

控制与决策

Control and Decision

大群体应急决策中考虑属性关联的偏好信息融合方法

徐选华, 余艳粉

引用本文:

徐选华, 余艳粉. 大群体应急决策中考虑属性关联的偏好信息融合方法[J]. *控制与决策*, 2021, 36(10): 2537–2546.

在线阅读 View online: <https://doi.org/10.13195/j.kzyjc.2020.0117>

您可能感兴趣的其他文章

Articles you may be interested in

[基于云模型和多层权重求解的多粒度语言大群体决策方法](#)

Multi-granularity linguistic large group decision-making based on cloud model and multi-layer weight determination

控制与决策. 2021, 36(9): 2257–2266 <https://doi.org/10.13195/j.kzyjc.2020.0102>

[基于群决策考虑属性效用一致性的DEA他评交叉效率公共权重排序法](#)

A common-weight ranking method for DEA peer-efficiency based on group decision-making and considering the consistency of attribute utility

控制与决策. 2021, 36(9): 2279–2289 <https://doi.org/10.13195/j.kzyjc.2019.1719>

[基于犹豫度和相似度的专家权重确定方法及其应用](#)

Expert weights determination method and application based on hesitancy degree and similarity measure

控制与决策. 2021, 36(6): 1482–1488 <https://doi.org/10.13195/j.kzyjc.2019.1382>

[考虑个体累积共识贡献的犹豫模糊语言自适应共识模型](#)

Adaptive consensus model with hesitant fuzzy linguistic information considering individual cumulative consensus contribution

控制与决策. 2021, 36(1): 187–195 <https://doi.org/10.13195/j.kzyjc.2019.0556>

[考虑时间序列的动态大群体应急决策方法](#)

Dynamic large group emergency decision-making method considering time series

控制与决策. 2020, 35(11): 2609–2618 <https://doi.org/10.13195/j.kzyjc.2019.0088>

大群体应急决策中考虑属性关联的偏好信息融合方法

徐选华[†], 余艳粉

(中南大学 商学院, 长沙 410083)

摘要: 针对大群体应急决策中属性之间相互关联的问题, 提出一种新的偏好信息融合方法. 首先, 利用最优离散拟合模型测度专家的风险偏好, 提出一种考虑专家风险偏好的犹豫模糊元补充方法; 其次, 运用 TF-IDF 算法获得相互关联的事件属性集; 再次, 结合传统的主成分分析法与误差理论, 提出基于犹豫模糊语言的主成分分析模型, 获得几个互不关联的主属性及其权重, 进而进行信息集结和方案择优; 最后, 通过广西洪涝灾害事件验证所提方法的可行性和有效性.

关键词: 大群体; 应急决策; 属性关联; 信息融合

中图分类号: C934

文献标志码: A

DOI: 10.13195/j.kzyjc.2020.0117

开放科学(资源服务)标识码(OSID):



引用格式: 徐选华, 余艳粉. 大群体应急决策中考虑属性关联的偏好信息融合方法 [J]. 控制与决策, 2021, 36(10): 2537-2546.

Preference information fusion method of large groups emergency decision-making based on attributes association

XU Xuan-hua[†], YU Yan-fen

(School of Business, Central South University, Changsha 410083, China)

Abstract: A new method of preference information fusion is proposed to solve the problem of attribute correlation in large group emergency decision making. Firstly, the optimal discrete fitting model is used to measure the risk preference of experts, and a hesitant fuzzy meta-supplement method considering the risk preference of experts is proposed. Secondly, TF-IDF algorithm is used to obtain the interrelated event attribute set. Then, combining the traditional principal component analysis method and error theory, a principal component analysis model (PCA) based on hesitating fuzzy language is proposed to obtain several unrelated principal attributes and their weights, and then information aggregation and alternatives selection are carried out. Finally, the feasibility and effectiveness of the proposed method are verified by the case of flood disaster in Guangxi.

Keywords: large groups; emergency decision making; attribute association; information fusion

0 引言

近年来我国重特大突发事件频发, 给社会带来了巨大的损失. 重特大突发事件具有高复杂性、高风险性和动态性, 需要专家大群体在较短时间内做出决策. 专家以犹豫模糊语言表达评估信息, 更能直观地体现其心理偏好, 而模糊语言变量中也包含专家个体的风险偏好. 在现实决策中, 属性之间存在关联性, 决策者在决策过程中, 参考不同的属性对方案集进行决策, 会受到属性关联的影响, 给出的偏好信息也会出现重叠, 专家给出的偏好信息便会失去区别性. 因此针对大群体应急决策问题, 处理属性之间的关联性才能更好地进行决策.

目前, 国内外关于属性关联背景下的群决策已有一些研究成果. 曹萍萍等^[1]结合属性期望和属性关联构建了综合感知价值矩阵, 借助 Choquet 积分对方案排序; Liu 等^[2]运用主成分分析法在属性关联背景下进行大群体决策; 王大澳等^[3]基于属性关联且属性值为区间灰数的背景, 提出了一种群决策方法; 陈振颂等^[4]定义了直觉梯形模糊平均算子, 提出基于直觉梯形模糊信息关联输入的改进群决策方法; Qu 等^[5]定义了 Shapley 双重犹豫模糊 Choquet 平均算子和 Shapley 双重犹豫模糊 Choquet 几何算子, 并结合 TOPSIS 进行群决策; Khan 等^[6]定义了毕达哥拉斯犹豫模糊 Choquet 积分几何算子、广义毕达

收稿日期: 2020-02-07; 修回日期: 2020-05-27.

基金项目: 国家自然科学基金项目(71971217); 国家自然科学基金重点项目(72091515, 71790615).

责任编辑: 李勇建.

[†]通讯作者. E-mail: xuxh@csu.edu.cn.

索犹豫模糊 Choquet 积分平均算子,并运用于多属性决策中;Wan 等^[7]运用模糊测度和 Choquet 积分,定义了三角直觉模糊 Choquet 积分算子;陈岩等^[8]根据专家给出的不确定信息,在属性关联背景下引入两个算子;Aggarwal 等^[9]在 Choquet 积分的基础上提出了态度 Choquet 积分算子,并进行属性关联下的决策行为;Ju 等^[10]定义了两个犹豫的模糊 Choquet 积分集合算子;王志良^[11]针对水资源管理准则决策问题,结合主成分分析法建立了决策理论方法;Tan 等^[12]考虑到偏好关联,定义了诱导 Choquet 积分算子.基于属性关联的大群体应急决策研究已有一些,但是处理方法大多是 Choquet 积分算子、区间数和传统 PCA 模型相结合^[13-16].犹豫模糊语言比较符合决策者的心理特征,能够很好地帮助决策者表达偏好信息,越来越多的决策者偏向于运用犹豫模糊语言表达个人偏好信息.但是目前在犹豫模糊语言的基础上,针对如何对多个关联属性进行降维处理的研究极少.

针对以上问题,本文基于属性之间相互关联的背景,从偏好信息融合的角度出发,考虑专家的风险偏好提出一种新的犹豫模糊元补充方法;建立基于犹豫模糊语言的主成分分析模型,降低属性集的维度,消除属性之间的关联性,由此获得几个互不关联的主属性,并定义主属性的权重,进而更好地进行信息融合,获得更加精确的决策结果.

1 问题描述及方法基础

1.1 问题描述

假设在大群体应急决策问题中, $X = \{x_1, x_2, \dots, x_l, \dots, x_p\}$ 表示 P 个备选应急方案的集合. 其中 x_l 表示第 l 个应急方案; $E = \{e_1, e_2, \dots, e_i, \dots, e_M\}$ ($M \geq 11$)^[17] 表示 M 个应急决策专家的集合, e_i 表示第 i 个决策专家,相应的专家权重为 $\omega = \{\omega_1, \omega_2, \dots, \omega_M\}$; $U = \{u_1, u_2, \dots, u_j, \dots, u_N\}$ 表示决策过程中需要考虑的 N 个属性集合, u_j 表示第 j 个决策属性,相应的属性权重为 $w = \{w_1, w_2, \dots, w_N\}$. 决策过程中,专家用犹豫模糊语言表达偏好,进而形成初始的偏好矩阵,矩阵 A^i 表示专家 e_i 对属性 u_j 给出的偏好,矩阵记为

$$A^i = \begin{matrix} & u_1 & u_2 & \cdots & u_j & \cdots & u_N \\ \begin{matrix} x_1 \\ x_2 \\ \vdots \\ x_l \\ \vdots \\ x_P \end{matrix} & \begin{bmatrix} a_1^{i1} & a_1^{i2} & \cdots & a_1^{ij} & \cdots & a_1^{iN} \\ a_2^{i1} & a_2^{i2} & \cdots & a_2^{ij} & \cdots & a_2^{iN} \\ \vdots & \vdots & \ddots & \vdots & \ddots & \vdots \\ a_l^{i1} & a_l^{i2} & \cdots & a_l^{ij} & \cdots & a_l^{iN} \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ a_P^{i1} & a_P^{i2} & \cdots & a_P^{ij} & \cdots & a_P^{iN} \end{bmatrix} \end{matrix},$$

其中 a_l^{ij} 表示专家 e_i 在属性 u_j 下针对方案 x_l 给出的犹豫模糊语言偏好信息.

