

基于云相似度的语言偏好信息多属性大群体决策方法

徐选华, 王佩[†], 蔡晨光

(中南大学 商学院, 长沙 410083)

摘要: 针对属性值为语言变量的多属性大群体决策问题, 提出一种基于云相似度的大群体决策方法. 首先, 将语言变量转换为云模型, 提出一种基于 α 截集的云距离和云相似度模型; 然后, 在此基础上提出云聚类算法, 从而对决策大群体进行聚类, 根据聚类结果求解聚集权重, 再利用云加权算术平均 (CWAA) 算子求出方案的群体偏好值, 进而计算方案群体偏好值的总分来确定方案的排序; 最后, 通过算例分析验证所提出方法的可行性和有效性.

关键词: 大群体决策; 云模型; 相似度; 语言偏好; 聚类算法

中图分类号: TP273

文献标志码: A

Linguistic multi-attribute large group decision-making method based on similarity measurement of cloud model

XU Xuan-hua, WANG Pei[†], CAI Chen-guang

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

Abstract: For multi-attribute large group decision-making problems with linguistic preference, a method based on the cloud clustering algorithm is proposed. Firstly, linguistic preferences are transformed to the cloud model, and a distance measure and similarity measure for cloud model are developed based on α -cut. Then, a cloud clustering algorithm is presented to classify decision makers into several clusters and then cluster weights are obtained. The cloud weighted arithmetic averaging operator is applied to aggregate collective group preferences and the final ranking results can be derived by computing the overall scores. Finally, an example is given to illustrate the practicality and effectiveness of the proposed method.

Keywords: large group decision-making; cloud model; similarity measurement; linguistic preference; clustering algorithm

0 引言

决策作为一项基本活动, 已成为人们工作和生活中必不可少的一部分. 根据决策问题复杂程度的不同, 往往对决策群体提出不同的要求. 一般来说, 当决策问题涉及属性数量较少、决策目标单一时, 单人决策或小群体决策足以满足决策要求, 但随着决策问题复杂程度的增加, 决策目标存在多重性, 决策属性呈现多维性、关联性以及随机性时, 传统单人决策模型^[1,2] 或小规模群决策模型^[3] 难以满足这类问题的实际需要, 需要大群体甚至特大群体专家共同参与, 如特大自然灾害应急决策、大型建设工程生态环境风险评估等所涉及的专家成员不仅规模比较庞大, 而且成员分布相当广泛, 这类问题被称为大群体决策问题^[4]. 大群体决策问题呈现以下特点^[5]: 1) 群体规模

比较庞大, 通常决策专家个数不少于 20 人^[4,6-10]; 2) 决策问题属性呈现多维性、复杂性和随机性; 3) 需要满足群体高一一致性要求.

目前, 大群体决策已经成为决策领域的一个热点研究领域^[7,11], 有关大群体决策的研究主要集中在偏好信息为效用值^[12]、实数值^[13]、区间直觉模糊集^[8-9] 等情形, 而针对偏好信息为语言变量的研究还比较少^[10]. 与量化的偏好信息相比, 语言变量^[14] 更加符合决策专家的表达习惯, 更能反映信息的模糊性和不确定性. 近年来, 云模型^[15] 开始被应用于语言类决策问题中. 与模糊数计算模型^[16]、有序语言计算模型^[17] 和二元语义模型^[18] 这 3 种方法相比, 云模型不仅可以有效刻画语言值的模糊性和随机性, 而且提供了一种更为客观的定性概念与定量值之间的转换模型.

收稿日期: 2016-02-11; 修回日期: 2016-04-27.

基金项目: 国家自然科学基金项目 (71171202)

作者简介: 徐选华 (1962—), 男, 教授, 博士生导师, 从事决策理论与方法、信息系统与决策支持系统等研究; 王佩 (1988—), 女, 博士生, 从事决策方法和应急管理决策的研究.

[†]通讯作者. E-mail: wangpeichn@csu.edu.cn

本文基于云模型理论对偏好信息为语言变量的大群体决策问题进行研究,定义一种基于 α 截集的云距离,并提出一种新的云相似度模型.在此基础上,给出一种基于云相似度的聚类方法,根据相似性程度的不同对大群体进行聚类形成聚集,利用云算子对各聚集的偏好信息进行集结得到各方案的综合属性值,根据云的总计分确定方案的排序.

1 云模型及相关概念

1.1 云模型及其运算

设 U 为给定论域, T 是与 U 相联系的某一定性概念,若元素 x 是定性概念 T 的一次随机实现,且满足 $x \sim N(\text{Ex}, \text{En}'^2)$, $\text{En}' \sim N(\text{En}, \text{He}^2)$, $y \in [0, 1]$ 表示 x 属于定性概念 T 的隶属程度,满足^[15]

$$y = e^{-\frac{(x-\text{Ex})^2}{2(\text{En}')^2}},$$

则 x 从论域 U 到区间 $[0, 1]$ 的映射,即 $y_T(X) : \forall x \in U \rightarrow [0, 1]$ 被定义为云,每一组 (x, y) 可看成一个云滴.

云的数字特征记为 $Y(\text{Ex}, \text{En}, \text{He})$,分别表示云的期望(Expected value)、熵(Entropy)和超熵(Hyper entropy).期望 Ex 是指云滴在论域空间分布的中心值;熵 En 表示定性概念的不确定性程度,包括随机性和模糊性两个方面;超熵 He 表示熵的不确定性程度.

