

国家自然科学基金项目资助（项目编号：71572123、71722004）

复杂模糊多属性大群体决策模型 与算法研究

◆刘炳胜 王安民 申映华 著

FUZA MOHU DUOSHUXING DAQUNTI
JUECE MOXING YU SUANFA YANJIU

国家自然科学基金项目资助(项目编号:71572123、71722004)

复杂模糊多属性大群体决策 模型与算法研究

刘炳胜 王安民 申映华 著



天津大学出版社

TIANJIN UNIVERSITY PRESS

图书在版编目(CIP)数据

复杂模糊多属性大群体决策模型与算法研究 / 刘炳胜, 王安民, 申映华著. —天津: 天津大学出版社, 2017. 12

国家自然科学基金项目资助(项目编号: 71572123、71722004)

ISBN 978-7-5618-6026-7

I. ①复… II. ①刘… ②王… ③申… III. ①群体 - 决策学 - 决策模型 - 研究 ②群体 - 决策学 - 算法 - 研究 IV. ①C934

中国版本图书馆 CIP 数据核字(2017)第 317157 号

出版发行 天津大学出版社
地 址 天津市卫津路 92 号天津大学内(邮编: 300072)
电 话 发行部: 022-27403647
网 址 publish.tju.edu.cn
印 刷 北京京华虎彩印刷有限公司
经 销 全国各地新华书店
开 本 169mm × 239mm
印 张 12
字 数 249 千
版 次 2017 年 12 月第 1 版
印 次 2017 年 12 月第 1 次
定 价 32.00 元

凡购本书, 如有缺页、倒页、脱页等质量问题, 烦请向我社发行部门联系调换

版权所有 侵权必究

序

决策问题已经成为管理科学研究领域的重要发展方向之一。目前,群决策领域主要倾向于“小群体”决策理论与方法的改进性研究,与之相比,“复杂大群体”决策技术与模型研究尚处于初期起步阶段,缺少较为成熟的方法体系支撑。近年来,随着我国社会民主化进程的不断推进,针对一些涉及普通公众切身利益的公共决策问题,亟须建立公众参与机制以求广泛征询、吸收和借鉴各方的意见,协调各方的关系和利益。因此,公共决策问题已经突破了传统的小群体决策范畴,决策环境和决策群体均发生了较大变化,复杂多属性大群体决策方法将成为解决该问题的重要技术。

公共决策问题的复杂性与多层性特点,决定了复杂大群体决策中决策主体、决策客体与决策环境同时呈现出复杂性。针对上述特点,本书按照“科学、客观确定决策主体聚集与权重,最大限度降低决策客体相关性,最大程度反映专家信息”的整体性研究思路,从决策主体聚类、决策主体定权、决策属性定权、决策信息集结四个方面开发设计了具体的模型与算法。

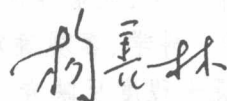
针对复杂大群体决策主体人数多、利益偏好不一致等特点,设计了“通过聚类方法将决策主体分类,形成若干个聚集,避免信息冗余”的研究思路,提出了“利益团体与实际偏好数据相结合、专家偏好聚类”的研究框架,开发了面向决策者聚类的直觉模糊环境下偏二叉树 DEA-DA 模型以及面向决策者聚类的语言环境下偏二叉树 DEA-DA 模型。

针对复杂大群体决策主体间若干利益聚集的特点,设计了“既区分组群特性,又考虑个体特性”的联动定权研究思路,提出了“区分上下层,联动定权”的研究框架,开发了直觉模糊环境下复杂多属性大群体决策者双层定权模型与算法、语言环境下复杂多属性大群体决策者双层定权模型、直觉模糊环境下复杂大群体决策属性偏最小二乘通径定权模型以及语言环境下复杂大群体决策属性偏量小二乘通径定权模型。

针对复杂大群体决策客体呈现出“属性相关、具有层次结构”的特点,

设计了“先保全面,再除相关,综合集结”的三阶段研究思路,提出了“一阶段建立全面性、层次性的决策属性体系,二阶段降维消相关,三阶段集结信息”的研究框架,开发了新型直觉模糊集相关度测量模型、区间直沉模糊环境下复杂大群体 IVIF-PCA 信息集结模型以及语言环境下复杂大群体主成分决策模型。

