Proceedings of Machine Learning Research vol 245:1–13, 2024 2024 International Conference on Machine Learning and Intelligent Computing

A Multi-attribute Large Group Decision-making Method Based on Interval-valued Pythagorean Fuzzy Number

Xinshang You
School of Economics and Management, Hebei University of Science and Technology, Shijiazhuang 050000, China

Zhe Zhou* ZHOUZHE0018@163.COM
School of Economics and Management, Hebei University of Science and Technology, Shijiazhuang 050000, China

Editors: Nianyin Zeng and Ram Bilas Pachori

Abstract

Firstly, a large group clustering algorithm based on data similarity is proposed, which can set different thresholds to cluster the decision results of expert groups. Secondly, the interval-valued pythagorean fuzzy number (IVPFN) hesitancy accuracy function and hesitancy score function are proposed. And based on the consideration of the distance between centroids and rectangular area in the geometric meaning of IVPFN, a calculation formula is proposed to distinguish different IVPFNs. Thirdly, use the above formula to construct a weight calculation model for evaluation criteria with adjustment coefficients. The decision matrix is weighted and its relative distance from the positive and negative ideal solutions is calculated to produce the final ranking. Finally, the cultural tourism project decision-making problem is analyzed as an arithmetic example and compared with the methods in the literature of other related fields to illustrate the rationality and scientificity of this paper.

Keywords: multi-attribute large group decision-making, clustering algorithm, interval-valued pythagorean fuzzy number, cultural tourism project

1. Introduction

Influenced by the complexity of the decision-making environment and the level of decision-makers, Pythagorean fuzzy number (PFN) was proposed and used for describing decision-making information (Yager, 2014). It has strong flexibility and a large description range, attracting many scholars to invest in decision-making related research, including emergency events (Zhao et al., 2023), venture capital project (Chang et al., 2022) evaluation of scheduling results (Li et al., 2022) and so on. Subsequently, the proposal of IVPFN extended it to the field of interval numbers, which can describe more fuzzy and complex information (Zhang, 2016). The research results related to IVPFN mainly focus on two aspects: synthesis operators (Li et al., 2019, 2021; Ibrahim et al., 2023) and method extensions (Yin et al., 2022; Ayyildiz et al., 2023; Yüksel and Dinçer, 2022; Al-Barakati et al., 2022), driving the development of the decision-making field greatly. With the expansion of the number of people involved in decision-making problems, some scholars have extended the content related to fuzzy numbers to the field of large group decision-making (Zhou et al., 2022; Liu et al., 2024). However, there are relatively few literature combining with IVPFN.

Therefore, a large group clustering algorithm based on data similarity is proposed, which is able to realize clustering of expert groups based on evaluation results under different thresholds. Subsequently, hesitant score function and hesitant accuracy function that can distinguish different all IVPFNs are proposed. Combining two new functions, a comprehensive differentiation formula for IVPFN is constructed, and its related properties are proved. An evaluation standard weight calculation model is constructed based on the newly proposed function and differentiation formula mentioned above. Finally, combining the idea of TOPSIS method, the final score and ranking can be given. To demonstrate the effectiveness and rationality of the algorithm, a cultural tourism project decision-making problem is introduced.

2. Related basic knowledge

2.1. Basic concepts

Definition 1 (Zhang, 2016). Let be a thesis, then an interval-valued Pythagorean fuzzy set (IVPFS) in the thesis can be denoted as $\bar{P} = \left\{\langle x, \bar{\mu}_p(x), \bar{\nu}_p(x)\rangle \,|x \in \bar{X}\right\}$, where: for any $x \in X$ in the set $\bar{P}$, the interval-valued mapping $\bar{\mu}_p : \bar{X} \to [0,1]$ such that $x \in \bar{X} \mapsto \bar{\mu}_p(x) = \left[\mu_p^L(x), \mu_p^U(x)\right] \subseteq [0,1]$, then $\bar{\mu}_p$ denotes its interval-valued affiliation function, and $\bar{\mu}_p(x)$ denotes its affiliation interval, interval-valued mapping $\bar{\nu}_p : \bar{X} \to [0,1]$ such that $x \in \bar{X} \mapsto \bar{\nu}_p(x) = \left[\nu_p^L(x), \nu_p^U(x)\right]$, then $\bar{\nu}_p$ denotes its interval-valued unaffiliated function and $\bar{\nu}_p(x)$ denotes its unaffiliated degree interval; $\pi_p = \left[\pi_p^L(x), \pi_p^U(x)\right] = \left[\sqrt{1-\left(\mu_p^U(x)\right)^2-\left(\nu_p^U(x)\right)^2}, \sqrt{1-\left(\mu_p^L(x)\right)^2-\left(\nu_p^L(x)\right)^2}\right]$ denotes the hesitancy interval; $\forall x \in \bar{P}$, $0 \le \mu_p^L(x) \le \mu_p^U(x) \le 1$, $0 \le \nu_p^L(x) \le \nu_p^U(x) \le 1$, $0 \le \left(\mu_p^U(x)\right)^2 + \left(\nu_p^U(x)\right)^2 \le 1$. Any element $\left\langle\left[\mu_p^L(x), \mu_p^U(x)\right], \left[\nu_p^L(x), \nu_p^U(x)\right]\right\rangle$ in an IVPFS is called an IVPFN, abbreviated as $\bar{P} = \left\langle\left[\mu_p^L, \mu_p^U\right], \left[\nu_p^L, \nu_p^U\right]\right\rangle$.

