



Improved cosine similarity measures for q -Rung orthopair fuzzy sets

Mehmet Ünver¹

¹ Ankara University

Funding: No specific funding was received for this work.

Potential competing interests: No potential competing interests to declare.

Abstract

In this short correspondence, we introduce some novel cosine similarity measures tailored for q -rung orthopair fuzzy sets (q -ROFSs), which capture both the direction and magnitude aspects of fuzzy set representations. Traditional cosine similarity measures focus solely on the direction (cosine of the angle) between vectors, neglecting the crucial information embedded in the lengths of these vectors. To address this limitation, we propose some improved cosine similarity measures, which extend the conventional cosine similarity by incorporating a length difference control term. These measures not only outperform traditional cosine similarity measures for q -ROFSs but also improve the existing cosine similarity measures for intuitionistic fuzzy sets and Pythagorean fuzzy sets, making them a valuable addition to the fuzzy set cosine similarity measures. These similarity measures are defined as the average and Choquet integral of two components: the first component, $\text{Cos}_{A,B}(x_j)$, quantifies the cosine similarity between q -ROFS A and B at each element x_j . The second component, $L_{A,B}(x_j)$, represents the difference in lengths between the vector representations of A and B at the same element x_j . This length-difference term ensures that the measures are sensitive to variations in both direction and magnitude, making them particularly suitable for applications where both aspects are significant. The measure derived through the Choquet integral also takes into account the interaction among the elements, thereby enhancing the sensitivity of solutions in various applications.

Mehmet Ünver

Ankara University, Faculty of Science,

Department of Mathematics, 06100 Ankara Türkiye

Email address: munver@ankara.edu.tr

Key words and phrases: q -rung orthopair fuzzy set, improved cosine similarity measure, Choquet integral.

1. Introduction

Fuzzy sets, as introduced by Zadeh [1], have proven to be highly effective in handling uncertainty and representing partial membership within a set. These sets are characterized by their membership functions. Extending this concept, Atanassov [2] introduced the theory of intuitionistic fuzzy sets (IFS), which includes a membership function $\mu_A: X \rightarrow [0, 1]$ and a non-membership function $\nu_A: X \rightarrow [0, 1]$, subject to the constraint $\mu_A(x) + \nu_A(x) \leq 1$ for all $x \in X$. Subsequently, Yager [3] introduced Pythagorean fuzzy sets (PFS), where the membership and non-membership functions satisfy the condition $\mu^2 A(x) + \nu^2 A(x) \leq 1$. When an element exhibits a membership degree of 0.6 alongside a non-membership degree of 0.5, this situation is valid within the context of PFSs, whereas it does not conform to the criteria for IFSs. However, when the membership degree is 0.7 and the non-membership degree is 0.9, such a case is valid for neither IFSs nor PFSs. This observation underscores the need for a further extension of the PFS concept. To address this need, Yager [4] extended the concept of PFS to the concept of q -rung orthopair fuzzy set (value) (q -ROFS(V)), introducing the condition $\mu_A^q(x) + \nu_A^q(x) \leq 1$ for each $x \in X$, where $q \geq 1$. These q -ROFSs have proven to be highly useful in decision-making problems due to their versatile structure, as evidenced by their frequent application in various contexts (e.g., [5][6][7][8]).

A similarity measure serves as a valuable tool for assessing the resemblance between two mathematical objects. Among these similarity measures, cosine similarity measures stand out as a specific variant. Cosine similarity measures are employed to quantify the similarity between two fuzzy sets by considering the cosine of the angle between the vector representations of their reciprocal components, and this concept has found successful applications in fuzzy set theory as well [9][10][11][12][13][14][15]. In many scenarios, the magnitude or length of a fuzzy set representation can be just as crucial as its orientation. For instance, when dealing with multi-criteria decision-making or image processing, the lengths of vector representations of the components of the fuzzy sets can convey significant information about the intensity or strength of membership. Traditional cosine similarity measures fall short in such situations, as they are incapable of distinguishing between fuzzy sets with similar orientations but significantly different magnitudes. To address this deficiency, we propose some improved cosine similarity measures, tailored specifically for q -ROFSs. Our measures combine the traditional cosine similarity with a length difference control term to provide a more comprehensive assessment of similarity. The q -