定义1 给定两朵云 $A_1 = Y(\text{Ex}_1, \text{En}_1, \text{He}_1)$ 和 $A_2 = Y(\text{Ex}_2, \text{En}_2, \text{He}_2)$, A_1 和 A_2 的运算定义为^[15]

$$\begin{aligned} A_1 + A_2 &= Y(\text{Ex}_1 + \text{Ex}_2, \sqrt{\text{En}_1^2 + \text{En}_2^2}, \sqrt{\text{He}_1^2 + \text{He}_2^2}); \\ A_1 - A_2 &= Y(\text{Ex}_1 - \text{Ex}_2, \sqrt{\text{En}_1^2 + \text{En}_2^2}, \sqrt{\text{He}_1^2 + \text{He}_2^2}); \\ \lambda A_1 &= Y(\lambda \text{Ex}_1, \sqrt{\lambda} \text{En}_1, \sqrt{\lambda} \text{He}_1), \lambda > 0. \end{aligned}$$

定义2 设 $A_i (i = 1, 2, \dots, n)$ 为一组云,且有 $\text{CWAA} : A^n \rightarrow A$,则定义 CWAA 为云的加权算术平均算子,且有^[19]

$$\text{CWAA}_w(A_1, A_2, \dots, A_n) = \sum_{i=1}^n w_i A_i.$$

其中: $w = (w_1, w_2, \dots, w_n)$ 是其对应的权重向量, $w_i \in [0, 1] (i = 1, 2, \dots, n)$ 且 $\sum_{i=1}^n w_i = 1$.

定理1 设 $A_i (i = 1, 2, \dots, n)$ 为一组云,其对应的权重向量 $w = (w_1, w_2, \dots, w_n)$ 满足 $w_i \in [0, 1] (i = 1, 2, \dots, n)$ 且 $\sum_{i=1}^n w_i = 1$,则由 CWAA 算子集结得到的结果仍然为一朵云,且有^[19]

$$\text{CWAA}_w(A_1, A_2, \dots, A_n) = \left(\sum_{i=1}^n w_i \text{Ex}_i, \sqrt{\sum_{i=1}^n w_i \text{En}_i^2}, \sqrt{\sum_{i=1}^n w_i \text{He}_i^2} \right). \quad (1)$$

定义3 对于云滴 (x, y) ,称 $Q = xy$ 为该云滴对概念 T 的一次计分.计分值 Q 的数学期望 \hat{Q} (或二分位数 Q_{mid})称为云对概念的总计分^[20].对于同一论域中的任意两朵云 A_1 和 A_2 ,当所对应的总计分 $\hat{Q}_1 \geq \hat{Q}_2$ (或 $Q_{\text{mid}1} \geq Q_{\text{mid}2}$)时,称 A_1 大于等于 A_2 ,记为 $A_1 \geq A_2$.

定义4 综合云用于将同类型的云进行合并,生成一朵更高层次的云.给定两朵云 $A = Y(\text{Ex}_1, \text{En}_1, \text{He}_1)$ 和 $B = Y(\text{Ex}_2, \text{En}_2, \text{He}_2)$,其对应的标准期望曲线为 $\text{Sl}_1(x)$ 和 $\text{Sl}_2(x)$,设得到的综合云为

$$\text{AB} = Y(\text{Ex}_{12}, \text{En}_{12}, \text{He}_{12}),$$

运算如下^[15]:

$$\begin{cases} \text{Ex}_{12} = \frac{\text{Ex}_1 \text{En}'_1 + \text{Ex}_2 \text{En}'_2}{\text{En}'_1 + \text{En}'_2}, \\ \text{En}_{12} = \text{En}'_1 + \text{En}'_2, \\ \text{He}_{12} = \frac{\text{He}_1 \text{En}'_1 + \text{He}_2 \text{En}'_2}{\text{En}'_1 + \text{En}'_2}. \end{cases} \quad (2)$$

如果有

$$\begin{aligned} \text{Sl}'_1(x) &= \begin{cases} \text{Sl}_1, & \text{Sl}_1(x) \geq \text{Sl}_2(x); \\ 0, & \text{else.} \end{cases} \\ \text{Sl}'_2(x) &= \begin{cases} \text{Sl}_2, & \text{Sl}_2(x) > \text{Sl}_1(x); \\ 0, & \text{else.} \end{cases} \end{aligned}$$

则有

$$\begin{aligned} \text{En}'_1 &= \frac{1}{\sqrt{2\pi}} \int \text{Sl}'_1(x) dx, \\ \text{En}'_2 &= \frac{1}{\sqrt{2\pi}} \int \text{Sl}'_2(x) dx. \end{aligned}$$

1.2 语言变量与云的转换模型

自然语言评价集 H 是指由奇数个语言变量组成的集合,设自然语言评级等级为

$$H = \{h_i | i = -g, \dots, 0, \dots, g, g \in N\},$$

则 H 满足以下性质^[14]: 1)有序性. $i > j \Leftrightarrow h_i > h_j$. 2)可逆性. 当 $i + j = 0$ 时, $h_i = \text{neg}(h_j)$, neg 表示逆算子.

定义5 给定语言评价集 $H = \{h_i | i = -t, \dots, 0, \dots, t, t \in N\}$,存在一个函数 f 可以将 h_i 转化为相应的数值 θ_i , $\theta_i \in [0, 1]$,即^[19]

$$\theta_i = \begin{cases} \frac{a^t - a^{-i}}{2a^t - 2}, & -t \leq i \leq 0; \\ \frac{a^t + a^i - 2}{2a^t - 2}, & 0 < i \leq t. \end{cases} \quad (3)$$

文献[21]给出了两种方法确定 a 的取值,即实验

法和主观法. 实验法表明 a 的取值范围为 $[1.36, 1.4]$, 主观法表明 $a \approx 1.37$.

定义6 给定论域 $[X_{\min}, X_{\max}]$, 设 $h_i \in H$ 转换后对应的云为 $A_i = Y(Ex_i, En_i, He_i)$, 其转换模型如下^[19]:

- 1) 利用式(3)计算 θ_i .
- 2) 计算 Ex_i .

$$Ex_i = X_{\min} + \theta_i(X_{\max} - X_{\min}).$$

- 3) 计算 En_i .

$$En'_i = \begin{cases} \frac{(1 - \theta_i)(X_{\max} - X_{\min})}{3}, & -t \leq i \leq 0; \\ \frac{\theta_i(X_{\max} - X_{\min})}{3}, & 0 < i \leq t. \end{cases}$$

$$En_{-i} = En_i =$$

$$\begin{cases} \frac{(\theta_{|i|-1} + \theta_{|i|} + \theta_{|i|+1})(X_{\max} - X_{\min})}{9}, & 0 < |i| \leq t - 1; \\ \frac{(\theta_{|i|-1} + \theta_{|i|})(X_{\max} - X_{\min})}{6}, & |i| = t; \\ \frac{(\theta_i + 2\theta_{i+1})(X_{\max} - X_{\min})}{9}, & i = 0. \end{cases}$$

- 4) 计算 He_i .

$$He_{-1} = He_i = \frac{En'^+ - En_i}{3},$$

$$En'^+ = \max_k \{En'_k\}.$$

2 基于云相似度的偏好聚类方法

聚类方法是一种群体分类的重要分析方法, 其原理是利用相似度将大群体聚类为若干个聚集, 同一聚集中的个体比不同聚集中的个体更相似. 目前常用的聚类算法包括 C -means、 K -means 等. 本文通过定义云距离和云相似度建模, 提出一种基于云相似度的聚类算法.

