

Uncertainty Measure of Pythagorean Fuzzy Sets

Xiaozhuan GAO^{a,d}, Lipeng PAN^a and Yong DENG^{a,b,c,d,1}

^a *Institute of Fundamental and Frontier Science, University of Electronic Science and Technology of China, Chengdu, 610054, China*

^b *School of Education, Shannxi Normal University, Xi'an, 710062, China*

^c *School of Knowledge Science, Japan Advanced Institute of Science and Technology, Nomi, Ishikawa 923-1211, Japan*

^d *Department of Management, Technology, and Economics, ETH Zurich, Zurich, Switzerland*

Abstract. Pythagorean fuzzy sets (PFS) can better express and handle the uncertainty information and has the more lager representation space. Hence, the reasonable and effective method to measure the uncertainty of PFS can better analyze information. From the view of Dempster-Shafer evidence theory, hesitancy degree can include the two focal elements (membership, non-membership). Hence, considering the number of focal elements for hesitancy degree to measure uncertainty is important. In addition, the difference between membership and non-membership degree plays an essential role in uncertainty measure. From the above views, the paper proposed the new uncertainty measure. Based on the new uncertainty measure, cross entropy and divergence of PFS can be presented. In addition, some numerical examples are used to explain the proposed methods by comparing other methods. Finally, the proposed divergence can be used in pattern recognition to verify its effectiveness.

Keywords. Pythagorean fuzzy sets, Dempster-Shafer evidence theory, Uncertainty measure, Cross entropy, Divergence

1. Introduction

Expressing information plays an essential role in knowledge representation. There are a lot of methodologies, such as Dempster-Shafer evidence theory [1,2], fuzzy sets [3] and other methodologies [4,5]. In those methodologies, fuzzy sets can handle the fuzziness information and have become the important methodology in decision-making and control. With the development of fuzzy sets, Atanassov proposed the intuitionistic fuzzy sets which can consider the membership, non-membership and hesitancy degree of information and is the generalization of fuzzy sets [6]. Yager expanded the intuitionistic fuzzy sets into the Pythagorean theme to propose Pythagorean fuzzy sets [7]. Pythagorean fuzzy

¹Corresponding Author: Yong Deng; E-mail: dengentropy@uestc.edu.cn.

sets (PFS) have the larger representation space than intuitionistic fuzzy sets and became the important math tool to handle uncertain information [8,9].

Due the strong space representation of PFS, how to measure the uncertainty of PFS is interesting. Thao and Smarandache proposed the new fuzzy entropy of PFS by exploiting the concept of probability [10]. Peng. et al proposed fuzzy Information Measures [11]. Xiao and Ding proposed the divergence by considering Jensen–Shannon divergence[12]. Yang. et al proposed the fuzzy entropy of PFS [13]. Uncertainty measure of PFS should consider all aspects of information, including membership, non-membership and hesitancy degree. Moreover, two PFSs with the same uncertainty of membership, non-membership and hesitancy degree maybe quite different. In other words, the difference between membership and non-membership degree should also be considered to measure the distribution of fuzziness [14]. Although there are various methods, they can not fully consider the uncertainty of PFS. Hence, how to measure the uncertainty of PFS is also an open issue.

Based on the above discussion, the paper proposed the new uncertainty measure which can consider all aspects of information as possible. Firstly, from the view of Dempster-Shafer evidence theory, the hesitant information can include the membership and non membership information which means there are two focal elements. Hence, the new uncertainty measure considers the number of focal elements. Secondly, the difference between membership and non-membership degree is computed by the score function. Based on the new uncertainty measure, the paper proposed the cross entropy and divergence of PFS. Moreover, some numerical examples can be used to explain the reasonableness of proposed methods by comparing with other methods. Finally, the new divergence can be applied to the pattern recognition to verify the effectiveness.

The structures of this article are as follows. The preliminaries of PFS and divergence are introduced in section 2. The new uncertainty measure, cross entropy and divergence of PFS are proposed in Section 3. Section 4 introduces the method of pattern recognition. In section 5, the paper can be concluded.

2. Preliminaries

In this section, we briefly recall some essential concepts of PFS and divergence.