ROFSs under consideration here are a versatile and expressive extension of classical fuzzy sets that can model complex and imprecise relationships more accurately. In addition to its applicability to q -ROFSs, it's worth noting that our novel cosine similarity measures extend their benefits to IFSs and PFSs, which share similarities with q -ROFSs in terms of the need to capture both direction and magnitude aspects. By introducing a length control term, these novel similarity measures enhance the assessment of similarity not only for q -ROFSs but also for IFSs and PFSs, making them a versatile and robust measure for various fuzzy set representations. Furthermore, the cosine similarity provided by the Choquet integral takes into account the interactions among elements, enhancing its sensitivity to variations and nuances in the data.

Some main contributions of the paper are listed below:

- The paper introduces some novel cosine similarity measures designed specifically for q -ROFSs. These measures address the limitations of traditional cosine similarity by incorporating a length difference control term.
- These measures offer a balanced assessment of similarity by simultaneously capturing both the direction, $\text{Cos}_{A,B}$, and the magnitude, $L_{A,B}$, aspects of q -ROFSs. This enhancement is crucial in applications where both factors play a significant role.
- The paper emphasizes that these measures are not limited to q -ROFSs but also improve similarity assessment for IFSs and PFSs. This broadens the scope of their utility across different fuzzy set representations.
- The utilization of the Choquet integral is demonstrated in the introduction of a novel cosine similarity measure, allowing for a more comprehensive assessment of similarity that accounts for element interactions, particularly in scenarios where interdependence among elements is of paramount importance.

2. Preliminaries

In this section, we revisit fundamental definitions employed within the scope of this paper. Following that, we proceed to introduce the anticipated enhanced cosine similarity measures tailored for q -ROFSs. In this section, we maintain the assumption that $X = \{x_1, \dots, x_n\}$ represents a finite set and $w = (w_1, \dots, w_n)$ is a weight vector, where $w_j \in [0, 1]$ for all

$$j = 1, \dots, n \text{ and } \sum_{j=1}^n w_j = 1.$$

Definition 1. [4] Let $q \geq 1$. A q -ROFS A in X is given by

$$A = \left\{ \left(x_j, \mu_A(x_j), \nu_A(x_j) \right) : j = 1, \dots, n \right\}$$

where $\mu_A, \nu_A: X \rightarrow [0, 1]$ are membership and non-membership functions, respectively, satisfying

$$\mu_A^q(x_j) + \nu_A^q(x_j) \leq 1.$$

When q takes on the value of 2, the fuzzy set is referred to as a PFS^[3], and when q equals 1, it is denoted as an IFS^[2].

In their work, Gerstenkorn and Manko^[16] introduced a correlation coefficient for IFSs. This concept serves as a foundation and motivation for Ye's cosine similarity measure^[15].

Definition 2.^[16] Let A and B be two IFSs. A correlation coefficient between A and B is defined by

$$k(A, B) := \frac{\sum_{j=1}^n (\mu_A(x_j)\mu_B(x_j) + \nu_A(x_j)\nu_B(x_j))}{\sqrt{\sum_{j=1}^n (\mu_A^2(x_j) + \nu_A^2(x_j)) \sum_{j=1}^n (\mu_B^2(x_j) + \nu_B^2(x_j))}}$$

In the work by Ye^[15], a cosine similarity measure and a weighted cosine similarity measure for IFSs were introduced as follows.

Definition 3.^[15] Let A and B be two IFSs. A cosine similarity measure between A and B is defined by

$$C_{IFS}(A, B) := \frac{1}{n} \sum_{j=1}^n \frac{\mu_A(x_j)\mu_B(x_j) + \nu_A(x_j)\nu_B(x_j)}{\sqrt{\mu_A^2(x_j) + \nu_A^2(x_j)} \sqrt{\mu_B^2(x_j) + \nu_B^2(x_j)}}$$

For $n = 1$, the cosine similarity measure C_{IFS} is equivalent to the correlation coefficient k .

