - d([1], [0]) \\ &= 2 - 1 \cdot 1 - 1 \cdot 1 = 0. \end{aligned}$$

For all $A, B \in F(X)$, inequality of (S3) is proved through

$$\begin{aligned} s(A, B) &= 2 - d((A \cap B), [1]) - d((A \cup B), [0]) \\ &\leqq 2 - d((D \cap D), [1]) - d((D \cup D), [0]) \\ &= s(D, D). \end{aligned}$$

In the above, inequality is satisfied from $d((A \cap B), [1]) \geqq d((D \cap D), [1])$ and $d((A \cup B), [0]) \geqq d((D \cup D), [0])$.

Finally, (S4) is satisfied from $\forall A, B, C \in F(X)$, $A \subset B \subset C$,

$$\begin{aligned} s(A, B) &= 2 - d((A \cap B), [1]) - d((A \cup B), [0]) \\ &= 2 - d(A, [1]) - d(B, [0]) \\ &\geqq 2 - d(A, [1]) - d(C, [0]) \\ &= s(A, C) \end{aligned}$$

also

$$\begin{aligned} s(B, C) &= 2 - d((B \cap C), [1]) - d((B \cup C), [0]) \\ &= 2 - d(B, [1]) - d(C, [0]) \\ &\geqq 2 - d(A, [1]) - d(C, [0]) \\ &= s(A, C). \end{aligned}$$

In the above inequality is also satisfied with

$$d(B, [0]) \leqq d(C, [0]) \text{ and } d(B, [1]) \leqq d(A, [1]).$$

Therefore proposed similarity measure (1) satisfy Definition 2.2. Similarly, we propose another similarity measure in the following theorem.

Theorem 3.2 For fuzzy set $A \in F(X)$ and distance measure d,

$$s(A, B) = 1 - d((A \cap B), [0]) - d((A \cup B), [1]) \quad (2)$$

is the similarity measure of fuzzy set A and B.

proof. Proofs are shown similarly as Theorem 3.1.

We have proposed the similarity measure that are induced from distance measure. By analyzing the similarity (1) and (2), similarity is proportional to the common area of two membership functions. Summation of areas C and D represent similarity. At this point we have a question how about non-convex membership functions are ? For the same area of between convex to convex and convex to non-convex, which pair has better similarity ? Even though two pairs have same similarity measure, their geometrical description may not be identical. Hence another measure is required to discriminate two pairs. Now we introduce non-convexity in next chapter and propose another measure for similarity.

4. Non-convex fuzzy membership function

First we introduce non-convex fuzzy membership function. Definition of non-convex fuzzy membership function can be found in reference [14]. Non-convex fuzzy sets are not common fuzzy membership function. Definition of non-convexity derived from convexity definitely.

Definition 4.1 [14] A fuzzy set A is convex if and only if for any x_1, $x_2 \in X$ and any $\lambda \in [0, 1]$,

$$\mu_A(\lambda x_1 + (1 - \lambda)x_2) \geq \min\{\mu_A(x_1), \mu_A(x_2)\} \quad (3)$$

Non-convexity fuzzy set is said if it is not convex. Non-convex membership functions can be notified naturally 3 sub classes [15].

· Elementary non-convex membership functions
· Time related non-convex membership functions
· Consequent non-convex membership functions

First, a discrete fuzzy set express elementary non-convex fuzzy membership functions. However continuous domain non-convex fuzzy set may be less common.

Next, time related non-convex membership functions can be found in energy supply by time of day or year, mealtime by time of day. This fuzzy set is interesting as it is also sub-normal and never has a membership of zero.

Finally, Mamdani fuzzy inferencing is a typical example of consequent non-convex sets. In a rule based fuzzy system the result of Mamdani fuzzy inferencing is a

non-convex fuzzy set where the antecedent and consequent fuzzy sets are triangular and/or trapezoidal.

Jang *et.al* insisted that the definition of convexity of a fuzzy set is not as strict as the common definition of onvexity of a function [14]. Then the definition of convexity of a function is

$$f(\lambda x_1 + (1-\lambda)x_2) \geq \lambda f(x_1) + (1-\lambda)f(x_2), \qquad (4)$$
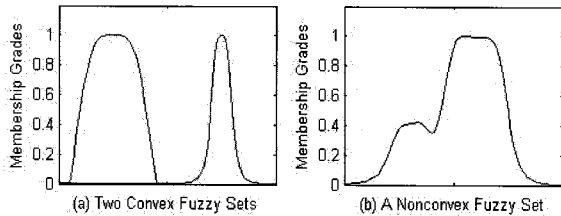
which is a tighter condition than (3).