本文主要解决的问题是:决策者在表达偏好信息时存在风险偏好,并且现实决策中属性间的关联性会影响决策结果.因此,如何考虑专家的风险偏好并在属性关联背景下进行决策是急需解决的问题.首先,对专家进行风险偏好测度,并运用 TF-IDF 算法获得属性集;其次,结合误差理论提出基于犹豫模糊语言的主成分分析模型;最后,结合主属性权重进行信息融合,进而对方案排序和择优.

1.2 传统的PCA模型

在多属性大群体决策中,备选方案经常由多个属性描述,并且这些属性表现出高度的依赖性和交互特征,主成分分析法能够很好地解决此问题.主成分分析 (principal component analysis, PCA) 的主要思想是将一组可能存在相关性的变量通过正交变换转换为几个线性不相关的综合变量,转换后的综合变量即为主成分 (PC),其主要作用是降低数据空间的维数.

传统 PCA 模型的降维过程是对应于方差-协方差矩阵的特征值和特征向量来完成的.假设原始样本数据有 p 个变量 X_1, X_2, \dots, X_p , 即 $X = (X_1, X_2, \dots, X_p)^T$; n 个数据,即矩阵 X 的第 i 行 $x_i = (x_{i1}, x_{i2}, \dots, x_{ip})^T, i = 1, 2, \dots, n$. 假设 μ 为随机变量 X 的均值, Σ 为协方差矩阵,则数据样本的第 k 个 PC 是 m 个变量的线性组合,PCA 即为协方差矩阵 Σ 的特征分解.传统主成分分析的主要步骤如下:

step 1: 首先对样本 X 标准化处理,得到标准化矩阵 Z :

$$Z_{ij} = \frac{x_{ij} - \bar{x}_j}{s_j}. \tag{1}$$

其中: $\bar{x}_j = \sum_{i=1}^n x_{ij} / n, s_j = \sqrt{\sum_{i=1}^n (x_{ij} - \bar{x}_j)^2 / (n - 1)}, i = 1, 2, \dots, n, j = 1, 2, \dots, p$.

step 2: 计算矩阵 Z 的协方差矩阵 Σ .

step 3: 计算协方差矩阵 Σ 的特征值和特征方程.假设协方差矩阵 Σ 的特征值为 $\lambda_1, \lambda_2, \dots, \lambda_m$, 特征向量为 e_1, e_2, \dots, e_m , 这 m 个特征向量即为 PC 系数.假设 y_1, y_2, \dots, y_p 为 X_1, X_2, \dots, X_p 的主成分,则 y_k 为得分矢量,有

$$y_k = X u_k. \tag{2}$$

step 4: 选择前 m 个最大的特征值对应的特征向量形成矩阵 T ,主成分的方差累积贡献率需超过 85% 确定 m 值,即

$$\sum_{j=1}^m \lambda_j / \sum_{j=1}^p \lambda_j \geq 0.85. \tag{3}$$

最后, 矩阵 $\tilde{V} = ZT$ 即为原始样本数据经过降维处理得到的主属性的得分矩阵. 矩阵 \tilde{V} 中 \tilde{v}_j^i 表示方案 x_i 在第 j 个主属性的得分值.

1.3 TF-IDF 算法

TF-IDF (term frequency-inverse document frequency) 即词频-逆文档频率, 用以评估一个字词对于一个文件或一个语料库中的其中一份文件的重要程度. TF (词频) 是某一词语在该文本中出现的次数, IDF (逆向文件频率) 是一个词语普遍重要性的度量. 基本原理是: 一个词条在某个文本出现的频率越高, 说明该词条能够越好代表这类的文本特征, 这样的词条应该给它们赋予较高的权重, 并选择作为该类文本的特征词以区别于其他类文档. 被提取词条的权重公式定义为

$$\text{TF-IDF}(T) = \text{tf}_d(T) \cdot \log(N/\text{df}(T)). \quad (4)$$

其中: $\text{tf}_d(T)$ 为词频, 表示词条 T 在文本 d 中出现的频率; N 表示所有文本数; $\text{df}(T)$ 表示在文本集合中含有词条 T 的文本数量.

特大突发事件发生后, 社会群体会在第一时间内通过微博等公众平台发表自己的意见, 运用 python 自然语言技术处理最新时间段内的微博文本, 以获得与事件相关的关键词, 并将关键词筛选归类得到大群体应急决策的属性集, 再计算关键词 TF-IDF 值可获得各个属性的权重.

1.4 误差理论

假设有 p 个直接测量值 x_1, x_2, \dots, x_p , 间接测量值为 y , 它们的函数关系为

$$y = f(x_1, x_2, \dots, x_p), \quad (5)$$

则 y 的误差与 x_1, x_2, \dots, x_p 的误差之间的关系成为误差传递.

假设 x_1, x_2, \dots, x_p 的随机误差为 $\delta_{x_1}, \delta_{x_2}, \dots, \delta_{x_p}$, y 的随机误差为 δ_y , 则误差传递的公式为 $\delta_{y \text{ lim}}^2 = \sum_{i=1}^p \left(\frac{\partial f}{\partial x_i}\right)^2 \delta_{x_i \text{ lim}}^2 + 2 \sum_{1 \leq i < j \leq p} \frac{\partial f}{\partial x_i} \frac{\partial f}{\partial x_j} \rho_{ij} \delta_{x_i \text{ lim}} \delta_{x_j \text{ lim}}$. 其中: δ_{lim} 为随机变量的极限误差, ρ_{ij} 为直接测量值 x_i 和 x_j 的误差相关系数. 特别地, 当随机误差 δ_{x_i} 和 δ_{x_j} 相互独立时, $\rho_{ij} = 0$. 则上述公式可转化为

$$\delta_{y \text{ lim}}^2 = \sum_{i=1}^p \left(\frac{\partial f}{\partial x_i}\right)^2 \delta_{x_i \text{ lim}}^2. \quad (6)$$

误差理论与 PCA 模型相结合主要是考虑犹豫模糊元补充时产生的误差, 与协方差矩阵结合获得与原数据最为接近的 PC 系数和 PC 得分矩阵.

2 基于犹豫模糊语言的 PCA 模型

2.1 专家风险偏好测度及其聚类

2.1.1 专家风险偏好测度

定义 1^[18] 广义离散语言术语集定义为 $S = \{s_\alpha | \alpha = -\tau, \dots, -1, 0, 1, \dots, \tau\}$. 其中: S 是一个离散且下标对称的语言术语集; τ 是一个正整数, $2\tau + 1$ 表示语言术语集 S 的粒度. 本文中专家的风险偏好以风险偏好因子 (θ_i^1, θ_i^2) 的形式体现, 当风险偏好因子取不同的值时, 专家表现为不同的风险偏好类型. θ_i^1 表示专家 e_i 对“收益”的敏感程度, θ_i^2 表示专家 e_i 对“损失”的敏感程度. 较强风险偏好的专家对收益敏感度较大, 对损失敏感度较小, 因此较强风险偏好专家的 θ_i^1 取值较大, θ_i^2 取值较小, 风险规避专家的取值则相反^[19].