目前国内外对于复杂大群体决策研究单纯侧重于决策主体的复杂性,对于决策客体与决策环境的复杂性研究甚少。本书将复杂大群体决策理论与技术进行了发展与推广,作为国内首部复杂多属性大群体决策的著作,在全面、系统梳理复杂大群体决策特点、发展演进趋势的基础上,围绕着决策主体聚类、决策主体定权、决策属性定权、决策信息集结四个方面开发了具体的模型与算法,形成了全面的决策方法体系。相信该著作在促进模糊与语言环境下复杂多属性大群体决策理论发展的同时,更将为我国社会民主化进程的推进提供技术支持,并更好地服务于我国的经济社会发展。



2017年1月10日

目 录

第 1 章 绪论	1
1.1 群决策理论发展回顾	1
1.2 复杂多属性大群体决策理论回顾	2
1.3 复杂多属性大群体决策环境	3
1.3.1 直觉模糊信息	3
1.3.2 语言标度信息	4
1.4 本书主要内容	5
第 2 章 复杂多属性大群体决策者聚类模型	8
2.1 面向决策者聚类的直觉模糊环境下偏二叉树 DEA-DA 模型	8
2.1.1 预备知识	9
2.1.2 区间直觉模糊样本向单值样本转化的 C-OWA 算子	14
2.1.3 偏二叉树 DEA-DA 循环分类模型	17
2.1.4 区间直觉模糊环境下复杂多属性大群体决策者分类流程	18
2.1.5 算例	19
2.2 面向决策者聚类的语言环境下偏二叉树 DEA-DA 模型	27
2.2.1 预备知识	28
2.2.2 复杂多属性大群体语言决策专家分类偏二叉树 DEA-DA 模型	32
2.2.3 算例	43
2.3 本章小结	48
第 3 章 复杂多属性大群体决策者定权模型	49
3.1 直觉模糊环境下复杂多属性大群体决策者双层定权模型	49
3.1.1 预备知识	50
3.1.2 直觉模糊环境下决策者双层定权算法	52
3.1.3 算例	55
3.2 语言环境下复杂多属性大群体决策者双层定权模型	58
3.2.1 预备知识	59
3.2.2 语言环境下专家双层定权方法	60
3.2.3 算例	65
3.3 本章小结	68

第4章 复杂大群体决策属性定权模型	70
4.1 直觉模糊环境下复杂大群体决策属性偏最小二乘通径定权模型	70
4.1.1 预备知识	71
4.1.2 区间直觉模糊样本向单值样本转化的两阶段法	73
4.1.3 算例	77
4.2 语言环境下复杂大群体决策属性偏最小二乘通径定权模型	82
4.2.1 预备知识	83
4.2.2 面向复杂大群体语言决策的多属性权重 PLS 通径模型构建	86
4.2.3 面向复杂大群体语言决策的多属性权重 PLS 通径模型确定方法	90
4.2.4 算例	93
4.3 本章小结	96
第5章 复杂多属性大群体信息集结模型	98
5.1 新型直觉模糊集相关度测量模型	98
5.1.1 预备知识	99
5.1.2 新型相关度测量模型	100
5.1.3 扩展:IVIFS 的相关系数计算方法	106
5.2 区间直觉模糊环境下复杂大群体 IVIF-PCA 信息集结:模型 I	109
5.2.1 预备知识	110
5.2.2 复杂多属性大群体决策的 IVIF-PCA 模型	111
5.2.3 算例	116
5.3 区间直觉模糊环境下复杂大群体 IVIF-PCA 信息集结:模型 II	121
5.3.1 预备知识	121
5.3.2 新概念的提出	122
5.3.3 IVIF-PCA 模型	127
5.3.4 基于 IVIF-PCA 模型的 CMALGDM 方法	131
5.3.5 算例与对比分析	134
5.4 语言环境下复杂大群体主成分决策方法	149
5.4.1 预备知识	150
5.4.2 复杂多属性大群体语言决策的属性主成分模型	153
5.4.3 复杂多属性大群体语言决策的属性主成分求取步骤	155
5.4.4 算例	158
5.5 本章小结	163
第6章 总结与展望	164
6.1 研究总结	164
6.2 进一步研究展望	165
参考文献	167

