

基于二元语义与粗糙集的多属性决策方法

徐天应^a, 千晓蓉^b

(昆明理工大学 a.理学院; b.城市学院, 昆明 650093)

摘要:文章针对多属性决策,提出一种基于二元语义和粗糙集的求解方法。基于依赖度和信息量来确定属性的权重。实例应用表明,该方法是合理科学的。

关键词:多属性决策;二元语义;粗糙集

中图分类号: O159 **文献标识码:** A **文章编号:** 1002-6487(2014)01-027-03

0 引言

在社会生活中,多属性决策已经成为决策科学,系统工程,管理学等领域中热点课题,多属性决策是利用某种决策方法对有限个决策方案在有限个属性下的属性值进行集成,从而可以对方案进行择优和排序的过程。由于人类的思维具有模糊性,以及客观事物具有复杂性,人们在对一些问题例如汽车的性能,企业的信用等进行评估时通常会给出一些定性的评估信息(如:好,中等,不好)。因此对属性值以语言变量或者不确定语言变量形式给出的多属性决策问题研究具有非常重要的理论意义和实际价值。语言多属性决策已经取得一定的研究成果。本文在此基础上提出了一种基于二元语义与粗糙集的多属性决策方法。二元语义以事先预备好的语言短语集合(例如:{很好,好,中等,差,很差})中的一个短语和一个实数值的二元形式来表达集成后所获得的语言评价信息,该方法可以减少语言评价信息集结和运算过程中出现的信息损失和信息的扭曲,同时在计算的可靠性和精确性等方面也明显优于其他的语言信息处理方法。粗糙集理论在处理离散数字信息功能上明显强于其他信息处理方法,对样本分布的要求不高,可以对一些并非按照常见分布规律分布的样本进行处理,同时可以很好地挖掘出指标之间潜在关系,定量分析出属性的重要性。然后再以属性的重要性进行属性约简,确定属性的权值,避免人为确定属性权重的主观性。通过实际例题应用,结果表明,该方法是合理的和科学的。

1 二元语义

二元语义信息是指某对象(或者目标、准则等)给出的评价结果用形如二元组 (s_k, α_k) 来表示。其中元素 s_k 和 α_k 的

含义描述如下:

(1) s_k 为预先定义好的语言评价集 S 中的第 k 个元素,它表示给出或者得到的语言信息与初始语言评价集合中最贴近的语言短语。例如一个由7个元素(即元素评价)构成的语言评价集 S 可定义为:

$$S = \{s_6 = FZ(\text{非常重要}), s_5 = HZ(\text{很重要}), \\ s_4 = Z(\text{重要}), s_3 = YB(\text{一般}), s_2 = C(\text{差}), \\ s_1 = HC(\text{很差}), s_0 = FZ(\text{非常差})\}.$$

S 具有如下性质:

有序性:当 $i \geq j$ 时,有 $s_i \geq s_j$;

存在逆运算算子: $Neg(s_j) = s_i$,其中 $j = T - i$;

极大化运算和极小化运算:当 $s_i \geq s_j$ 时,由 $\max s_i$,

$s_j = s_i; s_j = s_j$.

(2) α_k 称为符号转移值,且满足 $\alpha_k \in [-0.5, 0.5]$,它表示评价的结果与 s_k 的偏差。

定义1若 $s_k \in S$ 是一个语言短语,那么,相应的二元语义形式可以通过下面的函数 θ 获得:

$$\theta: S \rightarrow S \times [-0.5, 0.5]$$

$$\theta(s_i) = (s_i, 0), s_i \in S$$

定义2设实数 $\beta \in [0, T]$ 为语言评价集 S 经集结方法得到的实数,其中, T 为语言评价集 S 中元素的个数。则 β 可由如下的函数 Δ 表示二元语言信息

$$\Delta: [0, T] \rightarrow S \times [-0.5, 0.5]$$

$$\Delta(\beta) = \begin{cases} s_k & k = \text{round}(\beta) \\ \alpha_k = \beta - k & \alpha_k \in [-0.5, 0.5] \end{cases} \quad \text{其中, round 为四}$$

舍五入算子。

定义3设 (s_k, α_k) 是一个语言语义,其中 s_k 为 S 中第 k 个元素, $\alpha_k \in [-0.5, 0.5]$ 则存在一个逆函数 Δ^{-1} ,使其转换成相应的数值的 $\beta \in [0, T]$

$$\Delta^{-1}: S \times [-0.5, 0.5] \rightarrow [0, T]$$

基金项目:云南省教育厅科学研究基金资助项目(2012Y427)