Definition 2 (Liang et al., 2015). Let $\bar{P} = \left\langle\left[\mu_p^L, \mu_p^U\right], \left[\nu_p^L, \nu_p^U\right]\right\rangle$ be any IVPFN, and its score function and accuracy function are Eq. (1) and (2), respectively.

$$S\left(\bar{P}\right) = \frac{1}{2}\left(\left(\mu_p^U\right)^2 - \left(\nu_p^U\right)^2 + \left(\mu_p^L\right)^2 - \left(\nu_p^L\right)^2\right) \tag{1}$$

$$H\left(\bar{P}\right) = \frac{1}{2}\left(\left(\mu_p^U\right)^2 + \left(\nu_p^U\right)^2 + \left(\mu_p^L\right)^2 + \left(\nu_p^L\right)^2\right) \tag{2}$$

Definition 3 (Xu, 2007). Let $a_1 = \left\langle\left[\mu_{a_1}^L, \mu_{a_1}^U\right], \left[\nu_{a_1}^L, \nu_{a_1}^U\right]\right\rangle$ and $a_2 = \left\langle\left[\mu_{a_2}^L, \mu_{a_2}^U\right], \left[\nu_{a_2}^L, \nu_{a_2}^U\right]\right\rangle$ be any two IVPFNs, and their sizes can be compared by calculating score functions and accuracy functions. If $S(a_1) > S(a_2)$, then $a_1 > a_2$; if $S(a_1) < S(a_2)$, then $a_1 < a_2$; if $S(a_1) = S(a_2)$, then compare their accuracy functions; if $H(a_1) > H(a_2)$, then $a_1 > a_2$; if $H(a_1) < H(a_2)$, then $a_1 < a_2$; if $H(a_1) = H(a_2)$, then $a_1 = a_2$.

However, there are still some extreme cases that cannot be distinguished by the above formula, such as $\bar{P}_1 = \langle[0.3, 0.6], [0.2, 0.7]\rangle$ and $\bar{P}_2 = \left\langle\left[\sqrt{0.14}, \sqrt{0.31}\right], \left[\sqrt{0.13}, \sqrt{0.40}\right]\right\rangle$, where $S\left(\bar{P}_1\right) = S\left(\bar{P}_2\right) = -0.04$ and $H\left(\bar{P}_1\right) = H\left(\bar{P}_2\right) = 0.49$. In response to this issue, hesitation score function and hesitation accuracy score function are proposed to distinguish all IVPFNs.

Definition 4. Let $\bar{P} = \left\langle\left[\mu_p^L, \mu_p^U\right], \left[\nu_p^L, \nu_p^U\right]\right\rangle$ be any IVPFN, and its hesitation score function is shown in equation (3).

$$T\left(\bar{P}\right) = \frac{1}{2}\left(\left(\mu_p^U\right)^2 + \left(\nu_p^L\right)^2 - \left(\mu_p^L\right)^2 - \left(\nu_p^U\right)^2\right) \tag{3}$$

Definition 5. Let $\bar{P} = \left\langle\left[\mu_p^L, \mu_p^U\right], \left[\nu_p^L, \nu_p^U\right]\right\rangle$ be any IVPFN, and its hesitation accuracy function is shown in equation (4).

$$G\left(\bar{P}\right)=\frac{1}{2}\left(\left(\mu_p^U\right)^2+\left(\nu_p^U\right)^2-\left(\mu_p^L\right)^2-\left(v_p^L\right)^2\right) \tag{4}$$

Definition 6 (Xu, 2007). Let set be $A=\{a_1,a_2,\cdots a_n\}$, where $a_i=\left\langle\left[\mu_{a_i}^L,\mu_{a_i}^U\right],\left[v_{a_i}^L,v_{a_i}^U\right]\right\rangle(i=1,2,\cdots,n)$ is a IVPFN, then IVPFWA operator is shown in equation (5).

$$\mathrm{IVPFWA}\left(a_1,a_2,\cdots,a_n\right)=\left\langle\left[\sqrt{1-\prod_{i=1}^{n}\left(1-\left(\mu_{a_i}^L\right)^2\right)^{\omega_i}},\sqrt{1-\prod_{i=1}^{n}\left(1-\left(\mu_{a_i}^U\right)^2\right)^{\omega_i}}\right],\left[\prod_{j=1}^{m}\left(v_{a_i}^L\right)^{\omega_i},\prod_{j=1}^{m}\left(v_{a_i}^U\right)^{\omega_i}\right]\right\rangle \tag{5}$$

Where $\omega=(\omega_1,\omega_2,\cdots,\omega_n)$ is the weight of $A=(a_1,a_2,\cdots,a_n)$ and satisfies $\sum_{i=1}^{n}\omega_i=1$, $\omega_i\in[0,1]$.

Theorem 1 (Wang et al., 2009). Let $a_1=\left\langle\left[\mu_{a_1}^L,\mu_{a_1}^U\right],\left[\nu_{a_1}^L,\nu_{a_1}^U\right]\right\rangle$ and $a_2=\left\langle\left[\mu_{a_2}^L,\mu_{a_2}^U\right],\left[\nu_{a_2}^L,\nu_{a_2}^U\right]\right\rangle$ be any two IVPFNs, $a_1=a_2$ is satisfied iff $\mu_{a_1}^L=\mu_{a_2}^L$, $\mu_{a_1}^U=\mu_{a_2}^U$, $\nu_{a_1}^L=\nu_{a_2}^L$, $\nu_{a_1}^U=\nu_{a_2}^U$.

