

文章编号: 1001-0920(2012)02-0271-05

基于灰色聚类的大规模群体语言评价信息集结研究

王嵩华, 朱建军, 方志耕

(南京航空航天大学 经济与管理学院, 南京 210016)

摘要: 研究群体规模较大的情况下基于语言评价信息的决策方法, 提出了基于灰色关联聚类的方法以降低大规模群体的协调困难; 根据决策者所属群体类别差异, 基于语言信息转化函数提出了类内评价指标权重确定方法, 并测算了各群体类别的综合偏好; 研究群体类间权重确定模型, 建立了基于群体类别意见偏差最小的群体偏好集结模型; 进而提出了决策者判断一致度定义以及决策者偏差协调方法. 算例分析表明了该方法的应用步骤及可行性.

关键词: 群决策; 语言评价信息; 大规模; 集结

中图分类号: C943

文献标识码: A

Group aggregation method on large-scale linguistic evaluation information based on grey cluster

WANG He-hua, ZHU Jian-jun, FANG Zhi-geng

(College of Economics and Management, Nanjing University of Aeronautics and Astronautics, Nanjing 210016, China.

Correspondent: WANG He-hua, E-mail: wanghehua@nuaa.edu.cn)

Abstract: The aggregation method for a large-scale group decision making problem is studied. Firstly, the grey cluster method is used to reduce the negotiation difficulty of a group. Based on the transformation formula of a linguistic evaluation, the method on the weight of criteria inside of the group is developed according to the difference of the group. Then the integration preference of each category is measured. A method of the weight of category is proposed and the group preference aggregation model is established based on the smallest deviation object among the categories. The consensus degree of decision maker is defined. Furthermore, the coordination method is proposed to achieve the coherent result of the group. The example analysis shows the steps and the feasibility of the proposed method.

Key words: group decision making; linguistic evaluation information; large scale; aggregation

1 引言

在决策过程中, 由于客观事物具有不确定性、复杂性和模糊性, 为了准确表达评价意见, 决策者常以语言评价信息的方式给出自己对事物的认识与判断. 近年来, 很多学者致力于语言评价信息的决策方法研究. 在语言评价信息转化处理方面, 主要采用纯语言形式、模糊数、二元语义等方法^[1-4]; 在一致性测度方面, 文献[5]研究了语言判断矩阵的一致性; 在信息集结方面, [6]研究了信息不完全的多属性决策方法, [7]提出了基于语言偏好的多类型偏好集结方法, [8]提出了基于语言评价信息的多属性群决策方法, [9]研究了基于云模型偏好集结和方案优选方法.

随着决策民主化和科学化的要求, 参与决策的人数规模与日俱增, 群决策规模过于庞大导致现有一些决策模型无法适用. 对此, 文献[10]研究了复杂大群体决策方法, [11]研究了群体成员偏好信息以效用值形式给出的大群体决策问题, [12]提出了一种面向复杂大群体的群决策支持系统结构.

从现有研究看, 大规模群集结问题研究的难点主要集中在以下 4 个方面: 1) 决策者差异较大, 需鉴别各类决策者的真实能力和水平, 赋予相应权重以达到和谐决策效果; 2) 决策群体规模庞大使得现有决策模型无法应用, 模型求解效果不够理想, 甚至因模型过载而无法得到群体决策效果; 3) 当群体意见较为分散

收稿日期: 2010-09-01; 修回日期: 2011-02-23.

基金项目: 国家自然科学基金项目(70701017, 70971064, 71171112); 江苏省普通高校研究生科研创新计划项目(CXLX11.0221); 江苏省高校哲学社会科学基金项目(09SJD88033); 南京航空航天大学基本科研业务费专项科研基金项目(NS2010209).

作者简介: 王嵩华(1977-), 女, 博士生, 从事群体决策理论、复杂产品质量管理等研究; 方志耕(1962-), 男, 教授, 博士生导师, 从事灰色系统理论、博弈论等研究.

时, 现有方法不能有效协调众多决策者的差异偏好, 难以在短时间进行群体协调, 以至贻误战机; 4) 在群体决策过程中, 自然语言形式被广泛采用, 与数字形式相比, 这种形式加大了群体偏好集结和协调的难度, 而且尚没有文献报道相关研究方法.

鉴于此, 本文提出一种基于语言评价信息的大规模群集结方法. 首先将复杂大群体分解为若干小群体, 从而降低群决策的难度; 然后建立基于分类的类内指标权重确定模型, 提出了群体类内和类间的偏好集结方法; 最后, 根据群体分类特征提出了类内和类间决策者权重确定模型.