Definition 4.^[15] Let A and B be two IFSs. A weighted cosine similarity measure between A and B is defined by

$$W_{IFS}(A, B) := \sum_{j=1}^n w_j \frac{\mu_A(x_j)\mu_B(x_j) + \nu_A(x_j)\nu_B(x_j)}{\sqrt{\mu_A^2(x_j) + \nu_A^2(x_j)} \sqrt{\mu_B^2(x_j) + \nu_B^2(x_j)}}$$

In particular, if $w = (1/n, 1/n, \dots, 1/n)$, then W_{IFS} is reduced to C_{IFS} .

Liu et al.^[17] extended the concepts of C_{IFS} and W_{IFS} to encompass q -ROFSs. We now revisit the weighted one.

Definition 5.^[17] Let A and B be two q -ROFSs. A weighted cosine similarity measure between A and B is defined by

$$WC_{qROF}(A, B) := \sum_{j=1}^n w_j \frac{\mu_A^q(x_j)\mu_B^q(x_j) + \nu_A^q(x_j)\nu_B^q(x_j)}{\sqrt{\mu_A^{2q}(x_j) + \nu_A^{2q}(x_j)} \sqrt{\mu_B^{2q}(x_j) + \nu_B^{2q}(x_j)}}$$

3. Improved cosine similarity measures

While the existing cosine similarity measures have been valuable in the assessment of similarity among fuzzy sets, they often fall short of capturing the complete picture. These traditional measures primarily focus on the cosine of the angle between vector representations, overlooking the significance of vector magnitudes. This limitation can lead to inaccurate similarity assessments, especially when the lengths of vector representations vary significantly. In light of these

shortcomings, we present our enhanced cosine similarity measures for q -ROFSs. These measures not only address the deficiencies of traditional cosine similarity but also introduce a length difference control term to provide a more comprehensive and accurate evaluation of similarity. Our approach aims to bridge the gap left by previous measures, ensuring a robust and versatile tool for similarity assessment in various applications.

Definition 6. Let A and B be two q -ROFSs. An improved cosine similarity measure between A and B is defined by

$$ICSM_q(A, B) := \frac{1}{2^n} \sum_{j=1}^n \left(\text{Cos}_{A, B^{x_j}} + L_{A, B^{x_j}} \right)$$

where

$$\text{Cos}_{A, B^{x_j}} := \frac{\mu_A^q(x_j)\mu_B^q(x_j) + \nu_A^q(x_j)\nu_B^q(x_j)}{\sqrt{\mu_A^{2q}(x_j) + \nu_A^{2q}(x_j)}\sqrt{\mu_B^{2q}(x_j) + \nu_B^{2q}(x_j)}}$$

and

$$L_{A, B^{x_j}} := 1 - \left| \sqrt{\mu_A^{2q}(x_j) + \nu_A^{2q}(x_j)} - \sqrt{\mu_B^{2q}(x_j) + \nu_B^{2q}(x_j)} \right|.$$

In order to illustrate the effectiveness of our novel similarity measure, let us consider a practical example involving 3-ROFSs A and B . We compare the similarity results obtained using both the traditional cosine similarity measure WC_{qROF} and our enhanced measure $ICSM_q$.

Example 1. Let $X = \{x_1, x_2\}$. Consider 3-ROFSs A and B defined as follows:

$$A = \left\{ \langle x_1, 0.2, 0.4 \rangle, \langle x_2, 0.3, 0.6 \rangle \right\}$$

and

$$B = \left\{ \langle x_1, 0.4, 0.8 \rangle, \langle x_2, 0.1, 0.2 \rangle \right\}.$$

Then we have:

$$\begin{aligned} WC_{qROF}(A, B) &= w_1 \frac{0.2^3 0.4^3 + 0.4^3 0.8^3}{\sqrt{0.2^6 + 0.4^6} \sqrt{0.4^6 + 0.8^6}} + w_2 \frac{0.3^3 0.1^3 + 0.6^3 0.2^3}{\sqrt{0.3^6 + 0.6^6} \sqrt{0.1^6 + 0.2^6}} \\ &= w_1 + w_2 = 1 \end{aligned}$$

for any weight vector (w_1, w_2) such that $w_1 + w_2 = 1$. It is evident that the cosine similarity measure WC_{qROF} yields a similarity of 1 for A and B despite their substantial differences. Now, let's calculate the similarity between A and B using $ICSM_q$:

$$\begin{aligned}
 ICSM_q(A, B) &= \frac{1}{2} + \frac{1}{4} \left(2 - \left| \sqrt{0.2^6 + 0.4^6} - \sqrt{0.4^6 + 0.8^6} \right| - \left| \sqrt{0.3^6 + 0.6^6} - \sqrt{0.1^6 + 0.2^6} \right| \right) \\
 &= 0.83472.
 \end{aligned}$$

