

基于区间二型模糊集的多属性三支决策

王宁

(北京邮电大学理学院, 北京 100876)

摘要:三支决策为处理现实中的不确定性问题提供了一个良好的工具。区间二型模糊集在解决不确定性问题和描述不确定信息上有较好的优势。基于此,本文提出了一个区间二型模糊环境下的多属性三支决策模型。首先,本文定义了一个转换函数,将得到的区间二型模糊信息转换为数值;其次,考虑到各个属性的权重可能不同,定义了一个加权的条件概率的计算公式,进一步给出了区间二型模糊环境下的多属性三支决策模型。最后通过具体的案例证明了本文提出多属性三支决策模型的合理性和有效性。

关键词:三支决策;区间二型模糊集;条件概率

中图分类号:O225

文献标识码:A

文章编号:1004-7344(2022)27-0187-03

0 引言

1982年,波兰学者 PAWLAK^[1]提出了粗糙集理论,为处理不确定性提供了一个有效的工具。为了完善粗糙集理论,YAO^[2]提出了决策理论粗糙集的概念。基于决策理论粗糙集,YAO^[2]进一步提出了三支决策的理论。该理论将决策论域划分为三个互不相交的域,并给出了对应的解释。即用接受,拒绝和不做承诺三个决策行动来解释粗糙集的三个域。目前,三支决策已经成功应用到很多领域并取得了很好的结果^[3]。

目前,很多三支决策模型都是基于一型模糊集并取得了很好的结果,但随着决策环境不断地变复杂,一型模糊集已经不能满足决策的需求。二型模糊集作为一型模糊集的推广^[4-6],能更好地描述不确定信息。故 Xiao等^[7]提出了基于二型模糊集的三支决策。考虑到决策者的行动可能会受到不同风险态度的影响,Wang等^[8]提出了一个区间二型模糊环境下基于后悔理论的三支决策模型^[9]。在实际的决策问题中,往往涉及多个属性而不是单一的属性,所以多属性三支决策受到了很多学者的关注。区间二型模糊集在表达和计算上更加的简洁,因此我们将多属性三支决策推广到区间二型模糊集的环境下,来提高模型处理不确定信息的能力。

1 预备知识

1.1 区间二型模糊集

定义1 假设 U 是一个论域, \tilde{A} 是一个二型模糊集可以表示为:

$$\tilde{A} = \{[(x, u), \mu_{\tilde{A}}(x, u)] \mid \forall x \in X, \forall u \in J_x \subseteq [0, 1]\} \quad (1)$$

其中: $\mu_{\tilde{A}}(x, u)$ 是 \tilde{A} 的二型隶属函数,且 $0 \leq \mu_{\tilde{A}}(x, u) \leq 1$ 。当 $\mu_{\tilde{A}}(x, u) = 1$ 时, \tilde{A} 称作一个区间二型模糊集。特别的,我们有时用式(2)来表示 \tilde{A} :

$$\int \int_{x \in U, u \in J_x} \frac{\mu_{\tilde{A}}(x, u)}{(x, u)} \quad (2)$$

对 U 上的一个区间二型模糊集 \tilde{A} , 可以用式(3)表示:

$$\int \int_{x \in U, u \in J_x} \frac{1}{(x, u)} \quad (3)$$

为了方便计算,有时用 $[(a^+, b^+, c^+, d^+; m^+, n^+), (a^-, b^-, c^-, d^-; m^-, n^-)]$ 来表示区间二型模糊集,其中 V 是一个基本集, $a^+, b^+, c^+, d^+, a^-, b^-, c^-, d^- \in V, m^+, n^+, m^-, n^- \in [0, 1]$ 。

1.2 决策理论粗糙集和三支决策

由决策理论粗糙集的定义可知,一个模糊概念的上下近似可以通过条件概率和阈值来确定。在决策理论粗糙集中通常包含两个状态和三个行动,分别记作 $\Omega = \{C, \neg C\}, D = \{a_p, a_B, a_N\}$ 。 C 和 $\neg C$ 分别表示对象属于和不属于 C 。 a_p 表示对象 $x \in \text{pos}(C)$, 即接受决策; a_N 表示对象 $x \in \text{neg}(C)$, 即拒绝决策; a_B 表示对象 $x \in b_n(C)$, 即延迟决策; a_B 表示对象 $x \in b_n(C)$, 即延迟决策。每个行动的损失函数如表1所示,其中 $\lambda_{pp}, \lambda_{BP}, \lambda_{NP}$ 分别表示对象 x 属于状态 C 时执行行动 a_p, a_B, a_N 所产生的损失,