定义 2^[19] 若 R 是一个实数集, t 是一个实数, $\theta_i^1 \geq 0, \theta_i^2 \geq 0$, 广义 S 型语义函数定义为

$$g(t)_i = (1 + e^{-\theta_i^1 t})^{-1} \times I_{\{t \geq 0\}} + (1 + e^{-\theta_i^2 t})^{-1} \times I_{\{t < 0\}}. \quad (7)$$

其中

$$I_{\{t \in N\}} = \begin{cases} 1, & t \in \{N\}; \\ 0, & t \notin \{N\}. \end{cases}$$

$I_{\{t \in N\}}$ 是隶属于 R 的函数, N 是 R 的一个子集. $U_\mu = \{(s_t, u_t) | t = -\mu, \dots, 0, \dots, \mu; u_t = g(t)\}$ 为一个广义对称语言术语集, $S_1 = \{s_1, \dots, s_{\mu-1}, s_\mu\}$ 和 $S_2 = \{s_{-\mu}, \dots, s_{-2}, s_{-1}\}$ 分别为“收益”与“损失”两个不同方向的语言术语集, 语义函数 $g(t)_i = (1 + e^{-\theta_i^1 t})^{-1} \times 1_{\{t \geq 0\}} + (1 + e^{-\theta_i^2 t})^{-1} \times 1_{\{t < 0\}}$, 则可计算出每个语言术语所对应的语义值, $H = \{h(t_o), h(t_p), \dots, h(t_q)\}$ 是决策者提供的语言术语的主观值, 其中 $o, p, \dots, q \in [-t, t]$.

结合语义函数和专家提供的主观语义值, 利用欧氏距离函数建立一个模型

$$\begin{aligned} \min \{d_e(\theta_i^1, \theta_i^2, H)\} = & \min \{(|g(t_o) - h(t_o)|^2 + |g(t_p) - h(t_p)|^2 + \dots + |g(t_q) - h(t_q)|^2)^{1/2}\}; \\ \text{s.t. } & l_H \geq 3, 0 \leq \theta_i^1 \leq 3, 0 \leq \theta_i^2 \leq 3. \end{aligned} \quad (8)$$

其中 l_H 表示专家给出的主观语义值的个数. 求解上述问题获得专家的两个风险偏好因子值, 将该因子值代入式(7)可以得到专家的语义函数 $g(t)$.

2.1.2 专家聚类及权重求解

通过最优离散拟合方法得到每个专家两个不同方向的风险偏好因子值. 专家对“收益”和“损失”

两个方向的语言术语集至少提供3个主观语义值,将主观语义值以及相应的原始语义值代入模型(8)中,求解此模型即可得到专家的两个风险偏好因子值.例如专家给出的两个方向语言集的主观语义值分别为 $\{s_{-3} = 0.3, s_1 = 0.8, s_2 = 0.9\}$,代入模型(8)中计算得到专家的两个风险偏好因子值分别为 $\theta_1^1 = 1.26, \theta_1^2 = 0.10$.由文献[19]可知两个风险偏好因子的取值集中在 $[0, 3]$ 区间内,两个风险偏好因子的不同取值可对应不同风险偏好的专家.据此将决策者分为5类:强风险规避型(SA)、一般风险规避型(GA)、中立(N)、一般风险偏好型(GP)、强风险偏好型(SP).不同类型决策者的风险偏好因子区间范围如表1所示.

表1 不同风险偏好类型对应的风险偏好因子区间范围

风险偏好类型	θ_i^1	θ_i^2
SA	[0, 0.2]	(0.85, 3]
GA	(0.2, 0.4]	(0.55, 0.85]
N	(0.4, 0.55]	(0.4, 0.55]
GP	(0.55, 0.85]	(0.2, 0.4]
SP	(0.85, 3]	(0, 0.2]

根据群体成员偏好聚类的结果确定决策成员的权重.聚类的标准是决策成员的风险偏好,因此同一个聚集内决策成员的风险偏好相近,应赋予他们相同的权重.包含成员较多的聚集,其成员表达的决策信息代表了大多数成员的意见,根据多数原则将赋予其成员较大的权重.反之,决策成员较少的聚集将赋予其成员较小的权重.

聚集 $C^k (1 \leq k \leq K)$ 中包含 n_k 个成员,则聚集 C^k 中含 n_k 个成员的权重 ω_{n_k} 均相等.根据多数原则, ω_{n_k} 与 n_k 成正比,则成员 e_i 的权重 ω_i 可由下式确定:

$$\omega_i = \omega_{n_k} = n_k / \sum_{k=1}^K n_k^2, i = 1, 2, \dots, M. \quad (9)$$

决策专家的权重矢量为 $\omega = (\omega_1, \omega_2, \dots, \omega_M)$.

2.2 基于专家风险偏好的犹豫模糊元补充方法

在大群体应急决策中,犹豫模糊语言需补充较短的犹豫模糊元进行距离测度等运算.本文运用现有的3种常用的犹豫模糊元补充方法,分别是运用犹豫模糊元中的最大、均值、最小语言值进行补充.这些方法都存在一定的误差值,因此本文提出一种考虑专家风险偏好的犹豫模糊元补充方法,在一定程度上降低误差.

首先判断专家的风险偏好,若专家为风险偏好型,新的语言值从较大的语言值角度出发补充,则结合专家的风险偏好因子值,新的语言值为

$$s^* = \frac{\theta_i^1}{\sum_{i=1}^M \theta_i^1} s_{\max}; \quad (10)$$

若专家为风险规避型,新的语言值从较小的语言值角度出发补充,则结合专家的风险偏好因子值,新的语言值为

$$s^* = \frac{\theta_i^2}{\sum_{i=1}^M \theta_i^2} s_{\max}; \quad (11)$$

若专家为风险中立型,新的语言值从较大和较小语言值的均值角度补充,新的语言值为

$$s^* = \frac{s_{\max} + s_{\min}}{2}. \quad (12)$$

其中: s^* 为补充得到的语言值, s_{\max} 为犹豫模糊元中最大的语言值, s_{\min} 为犹豫模糊元中最小的语言值.

2.3 基于犹豫模糊语言的PCA模型

传统的PCA模型只适用于实数数据,而本文决策专家均使用犹豫模糊语言表达偏好信息,因此本文结合犹豫模糊元的补充方法和误差理论,建立一个新的基于犹豫模糊语言的PCA模型.本文将运用4种不同的犹豫模糊元补充方法,分别是均值法、最大值法、最小值法以及本文提出的考虑专家风险偏好的补充方法,这些方法在补充过程中会存在一定的误差,则考虑误差建立4种不同补充方法下的犹豫模糊语言PCA模型.

2.3.1 均值补充方法下的犹豫模糊语言PCA模型

首先使用均值方法补充犹豫模糊元,例如 $H(h_1) = \{s_1, s_3\}$ 和 $H(h_2) = \{s_{-2}, s_1, s_3\}$ 为两个犹豫模糊语言,为了方便比较 $H(h_1)$ 和 $H(h_2)$,需对 $H(h_1)$ 进行补充.相对应的语义值分别为 $H(h_1) = \{0.6, 0.8\}$, $H(h_2) = \{0.3, 0.5, 0.8\}$,则 $H(h_1)$ 中需要补充的语言值为 $s^* = \frac{1}{2}(0.6 + 0.8) = 0.7$,补充后的 $H(h_1)$ 为 $H(h_1)^* = \{0.6, 0.7, 0.8\}$.