Now, considering the result obtained using our novel similarity measure, $ICSM_q$, we find that the similarity between A and B is approximately 0.83472. It's worth noting that the improved cosine similarity measure provides a more rational and meaningful result.

Remark 1. In Example 1, the results of the similarity measures highlight an interesting observation. While WC_{qROF} assigns a perfect similarity score of 1 to A and B , indicating a high degree of similarity, $ICSM_q$ yields a value of 0.83472, suggesting a more nuanced assessment. This discrepancy arises from $ICSM_q$'s consideration of both direction and magnitude aspects. In this case, the lengths of vector representations play a significant role in determining the similarity, leading to a more comprehensive evaluation.

We now introduce a weighted variant of $ICSM_q$, denoted as $IWCSM_q$.

Definition 7. Let A and B be two q -ROFSs. An improved weighted cosine similarity measure between A and B is defined by

$$IWCSM_q(A, B) := \frac{1}{2} \sum_{j=1}^n w_j \left(\text{Cos}_{A, B^{X_j}} + L_{A, B^{X_j}} \right).$$

This newly introduced measure combines the benefits of $ICSM_q$ with weighted considerations, providing a more versatile similarity assessment for q -ROFSs.

Finally, we present an additional enhanced cosine similarity measure employing fuzzy measure theory. The classical versions of cosine similarity measures within the framework of fuzzy set theory are elaborated in [18]. Let us recall some basic definitions.

Definition 8. [19] Let $P(X)$ be the power set of X . If

- i.) $\sigma(\emptyset) = 0$,
- ii.) $\sigma(X) = 1$,
- iii.) $\sigma(U) \leq \sigma(V)$ for any $A, B \subset X$ such that $U \subseteq V$,

then the set function $\sigma: P(X) \rightarrow [0, 1]$ is called a fuzzy measure on X .

Definition 9. [19] Let σ be a fuzzy measure on X . The Choquet integral of a function $f: X \rightarrow [0, 1]$ with respect to σ is defined by

$$\int (C) \int_X f d\sigma := \sum_{k=1}^n \left(f(x_{(k)}) - f(x_{(k-1)}) \right) \sigma(E_{(k)}),$$

where the sequence $\{x_{(k)}\}_{k=0}^n$ is a permutation of the sequence $\{x_k\}_{k=0}^n$ such that $0 := f(x_{(0)}) \leq f(x_{(1)}) \leq f(x_{(2)}) \leq \dots \leq f(x_{(n)})$ and $E_{(k)} := \{x_{(k)}, x_{(k+1)}, \dots, x_{(n)}\}$.

Now, we are prepared to introduce the ultimate enhanced cosine similarity measure.

Definition 10. Let σ be a fuzzy measure on X and let A and B be two q -ROFSs. An improved Choquet cosine similarity measure between A and B is defined by

$$ICCSM_q(A, B) := (C) \int_X f_{A, B} d\sigma$$

where $f_{A, B}(x_j) = \frac{1}{2} \left(\text{Cos}_{A, B} x_j + L_{A, B} x_j \right)$ for any $j = 1, \dots, n$.

The enhancement of the cosine similarity measure by considering lengths and the heightened sensitivity afforded by the Choquet integral collectively establish $ICCSM_q$ as a significant tool within the realm of fuzzy decision-making theory.

4. Conclusion

In this paper, we have introduced an array of enhanced cosine similarity measures tailored for q -rung orthopair fuzzy sets. These novel measures address the limitations of traditional cosine similarity measures by simultaneously considering both the direction and magnitude aspects of q -rung orthopair fuzzy sets. Specifically:

- The *Arithmetic Mean-based Cosine Similarity Measure* employs the traditional cosine similarity formula enhanced with a length difference control term, offering a more comprehensive assessment of similarity.
- The *Weighted Arithmetic Mean-based Cosine Similarity Measure* extends the concept further by introducing a weight vector, allowing for customized emphasis on individual components and making it adaptable to diverse applications.
- The *Choquet Integral-based Cosine Similarity Measure* leverages the Choquet integral to consider element interactions, enhancing sensitivity and applicability in scenarios where interdependence among elements is critical.