2 主要方法和结果

2.1 基础概念

设有限方案集 $X = \{x_1, x_2, \dots, x_m\}$, 其中 x_i 为第 i 个决策方案; 方案的属性集 $A = \{a_1, a_2, \dots, a_n\}$, 其中 a_j 为第 j 个属性指标; $P = \{p^1, p^2, \dots, p^K\}$ 表示群体决策集, 其中 p^k 为第 k ($k = 1, \dots, K$) 个决策者. 大规模决策问题的决策者数目较大, 有时甚至可能多达上千人.

决策者以语言评价集 S 表达对方案属性的偏好程度^[13], 即决策者 k 从语言评价集 S 中选择一个元素表达对方案 i 中属性 j 的偏好程度, 记为 p_{ij}^k . 例如 7 个标度的语言评价集 S 为

$$\begin{cases} s_0 = N(\text{none}), s_1 = \text{VL}(\text{very low}), s_2 = L(\text{low}), \\ s_3 = M(\text{medium}), s_4 = H(\text{high}), \\ s_5 = \text{VH}(\text{very high}), s_6 = P(\text{perfect}). \end{cases}$$

为了更精确地处理语言偏好, 文献 [13] 等提出了二元语义的处理方式, 将决策者 k 的语言评价信息 p_{ij}^k 转化为 (r_{ij}^k, α^k) . 其中: α^k 为 p_{ij}^k 偏离 S 中某元素的距离, $\alpha^k \in (-1, 1)$. 语言评价集的语义转化函数 Δ, Δ^{-1} 及语义距离测度等性质见文献 [13-14] 等.

2.2 基于灰色关联聚类的大规模语言群集结方法

针对大规模群集结的难点, 采用聚类的方法将大规模群体进行适度分类, 使各类中决策者意见较为相近, 从而使众多大规模群体的难以协调问题转变为各类之间的协调, 降低了集结和协调的难度. 在这种情况下, 需要重点关注的问题有 4 个方面: 1) 采用何种方法基于哪种规则进行聚类. 2) 在聚类基础上, 如何确定指标权重和决策者权重. 两者的合理确定是问题关键所在, 这是学术界长期关注的问题, 研究成果丰富但略显复杂. 3) 各分类群体的权重如何确定, 如何有效集结各类意见. 从聚类的角度看, 类内决策者看法相近, 但是类间的差异较为显著. 4) 类间协调时, 需要找出一个类内决策者标杆, 这个标杆要求既能代表类, 又要真实存在, 而非虚拟决策者. 对此, 通过灰色

关联聚类方法将决策者进行聚类; 建立规划模型求解类内、类间权重, 集结群体综合偏好并查找类内核心决策者. 决策过程分为以下几个步骤:

Step 1 基于灰色关联度的大规模决策者聚类. 由于灰色关联度方法对聚类样本量的多少没有苛刻要求, 其通过数据之间的几何形状来分析关联情况, 计算简便, 一般很少出现量化结果与定性分析结果不符的情况, 本文将采用灰色关联聚类方法进行聚类分析. 实际上, 决策者的类别与属性指标权重有一定关系, 但考虑到指标权重获得的复杂性和主观性, 为方便起见, 在聚类时假设各指标重要性相当.

将各个决策者的评价信息作为观测对象, 得到如下序列:

$$\begin{cases} X_1 = (x_{11}, x_{12}, \dots, x_{1(mn)}), \\ X_2 = (x_{21}, x_{22}, \dots, x_{2(mn)}), \\ \vdots \\ X_k = (x_{k1}, x_{k2}, \dots, x_{k(mn)}). \end{cases}$$

经二元语义转化后作为特征数据, 可计算出 X_i 与 X_j 的灰色绝对关联度 ε_{ij} , 再根据灰色关联聚类方法对各决策者进行聚类^[15]. 选取临界值 $r \in [0, 1]$, 一般要求 $r > 0.5$. 当 $\varepsilon_{ij} \geq r$ ($i \neq j$) 时, X_j 与 X_i 具有同类特征, 可归为一类. 设得到反映群体意见的类为 Ω_s , $s = 1, 2, \dots, f$.