2.1 云相似度建模

α 截集是一种求解距离的有效方法, 在此提出一种基于 α 截集的距离公式. 由于云模型是由很多个云滴组成, 为了方便求解云模型的 α 截集, 利用云模型的几种期望曲线近似求解.

定义7 设任意一朵云 $A = Y(Ex, En, He)$, 其中 $x \sim N(Ex, En'^2)$, $En' \sim N(En, He^2)$, 其相对应的标准期望曲线SI表示为 $y = e^{-\frac{(x-Ex)^2}{2En'^2}}$, 其内期望曲线II表示为 $y = e^{-\frac{(x-Ex)^2}{2En^2}}$, 其外期望曲线OI表示为 $y = e^{-\frac{(x-Ex)^2}{2En^2}}$. 对于SI, 满足 $y \geq \alpha$ ($\alpha \in [0, 1]$) 的所有 x 构成的集合称为云 A 的标准期望曲线 α 截集, 如图1所示, 记为 $A_\alpha = [A_\alpha^L, A_\alpha^U]$; 对于OI, 满足 $y \geq \alpha$ ($\alpha \in [0, 1]$) 的所有 x 构成的集合称为云 A 的外期望曲线 α 截集, 记为 $AU_\alpha = [AU_\alpha^L, AU_\alpha^U]$; 对于II, 满足 $y \geq \alpha$ ($\alpha \in$

$[0, 1]$) 的所有 x 构成的集合称为云 A 的内期望曲线 α 截集, 记为 $AL_\alpha = [AL_\alpha^L, AL_\alpha^U]$. 其中, En 和 \overline{En} 可以根据 3σ 原则给出, 在本文中取 $En = En - 2He$, $\overline{En} = En + 2He$, 此时内外期望曲线所覆盖的云滴数占95%.

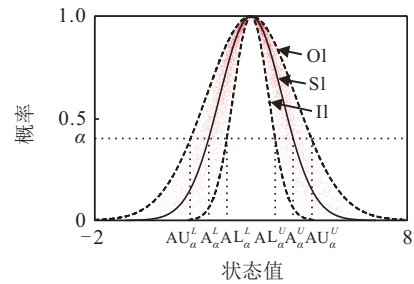


图1 云模型的 α 截集

利用云模型的期望曲线和其 α 截集对两朵云之间的距离进行求解, 具体求解思路如下: 设任意两朵云 $A = Y(Ex_1, En_1, He_1)$ 和 $B = Y(Ex_2, En_2, He_2)$. 当不考虑超熵时, 云模型的距离转化为对应标准期望曲线SI(图1)的距离, 并利用 α 截集进行求解; 当考虑超熵时, 可以将云模型转化为外期望曲线OI和内期望曲线II, 然后利用 α 截集进行求解. 设云模型 A 与 B 的距离表示为 $d(A, B) = Y(Ex_{12}, En_{12}, He_{12})$, 求解过程如下:

1) 对于 $\alpha \in [0, 1]$, A_α 与 B_α 的距离可以定义为 $d(A_\alpha, B_\alpha) = [d_\alpha^L, d_\alpha^U]$, 具体表示如下:

当 $Ex_1 \geq Ex_2$ 时, 有

$$[d_\alpha^L, d_\alpha^U] = A_\alpha - B_\alpha = [A_\alpha^L - B_\alpha^L, A_\alpha^U - B_\alpha^U];$$

当 $Ex_1 < Ex_2$ 时, 有

$$[d_\alpha^L, d_\alpha^U] = B_\alpha - A_\alpha = [B_\alpha^L - A_\alpha^L, B_\alpha^U - A_\alpha^U].$$

通过还原 $d(A_\alpha, B_\alpha)$ 的值求得 $d(A, B)$. 当 $Ex_1 \geq Ex_2$ 时, 有

$$Ex_{12} = \frac{A_\alpha^L - B_\alpha^L + A_\alpha^U - B_\alpha^U}{2} = Ex_1 - Ex_2,$$

即 $Ex_{12} = Ex_1 - Ex_2$.

同理, 当 $Ex_1 < Ex_2$ 时, $Ex_{12} = Ex_2 - Ex_1$.

综上, $Ex_{12} = |Ex_1 - Ex_2|$.

2) 由 $e^{-\frac{(A_\alpha^L - Ex_1)^2}{2En_1^2}} = e^{-\frac{(B_\alpha^L - Ex_2)^2}{2En_2^2}} = \alpha$ 可得

$$En_1 = \sqrt{\frac{(A_\alpha^L - A_\alpha^U)^2}{-8\ln\alpha}}, \quad En_2 = \sqrt{\frac{(B_\alpha^L - B_\alpha^U)^2}{-8\ln\alpha}}.$$

当 $Ex_1 \geq Ex_2$ 时, 有

$$e^{-\frac{(A_\alpha^L - B_\alpha^L - Ex_{12})^2}{2En_{12}^2}} = \alpha,$$

于是有

$$En_{12}^2 = \frac{(A_\alpha^L - A_\alpha^U - B_\alpha^L + B_\alpha^U)^2}{-8\ln\alpha} = (En_1 - En_2)^2,$$

即 $En_{12} = En_1 - En_2$.

同理, 当 $Ex_1 < Ex_2$ 时, $En_{12} = En_2 - En_1$.

综上, $En_{12} = |En_1 - En_2|$.