Through theoretical analysis and examples, we have demonstrated the advantages of these measures in providing a comprehensive and accurate assessment of similarity. Unlike traditional measures, these enhanced measures distinguish between fuzzy sets with varying vector lengths, making them invaluable tools in real-world applications where such distinctions matter significantly.

Furthermore, we emphasize that the concept of cosine similarity measures, which has gained substantial attention, as evidenced by its citation count of 4409 times in 2022 according to the Web of Science (see Figure 1), now gains an additional dimension with the introduction of these novel measures. We anticipate that our contributions will not only enhance the existing body of research but also attract further citations and exploration in the field of fuzzy set theory.

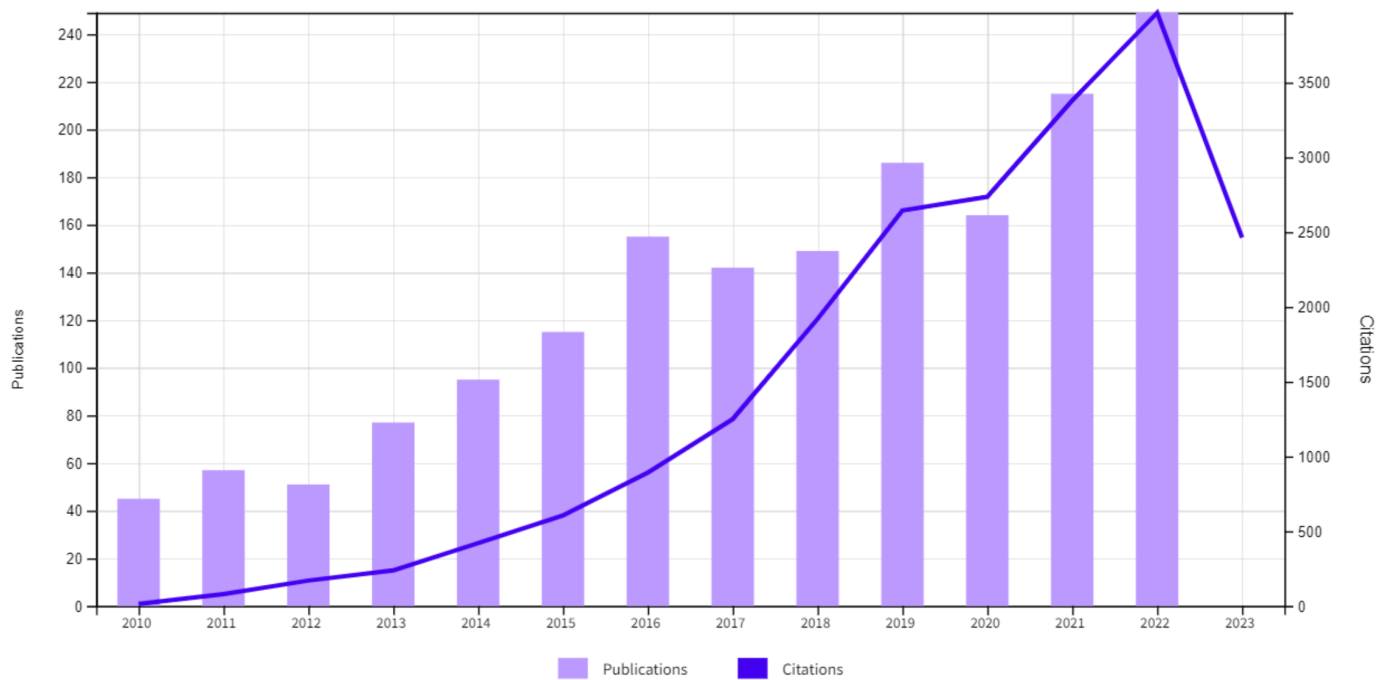


Figure 1. Citation Trend for Cosine Similarity Measures from Web of Science.