$\lambda_{PN}, \lambda_{BN}, \lambda_{NN}$ 分别表示对象 x 不属于状态 C 时执行行动 a_P, a_B, a_N 所产生的损失。假设 $P(C | [x])$ 是对象属于 $[x]$ 的情况下属于 C 的概率, 然后通过一对阈值 $0 \leq \beta \leq \alpha \leq 1$, 可以得到如下的决策规则:

接受: $\text{pos}(C) = \{x_i \in U \mid P(C | [x]) \geq \alpha\}$,

不做承诺: $\text{bn}(C) = \{x_i \in U \mid \beta < P(C | [x]) < \alpha\}$,

拒绝: $\text{neg}(C) = \{x_i \in U \mid P(C | [x]) \leq \beta\}$ 。

根据贝叶斯决策理论和最小风险原则, 我们可以用式(4)至式(6)计算出阈值参数 α, β, γ , 其中 $0 \leq \beta \leq \gamma \leq \alpha \leq 1$:

$$\alpha = \frac{\lambda_{PN} - \lambda_{BN}}{(\lambda_{PN} - \lambda_{BN}) + (\lambda_{BP} - \lambda_{PP})}; \quad (4)$$

$$\beta = \frac{\lambda_{BN} - \lambda_{NN}}{(\lambda_{BN} - \lambda_{NN}) + (\lambda_{NP} - \lambda_{PP})}; \quad (5)$$

$$\gamma = \frac{\lambda_{PN} - \lambda_{NN}}{(\lambda_{PN} - \lambda_{NN}) + (\lambda_{NP} - \lambda_{PP})}。 \quad (6)$$

表 1 损失函数的矩阵

	C	$\neg C$
a_P	λ_{PP}	λ_{PN}
a_B	λ_{BP}	λ_{BN}
a_N	λ_{NP}	λ_{NN}

2 基于区间二型模糊集的多属性三支决策

当决策环境更加复杂时, 使用一型模糊集有时不能很好的表达决策对象的评价信息。区间二型模糊集作为一型模糊集的拓展, 在表达不确定信息上优于一型模糊集, 所以我们提出了一个区间二型模糊环境下的多属性三支决策模型。

虽然区间二型模糊集在表达不确定信息上比一型模糊集更有优势, 但在表达形式和计算上会更加的复杂。当得到这些模糊评价信息时, 为了便于比较和计算, 我们首先要对这些区间二型模糊信息进行处理, 将区间二型模糊信息转化为数值的形式并得到一个数值的信息表。因此我们需要提出一个函数将区间二型模糊信息转换为数值。

定义 2 设 $\tilde{A} = [(a^+, b^+, c^+, d^+; m^+, n^+), (a^-, b^-, c^-, d^-; m^-, n^-)]$ 是 V 上的区间二型模糊集, 定义一个函数 $T: \tilde{A} \rightarrow [0, 1]$:

$$T(\tilde{A}) = \frac{b^+ \times m^+ + b^- \times m^- + c^+ \times n^+ + c^- \times n^-}{4 \times \max\{V\}}。 \quad (7)$$

通过定义 2 提出的函数, 我们可以将所有的区间二型模糊信息转化为数值, 下文将用一个案例来说明该定义的计算过程。

一所学校需要对一个学生进行评估, 假设 $V = \{1, 2, \dots, 10\}$, 班主任对该学生的评价为 $\tilde{A} = [(5, 7, 8, 9; 1, 0.9), (6, 7, 7, 8; 0.8, 0.8)]$, 其中 \tilde{A} 是 V 上的一个区间二型模糊集。假设 $V = \{1, 2, \dots, 10\}$, $\tilde{A} = [(5, 7, 8, 9; 1, 0.9), (6, 7, 7, 8; 0.8, 0.8)]$ 是 V 上的一个区间二型模糊集。利用函数 T 我们可以计算出:

$$\begin{aligned} T(\tilde{A}) &= \frac{b^+ \times m^+ + b^- \times m^- + c^+ \times n^+ + c^- \times n^-}{4 \times \max\{V\}} \\ &= \frac{7 \times 1 + 8 \times 0.9 + 7 \times 0.8 + 7 \times 0.8}{4 \times 10} \\ &= 0.635。 \end{aligned} \quad (8)$$