2.1. Pythagorean Fuzzy sets

Definition 2.1. (*Pythagorean fuzzy sets*)

Let $\Theta = \{\lambda_1, \lambda_2, \dots, \lambda_n\}$ be a universe of discourse, a Pythagorean fuzzy set A in Θ can be defined [7]:

$$A = \{ \langle \lambda_i, u_A(\lambda_i), v_A(\lambda_i) \rangle \mid \lambda_i \in \Theta \} \quad (1)$$

where $0 \leq u_A(\lambda_i) \leq 1$ and $0 \leq v_A(\lambda_i) \leq 1$, which represent the membership degree and non-membership degree respectively and should satisfy $u_A(\lambda_i)^2 +$

$v_A(\lambda_i)^2 \leq 1$. Besides, the hesitancy degree can be written as follows $h_A(\lambda_i) = \sqrt{1 - (u_A(\lambda_i)^2 + v_A(\lambda_i)^2)}$ [7].

2.2. Jensen–Shannon divergence

Definition 2.2. (Jensen–Shannon divergence measure)

There are two probability distributions $A = \{a_1, a_2, \dots, a_n\}$ and $B = \{b_1, b_2, \dots, b_n\}$. The Jensen–Shannon divergence between A and B is as follows [15]

$$JS(A, B) = \frac{1}{2} [S(A, \frac{A+B}{2}) + S(B, \frac{A+B}{2})] = H(\frac{A+B}{2}) - \frac{1}{2}H(A) - \frac{1}{2}H(B) \quad (2)$$

where $H(A)$ is the Shannon entropy.

3. The proposed method

This section mainly introduces the new uncertainty measure, cross entropy and divergence of PFS.

3.1. The Proposed Uncertainty Measure

Definition 3.1. (Uncertainty measure)

Let $\Theta = \{\lambda_1, \lambda_2, \dots, \lambda_n\}$ be a universe of discourse, the PFSs in Θ are $P(\lambda_i) = \langle \lambda_i, u(\lambda_i), v(\lambda_i) \rangle$. The uncertainty measure of PFSs is defined by

$$UM = - \sum_{i=1}^{i=n} (\sum_{j=1}^{j=3} \kappa_j(\lambda_i)^2 \log(\kappa_j(\lambda_i)^2)) + \pi(\lambda_i)^2 \log(\frac{\pi(\lambda_i)^2}{3}) \quad (3)$$

where $\kappa_1(\lambda_i) = u(\lambda_i)$, $\kappa_2(\lambda_i) = v(\lambda_i)$, $\kappa_3(\lambda_i)^2 = \triangle(\lambda_i)^2 = |u(\lambda_i)^2 - v(\lambda_i)^2|$, $\pi(\lambda_i)^2 = 1 - u(\lambda_i)^2 - v(\lambda_i)^2$. In addition, 3 is explained by the power set of Dempster-Shafer evidence theory which can be computed by $2^2 - 1$.

Definition 3.2. (Cross Entropy and Divergence)

There are two PFSs P and Q in Θ , where $P = \{\langle \lambda_1, u(\lambda_1), v(\lambda_1) \rangle, \dots, \langle \lambda_n, u(\lambda_n), v(\lambda_n) \rangle\}$, $Q = \{\langle \lambda_1, u(\lambda_1), v(\lambda_1) \rangle, \dots, \langle \lambda_n, u(\lambda_n), v(\lambda_n) \rangle\}$.

The new cross entropy between P and Q is defined by

$$NCE_{PQ} = \sum_{i=1}^{i=n} (\sum_{j=1}^{j=3} \kappa_{Pj}(\lambda_i)^2 \cdot \log(\frac{1}{\kappa_{Qj}(\lambda_i)^2})) + \pi_P(\lambda_i)^2 \cdot \log(\frac{3}{\pi_Q(\lambda_i)^2}) \quad (4)$$

The new divergence between P and Q is defined by

$$ND = \frac{1}{2} \sum_{i=1}^{i=n} \sum_{j=1}^{j=4} \left[\kappa_{Pj}(\lambda_i)^2 \log \frac{2\kappa_{Pj}(\lambda_i)^2}{\kappa_{Pj}(\lambda_i)^2 + \kappa_{Qj}(\lambda_i)^2} + \kappa_{Qj}(\lambda_i)^2 \log \frac{2\kappa_{Qj}(\lambda_i)^2}{\kappa_{Pj}(\lambda_i)^2 + \kappa_{Qj}(\lambda_i)^2} \right] \quad (5)$$

where $\kappa_4(\lambda_i) = \pi(\lambda_i)$. More importantly, the 0 can be regarded as the $1 \cdot 10^{-16}$ if denominator is 0.