References

- ^aZadeh, L. A., (1965). *Fuzzy sets. Information and Control.* 8(3), 338-353.
- ^{a, b}Atanassov, K., (1986). *Intuitionistic fuzzy sets. Fuzzy Sets and Systems.* 20(1), 87-96.
- ^{a, b}Yager, R. R., (2013). *Pythagorean fuzzy subsets. Proceeding of The Joint IFSA World Congress and NAFIPS Annual Meeting. Edmonton, Canada* 57-61.
- ^{a, b}Yager RR (2017) *Generalized orthopair fuzzy sets, IEEE Transactions on Fuzzy Systems.* 25(5).
- ^aGarg, H., (2021). *CN-q-ROFS: Connection number-based q-rung orthopair fuzzy set and their application to decision-making process. International Journal of Intelligent Systems,* 36, 3106-3143.
- ^aLiu, P. and Wang, P., (2018). *Some q-rung orthopair fuzzy aggregation operators and their applications to multiple-attribute decision making. International Journal of Intelligent Systems.* 33:259-280.
- ^aUnver, M., and Olgun, M., (2023). *Continuous Function Valued q-Rung Orthopair Fuzzy Sets and an Extended TOPSIS. International Journal of Fuzzy Systems,* doi.org/10.1007/s40815-023-01501-5.
- ^aWang, J., Zhang, R., Zhu, X., Zhou, Z., Shang, X., and Li, W., (2019). *Some q-rung orthopair fuzzy Muirhead means with their application to multiattribute group decision making. Journal of Intelligent and Fuzzy Systems.* 36(2), 1599-1161.
- ^aGarg, H., Olgun, M., Turkarslan, E., and Unver, M., (2022). *A Choquet Integral Based Cosine Similarity Measure for Interval-Valued Intuitionistic Fuzzy Sets and an Application to Pattern Recognition. Lobachevskii Journal of Mathematics,* 43(9), 2444-2452.
- ^aGupta, P., and Tiwari, P., (2016). *Measures of cosine similarity intended for fuzzy sets, intuitionistic and interval-valued intuitionistic fuzzy sets with application in medical diagnoses. In 2016 3rd International Conference on*

Computing for Sustainable Global Development (INDIACom), 1846-1849, IEEE.

11. ^a Kirisci, M., (2023). *New cosine similarity and distance measures for Fermatean fuzzy sets and TOPSIS approach. Knowledge and Information Systems, 65(2), 855-868.*
12. ^a Lahitani, A. R., Permanasari, A. E., and Setiawan, N. A., (2016). *Cosine similarity to determine similarity measure: Study case in online essay assessment. In 2016 4th International Conference on Cyber and IT Service Management, 1-6, IEEE.*
13. ^a Unver, M., Olgun, M., and Turkarslan, E., (2022). *Cosine and cotangent similarity measures based on Choquet integral for Spherical fuzzy sets and applications to pattern recognition. Journal of Computational and Cognitive Engineering, 1(1), 21-31.*
14. ^a Xia, P., Zhang, L., and Li, F., (2015). *Learning similarity with cosine similarity ensemble. Information sciences, 307, 39-52.*
15. ^{a, b, c, d, e} Ye, J. (2011). *Cosine similarity measures for intuitionistic fuzzy sets and their applications. Mathematical and computer modelling, 53(1-2), 91-97.*
16. ^{a, b} Gerstenkorn, T., and Manko, J., (1991). *Correlation of intuitionistic fuzzy sets. Fuzzy sets and systems, 44(1), 39-43.*
17. ^{a, b} Liu, D., Chen, X., and Peng, D., (2019). *Some cosine similarity measures and distance measures between q-rung orthopair fuzzy sets. International Journal of Intelligent Systems, 34(7), 1572-1587.*
18. ^a Olgun, M., Turkarslan, E., Unver, M., and Ye, J. (2021). *A cosine similarity measure based on the choquet integral for intuitionistic fuzzy sets and its applications to pattern recognition. Informatica, 32(4), 849-864.*
19. ^{a, b} Choquet, G. (1953). *Theory of capacities. In Annales de l'institut Fourier, 5, 131-295.*