第 1 章 绪论

1.1 群决策理论发展回顾

决策问题广泛地存在于社会的各个领域之中,其本质在于从众多的备选方案中选择出最为有效的方案,从而使得决策结果产生最大的效用。一个方案往往需要通过多个不同的属性或者准则进行评价,决策者通过给出自己对方案各个属性的偏好或者评价信息完成评价过程。通过对评价信息的集成处理或对比处理,可以比选出最优的决策方案,从而解决决策问题。学术界对决策问题的早期研究大多集中于单个决策者对多个方案的决策,然而随着社会民主化进程的加快,决策问题的复杂程度日益增加,决策问题的研究不再局限于单个决策者,而是推广至群体决策研究。群体决策的一个显著优点在于充分考虑了个体对于问题的认识和理解,集思广益从而避免了由于个人武断决策所带来的决策风险,确保了决策结果符合大多数人的期望。在决策过程中,显然个体之间给出的评价信息可能存在较大差异,这主要是因为评价信息往往带有主观性,评价信息很大程度上受限于决策者的知识结构、个人阅历、价值观等因素。因此如何科学合理地通过某种方式将多个决策个体对于多个方案的评价信息进行处理,从而获得最终群体决策结果的群体决策思想和方法受到学术界的广泛研究。Hwang、Desancis、陈珽等学者分别给出了群体决策的定义,提炼学者们的思想可以得出,群体决策是指在由多个决策个体组成的决策群体中,各决策个体首先根据各自的主观判断对备选方案给出偏好信息,然后决策群体依据某种决策规则将群体中各决策者的偏好信息集结成为群体偏好信息,最后再根据集结的群体偏好信息获得各备选方案的综合评价信息,并据此选择出对于群体最优或群体最满意的方案。

群体决策问题的研究已经趋于成熟(主要指小群体决策问题),目前成熟的群体决策方法在政治、经济、社会、军事、文化等各个领域得到了广泛的应用。然而,20世纪90年代以后随着信息技术的发展以及决策民主化的进一步深入,群体决策的方式也随之发生了较大的变化,主要呈现出以下几个特点:①通过互联网系统,群体中的决策个体可以在不同的地点不同的时间进行决策;②决策群体规模比较大(数量往往超过20人),群体成员分布广,时间和位置不固定,决策需要更加广泛的相关成员参与;③群体个体类别差异大且关系复杂,不仅局限于组织内部人员,而且扩展到组

织外部的有关人员,甚至会涉及社会公众,决策个体之间往往既有竞争又有合作;④由于决策问题的日益复杂,决策方案需要从更多不同的角度进行评价,换言之,决策属性的数目大量增加,且属性之间存在相互关联的现象;⑤决策者的评价信息通常表现出不确定性或者模糊性。陈晓红将具有这些特点的群体决策问题归结为复杂多属性大群体决策(CMALGDM)问题。

不难发现复杂多属性大群体决策的五个特点可以大致体现出该决策问题两个方面的复杂性,即决策群体的日趋复杂以及决策属性的日趋复杂。这两个方面的复杂性决定了复杂多属性大群体决策方法不同于常规的群体决策方法。后者多数面向小群体决策问题而难以支持这种复杂大群体决策问题,所以研究面向复杂多属性大群体决策的理论与方法显得尤为必要。

1.2 复杂多属性大群体决策理论回顾

目前关于复杂多属性大群体决策的研究还处于起步阶段,梳理现有的相关研究文献,可以发现该领域的研究内容主要表现在以下三个方面。

1. 复杂大群体的聚类与信息集成

徐选华和陈晓红基于矢量空间提出群体中个体间的偏好矢量相关性度量方法并进一步给出启发式算法,从而将群体成员划分为若干类别为群体决策提供必要支持。徐选华和范永峰基于蚁群聚类算法提出一种启发式聚类算法从而改进了基于矢量空间的聚类算法。徐选华和陈晓红等通过熵权法剔除提供较少信息量的决策个体并利用聚类方法对决策成员进行聚类从而确定出成员权重,最终结合该权重与效用矩阵获得最优方案。陈晓红和刘蓉提出基于全部最小连通支配集算法的聚类算法支持复杂大群体决策。刘蓉和陈晓红提出考虑决策个体学习进化能力的群体一致性修正方法尝试减少群体决策的失误概率。陈晓红等提出一种基于多属性多方案大群体决策方法,较好地解决了多属性多方案大群体决策问题。万树平首次针对属性值为区间数的多属性多方案大群体决策问题,提出一种区间数聚类算法,较好地避免聚类阈值选取的经验性,提高决策结果的可信度。