假设原始犹豫模糊语言数据矩阵为 $\mathbf{X} = (x_{ij})_{n \times p}$,运用均值方法补充犹豫模糊元之后的矩阵为 $\mathbf{C} = (c_{ij})_{n \times p}$,有

$$\mathbf{C} = \begin{bmatrix} c_{11} & c_{12} & \dots & c_{1j} & \dots & c_{1p} \\ c_{21} & c_{22} & \dots & c_{2j} & \dots & c_{2p} \\ \vdots & \vdots & \ddots & \vdots & \ddots & \vdots \\ c_{i1} & c_{i2} & \dots & c_{ij} & \dots & c_{ip} \\ \vdots & \vdots & \ddots & \dots & \ddots & \vdots \\ c_{n1} & c_{n2} & \dots & c_{nj} & \dots & c_{np} \end{bmatrix},$$

其中 c_{ij} 为 x_{ij} 运用均值法补充后的犹豫模糊语言.取误差值为均值与最大或最小语言值的差,构成误差矩阵 $\mathbf{D} = (d_{ij})_{n \times p}$,有

$$D = \begin{bmatrix} d_{11} & d_{12} & \dots & d_{1j} & \dots & d_{1p} \\ d_{21} & d_{22} & \dots & d_{2j} & \dots & d_{2p} \\ \vdots & \vdots & \ddots & \vdots & \ddots & \vdots \\ d_{i1} & d_{i2} & \dots & d_{ij} & \dots & d_{ip} \\ \vdots & \vdots & \ddots & \dots & \ddots & \vdots \\ d_{n1} & d_{n2} & \dots & d_{nj} & \dots & d_{np} \end{bmatrix}$$

其中: $d_{ij} = (x_{ij}^U - x_{ij}^L)/2$, x_{ij}^U 和 x_{ij}^L 分别为犹豫模糊语言 x_{ij} 中最大和最小的语言值. 将传统 PCA 模型运用于矩阵 $C = (c_{ij})_{n \times p}$, 假设选取了前 m 个主成分 PCs, 则矩阵 C 的主成分得分矩阵为

$$Y^c = (y_{ik}^c)_{n \times p} = (C - \bar{C})T = \begin{bmatrix} c_{11} - \bar{c}_1 & c_{12} - \bar{c}_2 & \dots & c_{1j} - \bar{c}_j & \dots & c_{1p} - \bar{c}_p \\ c_{21} - \bar{c}_1 & c_{22} - \bar{c}_2 & \dots & c_{2j} - \bar{c}_j & \dots & c_{2p} - \bar{c}_p \\ \vdots & \vdots & \ddots & \vdots & \ddots & \vdots \\ c_{i1} - \bar{c}_1 & c_{i2} - \bar{c}_2 & \dots & c_{ij} - \bar{c}_j & \dots & c_{ip} - \bar{c}_p \\ \vdots & \vdots & \ddots & \dots & \ddots & \vdots \\ c_{n1} - \bar{c}_1 & c_{n2} - \bar{c}_2 & \dots & c_{nj} - \bar{c}_j & \dots & c_{np} - \bar{c}_p \end{bmatrix} \cdot \begin{bmatrix} t_{11} & t_{12} & \dots & t_{1k} & \dots & t_{1m} \\ t_{21} & t_{22} & \dots & t_{2k} & \dots & t_{2m} \\ \vdots & \vdots & \ddots & \vdots & \ddots & \vdots \\ t_{j1} & t_{j2} & \dots & t_{jk} & \dots & t_{jm} \\ \vdots & \vdots & \ddots & \dots & \ddots & \vdots \\ t_{p1} & t_{p2} & \dots & t_{pk} & \dots & t_{pm} \end{bmatrix} \quad (13)$$

其中: $y_{ik}^c = \sum_{j=1}^p (c_{ij} - \bar{c}_j)t_{jk}$, $i = 1, 2, \dots, n, j = 1, 2, \dots, p, k = 1, 2, \dots, m$; y_{ik}^c 表示矩阵 C 中第 k 个主成分得分值; \bar{c}_j 表示矩阵 C 中第 j 列的均值, 则均值矩阵 \bar{C} 为第 j 列矢量; T 为 PC 系数矩阵, 是矩阵 C 的协方差矩阵的正交特征向量.

本文中假设 c_{ij} 的随机误差 $\delta_{c_{ij}}$ 之间是线性无关的, 使用式 (6) 估计误差传递关系, 则随机误差 y_{ik}^c 可以表示为

$$\delta_{y_{ik}^c \text{ lim}}^2 = \sum_{j=1}^m \left(\frac{\partial y_{ik}^c}{\partial c_{ij}} \right) \delta_{c_{ij} \text{ lim}}^2 = \sum_{j=1}^m d_{ij}^2 t_{jk}^2 \quad (14)$$

则可得出

$$\delta_{y_{ik}^c \text{ lim}} = \sqrt{\sum_{j=1}^m d_{ij}^2 t_{jk}^2} \quad (15)$$

因此矩阵的极限误差矩阵可以表示为

$$\delta_{Y_{\text{lim}}^c}^2 = (\delta_{y_{ik}^c \text{ lim}}^2)_{n \times m} = \hat{D}\hat{T} = \begin{bmatrix} d_{11}^2 & d_{12}^2 & \dots & d_{1j}^2 & \dots & d_{1p}^2 \\ d_{21}^2 & d_{22}^2 & \dots & d_{2j}^2 & \dots & d_{2p}^2 \\ \vdots & \vdots & \ddots & \vdots & \ddots & \vdots \\ d_{i1}^2 & d_{i2}^2 & \dots & d_{ij}^2 & \dots & d_{ip}^2 \\ \vdots & \vdots & \ddots & \dots & \ddots & \vdots \\ d_{n1}^2 & d_{n2}^2 & \dots & d_{nj}^2 & \dots & d_{np}^2 \end{bmatrix}$$

$$\begin{bmatrix} t_{11}^2 & t_{12}^2 & \dots & t_{1k}^2 & \dots & t_{1m}^2 \\ t_{21}^2 & t_{22}^2 & \dots & t_{2k}^2 & \dots & t_{2m}^2 \\ \vdots & \vdots & \ddots & \vdots & \ddots & \vdots \\ t_{j1}^2 & t_{j2}^2 & \dots & t_{jk}^2 & \dots & t_{jm}^2 \\ \vdots & \vdots & \ddots & \dots & \ddots & \vdots \\ t_{p1}^2 & t_{p2}^2 & \dots & t_{pk}^2 & \dots & t_{pm}^2 \end{bmatrix} \quad (16)$$

因此, 通过公式 $\delta_{y_{ik}^c \text{ lim}}^c = \sqrt{\sum_{j=1}^m d_{ij}^2 t_{jk}^2}$ 可获得 $\delta_{Y_{\text{lim}}^c}^2 = (\delta_{y_{ik}^c \text{ lim}}^2)_{n \times m}$ 的具体值.

基于以上分析, 可以获得考虑误差的主成分得分矩阵为

$$\tilde{Y}^c = (\tilde{y}_{ik}^c)_{n \times m} = [Y^c - \delta_{Y_{\text{lim}}^c}, Y^c + \delta_{Y_{\text{lim}}^c}] \quad (17)$$

2.3.2 考虑风险偏好的补充方法下的犹豫模糊语言 PCA 模型

假设运用本文 2.2 节中的方法补充犹豫模糊元后得到的矩阵为 $U = (u_{ij})_{n \times p}$, 补充后产生的误差构成误差矩阵 $\hat{D} = (\hat{d}_{ij})_{n \times p}$, 其中 $\hat{d}_{ij} = \min((x_{ij}^U - x_{ij}^*), (x_{ij}^U - x_{ij}^*))$, 因为补充值更加偏向于专家的风险偏好, 符合专家的心理态度. 因此矩阵 U 的主成分得分矩阵为 $Y^u = (y_{ik}^u)_{n \times p} = (U - \bar{U})\hat{T}$, 其中 $y_{ik}^u = \sum_{j=1}^p (u_{ij} - \bar{u}_j)\hat{t}_{jk}$, y_{ik}^u 表示矩阵 U 中第 k 个主成分得分值, \bar{u}_j 表示矩阵中第 j 列的均值. 则随机误差 y_{ik}^u 可以

表示为 $\delta_{y_{ik}^u \text{ lim}} = \sqrt{\sum_{j=1}^m \hat{d}_{ij}^2 \hat{t}_{jk}^2}$, 因此 Y^u 的极限误差可以表示为 $\delta_{Y_{\text{lim}}^u}^2 = (\delta_{y_{ik}^u \text{ lim}}^2)_{n \times m_1} = \hat{D}\hat{T}$. 基于以上分析, 可以获得考虑误差的主成分得分矩阵为

$$\tilde{Y}^u = (\tilde{y}_{ik}^u)_{n \times m_1} = [Y^u - \delta_{Y_{\text{lim}}^u}, Y^u + \delta_{Y_{\text{lim}}^u}]$$

对比均值补充法和考虑专家风险偏好的补充法可以发现, 两种补充方法下的 PCA 模型, 主要区别在于补充值和误差取值的不同. 4 种补充方法下的误差取值如表 2 所示.