Definition 7 (Xu and Yager, 2008). The positive and negative ideal solutions of the evaluation results (α^+, α^-) are:

$$\begin{aligned}\alpha^+&=\left\langle\left[\mu_+^L,\mu_+^U\right],\left[v_+^L,v_+^U\right],\left[\pi_+^L,\pi_+^U\right]\right\rangle\\&=\left\langle\left[\max\mu_{ij}^-,\max\mu_{ij}^+\right],\left[\min v_{ij}^-,\min v_{ij}^+\right],\right.\\&\left.\left[\sqrt{1-\left(\max\mu_{ij}^+\right)^2-\left(\min v_{ij}^+\right)^2},\sqrt{1-\left(\max\mu_{ij}^-\right)^2-\left(\min v_{ij}^-\right)^2}\right]\right\rangle\end{aligned} \tag{6}$$

$$\begin{aligned}\alpha^-&=\left\langle\left[\mu_-^L,\mu_-^U\right],\left[v_-^L,v_-^U\right],\left[\pi_-^L,\pi_-^U\right]\right\rangle\\&=\left\langle\left[\min\mu_{ij}^-,\min\mu_{ij}^+\right],\left[\max v_{ij}^-,\max v_{ij}^+\right],\right.\\&\left.\left[\sqrt{1-\left(\min\mu_{ij}^+\right)^2-\left(\max v_{ij}^+\right)^2},\sqrt{1-\left(\min\mu_{ij}^-\right)^2-\left(\max v_{ij}^-\right)^2}\right]\right\rangle\end{aligned} \tag{7}$$

2.2. Distinctness operator construction considering the geometric significance of IVPFNs

As shown in Figure 1, any IVPFN can represent the area of a rectangle, and the formula is as follows:

$$Q(\bar{P})=\left(\mu_p^U-\mu_p^L\right)\left(\nu_p^U-\nu_p^L\right) \tag{8}$$

If $S\left(\alpha_i\right)$ and $H\left(\alpha_i\right)$ are the same, the smaller the $Q\left(\bar{P}\right)$ value, the better the $\bar{P}$. Applying Eq. (1) and Eq. (2) cannot distinguish between $\bar{P}_1=\langle[0.3,0.6],[0.2,0.7]\rangle$ and $\bar{P}_2=\langle[\sqrt{0.14},\sqrt{0.31}],[\sqrt{0.13},\sqrt{0.40}]\rangle$, then Eq. (8) can be applied to calculate the area of the rectangle covered by the two IVPFNs, which is calculated as $Q\left(\bar{P}_1\right)=0.15$ and $Q\left(\bar{P}_2\right)=0.05$, then it can be seen that $\bar{P}_2$ is better.

Subsequently, by introducing the distance between the centers of gravity of the rectangle, the distance $N\left(\bar{P}_1,\bar{P}_2\right)$ between the centers of gravity of two different IVPFNs can be calculated.

Definition 8. The definition of the comprehensive differentiation between any two different IVPFNs $\alpha_1=\left\langle\left[\mu_{a_1}^L,\mu_{a_1}^U\right],\left[\nu_{a_1}^L,\nu_{a_1}^U\right]\right\rangle$ and $\alpha_2=\left\langle\left[\mu_{a_2}^L,\mu_{a_2}^U\right],\left[\nu_{a_2}^L,\nu_{a_2}^U\right]\right\rangle$ is:

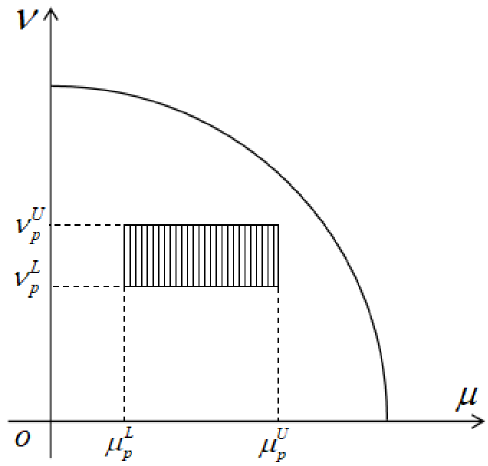

Figure 1: The geometric meaning image of PFN $\alpha = (\mu_\alpha, \nu_\alpha)$.

$$d(\alpha_1, \alpha_2) = \frac{1}{X}\left(|S(\alpha_1) - S(\alpha_2)| + |H(\alpha_1) - H(\alpha_2)| + |T(\alpha_1) - T(\alpha_2)| \right.$$
$$\left. + |G(\alpha_1) - G(\alpha_2)| + |Q(\alpha_1) - Q(\alpha_2)| + N(\alpha_1, \alpha_2)\right) \tag{9}$$

Where $X = \left(8 + \frac{\pi}{4} + \sqrt{2}\right)$.

Theorem 2. For any three IVPFNs $\alpha_i = \langle [a_i, b_i], [c_i, d_i] \rangle$ $(i = 1, 2, 3)$, the following condition can be satisfied.

1) $0 \le d(\alpha_1, \alpha_2) \le 1$, in particular $d(\alpha_1, \alpha_1) = 0$.

2) $d(\alpha_1, \alpha_2) = d(\alpha_2, \alpha_1)$.

3) $d(\alpha_1, \alpha_3) \le d(\alpha_1, \alpha_2) + d(\alpha_2, \alpha_3)$.

Proof. From the definitions given in this paper, $-1 \le S(\alpha_i) \le 1$, $0 \le H(\alpha_i) \le 2$, $-1 \le T(\alpha_i) \le 1$, $-1 \le G(\alpha_i) \le 1$, $0 \le Q(\alpha_i) \le \frac{\pi}{4}$, $0 \le N(\alpha_i, \alpha_j) \le \sqrt{2}$, $i = 1, 2, 3$.