Fig.2 Convex MF and Non-convex MF

Fig. 2 (a) show two convex fuzzy sets, the left fuzzy set satisfies both (3) and (4) while the right one satisfies (3) only. Whereas Fig. 2 (b) is a con-convex fuzzy set.

To obtain another measure discriminating two same similarity measure with (1) and (2), we have to consider another point of view. Commonly, different fuzzy membership function pair has different mean values of universe of discourse. Therefore we use "support" as measure for similarity to obtain implicit result.

$$support_A(x_i) = \frac{1}{n}\sum_{i=1}^{n}|x_i|, x_i \in A$$

Support between set A and B is represented as follows.

$$support(A,B) = |\frac{1}{n}\sum_{i=1}^{n}|x_{Ai}| - \frac{1}{m}\sum_{j=1}^{m}|x_{Bj}||,$$
$$x_{Ai} \in A, x_{Bj} \in B$$

Now we consider another similarity between set A and B is

$$s_2(A,B) = \frac{1}{support(A,B)+1}, \qquad (5)$$

Proofs of (5) can be obtain easily as follows. For (S1) it is natural from (5) itself. (S2) means similarity between fuzzy set and its complement satisfies minimum value. Hence the farther mean of universe of discourse, the bigger $support(A,B)$ become. If two fuzzy sets are the same, then $support(A,B)$ satisfies zero. Therefore (S3) is obtained easily. Finally, (S4) is also obtained easily from (S2) and (S3).

With the similarity properties of $s_1(A,B)$ and $s_2(A,B)$, next we combine similarity measure as follows

$$s(A,B) = \omega_1 s_1(A,B) + \omega_2 s_2(A,B). \qquad (6)$$

Where similarity measure (1) and (2) are replaced into $s_1(A,B)$, ω_1, ω_2 are weighting factors.

Now we consider the membership functions type 1 and 2 in Fig. 3 and 4. In the following figures, area between μ_A and μ_B are the same. Hence the similarity measure between μ_A and μ_B are same. As a result, two $s_1(A,B)$ has the same value. With the similarity measure $s_1(A,B)$ two fuzzy set pairs are exactly same similar. Now which case is more similar ?
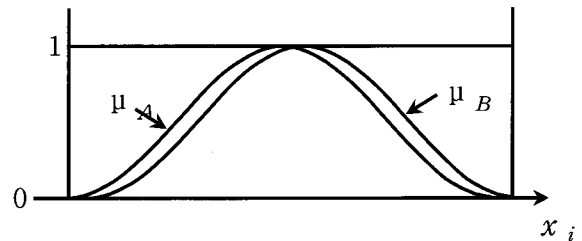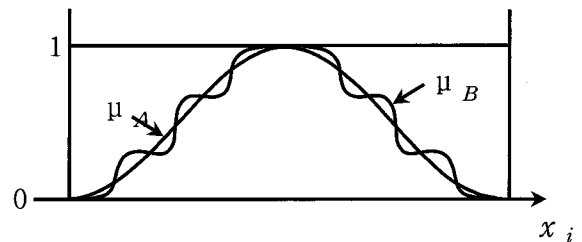

Fig.3 Membership functions type 1

Fig.4 Membership functions type 2

In the above case, the first part of (6) is the similarity measure based on distance measure, hence the value of Fig. 3 and Fig. 4 are the same. However the 2nd part of (6) represents the difference of average support. We have already verified similarity property of $s_1(A,B)$ and $s_2(A,B)$ separately. Determination of $s_2(A,B)$ is dependent on designer's point of view.

There are a lot of similarity measures besides of similarity measure $s_1(A,B)$. One of examples has the following formulation

$$s_2(A,B) = |max\{\mu_A(x_i)\} - max\{\mu_B(x_j)\}|.$$

Hence we can alter $s_2(A,B)$ with respect to each cases.