表 2 不同补充方法的误差取值对比

补充方法	误差取值
均值法	$d_{ij} = (x_{ij}^U - x_{ij}^L)/2$
本文方法	$\min((x_{ij}^U - x_{ij}^*), (x_{ij}^U - x_{ij}^L))$
最大值法	$d_{ij} = (x_{ij}^U - x_{ij}^L)$
最小值法	$d_{ij} = (x_{ij}^U - x_{ij}^L)$

综上所述, 基于犹豫模糊语言的 PCA 模型如图 1 所示.

2.4 偏好信息融合

通过犹豫模糊语言的 PCA 模型, 可获得 m 个不同的特征值, 则对应 m 个主属性, 运用方差贡献率来计算主属性的权重, 方差贡献率越大, 主属性越具有

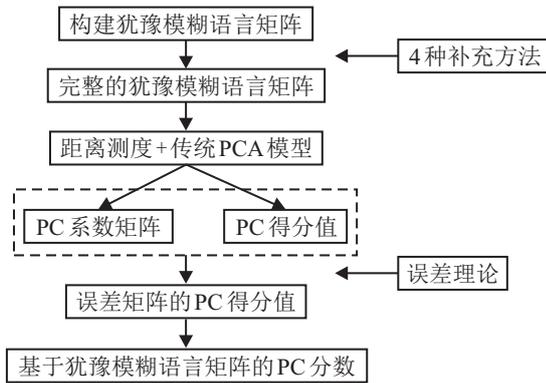


图1 基于犹豫模糊语言的PCA模型流程

代表性,权重越大,则主属性的权重可通过其相应的方差贡献率来表示,有

$$w_j = \frac{\lambda_j}{\sum_{j=1}^m \lambda_j} \quad (18)$$

结合主属性的权重,对方案*l*的主属性进行加权求和,即得专家群体对方案 Q_i 的最终评价值为

$$Q_l = \sum_{j=1}^m w_j \cdot \hat{v}_l^j \quad (19)$$

3 决策步骤

文中专家以犹豫模糊语言给出每个属性在各个方案的偏好信息,文中应用的语言术语为 $S = \{s_{-4} = \text{极差}, s_{-3} = \text{非常差}, s_{-2} = \text{差}, s_{-1} = \text{有点差}, s_0 = \text{平等}, s_1 = \text{有点好}, s_2 = \text{好}, s_3 = \text{非常好}, s_4 = \text{极好}\}$,语言粒度为9,相应的语义值分别为 $\{0.1, 0.2, 0.3, 0.4, 0.5, 0.6, 0.7, 0.8, 0.9\}$.则具体的大群体应急决策步骤如下.

step 1: 提取属性并获得权重.首先,收集突发事件发生后一段时间内的微博文本,将每个文本经过分词、清洗、词性标注与实体词识别预处理,获得与事件相关的关键词,并利用式(4)获得关键词的权值,将关键词筛选归类获得事件的属性,利用关键词TF-IDF值的均值获得属性权重.

step 2: 专家风险偏好测度及聚类.利用最优离散拟合方法,结合式(7)和模型(8)计算出专家的两个风险偏好因子值.然后根据表1将决策大群体专家 $E = e_1, e_2, \dots, e_p (P \leq 11)$ 分为5个聚集:强风险规避型(C^{SA}),一般风险规避型(C^{GA}),中立(C^N),一般风险偏好型(C^{GP}),强风险偏好型(C^{SP}).专家聚类后利用式(9)计算出每个专家的权重.

step 3: 运用4种犹豫模糊元补充方法补充偏好信息.通过均值法补充犹豫模糊元,获得完全偏好信息矩阵.运用本文的考虑专家风险偏好的补充方法以及最大值法、最小值法补充犹豫模糊元,分别获得

完整的犹豫模糊偏好信息矩阵 $U = (u_{ij})_{n \times p}$ 、 $Z = (z_{ij})_{n \times p}$ 、 $L = (l_{ij})_{n \times p}$.

step 4: 应用基于犹豫模糊语言的PCA模型.将通过4种方法得到的完全偏好信息矩阵 $C = (c_{ij})_{n \times p}$ 、 $U = (u_{ij})_{n \times p}$ 、 $Z = (z_{ij})_{m \times p}$ 、 $L = (l_{ij})_{n \times p}$ 应用于传统的PCA模型中,运用文献[20]中的犹豫模糊语言之间的距离公式以及式(18)获得完全偏好信息的PC系数和得分数 $Y^c = (y_{ik}^c)_{n \times m}$ 、 $Y^u = (y_{ik}^u)_{n \times m_1}$ 、 $Y^z = (y_{ik}^z)_{n \times m_2}$ 、 $Y^l = (y_{ik}^l)_{n \times m_3}$,并根据式(3)计算获得主属性(主成分)的个数.根据表2中4种补充方法的误差取值分别计算获得4个误差矩阵 $D = (d_{ij})_{n \times p}$ 、 $\dot{D} = (\dot{d}_{ij})_{n \times p}$ 、 $\vec{D} = (\vec{d}_{ij})_{n \times p}$ 、 $\overleftarrow{D} = (\overleftarrow{d}_{ij})_{n \times p}$.其中不需补充的犹豫模糊元的误差为0,将误差矩阵运用到传统的PCA模型中,运用式(15)和(16)获得误差矩阵的PC得分数 $\delta_{Y_{lim}^c}$ 、 $\delta_{Y_{lim}^u}$ 、 $\delta_{Y_{lim}^z}$ 、 $\delta_{Y_{lim}^l}$.

step 5: 偏好信息融合.将偏好信息矩阵的PC得分数和误差矩阵的PC得分数相结合,运用式(18)获得最终的犹豫模糊语言的PC分数 $\tilde{Y}^c = (\tilde{y}_{ik}^c)_{n \times m}$ 、 $\tilde{Y}^u = (\tilde{y}_{ik}^u)_{n \times m_1}$ 、 $\tilde{Y}^z = (\tilde{y}_{ik}^z)_{n \times m_2}$ 、 $\tilde{Y}^l = (\tilde{y}_{ik}^l)_{n \times m_3}$.

step 6: 方案排序.运用式(19)计算主成分也即主属性的权重,并结合PC分数与式(20)获得最终专家大群体对每个方案的综合评估值.

综上,大群体决策的具体流程如图2所示.

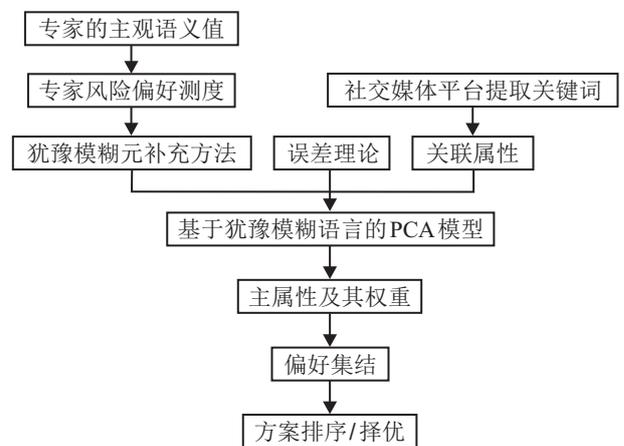


图2 大群体应急决策中考虑属性关联的信息融合方法流程

4 案例分析

4.1 案例描述与应用

以2019年5月广西省洪涝灾害事件为例,对本文所提出方法的可行性和有效性进行验证.5月25日以来,广西多地遭受强降雨袭击.截至28日17时,洪涝灾害已造成全区195673人受灾,因灾死亡7人,直接经济损失13358.99万元.事件发生时,政府立即成立事故紧急处理小组,相关部门迅速成立应急决策救援指挥中心,召集公安消防部门、医疗部门、供电局、环

保局等领域的专家, 结合实际情况迅速形成了 5 个应急方案 $X = \{x_1, x_2, x_3, x_4, x_5\}$.

方案 1(x_1): 及时了解交通、通信、供电、供水、医疗等设施的破坏情况, 掌握实时状况, 组织水上救援, 保障人员和财产的安全, 尽快恢复交通安全, 保障抗灾工作的顺利进行.

方案 2(x_2): 做好侦查检测工作, 查看河堰堤坝、道路、桥梁等是否有决堤和垮塌的危险. 组织水上救援突击队, 积极营救被困在水中、屋顶等险恶环境的遇险人员, 搜寻失踪人员, 加固堤坝、桥梁等重要建筑.