3) 由 $e^{-\frac{(AL_\alpha^L - Ex_1)^2}{2En_1^2}} = e^{-\frac{(BL_\alpha^L - Ex_2)^2}{2En_2^2}} = \alpha$ 可得

$$\underline{En}_1 = \sqrt{\frac{(AL_\alpha^L - AL_\alpha^U)^2}{-8\ln\alpha}}, \quad \underline{En}_2 = \sqrt{\frac{(BL_\alpha^L - BL_\alpha^U)^2}{-8\ln\alpha}},$$

同理可得

$$\overline{En}_1 = \sqrt{\frac{(AU_\alpha^L - AU_\alpha^U)^2}{-8\ln\alpha}}, \quad \overline{En}_2 = \sqrt{\frac{(BU_\alpha^L - BU_\alpha^U)^2}{-8\ln\alpha}}.$$

当 $Ex_1 \geq Ex_2$ 时, 云 A 与云 B 距离的内期望曲线的 α 截集表示为 $[AL_\alpha^L - BL_\alpha^L, AL_\alpha^U - BL_\alpha^U]$, 于是有

$$e^{-\frac{(AL_\alpha^L - BL_\alpha^L - Ex_{12})^2}{2En_{12}^2}} = \alpha,$$

化简可得

$$\underline{En}_{12} = \sqrt{\frac{(AL_\alpha^L - AL_\alpha^U + BL_\alpha^U - BL_\alpha^L)^2}{-8\ln\alpha}} = \underline{En}_1 - \underline{En}_2.$$

云 A 与云 B 距离的外期望曲线的 α 截集表示为 $[AU_\alpha^L - BU_\alpha^L, AU_\alpha^U - BU_\alpha^U]$, 于是有

$$e^{-\frac{(AU_\alpha^L - BU_\alpha^L - Ex_{12})^2}{2En_{12}^2}} = \alpha,$$

化简可得

$$\overline{En}_{12} = \sqrt{\frac{(AU_\alpha^L - AU_\alpha^U + BU_\alpha^U - BU_\alpha^L)^2}{-8\ln\alpha}} = \overline{En}_1 - \overline{En}_2.$$

因为 $\overline{En}_1 - \underline{En}_1 = 4He_1, \overline{En}_2 - \underline{En}_2 = 4He_2$, 故 $He_{12} = \frac{1}{4}(\overline{En}_{12} - \underline{En}_{12}) = He_1 - He_2$. 同理可得, 当 $Ex_1 < Ex_2$ 时, $He_{12} = He_2 - He_1$.

综上, $He_{12} = |He_1 - He_2|$. 因此, 云 A 与云 B 的距离为

$$d(A, B) = Y(|Ex_1 - Ex_2|, |En_1 - En_2|, |He_1 - He_2|). \quad (4)$$

定理 2 给出 3 朵云 $A = Y(Ex_1, En_1, He_1), B = Y(Ex_2, En_2, He_2)$ 和 $C = Y(Ex_3, En_3, He_3)$, 则云距离 d 满足以下性质:

- 1) $d(A, B) \geq 0$;
- 2) $d(A, B) = Y(0, 0, 0)$, 当且仅当 $A = B$;
- 3) $d(A, B) = d(B, A)$.

定理显然成立, 证明过程在此省略.

定义 8 设 $A = Y(Ex_i, En_i, He_i)$ 和 $B = Y(Ex_j, En_j, He_j)$ 为两朵云, 则 A 与 B 之间的相似度为

$$s(A, B) = 1 - \frac{|\hat{Q}(d(A, B))|}{\hat{Q}(A) + \hat{Q}(B)}. \quad (5)$$

其中: $d(A, B) = Y(|Ex_i - Ex_j|, |En_i - En_j|, |He_i - He_j|)$, $\hat{Q}(\cdot)$ 为云的总计分.

$s(A, B)$ 满足以下性质:

- 1) $0 \leq s(A, B) \leq 1$;
- 2) $s(A, B) = 1$, 当且仅当 $A = B$;
- 3) $s(A, B) = s(B, A)$.

证明 1) 由于满足

$$|\hat{Q}(d(A, B))| \geq 0, \quad 1 - \frac{|\hat{Q}(d(A, B))|}{\hat{Q}(A) + \hat{Q}(B)} \leq 1,$$

则 $s(A, B) \leq 1$ 得证.

由于满足

$$\hat{Q}(d(A, B)) \leq \hat{Q}(A) + \hat{Q}(B), \quad \frac{|\hat{Q}(d(A, B))|}{\hat{Q}(A) + \hat{Q}(B)} \leq 1,$$

即 $s(A, B) \geq 0$ 得证.

2) 当 $s(A, B) = 1$ 时, 有 $|\hat{Q}(d(A, B))| = 0$, 由此可得 $Ex_i = Ex_j, En_i = En_j, He_i = He_j$, 因此 $A = B$.

当 $A = B$ 时, $d(A, B) = Y(0, 0, 0)$, 于是有 $|\hat{Q}(d(A, B))| = 0$, 因此 $s(A, B) = 1$.

3) 性质 3) 显然成立. \square

2.2 不同云相似度比较

为了说明本文所提出的云相似度模型的有效性, 通过仿真实例进行数值实验, 然后将本文方法与 SCM 方法^[22]、LICM 方法^[23]、ECM 和 MCM 方法^[24]、CCM 方法^[25] 及 CFSM 方法^[26] 进行比较分析.

例 1 文献 [24] 给出了 3 朵云, $A_1 = Y(3, 3.123, 2.05), A_2 = Y(2, 3, 1)$ 和 $A_3 = Y(1.585, 3.556, 1.358)$, 分别利用本文所提出的方法与文献 [22-26] 的云相似度计算公式对 3 个云模型的相似度进行计算, 计算结果如表 1 所示.