作者简介:徐天应(1987-),男,江苏泰州人,硕士研究生,研究方向:决策分析。

千晓蓉(1962-),女,云南昆明人,教授,研究方向:统计模式识别。

$$\Delta^{-1}(s_k, \alpha_k) = k + \alpha_k = \beta$$

假设 (s_k, α_k) 和 (s_l, α_l) 为两个二元语言,关于二元语义的比较如下的规定:

若 $k < l$,则 $(s_k, \alpha_k) < (s_l, \alpha_l)$;

若 $k = l$,则

$\alpha_k = \alpha_l$,则 $(s_k, \alpha_k) = (s_l, \alpha_l)$;

$\alpha_k < \alpha_l$,则 $(s_k, \alpha_k) < (s_l, \alpha_l)$;

$\alpha_k > \alpha_l$,则 $(s_k, \alpha_k) > (s_l, \alpha_l)$.

定义4

设 $\{(s_1, \alpha_1), (s_2, \alpha_2), \dots, (s_n, \alpha_n)\}$ 是一组二元语义信息,则该组语言的算术平均算子定义为:

$$(\bar{s}, \bar{\alpha}) = \Delta \left(\frac{1}{n} \sum_{i=1}^n \Delta(s_i, \alpha_i) \right)$$

$$\bar{s} \in S, \bar{\alpha} \in [-0.5, 0.5]$$

定义5设:

$\{(s_1, \alpha_1), (s_2, \alpha_2), \dots, (s_n, \alpha_n)\}$

是一组二元语义信息,

$W = \{(w_1, \beta_1), (w_2, \beta_2), \dots, (w_n, \beta_n)\}$ 是对应的二元语义权重向量,则基于二元语义的加权算术平均算子定义为:

$$(\bar{s}, \bar{\alpha}) = \Delta \left(\frac{\sum_{i=1}^n \Delta^{-1}(w_i, \beta_i) \Delta^{-1}(s_i, \alpha_i)}{\sum_{i=1}^n \Delta^{-1}(s_i, \alpha_i)} \right) \bar{s} \in S, \bar{\alpha} \in [-0.5, 0.5]$$

定义6设 (s_i, α_i) 和 (s_j, α_j) 为任意两个二元语义信息,则它们之间的距离为:

$$d((s_i, \alpha_i), (s_j, \alpha_j)) = \Delta \left(\left| \Delta^{-1}(s_i, \alpha_i) - \Delta^{-1}(s_j, \alpha_j) \right| \right)$$

2 基于粗糙集的属性权重确定计算方法

在粗糙集中,设 $S = (U, A, V, f)$ 为一个信息系统,也称知识表示系统,其中 $U = \{x_1, x_2, \dots, x_m\}$ 是对象的非空有限集合,称为论域; $A = \{a_1, a_2, \dots, a_n\}$ 是属性的非空有限集合; V 是属性 A 所构成的域, V 是属性 a 的值域: $f: U \times A \rightarrow V$ 是一个信息函数, U 中任一元素取属性 a 在 V 中有唯一确定值,即 $V_a \in A, x \in U, f(x, a) \in V_a$.具有条件属性 C 和决策属性集 $D(A = C \cup D, C \cap D \neq \emptyset)$ 的信息系统称为决策信息系统,由条件属性和 C 和决策属性 D 定义的不可分辨关系对 U 产生的划分,分别表示为:

$$U/C = \{c_1, c_2, \dots, c_m\}$$

$$U/D = \{d_1, d_2, \dots, d_n\}$$

其中 $m = |U/C|, n = |U/D|$.

对 $p \in C, d_j$ 的上近似和下近似的分别定义为:

$$\underline{P}X(d_j) = \{x \mid [x]_p \subseteq d_j\}$$

$$\overline{P}X(d_j) = \{x \mid [x]_p \cap d_j \neq \emptyset\}$$

P 对于 D 的分类质量定义为:

$$\gamma(P, D) = \frac{\sum_{i=1}^n | \underline{P}X(d_i) |}{|U|}$$

式中 $\gamma(P, D)$ 表示 P 相对于 D 的重要程度。

2.1 属性权重确定

2.2 基于依赖度的属性权重确定方法

在决策表,不同的属性具有不同的重要性,为了找出某些属性的重要性,可以从表格中删去一些属性,在考虑没有该属性后分类的怎样变化。若去掉该属性相应分类变化较大,则说明该属性的重要性大;反之,说明该属性的重要性小。

在决策表 $S = (U, A, V, f)$ 中,对于属性 $a \in C$,关于 D 的重要性为:

$$sig(a) = \gamma(P, D) - \gamma(P - \{a\}, D) \quad (1)$$

2.3 基于信息量的属性权重确定方法

对于信息系统 S ,属性集 $P \subseteq C, U/P = \{C_1, C_2, \dots, C_l\}$,其中:

$l = |U/P|, P$ 的概率分布估计为 $p_i = \frac{|c_i|}{|U|}$ 。于是 P 的信息熵为:

$$H(P) = - \sum_{i=1}^l p_i \log p_i \quad (2)$$

条件属性集 P 已知决策属性集 D 的条件信息熵为:

$$H(D|P) = - \sum_{i=1}^l p_i \sum_{j=1}^n P(d_j|c_i) \log_2 P(d_j|c_i) \quad (3)$$

条件属性集 P 和决策属性集 D 的互信息量为:

$$I(D;P) = H(D) - H(D|P)$$

$$I(P;D) = H(P) - H(P|D)$$

两个信息源的互信息度反映了一个信息源到另外一个信息源获取的信息量。

当属性集删除了一个属性 a 后,把互信息度的变化定义为属性 a 对于决策集 D 的重要度,即:

$$SIG(a) = abs(I(P;D) - I(P - \{a\};D)) \quad (4)$$

在决策表 S 中,任意 $c \in C$,条件属性 c 的权重:

$$W(c) = \frac{sig(c) + SIG(c)}{\sum_{a \in C} sig(a) + SIG(a)} \quad (5)$$

依据上述理论分析和介绍,我们给出了此类多属性决策的步骤:

第一步:将决出者给出的语言评价信息转化。

第二步:利用粗糙集理论确定指标的综合权重。

第三步:计算各个评价对象的在所有指标下的综合评价价值。利用定义5的二元语义的算子加权算术算子对每个评价对象在不同指标下的群体评价价值进行二元语义集结。最终得到每个评价对象综合评价向量 S 。这些评价价值以二元语义的形式表示,即:

$$S = \{(s_1, \tilde{\alpha}_1), (s_2, \tilde{\alpha}_2), \dots, (s_m, \tilde{\alpha}_m)\}$$

3 企业信用的评估算例分析

考虑某一个地区企业信用评估,包括10个对象,每个对象具有5个条件属性,其中条件属性分别为创新能力 a_1 ,抵御市场风险能力 a_2 ,生产能力 a_3 ,财务质量 a_4 ,管理能力 a_5 ,决策属性 d 。现使用的语言评价集 S 是由5条语言评价信息构成的集合,即:

$$S = \{s_4 = HH(\text{很好}), s_3 = H(\text{好}), s_2 = Z(\text{中}), s_1 = C(\text{差}), s_0 = HC(\text{很差})\}。$$

表1 企业信用决策信息系统

| | s_1 | a_2 | a_3 | a_4 | a_5 | d |
|----------|-------|-------|-------|-------|-------|-----|
| u_1 | C | Z | Z | Z | C | H |
| u_2 | C | Z | C | C | Z | C |
| u_3 | C | Z | C | C | Z | C |
| u_4 | Z | H | H | Z | C | Z |
| u_5 | Z | Z | C | C | Z | Z |
| u_6 | Z | H | H | C | C | H |
| u_7 | C | Z | C | C | Z | C |
| u_8 | C | H | H | C | C | H |
| u_9 | C | Z | Z | C | C | H |
| u_{10} | Z | C | C | Z | Z | Z |

步骤1:根据评价语言信息集给出如表1的指标评价信息。

步骤2:约简和计算各个属性的依赖度和信息量

$$U/A = \{\{u_1\}, \{u_2, u_3, u_7\}, \{u_4\}, \{u_5\}, \{u_6\}, \{u_8\}, \{u_9\}, \{u_{10}\}\};$$

$$\{u_8\}, \{u_9\}, \{u_{10}\}\};$$

$$U/A - \{a_1\} = \{\{u_1\}, \{u_2, u_3, u_5, u_7\}, \{u_4\}, \{u_6\}, \{u_9\}, \{u_8\}, \{u_{10}\}\};$$

$$U/A - \{a_2\} = \{\{u_1\}, \{u_2, u_3, u_7\}, \{u_4\}, \{u_5\}, \{u_6\}, \{u_8\}, \{u_9\}, \{u_{10}\}\}$$

$$U/A - \{a_3\} = \{\{u_1\}, \{u_2, u_3, u_7, u_8\}, \{u_4\}, \{u_5\}, \{u_6\}, \{u_8\}, \{u_9\}, \{u_{10}\}\}$$

$$U/A - \{a_4\} = \{\{u_1, u_9\}, \{u_2, u_3, u_7\}, \{u_4\}, \{u_5\}, \{u_6\}, \{u_8\}, \{u_{10}\}\};$$

$$U/A - \{a_5\} = \{\{u_1\}, \{u_2, u_3, u_7\}, \{u_4\}, \{u_5\}, \{u_6\}, \{u_8, u_9\}, \{u_{10}\}\};$$