Step 2 类内权重 (指标权重和决策者权重) 求解及偏好集结模型. 为便于表述, 本节所指决策者均属于 Ω_s 类, 本步骤中除非必须, 否则类内各变量和已知量均未标注类号 s , 类内决策者数记为 d . 为集结群体偏好, 有效协调群体差异, 需要得到群体子类意见, 该意见与群体属性指标权重和类内决策者权重直接相关. 对此, 在无先验决策者权重信息时, 本文用试算法确定权重. 考虑到类内决策目标是得到最为接近的意见, 即类内决策者看法距离偏差最小, 故建立如下模型 M_1 以求解各决策者 k 的 j 属性指标权重 w_j^k :

$$M_1 : \begin{cases} \min \frac{1}{d} \sum_{i=1}^m \sum_{k,p \in \Omega_s, p > k} \left(\sum_{j=1}^n \Delta^{-1}(r_{ij}^k, \alpha^k) w_j^k - \sum_{j=1}^n \Delta^{-1}(r_{ij}^p, \alpha^p) w_j^p \right)^2; \\ \text{s.t.} \sum_{j=1}^n w_j^k = 1, w_j^k > 0. \end{cases}$$

定理 1 模型 M_1 必定存在最优解.

由上, 各决策者方案综合指标值可表示为

$$z_i^k(w) = \sum_{j=1}^n w_j^k \Delta^{-1}[r_{ij}^k, \alpha^k],$$

$$i = 1, 2, \dots, m, k \in \Omega_s. \quad (1)$$

设决策者权重均等 (为 $1/d$), 计算群体对各方案的综

合指标值为

$$z_i(w) = \sum_{k \in \Omega_s} z_i^k(w)/d. \quad (2)$$

将上述各决策者的方案综合指标值 $z_i^k(w)$ 与其所在群体的综合指标值 $z_i(w)$ 进行比较, 按各方案在引入决策者权重 λ_k 后的综合指标值与其群体综合指标值偏差最小的原则, 建立模型 M_2 , 即

$$M_2 : \begin{cases} \min \sum_{i=1}^m \sum_{k \in \Omega_s} (z_i^k(w)\lambda_k - z_i(w))^2; \\ \text{s.t. } \lambda_k \geq \zeta, \sum_{k=1}^d \lambda_k = 1. \end{cases}$$

其中: ζ 为常数, 为保证各决策者都有一定的发言权, 取 $\zeta = 1/a \times d$; d 为类内决策者数目; $a \geq 1$, 应根据决策情况而设定, 一般可取 $a = 2$ (表明最小决策者权重为平均值的一半).

定理2 模型 M_2 必定存在最优解.

根据求得的类内决策者权重 λ_k , 重新计算群体综合指标值

$$z_{si}(w) = \sum_{k \in \Omega_s} z_i^k(w)\lambda_k,$$

其中 s 为类序号.

Step3 类间权重求解及偏好集结模型. 基于各类内指标属性权重和决策者权重, 依据类间差异越小越好的原则, 构建类间权重 μ_s 求解模型 M_3 , 即

$$M_3 : \begin{cases} \min \sum_{i=1}^m \sum_{s,l=1, l>s}^f (z_{si}(w)\mu_s - z_{li}(w)\mu_l)^2; \\ \text{s.t. } \mu_s \geq 0, \sum_{s=1}^f \mu_s = 1. \end{cases}$$

其中: z_{si} 和 z_{li} 分别为 Step2 集结出的各类综合指标值, μ_s 和 μ_l 为各类的重要性权重.

定理3 模型 M_3 必定存在最优解.

一般情况下, 决策者数目越多的类, 其权重设置应适当偏大些, 表示更大的决策者代表范围. 考虑各类决策者数量的因素, 根据模型求解权重 μ_s , 并通过下式修正类间权重:

$$u_s \left(\sigma + \frac{\text{类内人数}}{\text{总人数}} \times (1 - \sigma) \right),$$

其中 σ 为参数 ($0 \leq \sigma \leq 1$). 该式可以对 μ_s 起到调节作用, 使 μ_s 更符合实际情况, 即当 $\sigma < 0.5$ 时, 经过调节, 人数多的类, 其权重将增大 ($\sigma = 0$ 时类内人数因素影响最大); 反之则减小 ($\sigma = 1$ 时不考虑类内人数影响). 利用 Step2 的群体综合指标值和 Step3 的类间修正权重, 依据下式集结大规模群体偏好:

$$z_i(w)' = \sum_{s=1}^f z_{si}(w)\mu_s, \quad i = 1, 2, \dots, m. \quad (3)$$

Step4 类内核心决策者的查找及偏好协调算

法. 类内核心决策者的查找和管理, 可有效协调群体分歧. 依据每类中各决策者综合指标值与每类综合指标值结果的距离大小, 寻找核心决策者, 规定各类中距离最短者为该类的核心决策者.