表 1 不同云相似度计算模型比较(例 1)

相似度	SM α	SCM ^[22]	LICM ^[23]	ECM ^[24]	CCM ^[25]	CFSM ^[26]
$s(A_1, A_2)$	0.8013	0.9627	0.9712	0.8728	0.9214	0.7790
$s(A_2, A_3)$	0.8841	0.9906	0.9850	0.9138	0.9757	0.9035
$s(A_3, A_1)$	0.6924	0.9614	0.9438	0.8336	0.8121	0.7223

由表 1 可知, A_2 和 A_3 的相似度最大(0.8841), A_1 和 A_3 的相似度最小(0.6924), 该结果与图 2 的直观印象一致. 本文方法与其他方法计算得到的结果一致, 但是本文方法更简单, 效果更好. SCM 方法在 $[Ex - 3En, Ex + 3En]$ 区域内随机选取一定数量的云滴, 然后计算其距离的平均值, 从而度量云模型之间的相似度, 该方法有两点不足: 一是判断两朵云是否相似依赖于相似度阈值, 二是只考虑部分云滴容易引起结果的不稳定. LICM 方法将云模型的 3 个数值特征视为一个数学向量, 通过计算向量之间的余弦夹角求解两个云模型之间的相似度, 该方法的不足是将云模型的期望、熵和超熵视为同一数量级进行计算, 从而产生较大的误差. ECM 是基于云模型期望曲线相交区域的面积计算相似度, ECM 方法没有考虑超熵的作

用;CCM方法利用概念跃升得到两朵云的综合云,然后根据综合云基础云的期望曲线相交面积求解相似度,该方法计算过程比较复杂,而且没有考虑超熵的作用;CFSM方法通过修正的正态云期望曲线的模糊贴近度对云模型的相似度进行测度,虽然充分利用了云模型的3个数字特征,但在熵占优的情况下容易削弱超熵对相似度的影响,导致计算结果的不准确性。

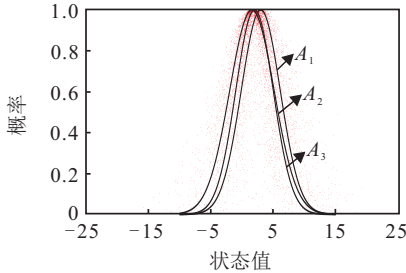


图2 3个云模型及其标准期望(例1)

例2 文献[25]给出3朵云, $A_1 = Y(1.41, 0.61, 0.31)$, $A_2 = Y(9.10, 0.66, 0.33)$ 和 $A_3 = Y(8.85, 0.75, 0.28)$, 分别利用本文所提出的方法和文献[22-26]的云相似度计算公式对3个云模型的相似度进行计算, 计算结果如表2所示。

表2 不同云相似度计算模型的比较(例2)

相似度	SM α	SCM ^[22]	LICM ^[23]	ECM ^[24]	CCM ^[25]	CFSM ^[26]
$s(A_1, A_2)$	0.2685	0.7626	0.9320	0.0001	0.0578	0.0000
$s(A_2, A_3)$	0.9861	0.9908	0.9999	0.8569	0.9265	0.8211
$s(A_3, A_1)$	0.2716	0.7540	0.9351	0.0000	0.0336	0.0000

由表2可知, A_2 和 A_3 的相似度最大(0.9861), A_1 和 A_2 的相似度最小(0.2685), 该结果与图3的直观印象一致. 本文方法与其他方法计算得到的结果都认为相似度最大的是(A_2, A_3). SCM方法和LICM方法计算(A_1, A_2)和(A_3, A_1)的相似度, 得到的结果与图3不相符, 明显不合理; 通过ECM方法和CFSM方法计算(A_1, A_2)和(A_3, A_1)的相似度分别为(0.001, 0.000)和(0.0000, 0.0000), 两组云模型相似度的差异太小, 无法比较或有效区分这两组云模型的差异性; CCM方法虽然能有效体现云模型相似度的差异, 但其未考虑超熵的作用。

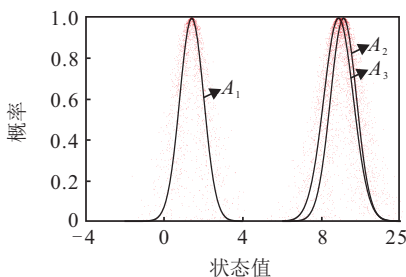


图3 3个云模型及其标准期望(例2)

显然, 本文所提出的方法计算过程更加简单, 不

仅考虑了云模型的所有数值特征, 而且更能体现不同类别中云模型的差异性, 且适应性较好。

2.3 基于云相似度的偏好聚类流程

假设通过聚类算法可以形成 t 个聚集, 并分别记为 G_1, G_2, \dots, G_t , 每个聚集相对应的成员个数为 g_1, g_2, \dots, g_t , 在聚类算法中决策者会给出一个阈值 δ 用于判断某一成员是否应该进入某一子聚类中, 具体算法如下。

Step 1: 初始化云计数器 $t = 1$, 并将需要聚类的云集合记为 Ω , 集合中云的个数记为 T , 设置聚类的阈值 δ 。

Step 2: 记 $i = 1$ 。

Step 3: 从集合 Ω 中选取一朵云, 并将该朵云从集合 Ω 中剔除. 如果 $g_i = 0$, 则将该朵云分配给 G_i ; 否则计算该朵云与 G_i 的相似度, 如果相似度大于阈值 δ , 则将这朵云分配给 G_i , 并计算这朵云与 G_i 的综合云, 最后执行Step 5; 否则, 执行Step 4。

Step 4: 记 $i = i + 1$, 并执行Step 3。

Step 5: 云计数器 $t = t + 1$, 如果 $t > T$, 则停止算法并记录所有聚集结果, 否则执行Step 2。

阈值 δ 的值由决策者根据实际决策问题给出. 一般地, 阈值 δ 越高, 聚集数越多; 阈值 δ 越低, 聚集数就越少。

3 决策方法

设决策问题有 N 个属性, 记为 $C = \{c_1, c_2, \dots, c_N\}$, 有 P 个方案, 记为 $X = \{x_1, x_2, \dots, x_P\}$. 决策群体记为 Ω , 由 M 个决策成员组成. 第 k 个决策成员对 i 个方案中第 j 个属性的决策值记为 $v_{ij}^k (i = 1, 2, \dots, P, j = 1, 2, \dots, N, k = 1, 2, \dots, M)$, 其中 v_{ij}^k 为语言变量, 记如下矩阵 V^k 为第 k 个决策者对所有方案的偏好矩阵:

$$V^k = \begin{bmatrix} v_{11}^k & v_{12}^k & \cdots & v_{1N}^k \\ v_{21}^k & v_{22}^k & \cdots & v_{2N}^k \\ \vdots & \vdots & \ddots & \vdots \\ v_{P1}^k & v_{P2}^k & \cdots & v_{PN}^k \end{bmatrix}$$

