

文章编号: 1001-0920(2005)03-0294-05

基于粗集理论的多属性决策分析

安利平^{1,2}, 陈增强², 袁著祉²

(1. 南开大学 国际商学院, 天津 300071; 2. 南开大学 信息技术科学学院, 天津 300071)

摘要: 针对多人多属性决策问题中多个决策者对同一类对象决策结果不一致的问题, 通过对多属性群决策问题的描述, 首次提出将不同决策者的不一致决策对象进行合并分析, 得到更加直观明确的规则。对规则集进行构成分析, 从而解释决策者之间的冲突所在。

关键词: 多属性决策分析; 粗集; 决策规则; 规则获取

中图分类号: TP18; N94

文献标识码: A

Multi-attribute decision analysis based on rough set theory

AN L i-ping^{1,2}, CH EN Zeng-qiang², YUAN Zhu-zhi²

(1. College of International Business, Nankai University, Tianjin 300071, China; 2. College of Information Technical Science, Nankai University, Tianjin 300071, China Correspondent: AN L i-ping, E-mail: anliping2000@sina.com)

Abstract: To the inconsistency of decision making in multi-person and multi-attribute decision problems, a multi-attribute group decision problem is described. Combination analysis is put forward in order to generate more explicit rules. And, the composition of the rule set is analyzed to explain the conflicts of the different agents.

Key words: multi-attribute decision analysis; rough sets; decision rules; rule generation

1 引言

粗集理论是一种处理模糊性和不确定性的数学方法, 该理论为解决多属性决策问题提供了有利的工具。利用粗集方法分析决策表, 可以评价特定属性的重要性, 建立属性集的约简、核以及从决策表中去除冗余属性, 从约简的决策表中产生分类规则并利用得到的规则进行决策。

许多情况下, 多个决策者对同一类对象进行决策时, 经常会产生不一致的情况。这种不一致性不能简单地视为噪声或误差, 它们可能传递着某些重要的决策信息, 因此应在建立决策者偏好模型时加以考虑。这就需要综合考虑各个决策者的经验决策数据, 从而发现他们的群决策规则, 找出冲突的原因所在, 为制定协调策略提供依据。这些问题也是群决策理论中的重要研究内容, 然而目前的研究成果还比

较少, 尚未形成一个完整的体系。

本文首先给出了基于粗集理论的多属性决策分析的一般方法; 然后描述了一种多属性群决策问题; 最后通过决策表的合并分析, 导出了意义明确的规则集, 并对规则集的构成进行分析。

2 基于粗集理论的多属性决策分析

本节对基于粗集的多属性决策分析给出一个研究框架和分析步骤。有关粗集理论的基本概念可参阅文献[1]。

2.1 约简与属性的重要性

多数情况下, 决策者关于决策对象形成的决策表中存在某些对决策任务无关或不重要的属性。如果去除这些冗余属性, 并找到最小的相关属性集(即约简), 具有与全部属性相同的分类决策能力, 就能简化决策表。

收稿日期: 2004-04-07; 修回日期: 2004-06-16

基金项目: 国家自然科学基金项目(60374037, 70471040)。

作者简介: 安利平(1971—), 男, 河北张家口人, 讲师, 博士后, 从事运筹技术、专家系统等研究; 陈增强(1964—), 男, 天津人, 教授, 博士生导师, 从事自适应控制、智能预测控制等研究。

粗集理论中衡量属性重要性的指标是从分类角度出发,即属性重要性可通过从属性集中去掉该属性后对决策表的分类能力的影响来测度。然而,这种重要性不能反映决策者对属性的先验权重或属性在实际问题中的作用。将两种重要性相结合作为选择属性或规则的准则,不失为一种合理的解决方案。例如, Hu 提出的计算约简的算法^[2]就利用核与用户指定的属性(反映了用户的先验权重)作为计算约简的出发点。

2.2 获取决策规则

获取规则最一般的方法是构造决策表 S 的决策矩阵,通过计算决策函数的主蕴涵得到每个对象 x

U 的约简,每个约简对应于 S 中的一条最小条件描述数量的决策规则。文献[3]提出一种规则获取算法,该算法以等价类为单位进行规则获取,可同时计算与规则相连的几个数量参数,如支持量、精度、适用度等。文献[4]在此基础上提出了规则更新的增量式算法。