方案 3(x_3): 确定救援路线, 交通部门及时修复被毁坏的的道路, 保证救援车队及装备安全抵达救援地点, 积极营救被困人员. 停止危险作业, 组织危险地区、危险地段人员安全转移. 减少群众恐慌, 对疏散撤离人员进行慰问, 平复他们的情绪. 排查危险作业, 清理灾后垃圾, 指导做好企业和农户灾后复产工作, 防范次生灾害.

方案 4(x_4): 组织水上救援部队紧急救援, 及时清点人员, 检查有无受伤和失踪的人员, 医疗队员及时抢救危、急伤员. 有效处置危险化学品泄露等次生灾害. 加强卫生防疫工作, 严防灾后疫情的传染和传播.

方案 5(x_5): 按照便于撤退的原则, 选择救援人员、车辆、物资的集结地. 在水域全面开展搜寻, 根据灾情的急缓程度灵活地组织实施受灾群众的解救工作, 将灾民群众转移至预定安置点.

为选取一个最为合理的方案, 特邀请 50 位来自不同领域的应急决策专家 $E = \{e_1, e_2, \dots, e_{50}\}$, 对案例实际背景的真实场景进行模拟, 根据当时洪涝灾害的实时情况对每个属性下的各个方案给出评估信息, 形成本文的案例数据. 为在最短时间内选择出最优方案, 决策具体步骤如下.

step 1: 提取事件属性并获得属性权重. 观察发现, 当微博针对此次广西洪涝灾害事件形成热点话题时, 文本数量会极速增长, 关键词会在相关微博中重复出现, 短时间内高频增长. 本文使用 2019 年 5 月 25 日 20 时至 28 日 15 时捕获的 50 000 条新浪微博文本数据, 作为获取关键词的原始数据. 将获得的所有实体词进行 TF-IDF 计算, 选取其中 600 个 TF-IDF 值最高的词条作为关键词, 再对所提取的关键词出现在同一文本中的词频进行关联度分析.

对关键词进行筛选归类, 关联性较强的归为一个属性类. 比如“伤亡”“祈福”“致哀”等关键词之间有很大的关联性, 则将这一类关键词归为属性“人员伤亡”; “洪水”“淹没”等关键词的关联度也很大, 将这

一类的关键词归为“环境损害”. 按照此种方法对关键词进行归类, 最终得到 10 个属性 $U = \{u_1, u_2, u_3, u_4, u_5, u_6, u_7, u_8, u_9, u_{10}\}$, 分别为“救援成本”“人员伤亡”“环境损害”“救援速度”“灾后重建”“经济损失”“疾病预防”“公众反应”“社会影响”“衍生灾害”. 同时, 关键词“伤亡”与“淹没”等关键词也有一定的关联, 因此所得的 10 个属性并不是相互独立的, 具有一定的关联性, 每个属性包含的关键词的数目分别为 30、102、36、90、18、72、54、60、60、78. 50 位专家在决策过程中以此 10 个属性为标准对 5 个备选方案进行评估. 将关键词对应的 TF-IDF 值归一化处理得到每个属性的权值, 如表 3 所示.

表 3 属性权重

属性 u_j	u_1	u_2	u_3	u_4	u_5	u_6	u_7	u_8	u_9	u_{10}
属性权重 w_j	0.05	0.17	0.06	0.15	0.03	0.12	0.09	0.10	0.10	0.13

step 2: 专家风险偏好测度及聚类. 专家风险偏好测度. 专家们分别给出至少 3 个主观语义值, 本文中取值 $\tau = 4$, 则语言术语为 $S = \{s_{-4}, s_{-3}, s_{-2}, s_{-1}, s_0, s_1, s_2, s_3, s_4\}$, 50 位专家给出的主观语义值分别如表 4 所示(只列出部分专家的语义值).

表 4 大群体专家的主观语义值

s_{-3}	s_{-2}	s_1	s_{-1}	s_0	s_2	s_{-4}	s_{-3}	s_1	s_{-4}	s_1	s_2				
e_1	0.3	0.4	0.7	e_2	0.1	0.2	0.5	e_3	0.2	0.3	0.7	e_4	0.1	0.6	0.7
⋮															
s_{-3}	s_1	s_2	s_{-4}	s_{-1}	s_2	s_{-2}	s_3	s_4	s_{-2}	s_{-1}	s_2				
e_{47}	0.4	0.7	0.9	e_{48}	0.1	0.5	0.7	e_{49}	0.1	0.4	0.6	e_{50}	0.4	0.5	0.8

将专家的主观语义值代入模型 (8) 中, 得出每个专家的风险偏好因子 θ_i^1 、 θ_i^2 , 具体如表 5 所示(部分值).

表 5 专家大群体的风险偏好因子值

	e_1	e_2	e_3	e_4	e_5	e_6	e_7	e_8	e_9	e_{10}
θ_i^1	0.84	0.01	0.84	0.42	1.26	1.39	0.20	0.01	0.74	0.14
θ_i^2	0.25	2.19	0.31	0.55	0.10	0.18	0.71	1.26	0.20	2.19

与风险偏好类型的因子值范围相对应, 将专家的风险偏好类型进行细分, 结果如表 6 所示.

表 6 专家聚类结果

风险偏好类型	成员个数	成员 e_i
SA	8	$e_{12}, e_{20}, e_{22}, e_{23}, e_{28}, e_{33}, e_{42}, e_{49}$
GA	8	$e_2, e_4, e_6, e_{24}, e_{25}, e_{26}, e_{30}, e_{43}$
N	16	$e_3, e_7, e_8, e_{10}, e_{13}, e_{29}, e_{32}, e_{35}, e_{37}, e_{38}, e_{40}, e_{41}, e_{46}, e_{48}$
GP	11	$e_1, e_5, e_{17}, e_{18}, e_{19}, e_{27}, e_{45}, e_{30}, e_{39}, e_{47}, e_{36}$
SP	7	$e_9, e_{11}, e_{15}, e_{16}, e_{21}, e_{44}, e_{50}$

将专家 e_1 的两个风险偏好因子值 $\theta_1^1 = 0.84$ 和 $\theta_1^2 = 0.25$ 代入式(7)中即可获得其语义函数为

$$g(t)_1 = (1 + e^{-0.84t})^{-1} \times 1_{\{t|t \geq 0\}} + (1 + e^{-0.25t})^{-1} \times 1_{\{t|t < 0\}}.$$

运用2.1节的方法进行风险偏好测度,进而将语言信息运用语义函数转化为实值,则专家的犹豫模糊语言矩阵可转化为实数矩阵. 专家 e_1 给出的偏好矩阵为 A_1 .

$$A^1 = \begin{matrix} x_1 \\ x_2 \\ x_3 \\ x_4 \\ x_5 \end{matrix} \begin{bmatrix} (s_{-4}, s_{-2}, s_{-1}) & (s_{-2}, s_2) & (s_1, s_2) \\ (s_0, s_2) & (s_1, s_4) & (s_3, s_4) \\ (s_{-2}, s_0, s_1) & (s_{-3}, s_1) & (s_2, s_4) \\ (s_{-1}, s_2) & (s_{-3}, s_1) & (s_1, s_3) \\ (s_{-2}, s_{-1}) & (s_{-1}, s_2) & (s_1, s_3) \end{bmatrix} \rightarrow \begin{matrix} (s_1, s_3) & (s_{-3}, s_0) & (s_{-2}, s_0) & (s_{-1}, s_1) \\ (s_0, s_2) & (s_{-2}, s_1) & (s_{-1}, s_1) & (s_{-1}, s_0) \\ (s_{-1}, s_1) & (s_{-4}, s_{-3}, s_0) & (s_{-4}, s_1) & (s_{-4}, s_{-2}) \\ (s_{-1}, s_1) & (s_{-4}, s_{-3}, s_0) & (s_{-4}, s_1) & (s_{-4}, s_{-2}) \\ (s_{-1}, s_2) & (s_{-3}, s_0) & (s_{-2}, s_0) & (s_0, s_2) \end{matrix} \leftarrow \begin{matrix} (s_0, s_1) & (s_1, s_3) & (s_{-2}, s_1) \\ (s_1, s_2) & (s_{-2}, s_2) & (s_0, s_2) \\ (s_{-4}, s_0) & (s_1, s_4) & (s_0, s_2) \\ (s_{-4}, s_0) & (s_1, s_4) & (s_0, s_2) \\ (s_1, s_4) & (s_1, s_3) & (s_{-2}, s_0, s_3) \end{matrix}.$$

根据表5中专家的两个风险偏好因子值,将专家的风险因子值代入式(7)即可得到语义函数,进而将犹豫模糊语言矩阵转化为实数矩阵. 根据2.1节专家

的权重计算方法得出专家的最终权重值.

step 3: 运用新的犹豫模糊元补充方法补充偏好信息,按照2.3节的方法获得4种犹豫模糊元补充方法下的PC得分矩阵,确定方案评估区间值,再运用文献[21]中区间数的排列公式分别对最终分数矩阵的区间数排序,获得方案排序结果如表7所示.