决策过程如下。

Step 1: 将各决策者的语言偏好转化为云, 设转换后第 k 个决策成员对 i 个方案中第 j 个属性的决策值记为 $r_{ij}^k (i = 1, 2, \dots, P, j = 1, 2, \dots, N, k = 1, 2, \dots, M)$, 其中 $R^k = (r_{ij}^k)_{P \times N}$ 记为转换后第 k 个决策者对各方案的偏好矩阵。

Step 2: 利用云模型聚类算法对决策者群体的偏好信息进行聚类. 由于聚集权重未知, 需要对聚集权重进行求解. 某一聚集内的专家人数越多, 该聚集对

群体决策的作用越大,说明该聚集的权重应越大;反之,某一聚集内的专家人数越少,说明该聚集的权重应越小.

设得到的聚类结果为 G_1, G_2, \dots, G_t , 各聚集对应的成员人数记为 g_1, g_2, \dots, g_t , 建立聚类权重求解模型求解聚集权重

$$w_i = \frac{g_i^2}{\sum_{j=1}^t g_j^2}, \quad i = 1, 2, \dots, t. \quad (6)$$

然后,利用式(2)计算各聚集的综合云,记为 $\tilde{G}_1, \tilde{G}_2, \dots, \tilde{G}_t$.

Step 3: 利用CWAA算子得到各方案 x_i 在属性 c_j 下的群体偏好值 r_{ij} .

Step 4: 利用CWAA算子得到各方案 x_i 的综合属性值 r_i , 利用定义3计算各方案的总计分 \hat{Q}_{r_i} ($i = 1, 2, \dots, P$), 然后根据总计分的大小对各方案进行排序.

4 算例

某公司需要选择一套网络安全系统,共有3个方案 $X = \{x_1, x_2, x_3\}$ 可供选择,在选择过程中主要考虑3个因素: c_1 (功能性), c_2 (有效性), c_3 (经济性), 属性权重分别为 $W = \{w_1 = 0.5, w_2 = 0.3, w_3 = 0.2\}$. 15位专家 $D = \{d_1, d_2, \dots, d_{15}\}$ 对各方案进行评估. 决策专家在如下语言评价集 $H = \{h_{-3} = \text{非常差}, h_{-2} = \text{差}, h_{-1} = \text{较差}, h_0 = \text{一般}, h_1 = \text{较好}, h_2 = \text{好}, h_3 = \text{非常好}\}$ 中选择语言变量对各方案进行评价,决策信息如表3所示.

表3 决策成员对各方案给出的决策信息

x_1	c_1	c_2	c_3	x_2	c_1	c_2	c_3	x_3	c_1	c_2	c_3
d_1	h_{-1}	h_0	h_{-2}	d_1	h_0	h_{-2}	h_{-1}	d_1	h_0	h_{-3}	h_{-3}
d_2	h_0	h_{-2}	h_0	d_2	h_{-2}	h_{-3}	h_{-1}	d_2	h_{-1}	h_0	h_{-2}
d_3	h_{-2}	h_{-3}	h_1	d_3	h_{-1}	h_{-2}	h_0	d_3	h_1	h_0	h_0
d_4	h_{-2}	h_2	h_{-1}	d_4	h_1	h_{-2}	h_0	d_4	h_{-1}	h_2	h_0
d_5	h_{-1}	h_{-2}	h_0	d_5	h_1	h_3	h_1	d_5	h_3	h_1	h_2
d_6	h_2	h_3	h_1	d_6	h_3	h_0	h_2	d_6	h_3	h_1	h_2
d_7	h_{-1}	h_{-2}	h_0	d_7	h_{-1}	h_{-2}	h_1	d_7	h_0	h_1	h_{-1}
d_8	h_{-2}	h_1	h_{-1}	d_8	h_1	h_{-2}	h_{-1}	d_8	h_{-1}	h_1	h_0
d_9	h_{-1}	h_{-2}	h_0	d_9	h_3	h_{-3}	h_{-2}	d_9	h_{-2}	h_{-1}	h_{-3}
d_{10}	h_1	h_{-1}	h_{-2}	d_{10}	h_{-2}	h_{-1}	h_2	d_{10}	h_0	h_{-1}	h_{-3}
d_{11}	h_3	h_1	h_{-3}	d_{11}	h_0	h_2	h_{-2}	d_{11}	h_1	h_3	h_3
d_{12}	h_2	h_1	h_2	d_{12}	h_{-1}	h_{-3}	h_{-3}	d_{12}	h_2	h_1	h_2
d_{13}	h_0	h_{-2}	h_{-1}	d_{13}	h_3	h_{-2}	h_3	d_{13}	h_{-2}	h_{-2}	h_1
d_{14}	h_1	h_2	h_{-1}	d_{14}	h_0	h_0	h_{-2}	d_{14}	h_3	h_{-1}	h_3
d_{15}	h_0	h_{-1}	h_{-2}	d_{15}	h_3	h_{-1}	h_1	d_{15}	h_1	h_2	h_{-2}
d_{16}	h_{-2}	h_{-2}	h_0	d_{16}	h_3	h_{-1}	h_{-2}	d_{16}	h_{-1}	h_{-1}	h_0
d_{17}	h_0	h_{-3}	h_0	d_{17}	h_2	h_{-3}	h_2	d_{17}	h_{-2}	h_{-2}	h_{-2}
d_{18}	h_{-1}	h_{-3}	h_{-1}	d_{18}	h_3	h_{-3}	h_2	d_{18}	h_{-2}	h_{-2}	h_{-3}
d_{19}	h_0	h_{-1}	h_{-1}	d_{19}	h_0	h_{-1}	h_2	d_{19}	h_0	h_{-1}	h_{-1}
d_{20}	h_2	h_2	h_3	d_{20}	h_0	h_{-2}	h_{-2}	d_{20}	h_2	h_2	h_3

Step 1: 将各决策者的语言偏好转化为云.