从决策表导出规则的主要问题是决定哪些属性应包括在规则的前件部分。利用粗集方法虽可计算最小决策规则,但导出的规则可能包含噪声,并且这种详尽的规则集容易造成过度匹配训练数据,对新对象分类质量变差。一般说,较少的规则虽对已知对象的分类不够完美(被噪声所影响),但规则集的泛化能力较强,对新对象有较高的分类质量。根据这一主导思想提出了基于不同测度的一些规则获取算法,如缩小边界域^[5]、在给定的阈值下保持正域^[6]、熵^[7]、动态约简规则^[8]等。

2.3 分类决策

从对象集中导出的规则集可用于对新对象的分类识别。当遇到一个新对象 x (只有条件属性,没有决策属性)需要作出决策时,则扫描规则集中的相关规则,将对象描述与决策规则相匹配,可能产生以下几种结果:

- 1) 新对象与一条确定性规则匹配;
- 2) 新对象与多条确定性规则匹配,且结论相同;
- 3) 新对象与一条或多条不确定性规则匹配,或与多条确定性规则匹配,但规则结论不同;
- 4) 没有规则与新对象匹配。

对于情况 1) 和 2),由于规则的结论只有一个,对象的分类是确定的,可直接根据规则的参数了解进一步的信息。

2.3.1 规则冲突的解决

对于情况 3),有以下几种策略:

选择适用度 $\text{coverage}(\alpha, \beta)$ 或精度 $\text{accuracy}(\alpha, \beta)$ 最大的规则结论,其中 α, β 为一条规则。

投票法^[9]:它是根据规则分类的专门技术,在实践中取得了良好的效果。

设需要分类的对象为 x ,扫描规则集 R ,找出规则前件匹配 x 的所有规则,即激活规则集 $R(x)$ 。每一规则 $r \in R(x)$ 为其后件的决策类指定一个数量作为 r 的得票数 $\text{votes}(r)$ 。 $\text{votes}(r)$ 有多种计算方法,一般取规则 r 的支持量,也可取更复杂的规则质量测度。

将结论为决策类 β 的所有规则的累计得票数除以标准化因素 $\text{nom}(x)$,便得到每一决策类的可信度 $\text{certainty}(x, \beta)$,即

$$\text{certainty}(x, \beta) = \frac{\text{votes}(\beta)}{\text{nom}(x)} \quad (1)$$

其中: $\text{votes}(\beta) = \sum_{r \in R_\beta} \text{votes}(r)$, $R_\beta = \{r \in R(x) : r$ 的结论为 $\beta\}$, $\text{nom}(x)$ 是标准化因素。一般地, $\text{nom}(x)$ 只作为比例因素,它是所有激活规则得票数的总和。

对象跟踪法^[9]:标准投票法虽然简单易行,但没有考虑 $R(x)$ 中导出规则的对象情况。如果两条或多条激活规则的对象集相交,则这些交集的对象在投票过程中将参与两次或多次。对象跟踪法就是为避免这种情况而提出的。

设从决策表 S 中导出一条规则 $\alpha, \beta \in R(x)$,定义规则的跟踪对象为 α_s ,即匹配规则前件的对象集合。所有激活规则的跟踪对象作为计算可信度的基础,即

$$\text{certainty}(x, d = v) = \frac{|\{y \in U : \exists (\alpha, \beta) \in R(x), \text{tracked}(R(x)) : d(y) = v\}|}{|\text{tracked}(R(x))|} \quad (2)$$

其中 $\text{tracked}(R(x)) = \{y \in U : \exists (\alpha, \beta) \in R(x), \text{使得 } y \in \alpha_s\}$ 。

考虑规则的前件长度:规则前件的长度表明了规则的详细和精确程度。虽然长的前件对规则适用性有影响,但这类规则的结论精度较高。在系统 LERS^[10]中考虑规则前件包含的属性信息可用于对新对象分类。