相当于班主任对这个学生的评价可以近似为 0.635。

处理完区间二型模糊信息, 则需要考虑三支决策的两个关键因素: 条件概率和阈值。首先, 考虑到不同的属性的重要程度可能是不同的, 如果将各个属性的权重设为一样的, 那么得到的决策结果可能会有偏差, 因此我们将权重加入条件概率的公式之中, 并给出了加权的条件概率的定义。

定义 3 设论域为 U , 属性集为 $At = \{a_1, a_2, \dots, a_n\}$, $F = \{f_1, f_2, \dots, f_n\}$ 为 $U \times At \rightarrow \overline{F(V)}$ 的映射的集合, 其中 V 是一个基本集, $\overline{F(V)}$ 是 V 上所有区间二型模糊集的集合。设 $(\omega_1, \omega_2, \dots, \omega_n)$ 为各个属性的权重, 我们定义对象 x 属于状态 C 的条件概率为:

$$P(C | x) = \sum_{k=1}^n \omega_k T[f(x, a_k)]。 \quad (9)$$

利用条件概率的公式计算出条件概率后, 则需要考虑三支决策的另一个关键因素: 阈值参数。阈值参数可以通过贝叶斯风险决策理论和最小风险原则计算得到, 为了减小模型的计算量, 本文的阈值参数是根据经验直接给出的。给定一对阈值 α 和 β ($0 \leq \beta \leq \alpha \leq 1$), 如下的三支决策规则被给出:

接受: $\text{pos}(C) = \{x_i \in U \mid P(C | x_i) \geq \alpha\}$,

不做承诺: $\text{bn}(C) = \{x_i \in U \mid \beta < P(C | x_i) < \alpha\}$,

拒绝: $\text{neg}(C) = \{x_i \in U \mid P(C | x_i) \leq \beta\}$ 。

通过以上讨论, 我们建立了一个基于区间二型模糊集的多属性三支决策模型。

3 算例分析

本节中笔者将通过一个数值的例子来验证所提出的基于区间二型模糊集的多属性三支决策的有效性和合理性。

假设一个科技公司需要招聘后端开发工程师，一共有 3 名求职者 $U=\{x_1, x_2, x_3\}$ 通过了笔试的筛选进入了面试的流程，面试官将通过专业知识、实践经验和综合分析能力三个方面对应聘者进行评估。这三个属性的权重分别为 $(0.3, 0.3, 0.4)$ ，基本集 $V=\{1, 2, \dots, 10\}$ ，状态集 $\Omega=\{C, \neg C\}$ ，其中状态 C 表示通过面试。得到的信息如表 2 所示。

表 2 一个关于面试的区间二型模糊信息表

U/At	专业知识	实践经验	综合分析能力
x_1	$[(5, 7, 8, 9; 1, 0.9), (6, 7, 7, 8; 0.8, 0.8)]$	$[(1, 3, 4, 5; 0.9, 0.9), (2, 3, 3, 4; 0.7, 0.7)]$	$[(7, 8, 9, 10; 1, 1), (8, 9, 9, 9.5; 0.9, 0.9)]$
x_2	$[(9, 10, 10, 10; 1, 1), (9.5, 10, 10, 10; 0.9, 0.9)]$	$[(5, 6, 7, 8; 0.9, 1), (6, 7, 7, 8; 0.8, 0.8)]$	$[(7, 8, 9, 10; 1, 0.9), (8, 9, 9, 10; 0.7, 0.7)]$
x_3	$[(1, 3, 5, 6; 0.8, 0.9), (2, 3, 4, 5; 0.7, 0.6)]$	$[(3, 5, 6, 7; 1, 0.9), (4, 5, 5, 6; 0.7, 0.7)]$	$[(5, 6, 7, 9; 1, 0.9), (6, 7, 7, 8; 0.8, 0.8)]$

首先，我们利用定义 2 中的函数 T 对得到的区间二型模糊信息进行处理，将所有的区间二型模糊信息转化为数值，得到的信息如表 3 所示。

表 3 一个关于面试的数值信息

U/At	专业知识	实践经验	综合分析能力
x_1	0.6350	0.2625	0.8300
x_2	0.9500	0.5900	0.7175
x_3	0.2850	0.4350	0.5857