利用式(3)可得

$$\theta_{-3} = 0, \theta_{-2} = 0.225, \theta_{-1} = 0.385, \theta_0 = 0.500,$$

$$\theta_1 = 0.615, \theta_2 = 0.775, \theta_3 = 1.000,$$

这里取 $a = 1.4$.

给定论域 $[X_{\min}, X_{\max}] = [2, 8]$, 利用定义6可将7个语言标度转换为如下7朵云: $Y_{-3}(2, 1.775, 0.075)$, $Y_{-2}(3.35, 1.593, 0.136)$, $Y_{-1}(4.31, 1.260, 0.247)$, $Y_0(5, 1.153, 0.282)$, $Y_1(5.69, 1.260, 0.247)$, $Y_2(6.65, 1.593, 0.136)$, $Y_3(8, 1.775, 0.075)$.

Step 2: 设置云模型聚类算法的阈值为0.9, 同时

对决策者群体的偏好信息进行聚类, 并利用式(6)求解聚类权重, 聚类结果和聚集偏好分别如表4和表5所示.

Step 3: 利用CWAA算子得到各方案 x_i 在属性 c_j 下的群体偏好值, 结果如表6所示.

表4 聚类结果

聚集	成员数	成员	聚集权重
G_1	5	$d_1, d_7, d_8, d_{10}, d_{15}, d_{19}$	0.30
G_2	3	d_2, d_3, d_4, d_{16}	0.20
G_3	3	d_5, d_6, d_{14}	0.15
G_4	2	$d_9, d_{13}, d_{17}, d_{18}$	0.20
G_5	2	d_{11}, d_{12}, d_{20}	0.15

表5 聚集偏好

		c_1	c_2	c_3
G_1	x_1	Y(4.83, 1.37, 0.27)	Y(4.58, 2.26, 0.22)	Y(3.80, 2.13, 0.19)
	x_2	Y(5.81, 3.38, 0.16)	Y(4.00, 1.84, 0.21)	Y(5.95, 2.51, 0.18)
	x_3	Y(5.16, 1.75, 0.26)	Y(4.02, 2.91, 0.17)	Y(3.25, 2.86, 0.16)
G_2	x_1	Y(3.64, 2.03, 0.16)	Y(4.18, 3.27, 0.12)	Y(4.89, 1.78, 0.26)
	x_2	Y(7.41, 2.48, 0.12)	Y(3.31, 2.40, 0.15)	Y(4.03, 1.00, 0.20)
	x_3	Y(4.66, 1.86, 0.25)	Y(5.33, 0.94, 0.22)	Y(4.46, 2.03, 0.23)
G_3	x_1	Y(5.62, 2.27, 0.19)	Y(5.93, 3.10, 0.11)	Y(4.89, 1.78, 0.26)
	x_2	Y(6.43, 2.58, 0.19)	Y(6.33, 2.63, 0.19)	Y(4.88, 2.72, 0.16)
	x_3	Y(8.00, 1.78, 0.08)	Y(5.00, 1.78, 0.25)	Y(6.99, 1.40, 0.12)
G_4	x_1	Y(4.57, 1.57, 0.26)	Y(2.39, 2.27, 0.09)	Y(4.57, 1.57, 0.26)
	x_2	Y(7.61, 2.28, 0.09)	Y(2.11, 1.60, 0.08)	Y(6.02, 3.12, 0.11)
	x_3	Y(3.35, 0.98, 0.14)	Y(3.42, 1.34, 0.15)	Y(2.97, 2.88, 0.12)
G_5	x_1	Y(6.98, 2.37, 0.15)	Y(6.03, 2.04, 0.20)	Y(3.45, 2.55, 0.16)
	x_2	Y(4.80, 1.46, 0.27)	Y(2.94, 2.18, 0.12)	Y(2.94, 2.18, 0.12)
	x_3	Y(6.43, 1.86, 0.16)	Y(6.88, 1.82, 0.14)	Y(7.65, 2.25, 0.09)

表6 集结后的群体偏好值

	c_1	c_2	c_3
x_1	Y(4.98, 1.87, 0.22)	Y(4.48, 2.60, 0.16)	Y(4.28, 1.98, 0.23)
x_2	Y(6.43, 2.65, 0.17)	Y(3.67, 2.10, 0.16)	Y(4.97, 2.42, 0.16)
x_3	Y(5.32, 1.67, 0.20)	Y(4.74, 2.01, 0.19)	Y(4.66, 2.45, 0.16)

Step4: 利用CWAA得到各方案 x_i 的综合属性值 r_i , 结果如下:

$$r_1 = Y(4.69, 2.14, 0.21), r_2 = Y(5.31, 2.45, 0.17), r_3 = Y(5.01, 1.95, 0.19).$$

利用定义3计算群体偏好值 r_i 的总计分, 可得

$$\hat{Q}_{r_1} = 3.319, \hat{Q}_{r_2} = 3.753, \hat{Q}_{r_3} = 3.543.$$

由于 $\hat{Q}_{r_2} > \hat{Q}_{r_3} > \hat{Q}_{r_1}$, 方案的排序为

$$x_2 \succ x_3 \succ x_1,$$

方案 x_2 最优, 故选择方案 x_2 .

5 结论

本文基于云模型对属性值为语言变量的多属性大群体决策问题进行了研究, 分别提出了云距离和云相似度模型, 在此基础上提出了一种基于 α 截集的云聚类算法, 最后构建了一种基于云聚类算法的大群体决策方法. 所提出的方法具有很好的操作性, 丰富并发展了大群体群决策理论. 然而, 本文的阈值选取具有一定的主观性, 且针对的是属性权重完全已知的情形, 在未来的研究中, 将对阈值的选取和权重信息不完全或完全未知的情形进行研究.