对于完全匹配的情况,设 $\alpha, \beta \in R(x)$,规则前件长度为 $|\alpha|$,支持量为 $\text{support}(\alpha, \beta)$,则定义决策类 β 的支持数为

$$s(\beta) = \frac{\text{support}(\alpha \rightarrow \beta) \times |\alpha|}{\alpha \beta R(x)} \quad (3)$$

选择最大的 $s(\beta)$ 所对应的结论作为决策结果, 这种方法适用于前件较长的规则. 对于简短的规则, LERS 系统不考虑 $|\alpha|$, 最终结果只取决于 support, 即 $s(\beta) = \frac{\text{support}(\alpha \rightarrow \beta)}{\alpha \beta R(x)}$. 这与投票法是一致的.

2.3.2 没有匹配规则

对于情况 4), 即无匹配规则, 目前主要利用以下两种方法:

部分匹配法: 在系统 LERS 中, 当没有任何规则与对象匹配时, 可考虑部分匹配的规则, 即对象匹配规则的某些条件属性. 设对象 x 所有被选择的部分匹配规则形成集合 $R_p(x)$, 规则 $\alpha \rightarrow \beta \in R_p(x)$ 被匹配的属性数为 $\text{mattri}(\alpha \rightarrow \beta)$. 定义匹配因素

$$\text{m fac}(\alpha \rightarrow \beta) = \frac{\text{mattri}(\alpha \rightarrow \beta)}{|\alpha|} \quad (4)$$

决策为 β 的支持数

$$s(\beta) = \frac{\text{m fac}(\alpha \rightarrow \beta) \times \text{support}(\alpha \rightarrow \beta)}{\alpha \beta R_p(x)} \quad (5)$$

对象被分类到 $s(\beta)$ 最大的决策类中, 该方法适于详细的规则.

相近规则法: Slowinski 提出一种与新对象的描述“最相近”的规则集^[11], 但这种方法需要决策者对算法有较透彻的了解, 因为在算法的实施中, 要求决策者对属性指定相对重要性的权重和属性的否决阈值.

3 多属性群决策分析

3.1 问题描述

在多属性决策问题中, 经常出现多个决策者对同一对象的决策结果不一致的情况, 即决策的策略互不相同. 这时, 可利用粗集理论的规则获取方法, 对每个决策者的决策对象分别导出规则进行比较分析, 发现并解释决策者策略之间一致和不一致的部分, 评价主体之间的冲突程度, 以促进主体之间的相互理解.

例 1 某信用卡申请决策表如表 1 所示^[12], 其中 account, balance, employed, monthly expense 为条件属性; decision 为决策属性. 各属性含义如下: account 为申请者是否有帐户, balance 为申请者的银行余额, employed 为申请者是否有工作, monthly expense 为申请者每月的开支.

文献 [12] 对数值属性 balance 和 monthly expense 进行特征化, 即

表 1 信用卡申请决策表

对象	account	balance	employed	monthly expense	decision
1	bank	700	yes	200	accept
2	other inst	1 200	yes	600	accept
3	other inst	1 600	yes	200	accept
4	bank	3 000	no	300	accept
5	bank	300	yes	600	reject
6	none	0	yes	400	reject
7	other inst	800	yes	600	reject
8	none	0	no	200	reject

low: balance < 500,
 medium: 500 < balance < 1 000,
 high: balance > 1 000;
 low: monthly expense < 250,
 medium: 250 < monthly expense < 500,
 high: monthly expense > 500

本文对各属性及其值进行编码, 得到表 2 的决策表

表 2 编码后的决策表

对象	account	balance	employed	monthly expense	decision
	a	b	c	d	e
1	1	1	1	1	1
2	2	2	1	2	1
3	2	2	1	1	1
4	1	2	2	3	1
5	1	3	1	2	2
6	3	3	1	3	2
7	2	1	1	2	2
8	3	3	2	1	2

设表 2 为决策者 1 的决策表(记为 T_1), 决策者 2 和决策者 1 在对象 8 的决策不一致, 即决策者 2 关于对象 8 的决策值为 *accept*. 设决策者 2 的决策表为 T_2 . 对 T_1 和 T_2 分别利用文献 [3] 提出的规则获取算法, 得到决策者 1 和决策者 2 的所有最简决策规则如表 3 和表 4 所示.