然后，通过定义 3 给出的条件概率的计算公式，得到各个求职者的条件概率如下：

$$P(C | x_1) = 0.601,$$

$$P(C | x_2) = 0.747,$$

$$P(C | x_3) = 0.450.$$

根据以往的经验，我们设阈值 $\alpha=0.7$ 和 $\beta=0.4$ ，根据推导出的三支决策规则可以得到：

$$\text{接受: } \text{pos}(C) = \{x_i \in U \mid P(C | x_i) \geq 0.7\} = \{x_2\},$$

$$\text{不做承诺: } \text{bn}(C) = \{x_i \in U \mid 0.4 < P(C | x_i) < 0.7\} = \{x_1, x_3\},$$

$$\text{拒绝: } \text{neg}(C) = \{x_i \in U \mid P(C | x_i) \leq 0.4\} = \emptyset.$$

通过上面的决策结果，我们可以知道求职者 x_2 通过了面试，得到了该科技公司的工作；针对求职者 x_1 和 x_3 ，该科技公司没有给出正面的评价，需要增加面试的次数再进一步考虑是否录用，决策的结果如图 1 所示。

4 结语

本文给出了一个基于区间二型模糊集的多属性三支决策模型。与基于一型模糊集的三支决策模型相比，该模型在表达模糊信息上更有优势。且能够处理更多的模糊信息和决策问题。条件概率和阈值是三支决策

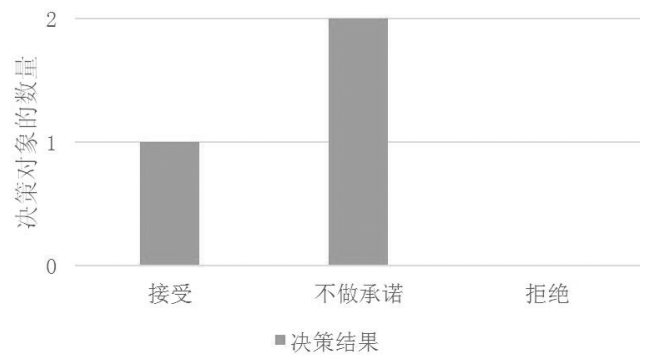


图 1 决策结果

的两个关键因素，本文给出了条件概率的计算公式，但阈值是根据经验直接给定的。为了完善我们的模型，在未来将研究区间二型模糊环境下阈值的确定方法。另外，随着决策环境不断地变复杂，决策过程中往往涉及多个决策者，所以我们将研究基于区间二型模糊集的多属性群三支决策。

参考文献

- [1] PAWLAK Z. Rough set[J]. International Journal of Computer and Information Science, 1982, 11 (5) : 341-356.
- [2] YAO Y Y. Three-way decisions with probabilistic rough sets[J]. Information Sciences, 2010, 180 (3): 341-353.
- [3] 于洪, 王国胤, 李天瑞, 等. 三支决策: 复杂问题求解方法与实践[M]. 北京: 科学出版社, 2015.
- [4] ZADEH L A. The concept of a linguistic variable and its applications in approximate reasoning- I [J]. Information Sciences, 1975, 8(3): 199-249.
- [5] ZADEH L A. The concept of a linguistic variable and its applications in approximate reasoning- II [J]. Information Sciences, 1975, 8(4): 301-357.
- [6] ZADEH L A. The concept of a linguistic variable and its applications in approximate reasoning- III [J]. Information Sciences, 1975, 9(1): 43-80.
- [7] XIAO Y C, HU B Q, ZHAO X R. Three-way decisions based on type-2 fuzzy sets and interval-valued type-2 fuzzy sets[J]. Journal of Intelligent & Fuzzy Systems, 2016, 31: 1385-1395.
- [8] WANG T X, LI H X, QIAN Y H, et al. A regret-based three-Way decision model under interval type-2 fuzzy environment[J]. IEEE Transactions on Fuzzy Systems, 2020(99): 175-189.

收稿日期: 2022-05-10

作者简介: 王宁(1997—), 男, 汉族, 河南驻马店人, 硕士研究生, 研究方向为代数与密码。