1)

$$d(\alpha_1, \alpha_2) = \frac{1}{X}\left(|S(\alpha_1) - S(\alpha_2)| + |H(\alpha_1) - H(\alpha_2)| + |T(\alpha_1) - T(\alpha_2)| \right.$$
$$\left. + |G(\alpha_1) - G(\alpha_2)| + |Q(\alpha_1) - Q(\alpha_2)| + N(\alpha_1, \alpha_2)\right) \ge 0$$

$$d(\alpha_1, \alpha_2) = \frac{1}{X}\left(|S(\alpha_1) - S(\alpha_2)| + |H(\alpha_1) - H(\alpha_2)| + |T(\alpha_1) - T(\alpha_2)| + |G(\alpha_1) - G(\alpha_2)| \right.$$
$$\left. + |Q(\alpha_1) - Q(\alpha_2)| + N(\alpha_1, \alpha_2)\right) \le \frac{1}{X}\left(2 + 2 + 2 + 2 + \frac{\pi}{4} + \sqrt{2}\right) \le 1$$
$$d(\alpha_1, \alpha_1) = \frac{1}{X}\left(|S(\alpha_1) - S(\alpha_1)| + |H(\alpha_1) - H(\alpha_1)| + |T(\alpha_1) - T(\alpha_1)| \right.$$
$$\left. + |G(\alpha_1) - G(\alpha_1)| + |Q(\alpha_1) - Q(\alpha_1)| + N(\alpha_1 \alpha_1)\right) = 0$$
$$i.e. 0 \le d(\alpha_1, \alpha_2) \le 1, d(\alpha_1, \alpha_1) = 0.$$

2)

$$d(\alpha_1, \alpha_2) = \frac{1}{X}\left(|S(\alpha_1) - S(\alpha_2)| + |H(\alpha_1) - H(\alpha_2)| + |T(\alpha_1) - T(\alpha_2)| \right.$$

$$+\left|G(\alpha_1)-G(\alpha_2)\right|+\left|Q(\alpha_1)-Q(\alpha_2)\right|+N(\alpha_1,\alpha_2))$$
$$=\frac{1}{X}\left(\left|S(\alpha_2)-S(\alpha_1)\right|+\left|H(\alpha_2)-H(\alpha_1)\right|+\left|T(\alpha_2)-T(\alpha_1)\right|+\left|G(\alpha_2)-G(\alpha_1)\right|\right.$$
$$\left.+\left|Q(\alpha_2)-Q(\alpha_1)\right|+N(\alpha_2,\alpha_1)\right)=d(\alpha_2,\alpha_1)$$

3)

$$d(\alpha_1,\alpha_3)=\frac{1}{X}\left(\left|\frac{1}{2}\left(b_1^2-d_1^2+a_1^2-c_1^2\right)-\frac{1}{2}\left(b_3^2-d_3^2+a_3^2-c_3^2\right)\right|\right.$$
$$+\left|\frac{1}{2}\left(a_1^2+b_1^2+c_1^2+d_1^2\right)-\frac{1}{2}\left(a_3^2+b_3^2+c_3^2+d_3^2\right)\right|+\left|\frac{1}{2}\left(b_1^2+c_1^2-a_1^2-d_1^2\right)-\frac{1}{2}\left(b_3^2+c_3^2-a_3^2-d_3^2\right)\right|$$
$$+\left|\frac{1}{2}\left(b_1^2+d_1^2-a_1^2-c_1^2\right)-\frac{1}{2}\left(b_3^2+d_3^2-a_3^2-c_3^2\right)\right|$$
$$\left.+\left|(b_1-a_1)(d_1-c_1)-(b_3-a_3)(d_3-c_3)\right|+N(\alpha_1,\alpha_3)\right)$$
$$=\frac{1}{X}\left(\left|\frac{1}{2}\left(b_1^2-d_1^2+a_1^2-c_1^2\right)-\frac{1}{2}\left(b_2^2-d_2^2+a_2^2-c_2^2\right)+\frac{1}{2}\left(b_2^2-d_2^2+a_2^2-c_2^2\right)-\frac{1}{2}\left(b_3^2-d_3^2+a_3^2-c_3^2\right)\right|\right.$$
$$+\left|\frac{1}{2}\left(a_1^2+b_1^2+c_1^2+d_1^2\right)-\frac{1}{2}\left(a_2^2+b_2^2+c_2^2+d_2^2\right)+\frac{1}{2}\left(a_2^2+b_2^2+c_2^2+d_2^2\right)-\frac{1}{2}\left(a_3^2+b_3^2+c_3^2+d_3^2\right)\right|$$
$$+\left|\frac{1}{2}\left(b_1^2+c_1^2-a_1^2-d_1^2\right)-\frac{1}{2}\left(b_2^2+c_2^2-a_2^2-d_2^2\right)+\frac{1}{2}\left(b_2^2+c_2^2-a_2^2-d_2^2\right)-\frac{1}{2}\left(b_3^2+c_3^2-a_3^2-d_3^2\right)\right|$$
$$+\left|\frac{1}{2}\left(b_1^2+d_1^2-a_1^2-c_1^2\right)-\frac{1}{2}\left(b_2^2+d_2^2-a_2^2-c_2^2\right)+\frac{1}{2}\left(b_2^2+d_2^2-a_2^2-c_2^2\right)-\frac{1}{2}\left(b_3^2+d_3^2-a_3^2-c_3^2\right)\right|$$
$$\left.+\left|\,(b_1-a_1)(d_1-c_1)-(b_2-a_2)(d_2-c_2)+(b_2-a_2)(d_2-c_2)-(b_3-a_3)(d_3-c_3)\,\right|+N(\alpha_1,\alpha_3)\right)$$
$$\leq\frac{1}{X}\left(\left|S(\alpha_1)-S(\alpha_2)\right|+\left|H(\alpha_1)-H(\alpha_2)\right|+\left|T(\alpha_1)-T(\alpha_2)\right|+\left|G(\alpha_1)-G(\alpha_2)\right|+\left|Q(\alpha_1)-Q(\alpha_2)\right|\right.$$
$$+\left(\left|S(\alpha_2)-S(\alpha_3)\right|+\left|H(\alpha_2)-H(\alpha_3)\right|+\left|T(\alpha_2)-T(\alpha_3)\right|\right.$$
$$\left.+\left|G(\alpha_2)-G(\alpha_3)\right|+\left|Q(\alpha_2)-Q(\alpha_3)\right|+N(\alpha_1,\alpha_2)+N(\alpha_2,\alpha_3)\right)=d(\alpha_1,\alpha_2)+d(\alpha_2,\alpha_3)$$

3. Evaluation criteria weighting model

The model in literature (You and Chen, 2017) is improved to construct the following model. Let C_j, $j=1,2,\ldots,n$ be the set of evaluation criteria and let A_i, $i=1,2,\ldots,m$ be the set of alternatives.