3.2. Numerical examples

Some numerical examples are used to explain and discuss the proposed uncertainty measure, cross entropy and divergence of PFS.

Example 1 Supposing there are two PFSs $A = \langle x, 0.45, 0.75 \rangle$, $B = \langle x, 0.55, 0.80 \rangle$, the corresponding uncertainty measures are as follows.

Table 1. Uncertainty measure of Example 1

	$E_1^{[11]}$	$E_2^{[11]}$	$E_3^{[11]}$	$E_4^{[11]}$	UM
A	0.5486	0.4706	0.6400	0.3600	1.6133
B	0.5161	0.4953	0.6625	0.4727	1.2413

Obviously, B should have the smaller uncertainty than A . Because the information of B is more certain due it has the lager membership and non membership than A . From this view, it can be seen that proposed uncertainty measure and E_1 is more reasonable than other methods. Besides, the UM can better express the difference between A and B than E_1 , which can show the effectiveness of new certainty measure.

Example 2 There are some PFSs $A = \langle x, 0.45, 0.75 \rangle$, $B = \langle x, 0.55, 0.80 \rangle$, $C = \langle x, 0.55, 0.78 \rangle$, the corresponding divergences are as follows.

$$D^{[12]}(A, B) = 0.0352, D^{[12]}(A, C) = 0.0225, D^{[12]}(B, C) = 0.0019$$
$$ND(A, B) = 0.0354, ND(A, C) = 0.0236, ND(B, C) = 0.0023$$

The D which is the divergence proposed by Xiao and ND show that B is more similar with the C than A which is reasonable.

Example 3 Supposing there are two PFSs $A = \langle x, 0.5, x \rangle$, $B = \langle x, 0.4, y \rangle$. The uncertainty, cross entropy and divergence of A, B can be discussed.

In this example, x can change from 0 to 0.8660, y can change from 0 to 0.9165. The uncertainty, cross entropy and divergence of *Example 3* are as Fig. 1.

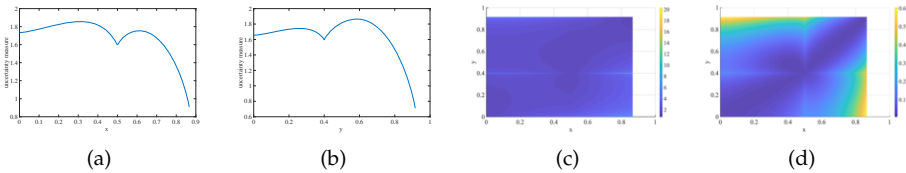


Figure 1. The uncertainty measure, cross entropy and divergence of *Example 3*. (a) uncertainty of A , (b) uncertainty of B , (c) cross entropy, (d) divergence

(a) The uncertainty of A and B can be discussed. For A , there are some conditions, as follows. (1) when x is about 0.3, the PFS is $A = \langle x, 0.5, 0.3 \rangle$. In this case, the hesitancy degree has the maximum value which has the maximum uncertainty. (2) when $x = 0.8660$, the PFS is $A = \langle x, 0.5, 0.8660 \rangle$. In this case, the hesitancy degree is 0 which has the minimum uncertainty. Next, the uncertainty

of B is analyzed. (1) when $y = 0.581$, the PFS is $B = \langle x, 0.4, 0.581 \rangle$. There is the maximum uncertainty. (2) when $y = 0.9165$, the PFS is $B = \langle x, 0.4, 0.9165 \rangle$, which has the minimum uncertainty.

(b) The cross entropy of A, B ($NCE(A, B)$) can be discussed as Fig. 1(c). (1) when $y = 0.4$, the $\Delta_B = 0$, the change of cross entropy can be shown in Fig. 2(a). It is obvious that there is the minimum cross entropy when $x = 0.5$. (2) when $x = 0.5$, the $\Delta_A = 0$. In this case, the change of cross entropy can be shown in Fig. 2(b). It can be seen that the change of cross entropy is symmetric about 0.5.

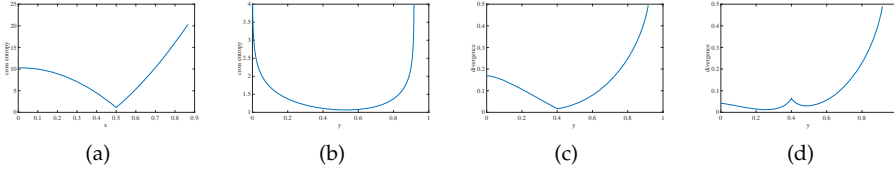


Figure 2. The cross entropy of Example. 3. (a) $y = 0.4$, (b) $x = 0.5$. The divergence of Example. 3. (a) $x = 0.5$, (b) $x = 0.3$.