参考文献(References)

- [1] Li D F, Sun T. Fuzzy linear programming approach to multi-attribute decision making with linguistic variables and incomplete information[J]. Advances in Complex Systems, 2007, 10(4): 505-525.
- [2] Wan S P, Li D F. Fuzzy mathematical programming approach to heterogeneous multi-attribute decision-making with interval-valued intuitionistic fuzzy truth degrees[J]. Information Sciences, 2015, 325: 484-503.
- [3] Li D F, Sun T. Fuzzy LINMAP method for multi-attribute group decision making with linguistic variables and incomplete information[J]. Int J of Uncertainty, Fuzziness and Knowledge-Based Systems, 2007, 15(2): 153-173.
- [4] 陈晓红, 刘蓉. 改进的聚类算法及在复杂大群体决策中的应用[J]. 系统工程与电子技术, 2006, 28(11): 1695-1699.
(Chen X H, Liu R. Improved clustering algorithm and its application in complex huge group decision-making[J]. Systems Engineering and Electronics, 2006, 28(11): 1695-1699.)
- [5] 徐选华. 面向特大自然灾害复杂大群体决策模型及应用[M]. 北京: 科学出版社, 2012: 2-3.
(Xu X H. Complex large group decision making models and its application oriented outside nature disasters[M]. Beijing: Science Press, 2012: 2-3.)
- [6] Liu B S, Shen Y H, Chen X H, et al. A complex multi-attribute large-group pls decision-making method in the interval-valued intuitionistic fuzzy environment[J]. Applied Mathematical Modelling, 2014, 38(17): 4512-4527.
- [7] 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 J of Operational Research*, 2015, 245 (1): 209-225.
- [8] Liu B S, Chen Y, Shen Y H, et al. A complex multi-attribute large-group decision making method based on the interval-valued intuitionistic fuzzy principal component analysis model[J]. *Soft Computing*, 2014, 18 (11): 2149-2160.
- [9] 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.
- [10] Liu B S, Shen Y H, Chen Y, et al. A two-layer weight determination method for complex multi-attribute large-group decision-making experts in a linguistic environment[J]. *Information Fusion*, 2015, 23: 156-165.
- [11] Palomares I, Martinez L, Herrera F. A consensus model to detect and manage noncooperative behaviors in large-scale group decision making[J]. *IEEE Trans on Fuzzy Systems*, 2014, 22 (3): 516-530.
- [12] 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): 440-450. (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): 440-450.)
- [13] 徐选华, 王敏赛, 陈晓红. 偏好冲突优化的多属性多阶段大群体决策方法[J]. *系统工程学报*, 2014, 29 (1): 48-55. (Xu X H, Wang M S, Chen X H. Multi-attribute & multi-stage large group decision-making method for preference conflict optimization[J]. *J of Systems Engineering*, 2014, 29(1): 48-55.)
- [14] Delgado M, Verdegay J L, Vila M A. A model for linguistic partial information in decision making problem[J]. *Int J of Intelligent Systems*, 1994, 9(4): 365-378.
- [15] 李德毅, 杜鹁. 不确定性人工智能[M]. 北京: 国防工业出版社, 2005: 57-383. (Li D Y, Du Y. *Artificial intelligence with uncertainty*[M]. Beijing: National Defence Industry Press, 2005: 57-383.)
- [16] Jiang Y P, Fan Z P, Ma J. A method for group decision making with multi-granularity linguistic assessment information[J]. *Information Sciences*, 2008, 178(4): 1098-1109.
- [17] Xu Y J, Merigó J M, Wang H M. Linguistic power aggregation operators and their application to multiple attribute group decision making[J]. *Applied Mathematical Modelling*, 2012, 36(11): 5427-5444.
- [18] Merigo J M, Gil-Lafuente A M. Induced 2-tuple linguistic generalized aggregation operators and their application in decision-making[J]. *Information Sciences*, 2013, 236: 1-16.
- [19] Wang J Q, Peng L, Zhang H Y, et al. Method of multi-criteria group decision-making based on cloud aggregation operators with linguistic information[J]. *Information Sciences*, 2014, 274: 177-191.
- [20] 王坚强, 杨恶恶. 基于蒙特卡罗模拟的直觉正态云多准则群决策方法[J]. *系统工程理论与实践*, 2013, 33 (11): 2859-2865. (Wang J Q, Yang W E. Multiple criteria group decision making method based on intuitionistic normal cloud by Monte Carlo simulation[J]. *Systems Engineering—Theory & Practice*, 2013, 33(11): 2859-2865.)
- [21] 鲍广宇, 连向磊, 何明, 等. 基于新型语言评估标度的二元语义改进模型[J]. *控制与决策*, 2010, 25(5): 780-784. (Bao G Y, Lian X L, He M, et al. Improved two-tuple linguistic representation model based on new linguistic evaluation scale[J]. *Control and Decision*, 2010, 25(5): 780-784.)
- [22] 张勇, 赵东宁, 李德毅. 相似云及其度量分析方法[J]. *信息与控制*, 2004, 33(2): 129-132. (Zhang Y, Zhao D N, Li D Y. The similar cloud and the measurement method[J]. *Information and Control*, 2004, 33(2): 129-132.)
- [23] 张光卫, 李德毅, 李鹏, 等. 基于云模型的协同过滤推荐算法[J]. *软件学报*, 2007, 18(10): 2403-2411. (Zhang G W, Li D Y, Li P, et al. A collaborative filtering recommendation algorithm based on cloud model[J]. *J of Software*, 2007, 18(10): 2403-2411.)
- [24] 李海林, 郭崇慧, 邱望仁. 正态云模型相似度计算方法[J]. *电子学报*, 2011, 39(11): 2561-2567. (Li H L, Guo C H, Qiu W R. Similarity measurement between normal cloud models[J]. *Acta Electronic Sinica*, 2011, 39(11): 2561-2567.)
- [25] 查翔, 倪世宏, 谢川, 等. 云相似度的概念跃升间接计算方法[J]. *系统工程与电子技术*, 2015, 37(7): 1676-1682. (Zha X, Ni S H, Xie C, et al. Indirect computation approach of cloud model similarity based on conception skipping[J]. *Systems Engineering and Electronics*, 2015, 37(7): 1676-1682.)
- [26] 龚艳冰, 蒋亚东, 梁雪春. 基于模糊贴近度的正态云模型相似度度量[J]. *系统工程*, 2015, 33(9): 133-137. (Gong Y B, Jiang Y D, Liang X C. Similarity measurement for normal cloud models based on fuzzy similarity measure[J]. *Systems Engineering*, 2015, 33(9): 133-137.)