For an alternative A_i, the larger the values of $S(\alpha_i)$ and $H(\alpha_i)$, the better A_i is. The smaller the values of $T(\alpha_i)$, $G(\alpha_i)$ and $Q(\bar{P})$, the more explicit their decision information is. Construct the model M_1 from the perspective of evaluation results.

$$\begin{cases}\max f_1(\omega)=\sum_{i=1}^{m}\sum_{j=1}^{n}\omega_j\frac{S(a_{ij})+H(a_{ij})}{S(a_{ij})+H(a_{ij})+T(a_{ij})+G(a_{ij})+Q(a_{ij})};\\ \text{s.t. }\sum_{j=1}^{n}\omega_j=1;\\ 0\leq\omega_j\leq 1\end{cases}$$

If the difference in the evaluation results given by the decision-making experts under C_j is large, it means that C_j is more capable of distinguishing between alternatives and should be given a higher weight, and vice versa. Construct the model M_2 from the ability to distinguish evaluation criteria, with the use of Eq. (9).

$$\begin{cases} \max f_2(\omega) = \sum_{i=1}^{m} \sum_{j=1}^{n} \sum_{1 \le i \le k \le n} \omega_j d(a_{ij}, a_{kj}) \\ \text{s.t. } \sum_{j=1}^{n} \omega_j = 1; \\ 0 \le \omega_j \le 1 \end{cases}$$

Considering the above two aspects together, the adjustment coefficient α is introduced and the model M is constructed.

$$\begin{cases} \max f(\omega) = \alpha f_1(\omega) + (1-\alpha) f_2(\omega) \\ \text{s.t. } \sum_{j=1}^{n} \omega_j = 1 \\ 0 \le \omega_j \le 1; \\ 0 \le \alpha \le 1 \end{cases}$$

where α denotes a pre-given adjustment factor by the decision maker. The model M is solved below by constructing a Lagrangian auxiliary function $L(\omega, \lambda)$.

$$L(\omega, \lambda) = f(\omega) + \lambda \left(\sum_{j=1}^{n} \omega_j - 1 \right)$$

Find the partial derivatives of ω_j and λ, making them equal to 0.

$$\begin{cases} \frac{\partial L}{\partial \omega_j} = \sum_{i=1}^{m} \left(\alpha \frac{S(a_{ij})+H(a_{ij})}{S(a_{ij})+H(a_{ij})+T(a_{ij})+G(a_{ij})+Q(a_{ij})} + \sum_{1 \le i \le k \le n} (1-\alpha) d(a_{ij}, a_{kj}) \right) + \lambda = 0; \\ \frac{\partial L}{\partial \lambda} = \sum_{j=1}^{n} \omega_j - 1 = 0. \end{cases}$$

Solving the above equation gives:

$$\omega_j = \frac{\sum_{i=1}^{m} \left(\alpha \frac{S(a_{ij})+H(a_{ij})}{S(a_{ij})+H(a_{ij})+T(a_{ij})+G(a_{ij})+Q(a_{ij})} + \sum_{1 \le i \le k \le n} (1-\alpha) d(a_{ij}, a_{kj}) \right)}{\sum_{j=1}^{n} \sum_{i=1}^{m} \left(\alpha \frac{S(a_{ij})+H(a_{ij})}{S(a_{ij})+H(a_{ij})+T(a_{ij})+G(a_{ij})+Q(a_{ij})} + \sum_{1 \le i \le k \le n} (1-\alpha) d(a_{ij}, a_{kj}) \right)} \tag{10}$$

Using Eq. (11) can calculate the final weights of evaluation criteria.

$$\omega_{\text{end}} = \frac{\sum \text{ The i-th evaluation criterion } \omega_i}{\sum \text{ All evaluation criteria } \omega_i} \tag{11}$$

4. Similarity-based clustering algorithm

Referring to the data "similarity" in (Hu and Luo, 2007) and "correlation" in (Xu and Chen, 2005), the following definitions are given.

Definition 9. In an $m \times n$ dimensional linear space, $\rho(u, v) = \frac{|u-\bar{u}| \cdot |v-\bar{v}|^{\mathrm{T}}}{\|u-\bar{u}\|_2 \cdot \|v-\bar{v}\|_2}$ is said to be the similarity of evaluation matrices u and v, where the mean vector of the set of vectors composed of vectors v_{ij} is $\bar{v} = (\overline{v_1}, \overline{v_2}, \cdots, \overline{v_m})^{\mathrm{T}}$, $\overline{v_i} = \frac{1}{n} \sum\limits_{j=1}^{n} v_{ij}$. The definition of $\bar{u}$ is the same as above. $\| v\|_2 = \left(\sum\limits_{i=1}^{m} \sum\limits_{j=1}^{n} |v_{ij}|^2 \right)^{\frac{1}{2}}$ represents the calculation of norm, and $u \cdot v = \sum\limits_{i=1}^{n} \sum\limits_{j=1}^{n} u_{ij} v_{ij}$ represents dot product. The similarity between $V$and $\bar{v}$ is $\rho(v, V) = \frac{1}{T}\sum_{t=1}^{T}\rho\left(v, v^t\right)$, where T satisfying $1 \leq t \leq T$, $T \geq 2$ is the number of all elements in the set of decision matrices. $\rho(v, V) = \frac{1}{C_T^2}\sum_{1\leq i<j\leq T}\rho\left(v^i, v^j\right)$ is the average similarity between all vectors in the vector set V, and it is also the average consistency of clustering, where $C_T^2 = \frac{T(T-1)}{2}$. If $T = 1$, let $\rho(V) = 0$.