(c) The divergence of A and B can be discussed, as shown in Fig. 1(d). (1) when $x = 0.5, y = 0.4, A = \langle x, 0.5, 0.5 \rangle, B = \langle x, 0.4, 0.4 \rangle$. In this case, $\Delta_A = 0, \Delta_B = 0$. Hence, the score function can not express the difference between A and B . However, hesitancy degree can help us distinguish the difference. (2) Fig. 2(c) and (d) discussed the change of divergence with y when $x = 0.5$ and $x = 0.3$. When $y = 0.4$, the divergence of $x = 0.5$ is more than $x = 0.3$, which reflects the importance of score function.

4. Application

This section introduced the method of pattern recognition based on the proposed divergence to explain its effectiveness.

4.1. Method of pattern recognition based on proposed method

Problem Statement : There are some patterns $P = \{p_1, p_2, \dots, p_m\}$ which can be expressed by PFS $p_j = \{x_i, \mu_j(x_i), v_j(x_i)\}$. Given n samples $S = \{s_1, s_2, \dots, s_n\}$ which can be expressed by PFSs $s_t = \{x_i, \mu_t(x_i), v_t(x_i)\}$. The samples should be recognised to the certain pattern. The specific steps are as follows.

Step 1 : Compute the distance between samples and patterns

The distance can be computed by using the view of Xiao [12], as follows.

$$d_{jt}(x_i) = \sqrt{\frac{1}{n} \sum ND(p_j, s_t)}$$

Step 2 : Determine the minimum value of distance

There are some distances between every sample and all patterns. The minimum distance can be determined as

$$\tilde{D}(p_\alpha, s_t) = \min_{1 \leq j \leq m} D(p_j, s_t)$$

Step 3 : Obtain the results of recognition
The sample can be classified as the pattern with minimum distance.

$$s_t \leftarrow p_\alpha$$

4.2. Application of medical diagnosis

Medical diagnosis can be used to explain the proposed pattern recognition which can be introduced simply. There are four patients $P = \{p_1, p_2, p_3, p_4\}$ with the five symptoms $S = \{s_1 : \textit{Temperature}, s_2 : \textit{Headache}, s_3 : \textit{Stomachpain}, s_4 : \textit{Cough}, s_5 : \textit{Chestpain}\}$. The possible diagnosis $D = \{d_1 : \textit{Viral fever}, d_2 : \textit{Malaria}, d_3 : \textit{Typhoid}, d_4 : \textit{Stomachproblem}, d_5 : \textit{Chestproblem}\}$. The more related information can be obtained in [12]. The results of medical diagnosis by using proposed method are as Tab. 2.

Table 2. The results of pattern recognition

	D_1	D_2	D_3	D_4	D_5	Results
p_1	0.3055	0.2501	0.3885	0.4776	0.5337	D_2
p_2	0.4388	0.5223	0.4178	0.1864	0.4734	D_4
p_3	0.3878	0.4223	0.3888	0.4797	0.5068	D_3
p_2	0.3296	0.4030	0.4528	0.4519	0.5466	D_1

To further explore the reasonableness and effectiveness of proposed method, there are some comparisons with other methods as Tab. 3. It can be seen that the results of proposed method are similar to those of most methods which proves the proposed method is practical in dealing with medical diagnosis.

Table 3. The results of different methods

	Own ^[16]	Deetal. ^[17]	Szmidtetal. ^[18]	Weietal. ^[19]	Mondaletal. ^[20]	Xiao ^[12]	Proposed
p_1	D_2	D_2	D_2	D_2	D_2	D_2	D_2
p_2	D_4	D_4	D_4	D_4	D_4	D_4	D_4
p_3	D_3	D_2	D_3	D_3	D_3	D_3	D_3
p_4	D_2	D_2	D_1	D_1	D_1	D_1	D_1

5. Conclusion

In Pythagorean fuzzy sets (PFS), how to measure uncertainty is important and an open issue. PFS can include the membership, non-membership and hesitancy degree. Hence considering them comprehensively can help us better analyze the uncertainty of PFS. The paper proposed the uncertainty measure of PFS by considering the membership, non-membership, hesitancy degree and difference

between membership and non-membership degree. Besides, from the view of Dempster-Shafer evidence theory, the number of focal elements of hesitancy information which can include the membership and non-membership information is considered. Based on the proposed uncertainty measure, the cross entropy and divergence of PFS can be also presented. In addition, the new uncertainty measure can be compared with the other methods, which shows the new uncertainty measure can better reflect the difference between PFSs and is reasonable. Finally, the medical diagnosis can be applied to the method of proposed the pattern recognition based on the new divergence. The results are used to compare with other methods to verify reasonableness and effectiveness of the new method.

