Relation between Certainty and Uncertainty with Fuzzy Entropy and Similarity Measure

Sanghyuk Lee, Yujia Zhai
Department of Electrical and Electronic Engineering, Xi'an Jiaotong-Liverpool University

Abstract We survey the relation of fuzzy entropy measure and similarity measure. Each measure represents features of data uncertainty and certainty between comparative data group. With the help of one-to-one correspondence characteristics, distance measure and similarity measure have been expressed by the complementary characteristics. We construct similarity measure using distance measure, and verification of usefulness is proved. Furthermore analysis of similarity measure from fuzzy entropy measure is also discussed.

• Key Words : fuzzy entropy, comparative data

1. Introduction

A view point of data analysis among data groups with similarity and dissimilarity measure has been emphasized on the design distance measure and similarity measure [1,2]. When information analysis of two scattered data group is carrying on, then the scattered data groups can be considered by fuzzy set. In fuzzy set, entropy is related with data uncertainty, and information is limited to the certainty with respect to deterministic data, *i.e.* similarity to the deterministic one. Hence fuzzy entropy and similarity measure are both used for the quantifying uncertainty and similarity measure of data [1,2]. Data uncertainty and certainty are usually expressed through probability point of view, probability of event denotes, which lies within. That probability value has the meaning of certainty and uncertainty simultaneously. Degree of similarity between two or more data has central role for the fields of decision making, pattern classification, or etc. [3,8]. Until now the research of designing similarity measure has been made by numerous researchers [8-12]. Two design methods are introduced through fuzzy number approach [8-11] and distance measure [12]. Method by fuzzy number make easy to design similarity measure.

However considering similarity measures are restricted within triangular or trapezoidal membership functions [8,11]. Whereas similarity measure based on the distance measure is applicable to general fuzzy membership function including non-convex fuzzy membership function [12].

For fuzzy set, uncertain knowledge is contained in fuzzy set itself. Hence uncertainty of the data can be also obtained from analyzing the fuzzy membership function. Mentioned uncertainty is described fuzzy entropy. Characterization and quantification of fuzziness are important issues that affect the management of uncertainty in many system models and designs. The fact that the entropy of a fuzzy set is a measure of the fuzziness of that fuzzy set has been established by previous researchers [14,16]. Liu proposed the axiomatic definitions of entropy, distance measure, and similarity measure, and discussed the relations between these three concepts. Kosko considered the relation between distance measure and fuzzy entropy. Bhandari and Pal provided a fuzzy information measure for discrimination of a fuzzy set

*교신저자 : Sanghyuk Lee(Sanghyuk.Lee@xjtlu.edu.cn)
접수일 2014년 9월 12일 수정일 2014년 11월 26일 게재확정일 2014년 12월 21일

relative to some other fuzzy set. Pal and Pal analyzed classical Shannon information entropy. In this paper, we try to analyze relations between fuzzy entropy and similarity measure. With the help of distance measure, we design the similarity measure. Obtained similarity measure produce fuzzy entropy based on one-to-one correspondence between distance measure and similarity measure. Fuzzy entropy, from similarity measure, is proved by verifying definition of fuzzy entropy. We also continue discussion of similarity measure from fuzzy entropy.

In the following chapter, we discuss the definition of fuzzy entropy and similarity measure of fuzzy set. We also introduce the previous obtained fuzzy entropy and similarity measure. In Chapter 3, fuzzy entropy is induced from similarity measure and vise versa. We proved the usefulness of those properties, and discussed correlation of two measures. Finally, conclusions are followed in Chapter 4.

2. Fuzzy Entropy and Similarity Measure Analysis

Fuzzy entropy represents the fuzziness of fuzzy set. Fuzziness of fuzzy set is represented through degree of ambiguity, hence the entropy is obtained from fuzzy membership function itself. Liu presented the axiomatic definitions of fuzzy entropy and similarity measure [13], and these definitions have the meaning of difference or closeness for different fuzzy membership functions. First we introduce fuzzy entropy. We design fuzzy entropy based on distance measure satisfying definition of fuzzy entropy. Notations of Liu are used in this paper [13].

2.1 Definition

[13] A real function: $e: F(X) \to R^+$ is called an entropy on $F(X)$, if e has the following properties:

(E1) $e(D)=0$, $\forall D \in P(X)$

(E2) $e([1/2]) = \max_{A \in F(X)} e(A)$

(E3) $e(A^*) \le e(A)$,for any sharpening A^* of A

(E4) $e(A)=e(A^c)$, $\forall A \in F(X)$

where $[1/2]$ is the fuzzy set in which the value of the membership function is $1/2$, $R^+ = [0, \infty)$, X is the universal set, $F(X)$ is the class of all fuzzy sets of X, $P(X)$ is the class of all crisp sets of X and D^c is the complement of D.