表7 不同补充方法的方案排序结果

补充方法	方案排序结果
均值法	$x_3 \succ x_4 \succ x_1 \succ x_5 \succ x_2$
本文方法	$x_3 \succ x_4 \succ x_1 \succ x_2 \succ x_5$
最大值法	$x_1 \succ x_3 \succ x_4 \succ x_5 \succ x_2$
最小值法	$x_4 \succ x_3 \succ x_1 \succ x_2 \succ x_5$

根据均值法和本文的犹豫模糊元补充方法所得,选择方案 x_3 进行决策.

4.2 案例结果分析

本节主要从3个方面进行结果分析,分别是:1)4种犹豫模糊元补充方法的偏好信息对比;2)4种误差取值矩阵对比;3)最终方案排序结果的对比分析.

1)采用4种犹豫模糊元补充方法的偏好信息对比(为方便比较,只列出专家 e_1 针对方案 x_1 在各属性下的偏好信息).

从表8可以清晰地看出:针对属性 u_2 下的偏好信息,均值法补充的值为0.61,本文方法为0.67,最大值法为0.84,最小值法为0.38. 显而易见,最大值、最小值法获得的补充值和均值法、本文方法的补充值相差较大. 从其余属性的偏好信息数据也可以看到这一特点,这也与最后方案排序结果差别相一致.

表8 4种补充方法下方案 x_1 的偏好信息矩阵

	u_1	u_2	u_3	u_4	u_5
均值法	(0.27, 0.38, 0.44)	(0.38, 0.61, 0.84)	(0.40, 0.62, 0.84)	(0.70, 0.80, 0.90)	(0.38, 0.44, 0.50)
本文方法	(0.27, 0.38, 0.44)	(0.38, 0.67, 0.84)	(0.40, 0.67, 0.84)	(0.70, 0.84, 0.90)	(0.32, 0.46, 0.50)
最大值法	(0.27, 0.38, 0.44)	(0.38, 0.84, 0.84)	(0.40, 0.84, 0.84)	(0.70, 0.90, 0.90)	(0.32, 0.50, 0.50)
最小值法	(0.27, 0.38, 0.44)	(0.38, 0.38, 0.84)	(0.40, 0.40, 0.84)	(0.70, 0.70, 0.90)	(0.32, 0.32, 0.50)
	u_6	u_7	u_8	u_9	u_{10}
均值法	(0.32, 0.41, 0.50)	(0.44, 0.57, 0.70)	(0.50, 0.60, 0.70)	(0.70, 0.80, 0.90)	(0.38, 0.54, 0.70)
最大算法	(0.38, 0.46, 0.50)	(0.44, 0.63, 0.70)	(0.50, 0.63, 0.70)	(0.70, 0.84, 0.90)	(0.38, 0.63, 0.70)
本文方法	(0.38, 0.50, 0.50)	(0.44, 0.70, 0.70)	(0.50, 0.70, 0.70)	(0.70, 0.90, 0.90)	(0.38, 0.70, 0.70)
最小算法	(0.38, 0.38, 0.50)	(0.50, 0.50, 0.70)	(0.70, 0.70, 0.90)	(0.38, 0.38, 0.70)	(0.38, 0.38, 0.70)

2)采用4种补充方法相对应的误差取值对比(为方便比较,只列出专家 e_1 针对方案 x_1 在各属性下的误差取值).

本文基于犹豫模糊语言的PCA模型和传统PCA的主要差别是协方差矩阵的获取,因此在犹豫模糊元距离求解过程中会产生一定的误差. 由表9可知,专家 e_1 针对方案 x_1 在属性 u_1 下给出的犹豫模糊元长

度为3,其余长度为2,因此属性 u_1 下的犹豫模糊元不需要补充误差取值为0. 分析属性下 u_2 的误差取值,均值法误差值为0.23,本文方法误差值为0.17,最大值法误差值为0.46,最小值法误差值为0.46. 由此可知,均值法和本文方法误差取值较小,最大值和最小值法误差取值都相对偏大,而本文方法相对于均值法误差取值更小,说明了本文补充方法的有效性.

表 9 4 种补充方法下方案 x_1 的误差取值矩阵

	u_1	u_2	u_3	u_4	u_5	u_6	u_7	u_8	u_9	u_{10}
均值法	0	0.23	0.22	0.10	0.09	0.06	0.13	0.10	0.10	0.16
本文方法	0	0.17	0.17	0.06	0.04	0.04	0.07	0.07	0.06	0.07
最大值法	0	0.46	0.44	0.20	0.18	0.12	0.26	0.20	0.20	0.32
最小值法	0	0.46	0.44	0.20	0.18	0.12	0.26	0.20	0.20	0.32

3) 最终方案排序结果对比分析.

由表 7 可知, 4 种不同的犹豫模糊元补充方法下决策得到 4 种不同的方案排序结果. 针对本文方法得出的方案结果可以看出, 本文所提出的犹豫模糊元补充方法和均值法下得出的最优结果一致, 均为方案 x_3 , 最大值法和最小值法得出的误差最大, 方案排序很不相同, 因此也说明本文提出的犹豫模糊元的有效性, 以及属性关联背景下对多维属性降维的必要性.

运用最大值法或者最小值法进行犹豫模糊元补充时, 误差为最大值与最小值之差, 相对于均值法补值产生的误差是偏大的, 因此最终的决策结果是不具有信服力的. 最终的排序结果包括犹豫模糊元补充后的偏好信息矩阵的 PC 得分数和误差矩阵的 PC 得分数, 基于上述 1) 和 2) 的对比分析, 均值法和本文方法补充模糊元得到的排序结果比最大值法和最小值法得到的排序结果更具有可靠性, 而本文方法相对于均值方法的排序结果更加符合专家群体的风险偏好, 更具有说服力, 误差也更小.

4.3 对比分析

此外, 为了说明本文方法的优势和有效性, 将本文的方法与文献 [1] 和文献 [22] 进行对比, 将本文中的数据运用文献 [1] 和文献 [22] 中所提方法进行计算, 得出的方案排序结果如表 10 所示.

表 10 方案排序结果对比

	方案排序
本文方法	$x_3 \succ x_4 \succ x_1 \succ x_2 \succ x_5$
文献 [1] 方法	$x_3 \succ x_4 \succ x_2 \succ x_1 \succ x_5$
文献 [22] 方法	$x_1 \succ x_3 \succ x_5 \succ x_4 \succ x_2$

文献 [1] 方法获得最佳方案为 x_3 , 与本文的最佳方案一致, 说明了本文所提方法的合理性和有效性. 但是方案 x_1 与方案 x_2 排序不一致, 主要原因有以下几点:

1) 文献 [1] 给出了一种具有参考点且考虑属性关联的突发事件应急风险决策方法, 文中运用离散 Choquet 积分处理属性关联的情况. 但是文献 [1] 中属性之间的冗余关系是直接假定的, 未给出任何参考依据, 其案例分析中也是直接假定属性的权重值. 但是本文认为属性之间的关联程度应该由突发事件的具体情况而定, 从事件中提取属性计算属性之间的关联性才具有实际意义和价值, 直接假定属性间的关联性

是偏离实际且不具有客观性和合理性的.