定义1 设类 Ω_s 的综合指标值为 $z_{si}(w)$ ($i = 1, 2, \dots, m$), 类中各决策者的综合指标值为 $z_{si}^k(w)$ ($k = 1, 2, \dots, d_s$), 则称类中各决策者综合指标值与类的综合指标值结果的距离最短者为该类的核心决策者 P_s^* , 即

$$P_s^* = \min_{k=1, \dots, d_s} \sum_{i=1}^m |z_{si}^k(w) - z_{si}(w)|.$$

定义2 设有 k ($k = 1, 2, \dots, K$) 个决策者对 $X = \{x_1, x_2, \dots, x_m\}$ 个方案进行评价, 则称第 k 个决策者与群体判断的偏差 γ^k 为决策者的判断一致度, 即

$$\gamma^k = \left(r - \sum_{i=1}^m |z_i^k(w) - z_i(w)'| \right) / r. \quad (4)$$

其中: r 为语言评价集 S 的总标度, 如语言评价集 S 由 $S_0 \rightarrow S_6$ 表示, 则 $r = 6$; $z_i^k(w)$ 为各决策者对方案的综合指标值, $z_i(w)'$ 为群体对方案的综合指标值. 若 $\gamma^k \geq \rho$, 则表示决策者与群体基本达成一致. 一般可以根据具体情况设置不同的偏差阈值: 当决策精度要求高时, 偏差值 ρ 可适当设大; 反之则小.

定义3 设有 k ($k = 1, 2, \dots, K$) 个决策者对 $X = \{x_1, x_2, \dots, x_m\}$ 个方案的 $A = \{a_1, a_2, \dots, a_n\}$ 个属性指标进行评价, 若第 k 个决策者的判断与群体的平均偏离程度

$$\varepsilon^k = \frac{\sum_{i,j} \left[\Delta^{-1}(r_{ij}^k, \alpha^k) - \sum_{k=1}^K \Delta^{-1}(r_{ij}^k, \alpha^k) / K \right]}{mn} \geq b[3], \quad (5)$$

则称第 k ($k = 1, 2, \dots, K$) 个决策者为严重分歧决策者. 式(5)表示有 mn 个评价值, b 为临界值, 方括号中为7粒度评价语言时的建议值.

对于与群体判断偏差较大的决策者和严重分歧决策者, 根据造成的原因有两种考虑: 1) 经慎重分析后确保非集中犯错现象; 2) 如果由于决策者认识等问题造成严重分歧, 则可暂不考虑该决策者的判断, 以谋求群体决策的快速一致, 或协调后由决策者适当修改偏好以达到较好的决策效果.

当 $\gamma^k < \rho$ 且 $\varepsilon^k < b[3]$ 时, 由决策者 k 调整其偏好, 决策者判断修改为

$$P'_k = (1 - p)P_k + pP_s^*.$$

其中: P_k 为第 k 位决策者的原始评价矩阵; P_s^* 为第 k 位决策者所在类的核心决策者; p 为协调幅度, 每次协调幅度视具体情况而定, 直至 $\gamma^k \geq \rho$ 且 $\varepsilon^k < b[3]$ (具体方法将另文详述).

各个方案的综合偏好并做归一化处理, 可得 (0.277, 0.242, 0.241, 0.233).

Step 4 经计算无严重分歧决策者, 对类内核心专家进行查找. 可得第2类中的核心专家为 P_2 , 第3类中为 P_6 , 第4类中为 P_8 . 第1和第5类人数少于3, 故无需查找核心专家.

计算决策一致性, 结果如下:

$$\begin{aligned} \gamma^1 &= 0.453, \gamma^2 = 0.684, \gamma^3 = 0.855, \gamma^4 = 0.670, \\ \gamma^5 &= 0.714, \gamma^6 = 0.778, \gamma^7 = 0.855, \gamma^8 = 0.730, \\ \gamma^9 &= 0.688, \gamma^{10} = 0.715, \gamma^{11} = 0.871, \gamma^{12} = 0.633, \\ \gamma^{13} &= 0.730, \gamma^{14} = 0.601, \gamma^{15} = 0.587. \end{aligned}$$

依据定义2, 取 $\rho = 0.6$, 专家 P_1 需要调整其偏好信息. 设 $p = 0.8$, P_1 修改后为

$$\begin{bmatrix} (s_2, 0.4) & (s_1, 0.4) & (s_4, 0) & (s_3, -0.4) \\ (s_5, 0) & (s_1, -0.2) & (s_4, -0.4) & (s_3, 0) \\ (s_4, -0.4) & (s_3, -0.4) & (s_2, 0) & (s_2, 0.4) \\ (s_3, 0) & (s_3, -0.2) & (s_4, 0.2) & (s_1, 0) \end{bmatrix}.$$

经过修改, 得到 $\gamma^1 = 0.861$, 符合群体一致性要求, 最终群体综合偏好变化为 (0.271, 0.238, 0.260, 0.231). 因此, 在这次方案选择过程中, 该公司应重点考虑方案 X_1 .