A lot of fuzzy entropy satisfying Definition 2.1 can be formulated. We have designed fuzzy entropy in our previous literature [1]. Now two fuzzy entropies are illustrated without proofs.

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.

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.

Exact meaning of fuzzy entropy of fuzzy set A is fuzziness of fuzzy set A with respect to crisp set. We commonly consider crisp set as A_{near} or A_{far} .In the above fuzzy entropies, one of well known Hamming distance is commonly used as distance measure between fuzzy sets A and B,

$$d(A,B)=\frac{1}{2}\sum_{i=1}^{n}\left|\mu_A(x_i)-\mu_B(x_i)\right|$$

where $X=\{x_1,x_2,\cdots x_n\}$, $|k|$ is the absolute value of k . $\mu_A(x)$ is the membership function of $A\in F(X)$.

Basically fuzzy entropy means the difference between two fuzzy membership functions. Next we will introduce the similarity measure, and it describes the degree of closeness between two fuzzy membership functions. It is also found in literature of Liu.

2.2 Definition

[13] A real function $s : F^2 \to R^+$ 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) $s(D,D^c)=0$, $\forall D\in P(X)$

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

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

With Definition 2.2, we propose the following theorem as the similarity measure.

Similarity Measure 1. For any set $A,B\in F(X)$,if d satisfies Hamming distance measure and $d(A,B)=d(A^C,B^C)$, then

$$s(A,B)=1-d\left(\left(A\cap B^C\right),[0]\right)-d\left(\left(A\cup B^C\right),[1]\right) \tag{1}$$

is similarity measure between set A and set B .

We have proposed the similarity measure that is induced from distance measure. The similarity is useful for the non interacting fuzzy membership function pair. Another similarity is also obtained, and it can be found in our previous literature [2].

Similarity Measure 2. For any set $A,B\in F(X)$ if d satisfies Hamming distance measure, then

$$s(A,B)=2-d\left(\left(A\cap B\right),[1]\right)-d\left(\left(A\cup B\right),[0]\right) \tag{2}$$

is also similarity measure between set A and set B .

To be a similarity measure, similarity (1) and (2) do not need assumption $d(A,B)=d(A^C,B^C)$. Liu also pointed out that there is an one-to-one relation between all distance measures and all similarity measures, $d+s=1$. In the next chapter, we derive similarity measure that is generated by distance measure. Furthermore entropy is derived through similarity measure by the properties of Liu. It is obvious that Hamming distance is represented as

$$d(A,B)=d\left(\left(A\cap B\right),[1]\right)-(1-d\left(\left(A\cup B\right),[0]\right) \tag{3}$$

Where $A\cap B=\min\left(\mu_A(x_i),\mu_B(x_i)\right)$ and $A\cup B=\max\left(\mu_A(x_i),\mu_B(x_i)\right)$ are satisfied. With the Proposition 3.4 of Liu [13], we generate the similarity measure or distance measure from distance measure or similarity measure [13].

Proposition 2.1 [13] There exists an one-to-one correlation between all distance measures and all similarity measures, and a distance measure d and its corresponding similarity measure s satisfy $s+d=1$. With the property of $s=1-d$, we can construct the similarity measure with distance measure d ,that is $s<d>$. From (3) it is natural to obtain following result.

$$\begin{aligned} d(A,B)&=d\left(\left(A\cap B\right),[1]\right)+d\left(\left(A\cup B\right),[0]\right)-1\\ &=1-s(A,B)\end{aligned}$$

There fore we propose the similarity measure with above expression.

$$s<d> = 2 - d\left(\left(A \cap B\right),[1]\right) - d\left(\left(A \cup B\right),[0]\right) \qquad (4)$$

This similarity measure is exactly same with (2). At this point, we verified the one-to-one relation of distance measure and similarity measure. In the next chapter, we verify that the fuzzy entropy is derived through similarity (2).

3. Entropy Derivation with Similarity Measure

Liu also suggested propositions about entropy and similarity measure. He also insisted that the entropy can be generated by similarity measure and distance measure, those are denoted by $e<s>$ and $e<d>$.

3.1 Entropy generation by similarity

Proposition 3.5 and 3.6 of reference [13] are summarized as follows.

Proposition 3.1 [13] If s is a similarity measure on F, define

$$e(A) = s\left(A, A^C\right), \quad \forall A \in F.$$

Then e is an entropy on F.