可以看出, 两个决策者的规则集有不一致之处. 例如表 4 中的规则 (d, 1) (e, 1), 表示只要 d 的取值为 1, 则决策者 2 的决策为 1, 这与表 3 中的规则 (c, 2) (d, 1) (e, 2) 不一致. 表 3 中包含 (d, 1) 的其他规则, 即前件为 (b, 1) (d, 1), (a, 1) (d, 1), (a, 2) (d, 1) 和 (c, 1) (d, 1) 的规则, 表示决策者需要考虑其他属性才能根据 (d, 1) 作出决

表 3 T1 的决策规则

对象	规 则	前 件	后件
1	(a, 1)	(b, 1), (b, 1)	(d, 1), (a, 1) (d, 1), (c, 1) (d, 1)
2	(b, 2)		(e, 1)
3	(b, 2), (a, 2)	(d, 1), (c, 1)	(d, 1)
4	(b, 2), (a, 1)	(c, 2), (c, 2)	(d, 3), (a, 1) (d, 3)
5	(b, 3), (a, 1)	(d, 2)	
6	(a, 3), (b, 3), (c, 1)	(d, 3)	(e, 2)
7	(a, 2)	(b, 1), (b, 1)	(d, 2)
8	(a, 3), (b, 3), (c, 2)	(d, 1)	

表 4 T2 的决策规则

对象	规 则	前 件	后件
1	(a, 1)	(b, 1), (d, 1)	
2	(b, 2)		(e, 1)
3	(b, 2), (d, 1)		(e, 1)
4	(b, 2), (c, 2), (a, 1)	(d, 3)	
8	(c, 2), (d, 1)		
5	(a, 1)	(b, 3), (b, 3)	(c, 1), (b, 3) (d, 2), (a, 1) (d, 2)
6	(a, 3)	(c, 1), (a, 3)	(d, 3), (b, 3) (c, 1), (e, 2)
	(b, 3)	(d, 3), (c, 1)	(d, 3)
7	(a, 2)	(b, 1), (b, 1)	(d, 2)

策为 1. 当对象和属性增多时, 导出规则的数量随之增加, 决策者之间的不一致分析也更加复杂

3.2 合并分析

粗集理论可在不一致的决策表中获取规则. 将多个决策者的决策表(一致的或不一致的)合并, 可得到可能不一致的决策表, 利用规则获取方法可得到信息更加明确的规则集

例 2 在例 1 中, 将决策者 2 关于对象 8 的决策加入表 2, 视为对象 9, 得到不一致的决策表. 利用文

表 5 合并决策表的决策规则

对象	规 则	前 件	后件
1	(a, 1)	(b, 1), (b, 1)	(d, 1), (a, 1) (d, 1), (c, 1) (d, 1)
2	(b, 2)		(e, 1)
3	(b, 2), (a, 2)	(d, 1), (c, 1)	(d, 1)
4	(b, 2), (a, 1)	(c, 2), (c, 2)	(d, 3), (a, 1) (d, 3)
5	(a, 1)	(b, 3), (b, 3)	(c, 1), (b, 3) (d, 2), (a, 1) (d, 2)
6	(a, 3)	(c, 1), (a, 3)	(d, 3), (b, 3) (e, 2)
	(c, 1), (b, 3)	(d, 3), (c, 1)	(d, 3)
7	(a, 2)	(b, 1), (b, 1)	(d, 2)
8, 9	(a, 3)	(c, 2), (a, 3)	(d, 1), (b, 3) (e, 1) (e, 2)
	(c, 2), (b, 3)	(d, 1), (c, 2)	(d, 1)

献[3]的规则获取算法得到的 ∂ 决策矩阵和规则如表 5 所示

为与例 1 对比, 现来观察包含 (d, 1) 的规则. 由对象 8 和 9 导出的不确定规则的前件 (a, 3) (d, 1), (b, 3) (d, 1), (c, 2) (d, 1) 表示两个决策者根据 (d, 1) 作决策时, 在属性及其值 (a, 3), (b, 3), (c, 2) 上产生冲突, 在其他属性上不产生冲突; 而在例 1 中, 只能得出与规则 (c, 2) (d, 1) (e, 2) 有冲突. 可以看出, 这样得到的规则集的信息更加明确, 分析更加简单