$$POS_A(d) = \{\{u_1, u_6, u_8\}, \{u_2, u_3, u_5, u_7\}, \{u_4, u_9, u_{10}\}\}$$

$$POS_{A-a_1}(d) = \{u_1, u_4, u_8, u_{10}\}$$

$$POS_{A-a_2}(d) = \{\{u_1\}, \{u_2, u_3, u_7\}, \{u_4\}, \{u_5\}, \{u_6\}, \{u_8\}, \{u_9\}, \{u_{10}\}\}$$

$$\{u_{10}\}$$

$$POS_{A-a_3}(d) = \{\{u_1\}, \{u_4\}, \{u_5\}, \{u_6\}, \{u_8\}, \{u_9\}, \{u_{10}\}\}$$

$$POS_{A-a_4}(d) = \{\{u_2\}, \{u_3\}, \{u_4\}, \{u_5\}, \{u_7\}, \{u_8\}, \{u_{10}\}\}$$

$$POS_{A-a_5}(d) = \{\{u_1\}, \{u_2\}, \{u_3\}, \{u_4\}, \{u_5\}, \{u_6\}, \{u_7\}, \{u_{10}\}\}$$

从以上叙述可知,A的 d 约简为 $\{a_1, a_2, a_3, a_4, a_5\}$ 。

于是根据(1)式和(2)式有:

$$\text{Sig}(a_1) = 0.6, \text{Sig}(a_2) = 0, \text{Sig}(a_3) = 0.4, \text{Sig}(a_4) = 0.2,$$

$$\text{Sig}(a_5) = 0.2$$

于是根据(3)式和(4)式有:

$$\text{SIG}(a_1) = 0.5, \text{SIG}(a_2) = 0.5, \text{SIG}(a_3) = 0.2, \text{SIG}(a_4) = 0.1,$$

$$\text{SIG}(a_5) = 0.2$$

于是根据式(5)可得:

$$W(a_1) = 0.4, W(a_2) = 0.2, W(a_3) = 0.2, W(a_4) = 0.1, W(a_5) =$$

$$0.1。$$

步骤3:综合评估

首先对数据进行无量纲化出来,利用定义5的二元语义的算子加权算术算子对每个评价对象在不同指标下的群体评价价值进行二元语义集结。最终得到每个评价对象综合评价向量 S 。利用定义5的二元语义的算子加权算术算子对每个评价对象在不同指标下的群体评价价值进行二元语义集结。最终得到每个评价对象综合评价向量 S 。这些评价价值以二元语义的形式表示,即:

$$S = \{(s_1, 0.15), (s_1, 0.13), (s_1, 0.13),$$

$$(s_1, 0.19), (s_1, 0.17), (s_1, 0.22),$$

$$(s_1, 0.13), (s_1, 0.17), (s_1, 0.13),$$

$$(s_1, 0.16)\}$$

因此企业信用最好是 u_6 。

4 结论

近年来,有关语言决策理论已经引起了许多学者的高度关注,它在决策科学中具有重要的应用。确定属性的权重的方法是当前管理决策领域的研究的热点之一。本文提出了一种基于二元语义和粗糙集的多属性决策方法。通过粗糙集理论确定属性权重,进行二元语义信息集结。通过实际例题应用,证明该方法是合理的和科学的。该方法在企业战略决策,项目评估,企业质量信用评估,企业信用评估,企业知识管理绩效评价等诸多领域将有着广泛的应用前景。

参考文献:

- [1] Herrera F, Herrera-Viedma E, Verdegay L.A Sequential Selection Process in Group Decision Making with a Linguistic Assessment Approach[J]. Information Sciences, 1995, 85(4).
- [2] Herrera F, Martinez L.A 2-Tuple Linguistic Representation Model for Computing with Words[J]. IEEE Trans on Fuzzy Systems, 2000, 8(6).
- [3] Herrera F, Herrera-Viedma E. Linguistic Decision Analysis Steps for Solving Decision Problems under Linguistic Information[J]. Fuzzy Sets and Systems, 2000, 115(1).
- [4] 姜艳萍, 樊治平. 基于二元符号运算的群决策方法[J]. 系统工程与电子技术, 2003, 25(11).
- [5] 李洪燕, 樊治平. 一种基于二元语义的多指标群决策方法[J]. 东北大学(自然科学版), 2003, 24(5).
- [6] 刘培德, 关忠良. 一种基于二元语义的混合型多属性决策方法[J]. 控制与决策, 2009, 24(7).
- [7] 张文修, 吴伟志, 梁吉业等. 粗糙集理论与方法[M]. 北京: 科学技术出版社, 2001.
- [8] 万树平. 基于分布规划的区间梯形直觉模糊数多属性决策方法[J]. 控制与决策, 2012, 27(3).

(责任编辑/亦 民)