The specific simulation steps are as follows:

Step 1. Determine the group Ω, and label the evaluation vectors given by the members therein from 1 to T in order after randomly ordering them to form a set Ψ. At the same time, let Q denote a temporary set, which is initially set to be the empty set.

Step 2. Choose the appropriate threshold γ, where $0 \leq \gamma \leq 1$.

Step 3. Assuming the initial number of clusters is $k = 1$, and the evaluation vector selected from set Ψ has the sequence number $i = 1$.

Step 4. Select the vector V^i from the set Ψ in numbered order and place it in the cluster Ω_k. The number of decision members in the cluster Ω_k at this point is $n_k = 1$.

Step 5. If the set Ψ is non-empty, select the next vector V^i, $i = i + 1$ from it in turn, if the set Ψ is empty then go to Step 7.

Step 6. The $\rho\left(V^i, \Omega_k\right)$ between V^i and the set composed of vectors in the cluster Ω_k is calculated. If $\rho\left(V^i, \Omega_k\right) > \gamma$ then V^i is assigned to that cluster Ω_k; if $\rho\left(V^i, \Omega_k\right) \leq \gamma$ then it is temporarily assigned to the set of temporary vectors Q. In the meantime V^i needs to be removed from the set Ψ. Go to Step 5.

Step 7. If Q is non-empty, let the set $\Psi = Q$; Q be the empty set, at this time the count of clustering is $k = k + 1$, go to Step 4; if Q is the empty set, go to Step 8.

Step 8. Output the result and end the calculation.

5. Calculus analysis

Organize and analyze the existing literature to give a set of alternatives: A_1 building a modern cultural tourism project based on ethnic culture, A_2 building a cultural tourism expansion project based on folklore and folk villages, A_3 building a cultural tourism project based on "culture+specialty", A_4 building an internet celebrity cultural tourism project, A_5 building a cultural tourism project based on a small portion of the relevant local ethnic cultures and a set of evaluation criteria: C_1 agreeableness, C_2 effectiveness of the previous construction program budget, C_3 degree of influence and C_4 ecological and economic benefits.

20 experts in relevant fields will be invited and involved in the decision-making, as follows:

Step 1. Under the evaluation criterion $C = (C_1, C_2, C_3, \cdots, C_n)$, the expert community gives the evaluation results of the set of alternatives $A = (A_1, A_2, A_3, \cdots, A_m)$ to form a decision matrix.

Step 2. It was processed and analyzed using the large population clustering algorithm described above. The initial clustering of 20 expert decision matrices was carried out and brought into the computational software Netlogo for 5000 evolutions in the following process.

Setting different thresholds, 20 decision data are subjected to clustering operations. Due to the selection of 20 samples in this article, small adjustments to the threshold will not result in significant changes in clustering results. Therefore, this article sets thresholds of 0.70, 0.75, 0.80, and 0.85 for clustering operations. From the results, it can be seen that the optimal result is achieved when the threshold is 0.85, where the number of clustered groups is 4, and the number of internal groups in each group is relatively uniform. In addition, there is a high and stable average consistency in this result, as shown in Figure 2.

Next, the clustering results at the threshold value of 0.85 are processed, and the average consistency is ranked from high to low under the consideration of the number of members within the clustered group, and it can be seen that the 502nd evolution is optimal, as shown in Table 1. The classification of cluster members is shown in Table 2. The clustering results are shown in Figure 3.

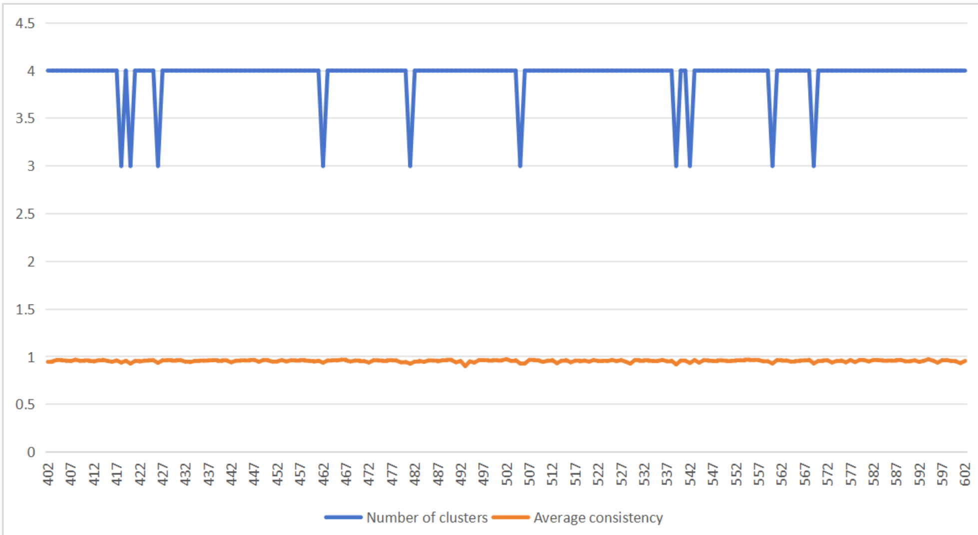

Figure 2: Number of clusters with a threshold of 0.85 and mean consistency.

Table 1: Sorting of clustering results at the threshold of 0.85.