Now we check whether our similarities (1) and (2) satisfy Proposition 3.1. Proof can be obtained by checking whether

$$s(A, A^C) = 2 - d\left(\left(A \cap A^C\right),[1]\right) - d\left(\left(A \cup A^C\right),[0]\right)$$

satisfy from (E1) to (E4) of Definition 2.1.

For (E1), $\forall D \in P(X)$,

$$s(D, D^C) = 2 - d\left(\left(D \cap D^C\right),[1]\right) - d\left(\left(D \cup D^C\right),[0]\right)$$
$$= 2 - d\left([0],[1]\right) - d\left([1],[0]\right) = 0$$

(E2) represents that crisp set $1/2$ has the maximum entropy value. Therefore, the entropy $e([1/2])$ satisfies

$$s([1/2],[1/2]^C) = 2 - d\left(\left([1/2] \cap [1/2]^C\right),[1]\right)$$
$$- d\left(\left([1/2] \cup [1/2]^C\right),[0]\right)$$
$$= 2 - d\left([1/2],[1]\right) - d\left([1/2,[0]]\right)$$
$$= 2 - 1/2 - 1/2 = 1$$

In the above equation, $[1/2]^C = [1/2]$ is satisfied.

(E3) shows that the entropy of the sharpened version of fuzzy set A, $e\left(A^*\right)$, is less than or equal to $e(A)$.

$$s(A^*, A^{*C}) = 2 - d\left(\left(A^* \cap A^{*C}\right),[1]\right) - d\left(\left(A^* \cup A^{*C}\right),[0]\right)$$
$$\le 2 - d\left(\left(A \cap A^C\right),[1]\right) - d\left(\left(A \cup A^C\right),[0]\right)$$

Finally, (E4) is proved directly

$$s(A, A^C) = 2 - d\left(\left(A \cap A^C\right),[1]\right) - d\left(\left(A \cup A^C\right),[0]\right)$$
$$= 2 - d\left(\left(A^C \cap A\right),[1]\right) - d\left(\left(A^C \cup A\right),[0]\right) = s(A^C, A)$$

From the above proof, our similarity measure

$$s(A, A^C) = 2 - d\left(\left(A \cap A^C\right),[1]\right) - d\left(\left(A \cup A^C\right),[0]\right)$$

generate fuzzy entropy.

Next another similarity (1) between A and A^C

$$s(A,A^C)=1-d\left(\left(A\cap A\right),[0]\right)-d\left(\left(A\cup A\right),[1]\right)$$
$$=1-d\left(\left(A\right),[0]\right)-d\left(A,[1]\right).$$

is also satisfied and proved easily.

3.2 Relation of similarity and distance

With the property of one-to-one correspondence between similarity and distance, we have derived similarity measure with distance measure. Furthermore with the similarity measure we also obtained fuzzy entropy. For the derivation of similarity measure, $s=1-d$ is also used. If we use distance measure (3)

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

We obtain the corresponding similarity measure

$$s<d>=2-d\left(\left(A\cap B\right),[1]\right)-d\left(\left(A\cup B\right),[0]\right).$$

then this similarity is identical to (2).

From another similarity (1)

$$s(A,B)=1-d\left(\left(A\cap B^C\right),[0]\right)-d\left(\left(A\cup B^C\right),[1]\right),$$

is $d(A,B)=d\left(\left(A\cap B^C\right),[0]\right)+d\left(\left(A\cup B^C\right),[1]\right)$ satisfied ?

By the definition of distance measure of Liu [13],

$$d(A,B)=d\left(\left(A\cap B^C\right),[0]\right)+d\left(\left(A\cup B^C\right),[1]\right)$$
$$=d\left(\left(A\cap B^C\right)^C,[0]^C\right)+d\left(\left(A\cup B^C\right)^C,[1]^C\right)$$
$$=d\left(\left(A^C\cup B\right),[1]\right)+d\left(\left(A^C\cap B\right),[0]\right)$$
$$=d(B,A).$$

$$d(A,A)=d\left(\left(A\cap A^C\right),[0]\right)+d\left(\left(A\cup A^C\right),[1]\right)$$
$$=d\left([0],[0]\right)+d\left([1],[1]\right)=0.$$

For $d(A,B)=d\left(\left(A\cap B^C\right),[0]\right)+d\left(\left(A\cup B^C\right),[1]\right)$

$$\le d\left(\left(D\cap D^{CC}\right),[0]\right)+d\left(\left(D\cup D^{CC}\right),[1]\right)$$
$$=d\left(D,[0]\right)+d\left(D,[1]\right)=1.$$

Hence it is natural that distance between crisp set and its complement become maximal value. Finally,

$$d(A,B)=d\left(\left(A\cap B^C\right),[0]\right)+d\left(\left(A\cup B^C\right),[1]\right)$$
$$\le d\left(\left(A\cap C^C\right),[0]\right)+d\left(\left(A\cup C^C\right),[1]\right)$$
$$=d(A,C)$$

and

$$d(B,C)=d\left(\left(B\cap C^C\right),[0]\right)+d\left(\left(B\cup C^C\right),[1]\right)$$
$$\le d\left(\left(A\cap C^C\right),[0]\right)+d\left(\left(A\cup C^C\right),[1]\right)$$
$$=d(A,C)$$

are satisfied because of inclusion property, $A\subset B\subset C$.