4 结 论

针对具有语言评价信息的大规模群集结问题, 本文给出了一种基于灰色关联聚类的大规模语言评价信息群集结方法. 一方面解决了大规模决策者之间意见较为分散时的决策问题; 另一方面通过对属性权重、决策者权重和类间权重的模型计算求解, 解决了现有决策模型对大规模群集结求解效果不理想问题. 最后对群集结一致性进行了度量. 该方法简单可行、易于理解. 下一步将重点研究引入时间维度的大规模群集结模型及其相应的集结模型性质.

参考文献(References)

- [1] Dong Y C, Xu Y F, Yu Shui. Linguistic multiperson decision making based on the use of multiple preference relations[J]. *Fuzzy Sets and Systems*, 2009, 160(5): 603-623.
- [2] José M Merigó. Fuzzy decision making with immediate probabilities[J]. *Computers & Industrial Engineering*, 2010, 58(4): 651-657.
- [3] 汪洋, 孙林岩. 李克特式量表与模糊语言量表计分的差异比较——以梯形模糊数仿真为例[J]. *运筹与管理*, 2008, 17(1): 48-52.
(Wang Y, Sun L Y. Comparing the score difference between the Likert scale and the fuzzy linguistic questionnaire: The simulating research of trapezoid fuzzy number[J]. *Operations Research and Management Science*, 2008, 17(1): 48-52.)
- [4] 张可, 刘思峰, 朱建军. 基于二元语义的八类偏好信息的群决策方法[J]. *系统工程*, 2008, 26(9): 116-121.
(Zhang K, Liu S F, Zhu J J. Group decision making method based on two-tuple linguistic for eight types of preference information[J]. *Systems Engineering*, 2008, 26(9): 116-121.)
- [5] 唐燕武, 陈华友. 语言判断矩阵的满意一致性判定方法[J]. *合肥工业大学学报: 自然科学版*, 2010, 33(4): 621-624.
(Tang Y W, Chen H Y. A method for determining satisfying consistency of linguistic judgment matrices[J]. *J of Hefei University of Technology: Natural Science*, 2010, 33(4): 621-624.)
- [6] 梁昌勇, 张恩桥, 戚筱雯. 一种评价信息不完全的混合型多属性群决策方法[J]. *中国管理科学*, 2009, 17(4): 126-132.
(Liang C Y, Zhang E Q, Qi X W. A method of multi-attribute group decision making with incomplete hybrid assessment information[J]. *Chinese J of Management Science*, 2009, 17(4): 126-132.)
- [7] Xu Z S. Interactive group decision making procedure based on uncertain multiplicative linguistic preference relations[J]. *J of Systems Engineering and Electronics*, 2010, 21(3): 408-415.
- [8] 廖貅武, 李垣, 董广茂. 一种处理语言评价信息的多属性群决策方法[J]. *系统工程理论与实践*, 2006, 9(9): 91-97.
(Liao X W, Li Y, Dong G M. A multi-attribute group decision-making approach dealing with linguistic assessment information[J]. *Systems Engineering-Theory & Practice*, 2006, 9(9): 91-97.)
- [9] 王洪利, 冯玉强. 基于云模型具有语言评价信息的多属性群决策研究[J]. *控制与决策*, 2005, 20(6): 679-681.
(Wang H L, Feng Y Q. On multiple attribute group decision making with linguistic assessment information based on cloud model[J]. *Control and Decision*, 2005, 20(6): 679-681.)
- [10] 陈晓红. 复杂大群体决策方法及应用[M]. 北京: 科学出版社, 2009.
(Chen X H. *Complex large group decision making and application*[M]. Beijing: Science Press, 2009.)
- [11] 徐选华, 陈晓红, 王红伟. 一种面向效用值偏好信息的大群体决策方法[J]. *控制与决策*, 2009, 24(3): 440-445.
(Xu X H, Chen X H, Wang H W. A kind of large group decision-making method oriented utility valued preference information[J]. *Control and Decision*, 2009, 24(3): 440-445.)