3.3 规则集构成分析

在例 2 中, 合并的决策表可视为两个决策者之间的不一致对象 8 作为新对象(决策属性为 1), 加入决策者 1 的决策表中得到的, 原决策表的规则因此发生变化

设决策表 T1 的条件等价类为 $X_i (i = 1, 2, \dots, n)$, 决策类为 $Y_j (j = 1, 2, \dots, m)$, X_i 的决策函数为 B_i^0 , 并设 $U/I(\partial = \{Z_1, Z_2, \dots, Z_p\})$. 加入不一致对象 x 后, $X_i \cap Y_j$ 的规则形式发生变化. 具体有以下两种情况:

1) $x \in X_i, x \in Y_k, X_i \cap Y_k = \emptyset, X_i \subseteq Z_g$. 加入 x 后, 使 $X_i \subseteq Z_f, Z_f \subset U/I(\partial, f = g)$. 设 $Z_f = X_{i_f}, Z_g = X_{i_g}$.

规则变化过程为: 在 ∂ 决策矩阵中删除元素 $m_{i, i_f} (i < i_f)$ 或 $m_{i_f, i} (i_f < i)$, 增加元素 $m_{i, i_g} (i < i_g)$ 或 $m_{i_g, i} (i_g < i)$. 计算等价类 X_i, X_{i_f} 和 X_{i_g} 的决策函数 $B_i^1, B_{i_f}^1$ 和 $B_{i_g}^1$, 得到相应的规则, 而其他等价类的规则形式不变

2) $x \in X_i, x \in Y_k, X_i \cap Y_k = \emptyset, X_i \subseteq Z_g$. 加入 x 后, $X_i \subseteq Z_f$ 且 $Z_f \not\subset U/I(\partial)$. 设 $Z_g = X_{i_g}$.

规则变化过程为: 在 ∂ 决策矩阵第 i 行(列)增加元素 $m_{i, i_g} (i < i_g)$ 或 $m_{i_g, i} (i_g < i)$, 计算等价类 X_i 和 X_{i_g} 的决策函数 B_i^1 和 $B_{i_g}^1$, 得到相应的规则, 而其他等价类的规则形式不变

上述规则构成分析适用于合并前的单个决策表不一致的情况

例 3 设 T1 的条件等价类为 X_1, X_2, \dots, X_8 , 决策类 $Y_1 = \{x: e(x) = 1\}, Y_2 = \{x: e(x) = 2\}$, ∂ 决策类 $Z_1 = Y_1, Z_2 = Y_2$. 将决策者 2 关于对象 x_8 的决策加入 T1, 视为 x_9 , 得到不一致的决策表, 其中 $x_9 \in X_8, x_9 \in Y_1, X_8 \cap Y_1 = \emptyset, X_8 \subseteq Z_2$. 加入 x_9 后, 产生 ∂ 决策类 $Z_3 = \{x_8, x_9\}$, 且 $Z_3 \not\subset U/I(\partial)$. 在 T1 中, $Z_2 = X_5 \cap X_6 \cap X_7 \cap X_8$, 所以规则变化过程为:

在表4的决策矩阵中增加元素 $m_{58} = acd, m_{68} = cd, m_{78} = abcd$, 并更新决策函数 $B_5 \sim B_8$, 而等价类 $X_1 \sim X_4$ 的规则不变

如果令T2不变, 将决策者1关于 x_8 的决策加入T2, 则得到的结果是等价类 X_5, X_6, X_7 的规则不变, 而其余等价类的规则需要更新

综上所述, 在已知决策者1和决策者2的决策规则的情况下, 只需对 X_8 的规则进行更新即可得到合并后决策表的所有规则

利用上述规则更新分析, 表5中的规则由以下3部分组成:

- 1) 决策者1关于 X_i ($e, 1$)的规则, 即表3中从对象1~4导出的规则;
- 2) 决策者2关于 X_i ($e, 2$)的规则, 即表4中从对象5~7导出的规则;
- 3) 表5中由对象8和9得到的不确定规则 $[x_8]$ ($e, 1$) ($e, 2$).