4. Conclusions

We have discussed the similarity measure that is derived from distance measure. The proposed similarity usefulness is proved. Furthermore with the relation between fuzzy entropy and similarity measure, we also verified that the fuzzy entropy is induced through similarity measure. In this paper our proposed similarity measures are provided for the design of fuzzy entropy. Among the proposed similarity measure, a similarity satisfies fuzzy entropy trivially. Even though there are similarity measure satisfying similarity definition, there can exist trivial fuzzy entropy. Finally, proposed similarity measure can be applied to the general types of fuzzy membership functions.

References

[1] S.H. Lee, S.P. Cheon, and Jinho Kim, "Measure of certainty with fuzzy entropy function", LNAI, Vol. 4114, pp. 134-139, 2006.

[2] S.H. Lee, J.M. Kim, and Y.K. Choi, "Similarity measure construction using fuzzy entropy and distance measure", LNAI, Vol. 4114, pp. 952-958, 2006.

[3] Ronald R. Yager, "Monitored heavy fuzzy measures and their role in decision making under uncertainty," Fuzzy Sets and Systems, Vol. 139, No. 3, pp. 491-513, 2003.

[4] Y. Rébillé, "Decision making over necessity measures through the Choquet integral criterion", Fuzzy Sets and Systems, Vol. 157, No. 23, pp. 3025-3039, 2006.

[5] V. Sugumaran, G.R. Sabareesh and K.I. Ramachandran, "Fault diagnostics of roller bearing using kernel based neighborhood score multi-class support vector machine," Expert Systems with Applications, Vol. 34, No. 4, pp. 3090-3098, 2008.

[6] W. S. Kang and J. Y. Choi, "Domain density description for multiclass pattern classification with reduced computational load", Pattern Recognition, Vol. 41, No. 6, pp. 1997-2009, 2008.

[7] F. Y. Shih and K. Zhang, "A distance-based separator representation for pattern classification," Image and Vision Computing, Vol. 26, No. 5, pp. 667-672, 2008.

[8] S.M. Chen, "New methods for subjective mental workload assessment and fuzzy risk analysis", Cybern. Syst. : Int. J., Vol. 27, No. 5, pp. 449-472, 1996.

[9] 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, pp 551-555, 1999.

[10] H.S. Lee, "An optimal aggregation method for fuzzy opinions of group decision," Proc. 1999 IEEE Int. Conf. Systems, Man, Cybernetics, Vol. 3, pp. 314-319, 1999.

[11] 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, pp. 45-56, 2003.

[12] S. H. Lee, Y. T. Kim, S. P. Cheon and S. S. Kim, "Reliable data selection with fuzzy entropy," LNAI, Vol. 3613, pp. 203-212, 2005.

[13] X. Liu, "Entropy, distance measure and similarity measure of fuzzy sets and their relations", Fuzzy Sets and Systems, Vol. 52, pp. 305-318, 1992.

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

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

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

저자소개

Sanghyuk Lee [정회원]

- 1988. Feb. : Chungbuk National University, Korea, Electrical Engineering, (B. Eng.)
- 1991. Feb : Seoul National University, Korea, Electrical Engineering, (M.S.)
- 1998. Feb : Seoul National University, Korea, Electrical Engineering, (Ph.D.)
- 2011. Aug. ~ Present: Professor of Dept. of Electrical Engineering, XI'anJiaotong-Liverpool University, China
- E-Mail : Sangyuk.Lee@xjtlu.edu.cn

<Research Interest> : Nonlinear Control and Robustness, Automotive Engine Modeling and Dynamics, Analysis Artificial Intelligence

Yujia Zhai [정회원]

- 2001. Department of Electrical Engineering, Changchun University (Bachelor of Eng.)
- 2004. Department of Electrical Engineering and Electronics, University of Liverpool (UoL) (Master in Information and Intelligence Engineering.)
- 2009. Liverpool John Moores University (LJMU) (Ph.D. in Control Engineering.)
- E-Mail : yujia.zhai@xjtlu.edu.cn

<Research Interest> : Nonlinear Control and Robustness, Automotive Engine Modeling and Dynamics, Analysis Artificial Intelligence