References

- [1] A. P. Dempster, Upper and lower probabilities induced by a multivalued mapping, *The Annals of Mathematical Statistics* 38 (2) (1967) 325–339.
- [2] G. Shafer, *A mathematical theory of evidence*, Vol. 42, Princeton university press, 1976.
- [3] L. A. Zadeh, Fuzzy sets, in: *Fuzzy sets, fuzzy logic, and fuzzy systems: selected papers by Lotfi A Zadeh*, World Scientific, 1996, pp. 394–432.
- [4] K. Ullah, T. Mahmood, N. Jan, Similarity measures for t-spherical fuzzy sets with applications in pattern recognition, *Symmetry* 10 (6) (2018) 193.
- [5] X. Gao, Y. Deng, Quantum model of mass function, *International Journal of Intelligent Systems* 35 (2) (2020) 267–282.
- [6] K. T. Atanassov, *On intuitionistic fuzzy sets theory*, Vol. 283, Springer, 2012.
- [7] R. R. Yager, Pythagorean fuzzy subsets, in: *2013 joint IFSA world congress and NAFIPS annual meeting (IFSA/NAFIPS)*, IEEE, 2013, pp. 57–61.
- [8] H. Garg, Linguistic pythagorean fuzzy sets and its applications in multiattribute decision-making process, *International Journal of Intelligent Systems* 33 (6) (2018) 1234–1263.
- [9] L. Pan, X. Gao, Y. Deng, K. H. Cheong, The constrained pythagorean fuzzy sets and its similarity measure, *IEEE Transactions on Fuzzy Systems* (2021) DOI: 10.1109/TFUZZ.2021.3052559.
- [10] N. X. Thao, F. Smarandache, A new fuzzy entropy on pythagorean fuzzy sets, *Journal of Intelligent & Fuzzy Systems* 37 (1) (2019) 1065–1074.
- [11] X. Peng, H. Yuan, Y. Yang, Pythagorean fuzzy information measures and their applications, *International Journal of Intelligent Systems* 32 (10) (2017) 991–1029.
- [12] F. Xiao, W. Ding, Divergence measure of pythagorean fuzzy sets and its application in medical diagnosis, *Applied Soft Computing* 79 (2019) 254–267.
- [13] M.-S. Yang, Z. Hussain, Fuzzy entropy for pythagorean fuzzy sets with application to multi-criterion decision making, *Complexity* 2018 (2018).
- [14] Y. Song, Q. Fu, Y.-F. Wang, X. Wang, Divergence-based cross entropy and uncertainty measures of atanassov's intuitionistic fuzzy sets with their application in decision making, *Applied Soft Computing* 84 (2019) 105703.
- [15] J. Lin, Divergence measures based on the shannon entropy, *IEEE Transactions on Information theory* 37 (1) (1991) 145–151.
- [16] C.-M. Own, Switching between type-2 fuzzy sets and intuitionistic fuzzy sets: an application in medical diagnosis, *Applied Intelligence* 31 (3) (2009) 283.
- [17] S. K. De, R. Biswas, A. R. Roy, An application of intuitionistic fuzzy sets in medical diagnosis, *Fuzzy sets and Systems* 117 (2) (2001) 209–213.
- [18] E. Szmidt, J. Kacprzyk, Intuitionistic fuzzy sets in intelligent data analysis for medical diagnosis, in: *International conference on computational science*, Springer, 2001, pp. 263–271.
- [19] C.-P. Wei, P. Wang, Y.-Z. Zhang, Entropy, similarity measure of interval-valued intuitionistic fuzzy sets and their applications, *Information Sciences* 181 (19) (2011) 4273–4286.
- [20] K. Mondal, S. Pramanik, Intuitionistic fuzzy similarity measure based on tangent function and its application to multi-attribute decision making, *Global journal of advanced research* 2 (2) (2015) 464–471.