Evolutionary number	Number of clusters	Average consistency
502	4	0.97347
594	4	0.97252
4866	4	0.97213
1225	4	0.97209
166	4	0.97156

Table 2: Member partitioning results of the 502nd clustering population with the threshold of 0.85.

Serial number of the cluster	Number of members	Serial numbers of members	Average consistency
1	7	5, 6, 13, 14, 15, 16, 19	0.96399
2	5	7, 8, 9, 10, 12	0.98267
3	4	1, 2, 3, 20	0.98672
4	4	4, 11, 17, 18	0.96052

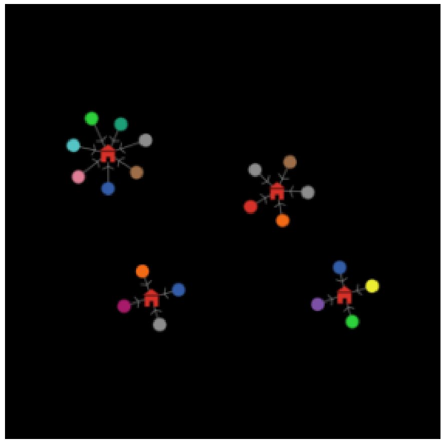

Figure 3: The 502nd clustering result graph with the threshold of 0.85.

Subsequently, the decision matrix of the experts included in each clustering group is averaged, and the decision matrices of clusters 1-4 can be obtained separately to represent the decision results of the four clustering groups. The weight of each clustering group is assigned according to the number of experts it contains to obtain $\omega_j = (0.35, 0.25, 0.20, 0.20)$. A weighting operation can be performed on the 4 matrices based on this weight.

Step 3. By inputting the matrix obtained above into model M and adjusting to different values for calculation, the evaluation criteria weights for all cases can be obtained, as shown in Table 3.

Table 3: Weight of evaluation criteria under different adjustment coefficients α.

	C_1	C_2	C_3	C_4
$\alpha = 0.1$	0.2634	0.2461	0.2427	0.2478
$\alpha = 0.3$	0.2572	0.2485	0.2442	0.2501
$\alpha = 0.5$	0.2553	0.2494	0.2447	0.2506
$\alpha = 0.7$	0.2543	0.2499	0.2449	0.2509
$\alpha = 0.9$	0.2537	0.2502	0.2450	0.2511

Step 4. The α^+, α^- of the four clustering decision matrices and their weighted clustering decision matrices are found according to Definition 7, and the weights obtained in Step 3 are brought into Eq. (5) for calculation. Next, apply Eq. (9) to the integrated distance $d^+(A_i, \alpha^+)$ and $d^-(A_i, \alpha^-)$ between each decision matrix and α^+, α^-. Finally, using $M_i = \frac{d^-(A_i,\alpha^-)}{d^+(A_i,\alpha^+)+d^-(A_i,\alpha^-)}$ to find and order the relative distances.

$\alpha = 0.5$ is selected for detailed analysis, so that the expert evaluation results and the distinguishing ability of evaluation criteria can be considered in a balanced way. The results are shown in Table 4.

Step 5. Standardize the results obtained from Step 4 to obtain the final result shown in Figure 4. From this graph, it can be seen that A_2 is optimal, and the final ranking is $A_2 > A_1 > A_3 > A_5 > A_4$.

As can be seen from Figure 4, the final ranking of alternatives given by different clustering groups is consistent, but has different final scores for a specific option. In addition, the method can analyze the situation that each clustering group contains members, which is of strong practical significance. Members belonging to the same clustering group usually have more similar features. Dividing decision-makers into multiple clustering groups facilitates considering the needs of different groups of people when making decisions, thereby optimizing decision-making. The principles of different clustering algorithms vary greatly and their purposes are not the same. Therefore, this article only compares and analyzes the content other than clustering algorithms.

Table 4: Weight of evaluation criteria under different adjustment coefficients.

$\alpha = 0.5$	Cluster 1	Cluster 2	Cluster 3	Cluster 4	Weighted cluster
A_1	0.8183	0.7681	0.7972	0.8287	0.7883
A_2	0.8834	0.9668	0.8435	0.9240	0.9582
A_3	0.3854	0.4360	0.3434	0.4360	0.3993
A_4	0.0411	0.0417	0.1150	0.0690	0.0383
A_5	0.3072	0.1773	0.2982	0.1327	0.1811

Figure 4: Comparison of combined distances between different clustered groups plot.

Let $\alpha = 0.5$, and import the data from (Li et al., 2019), (Li et al., 2021) and (You and Chen, 2017) into the model of this paper for calculation. Compare the calculated results with those of the original text, as shown in Table 5. From the final result, it can be seen that although the focus of different methods may vary slightly, it does not affect the selection of the optimal result.

Table 5: Comparison with the results of (Li et al., 2019), (Li et al., 2021) and (You and Chen, 2017).

Method	Arrange in order
Method from the (Li et al., 2019)	$A_2 > A_4 > A_1 > A_3$
Method in this paper	$A_2 > A_4 > A_1 > A_3$
Method from the literature (Li et al., 2021)	$A_2 > A_3 > A_1$
Method in this paper	$A_2 > A_3 > A_1$
Method from the literature (You and Chen, 2017)	$A_5 > A_4 > A_2 > A_3 > A_1$
Method in this paper	$A_5 > A_4 > A_2 > A_3 > A_1$

6. Concluding remarks

The multi-attribute large group decision-making method system based on interval Pythagorean fuzzy numbers proposed in this paper not only contains the basic tools for all-around description of the evaluation results and the tools for information mining of the decision results, but also contains the computational tools based on the weighted integration of decision matrices as well as the model tools based on these contents. The methodology system has strong practical significance and reference value for the decision-making of cultural tourism projects and other types of decision-making problems.