5. Conclusions

We introduce the distance measure and similarity measure, similarity measure can be represented by the function of distance measure. By the one to one correspondence of distance measure and similarity measure, we construct the similarity measure using distance measure. With the proposed similarity measure we analyze the similarity between fuzzy membership function

pair, especially non-convex fuzzy membership function. Furthermore modified similarity measure is constructed through "support" characteristic. We verify that the proposed measure is also satisfied as the similarity measure.

References

[1] D. Bhandari, N. R. Pal, "Some new information measure of fuzzy sets", Inform. Sci. 67, 209-228, 1993.

[2] A. Ghosh, "Use of fuzziness measure in layered networks for object extraction: a generalization", *Fuzzy Sets and Systems*, 72, 331-348, 1995.

[3] B. Kosko, *Neural Networks and Fuzzy Systems*, Prentice-Hall, Englewood Cliffs, NJ, 1992.

[4] Liu Xuecheng, "Entropy, distance measure and similarity measure of fuzzy sets and their relations", *Fuzzy Sets and Systems*, 52, 305-318, 1992.

[5] N.R. Pal, S.K. Pal, "Object-background segmentation using new definitions of entropy", *IEEE Proc.* 36, 284-295, 1989.

[6] J. L. Fan, W. X. Xie, "Distance measure and induced fuzzy entropy", *Fuzzy Set and Systems*, 104, 305-314, 1999.

[7] J. L. Fan, Y. L. Ma, and W. X. Xie, "On some properties of distance measures", *Fuzzy Set and Systems*, 117, 355-361, 2001.

[8] S. Aksoy and R.M. Haralick, "Feature normalization and likelihood-based similarity measures for image retrieval", *Pattern Recognition Letters*, 22, 563-582, 2001.

[9] S.H. Lee, S.P. Cheon and J.H. Kim, "Measure of certainty with fuzzy entropy function", *LNAI*, vol. 4114, 134-139, 2006.

[10 S.M. Chen, "New methods for subjective mental workload assessment and fuzzy risk analysis", *Cybern. Syst. : Int. J.*, vol 27, no. 5, 449-472, 1996.

[11] C.H. Hsieh and S.H. Chen, "Similarity of generalized fuzzy numbers with graded mean integration representation," in *Proc. 8th Int. Fuzzy Systems Association World Congr.*, vol 2, 551-555, 1999.

[12] HS. Lee, "An optimal aggregation method for fuzzy opinions of group decision," *Proc. 1999 IEEE Int. Conf. Systems, Man, Cybernetics*, vol. 3, 314-319, 1999.

[13] S.J. Chen and S.M. Chen, "Fuzzy risk analysis based on similarity measures of generalized fuzzy numbers," *IEEE Trans. on Fuzzy Systems*, vol. 11, no. 1, 45-56, 2003.

[14] J.S.R. Jang, C.T.Sun and E. Mizutani, *Neuro-Fuzzy and Soft Computing*, Prentice Hall, Upper Saddle River, 1997.

[15] J.M. Garibaldi, Musikasuwan, S. Ozen, T. John, R.I., "A case study to illustrate the use of non-convex membership functions for linguistic terms", *2004 IEEE International Conference on Fuzzy Systems*, vol.3, 1403- 1408, 2004.

저 자 소 개

박현정(Park, Hyun-Jeong)
She received the B.S. in Mathematics from Kyunghee University, in 1992, M.S. and Ph.D. degrees in Mathematics Education from Ewha Womans University, in 2001 and 2007, respectively. She served as a Researcher of Mathematics from Mar. 2002 to Dec. 2002 in The Korean Institute of Curriculum and Evaluation. Currently, she is an Instructor of Teaching Psychology of Mathematical Reasoning, Research and guiding curriculum planning in Ewha Womans University and Kyunghee University. Her research interests include fuzzy theory, In-depth examination on thinking processes in students when solving mathematical problems, Evaluation of thinking process using qualitative methods during the process of solving problem or understanding concepts.

E-mail : hyunjp@ewhain.net

김성신(Sungshin Kim)
제 17권 4호 (2007년 8월호) 참조

E-mail : sskim@pusan.ac.kr

이상혁(Sang-Hyuk Lee)
제 17권 1호 (2007년 2월호) 참조
E-mail : leehyuk@changwon.ac.kr