对于规则1)和2), 虽然决策者1和决策者2关于各决策类之间的决策规则不完全相同, 但没有冲突; 对于规则3), 由于决策者1和决策者2关于对象8的决策不一致而导致冲突。可以看出, 利用合并后的决策表得到的规则集更加明确地表示了决策者之间的关系

对于多个决策者之间存在多个不一致对象的情况, 经合并分析后可得到决策群体中一致的规则和不一致的规则, 便于进行比较分析

4 结 语

粗集理论可用于处理含有不确定、不一致信息的多属性决策问题。国内外关于这方面的研究刚刚兴起。本文结合近年来该理论的一些研究成果, 提出了根据粗集理论的有关概念和方法解决多属性决策问题的一般过程, 即在构造决策表的基础上进行属性的约简与评价、获取决策规则、分类决策。针对多人多属性决策问题中出现的多个决策者对同一类对象的决策结果不一致, 即决策的策略互不相同的情况, 提出对不同决策者的经验决策数据进行综合分析, 导出决策规则集并对其构成进行分析的方法, 以便解释决策者之间决策策略不同、决策结果不一致的现象, 评价决策者之间的冲突程度, 以促进他们之间的相互了解

在上述研究成果的基础上, 本文将在以下方面进一步研究: 1) 如何对得到的群决策规则进行评价和筛选, 最终得到有效的决策规则; 2) 如何利用得到的群决策规则进行决策, 如分类决策的策略; 3)

考虑群体成员的不同决策能力以及群体相关性对决策规则、规则参数和决策结果的影响

参考文献(References)

- [1] Pawlak Z. *Rough sets: Theoretical aspects of reasoning about data* [M]. Dordrecht: Kluwer Academic Publishers, 1991.
- [2] Hu X H. Knowledge discovery in databases: An attribute-oriented rough set approach [D]. Regina: University of Regina, 1995.
- [3] 安利平, 吴育华, 仝凌云. 一种基于粗糙集理论的规则获取算法[J]. *管理科学学报*, 2001, 4(5): 74-78.
(An L P, Wu Y H, Tong L Y. An approach for rule generation based on rough set theory [J]. *J of Management Science*, 2001, 4(5): 74-78.)
- [4] 安利平, 吴育华, 仝凌云. 增量式获取规则的粗糙集方法[J]. *南开大学学报*, 2003, 36(2): 98-103.
(An L P, Wu Y H, Tong L Y. Rough set approach to incremental acquisition of rules [J]. *J of Nankai University*, 2003, 36(2): 98-103.)
- [5] Ziarko W. Variable precision rough set model[J]. *J of Computer and System Sciences*, 1993, 46: 39-59.
- [6] Skowron A. Synthesis of adaptive decision systems from experimental data[A]. *Proc Fifth Scandinvian Conf on Artificial Intelligence* [C]. Trondheim, 1995: 220-238.
- [7] Slezak D. Approximate reducts in decision tables[A]. *Proc Sixth Int Conf Information Management of Uncertainty in Knowledge-based Systems* [C]. Granada, 1996: 1159-1164.
- [8] Bazan G J, Skowron A, Synak P. Dynamic reducts as a tool for extracting laws from decision tables[A]. *Proc Symp on Methodologies for Intelligent Systems* [C]. Charlotte: Springer-Verlag, 1994: 346-355.
- [9] Øhrn A. Discernibility and rough sets in medicine: Tools and applications [D]. Trondheim: Norwegian University of Science and Technology, 1999.
- [10] Grzymala W J. LERS — A system learning from examples based on rough sets[A]. *Intelligent Decision Support* [C]. Dordrecht: Kluwer Academic Publishers, 1992: 3-18.
- [11] Slowinski R. Rough set learning of preferential attribute in multi-criteria decision making [A]. *Proc Seventh Int Symp on Methodology for Intelligent Systems* [C]. Trondheim, 1993: 642-651.
- [12] Pawlak Z, Slowinski R. Rough set approach to multi-attribute decision analysis [J]. *European J of Operational Research*, 1994, 72: 443-459.