2) 本文方法在专家权重和偏好信息方面都考虑了专家的风险偏好, 而文献 [1] 中风险主要体现在决策矩阵上, 在构建决策者的感知价值矩阵时, 直接自定义所有专家的失望参数和欣喜参数均为 0.8, 认为当偏好值低于失望参数时感知的失望较小, 高于欣喜参数时感知的欣喜较小, 表明文献 [1] 认为所有专家对于失望和欣喜的感知是一样的. 但是本文中认为专家对于失望和欣喜两个方向的风险感知态度是不同的, 本文中 θ_i^1 表示专家 e_i 对“收益”的敏感程度, θ_i^2 表示专家 e_i 对“损失”的敏感程度, 这两个风险偏好因子值是不相同的. 此外, 文献 [1] 中参数值的给定缺乏仿真实验和严格的理论依据, 但是本文中每个专家的风险态度都可以准确地衡量, 不需要自定义参数值, 且更加贴近专家的心理风险偏好.

3) 文献 [1] 中忽略了专家权重的确定, 为统一方案排序的标准, 计算时赋予专家相同的权重. 但是现实决策中决策专家的风险偏好、专业程度和经验均不同, 其权重也应不同. 此外, 文献 [1] 中 Choquet 积分算子并不能适用于犹豫模糊语言在运用本文数据进行计算时, 需在最初时即运用本文方法将犹豫模糊语言转化为实数, 进而进行计算, 此过程会造成更大的数据丢失.

为了再次说明本文方法的有效性, 将本文方法和文献 [22] 方法进行对比, 文献 [22] 方法获得最佳方案为 x_1 , 且方案排序结果与本文方法基本不一致. 出现这种排序的主要原因在于文献 [22] 在犹豫模糊语言环境下运用目标规划模型确定了属性权重, 但是文中过程是假设属性之间是相互独立的, 而本文中考虑了属性之间的关联性, 出现方案排序的变化是合理的. 但是在现实决策中属性之间会存在一定的关联性, 文献 [22] 的决策方法会与准确的决策结果产生很大误差, 也直接说明在群决策中考虑属性关联十分必要.

5 结论

本文针对大群体决策中出现的属性关联问题提出了一种大群体应急决策信息融合方法, 主要创新点为: 1) 考虑了专家的风险偏好, 并据此提出了一种新的犹豫模糊元补充方法; 2) 提出了基于犹豫模糊语言的 PCA 模型.

对于属性关联下的信息融合问题, 基于犹豫模糊语言的 PCA 模型能够处理基于犹豫模糊语言的偏好信息, 现有研究中 PCA 模型只拓展到区间数, 其他偏好表达方式下如何运用 PCA 模型尚未研究, 之后的研究会更多方面地考虑多种偏好表达方式下的 PCA 模型.

参考文献(References)

- [1] 曹萍萍, 李铭洋, 郭燕青. 考虑参照点且属性关联的应急风险决策方法[J]. 信息与控制, 2018, 47(4): 473-478.
(Cao P P, Li M Y, Guo Y Q. An emergency risk decision method that considers reference points and attribute associations[J]. Information and Control, 2018, 47(4): 473-478.)
- [2] Liu B S, Shen Y H, Zhang W, et al. An interval-valued intuitionistic fuzzy principal component analysis model-based method for complex multi-attribute large-group decision-making[J]. European Journal of Operational Research, 2015, 245(1): 209-225.
- [3] 王大澳, 菅利荣, 刘思峰. 基于 Choquet 积分的多属性灰靶群决策方法[J]. 控制与决策, 2017, 32(7): 1286-1292.
(Wang D O, Jian L R, Liu S F. Multi-attribute grey target group decision making method based on Choquet integral[J]. Control and Decision, 2017, 32(7): 1286-1292.)
- [4] 陈振颂, 李延来. 基于 ITFN 信息关联输入的改进群体 MULTIMOORA 决策方法[J]. 系统工程理论与实践, 2014, 34(12): 3138-3156.
(Chen Z S, Li Y L. Improved MULTIMOORA decision making method based on ITFN information correlation input[J]. System Engineering—Theory & Practice, 2014, 34(12): 3138-3156.)
- [5] Qu G H, Li Y L J, Qu W, et al. Some new Shapley dual hesitant fuzzy Choquet aggregation operators and their applications to multiple attribute group decision making-based TOPSIS[J]. Journal of Intelligent & Fuzzy Systems, 2017, 33(4): 2463-2483.
- [6] Khan M S A, Abdullah S, Ali A, et al. Pythagorean hesitant fuzzy Choquet integral aggregation operators and their application to multi-attribute decision-making[J]. Soft Computing, 2019, 23(1): 251-267.
- [7] Wan S P, Dong J Y. A multi-attribute decision making method based on choquet integral operator with triangular intuitionistic fuzzy numbers[J]. Chinese Journal of Management Science, 2014, 22(3): 121-129.
- [8] 陈岩, 李庭. 基于 Choquet 积分的直觉不确定语言信息集结算子及其应用[J]. 控制与决策, 2016, 31(5): 842-852.
(Chen Y, Li T. An intuitionistic uncertain language information aggregation operator based on Choquet integrals and its application[J]. Control and Decision, 2016, 31(5): 842-852.)
- [9] Aggarwal M, Fallah Tehrani A. Modeling human decision behavior with preference learning[J]. Inform Journal on Computing, 2019, 31(2): 318-334.
- [10] Ju Y, Yang S, Liu X. Some new dual hesitant fuzzy aggregation operators based on Choquet integral and their applications to multiple attribute decision making[J]. Journal of Intelligent & Fuzzy Systems, 2014, 27(6): 2857-2868.
- [11] 王志良. 水资源管理多属性决策与风险分析理论方法及应用研究[D]. 成都: 四川大学, 2003.
(Wang Z L. The theoretical method and application of multi-attribute decision making and risk analysis in water resources management[D]. Chengdu: Sichuan University, 2003.)
- [12] Tan C Q, Ma B J, Chen X H. Group decision making method based on Choquet integration operator[J]. Journal of Industrial Engineering and Engineering Management, 2010, 24(4): 155-160.
- [13] Gioia F, Lauro C N. Principal component analysis on interval data[J]. Computational Statistics, 2006, 21(2): 343-363.
- [14] Liu B S, Shen Y H, Chen X H, et al. A partial binary tree DEA-DA cyclic classification model for decision makers in complex multi-attribute large-group interval-valued intuitionistic fuzzy decision-making problems[J]. Information Fusion, 2014, 18: 119-130.
- [15] Lauro C H, Palumbo F. Principal components analysis of interval data: Asymbolic data analysis approach[J]. Computational Statistics, 2000, 15(1): 73-87.
- [16] Douzal-Chouakria A, Billard L, Diday E. Principal component analysis for interval-valued observations[J]. Statal Analysis and Data Mining, 2011, 4(2): 229-246.
- [17] Song G X, Yang H. Analysis of decision behavior in group decision making[J]. Academic Research, 2000(3): 48-49.
- [18] Zhou W, Xu Z S. Generalized asymmetric linguistic term set and its application to qualitative decision making involving risk appetites[J]. European Journal of Operational Research, 2016, 254(2): 610-621.
- [19] Zhou W, Xu Z S. Asymmetric fuzzy preference relations based on the generalized sigmoid scale and their application in decision making involving risk appetites[J]. IEEE Transactions on Fuzzy Systems, 2016, 24(3): 741-756.
- [20] Liao H C, Xu Z S, Zeng X J, et al. Qualitative decision making with correlation coefficients of hesitant fuzzy linguistic term sets[J]. Knowledge-Based Systems, 2015, 76: 127-138.
- [21] Torra V. The weighted OWA operator[J]. International Journal of Intelligent Systems, 1997, 12: 153-166.
- [22] 陈秀明, 刘业政. 多粒度犹豫模糊语言环境下未知权重的多属性群推荐方法[J]. 控制与决策, 2016, 31(9): 1631-1637.
(Chen X M, Liu Y Z. Method of group recommender systems with unknown attribute weights in a multi-granular hesitant fuzzy linguistic term environment[J]. Control and Decision, 2016, 31(9): 1631-1637.)

作者简介

徐选华(1962—), 男, 教授, 博士生导师, 从事大数据决策理论与方法、信息系统与决策支持系统等研究, E-mail: xuxh@csu.edu.cn;

余艳粉(1993—), 女, 硕士生, 从事大数据决策理论与方法、应急管理决策的研究, E-mail: 2733983727@qq.com.

(责任编辑: 齐 霖)