Acknowledgments

This paper here was supported by Hebei Provincial Department of Education Humanities and Social Science Research Major Project Research Project ZD202411.

References

Abdullah Al-Barakati, Arunodaya Raj Mishra, Abbas Mardani, and Pratibha Rani. An extended interval-valued pythagorean fuzzy waspas method based on new similarity measures to evaluate the renewable energy sources. *Applied Soft Computing*, 120:108689, 2022. doi: 10.1016/j.asoc.2022.108689.

Ertugrul Ayyildiz, Aslihan Yildiz, Alev Taskin, and Coskun Ozkan. An interval valued pythagorean fuzzy ahp integrated quality function deployment methodology for hazelnut production in turkey. *Expert Syst. Appl.*, 231(C), nov 2023. doi: 10.1016/j.eswa.2023.120708.

Juan Chang, Yingxue Du, and Weifeng Liu. Pythagorean hesitant fuzzy risky multi-attribute decision making method based on cumulative prospect theory and vikor. *Operations Research and Management Science*, 31:50–56, 2022.

Lihui Hu and Guosong Luo. Improved algorithm for group clustering based on vector space. *Systems Engineering and Electronics*, pages 472–474, 2007.

Hassan Abdulsattar Ibrahim, Aws Alaa Zaidan, Sarah Qahtan, and Bilal Bahaa Zaidan. Sustainability assessment of palm oil industry 4.0 technologies in a circular economy applications based on interval-valued pythagorean fuzzy rough set-fwzic and edas methods. *Appl. Soft Comput.*, 136 (C), mar 2023. doi: 10.1016/j.asoc.2023.110073.

Jinjun Li, Yue Bao, and Chensong Li. Interval-valued pythagorean fuzzy heronian means operators and their application in decision making. *Fuzzy Systems and Mathematics*, 33:134–148, 2019.

Jinjun Li, Li Tingting, Yue Bao, and Minghao Chen. Interval-valued pythagorean fuzzy power geometric-geometric heronian means operators and their application in multiple attribute group decision making. *Operations Research and Management Science*, 30:78–83, 2021.

Yanhe Li, Ran Wang, Wanpeng Zhou, and Bingtuan Gao. Evaluation of dispatching results of power system with high penetration of renewable energy based on pythagorean fuzzy set and topsis. *Energy Reports*, 8:524–532, 2022. doi: 10.1016/j.egyr.2022.08.134. 2022 The 5th International Conference on Electrical Engineering and Green Energy.

Wei Liang, Xiaolu Zhang, and Manfeng Liu. The maximizing deviation method based on interval-valued pythagorean fuzzy weighted aggregating operator for multiple criteria group decision analysis. *Discrete Dynamics in Nature and Society*, 2015(1):746572, 2015. doi: 10.1155/2015/746572.

Peide Liu, Xin Dong, Peng Wang, and Runyu Du. Managing manipulation behavior in hydrogen refueling station planning by a large group decision making method with hesitant fuzzy linguistic information. *Inf. Sci.*, 652(C), jan 2024. doi: 10.1016/j.ins.2023.119741.

Zhoujing Wang, Kevin W. Li, and Weize Wang. An approach to multiattribute decision making with interval-valued intuitionistic fuzzy assessments and incomplete weights. *Inf. Sci.*, 179(17): 3026–3040, aug 2009. doi: 10.1016/j.ins.2009.05.001.

Xuanhua Xu and Xiaohong Chen. Research on the group clustering method based on vector space. *Systems Engineering and Electronics*, pages 1034–1037, 2005.

Zeshui Xu. Intuitionistic fuzzy aggregation operators. *IEEE Transactions on Fuzzy Systems*, 15(6): 1179–1187, 2007. doi: 10.1109/TFUZZ.2006.890678.

Zeshui Xu and Ronald R. Yager. Dynamic intuitionistic fuzzy multi-attribute decision making. *International Journal of Approximate Reasoning*, 48(1):246–262, 2008. doi: 10.1016/j.ijar.2007.08.008. Special Section: Perception Based Data Mining and Decision Support Systems.

Ronald R. Yager. Pythagorean membership grades in multicriteria decision making. *IEEE Transactions on Fuzzy Systems*, 22(4):958–965, 2014. doi: 10.1109/TFUZZ.2013.2278989.

Dongliang Yin, Guoheng Cui, Xiaoying Huang, and Huan Zhang. Interval-valued pythagorean fuzzy multi-attribute decision-making based on improved score function and prospect theory. *Systems Engineering and Electronics*, 44:3463–3469, 2022.

appreciate You and Tong Chen. Research on selecting optimum construction plan for public cultural facilities under the interval-valued intuitionistic fuzzy environment. *Systems Science and Mathematical Sciences*, 37:1468–1479, 2017.

Serhat Yüksel and Hasan Dinçer. Identifying the strategic priorities of nuclear energy investments using hesitant 2-tuple interval-valued pythagorean fuzzy dematel. *Progress in Nuclear Energy*, 145:104103, 2022. doi: 10.1016/j.pnucene.2021.104103.

Xiaolu Zhang. Multicriteria pythagorean fuzzy decision analysis: A hierarchical qualiflex approach with the closeness index-based ranking methods. *Information Sciences*, 330:104–124, 2016. doi: 10.1016/j.ins.2015.10.012. SI Visual Info Communication.

Jinghua Zhao, Haiying Rong, Xiting Lyu, and et al. Pythagorean fuzzy multi-attribute emergency decision-making method based on dematel and trust network. *Application Research of Computers*, 1(40):165–171, 2023.

Yuanyuan Zhou, Chengli Zheng, and Mark Goh. Statistics-based approach for large-scale group decision-making under incomplete pythagorean fuzzy information with risk attitude. *Know.-Based Syst.*, 235(C), jan 2022. doi: 10.1016/j.knosys.2021.107654.