2. 复杂多属性的处理和系统集成

张晓和樊治平针对多属性多标度大群体决策问题,提出一种基于随机占优准则的决策方法,首先将决策者的评价信息转化为关于指标评价标度的概率分布,然后确定方案之间的随机占优关系矩阵,最后运用 PROMETHEE II 方法得到方案的排序结果。徐选华和万奇锋针对属性权重信息完全未知和属性值为连续型随机变量,提出了一种大群体决策方法。该方法将随机属性值转化成区间数,然后给出区间数模糊聚类算法,在此基础上定义并计算群体中各个聚集和群体的区间评价矩阵,最后

利用不确定性有序加权平均算子获得最优方案。徐选华和张丽媛等基于二元关系形成群体成员偏好矢量属性关系矩阵,构建了两个决策个体偏好信息相聚性度量模型,基于该模型提出了一种偏好信息集结和决策方案排序的方法。

3. 复杂大群体决策支持系统构建

陈晓红和陈志阳等为了解决复杂大群体决策问题求解过程中产生的冲突加剧、信息超载、效率低下和共识难以达成等问题,总结并提出一种适合软件实施需求,实现人机结合的决策问题分解、决策群体分解、并行求解、求解结果合成和层级驱动的框架。徐选华和陈晓红针对复杂大群体决策支持系统开发的复杂性,提出了以决策问题求解为导向的一种群决策支持系统结构。该系统结构首先设计了复杂决策问题求解流程和方法,然后提出了系统处理流程,在此基础上构建了系统的层次结构和功能结构,以期辅助面向复杂大群体的群决策支持系统的开发。

总结以上研究可以看出,目前复杂多属性大群体决策的研究仍然处于起步阶段,相关研究成果主要集中在复杂大群体的聚类及信息集结方面,对于属性复杂性(主要包括属性数目繁多和属性之间相关性强的)的研究几乎没有涉及;决策者对于方案的评价信息以单个数值信息为主,因而不能准确地反映决策者对方案属性的偏好;对于决策支持系统的构建也只是给出了一些笼统的框架,不够细致深入。

1.3 复杂多属性大群体决策环境

由于复杂多属性大群体决策的评价信息往往具有不确定性或者模糊性,因此如何选择决策者评价信息的表达形式(决策环境)是首先要解决的问题。本节主要介绍目前广泛研究和使用的两种不确定信息的表示方法:直觉模糊信息和语言标度信息。

1.3.1 直觉模糊信息

自从 Zadeh 于 1965 年提出模糊集理论以来,该理论已在社会的各个领域得到广泛的应用。模糊集理论的关键思想在于把取值为 1 或 0 的特征函数推广到可在 $[0, 1]$ 中任意取值的隶属函数。然而,美中不足的是模糊集的隶属函数值仅是一个单一的值,它无法同时表示支持、反对以及犹豫这三个方面的程度。鉴于此,Atanassov 于 1983 年对传统的模糊集进行了拓展,提出了直觉模糊集的概念。由于直觉模糊集同时考虑了隶属度、非隶属度和犹豫度这三个方面的信息,因此,它比传统的模糊集在处理模糊性和不确定性等方面更具有灵活性和通用性。在此基础之上,Atanassov 和 Gargov 于 1989 年进一步引入区间直觉模糊集的概念,将隶属度、非隶属度表示为 $[0, 1]$ 区间的闭子区间,即用隶属度区间取代原有的隶属度函数,用非隶属度区间取

代原有的非隶属度函数,有效地拓展了直觉模糊集处理不确定性信息的能力,从而更好地解决了实际决策问题。本节进一步给出直觉模糊集的数学定义。

定义 1.1 假设 X 为非空集合,称

$$A = \{ \langle x, \mu_A(x), v_A(x) \rangle \mid x \in X \} \quad (1.1)$$

为直觉模糊集(IFS)。其中, $\mu_A(x)$ 表示元素 x 属于 A 的程度, $v_A(x)$ 表示元素 x 不属于 A 的程度,满足 $\mu_A: X \rightarrow [0, 1], x \in X \rightarrow \mu_A(x) \in [0, 1], v_A: X \rightarrow [0, 1], x \in X \rightarrow v_A(x) \in [0, 1], 0 \leq \mu_A(x) + v_A(x) \leq 1$ 。为了方便表述, Xu 和 Yager 称二元组 $\alpha = (\mu_\alpha, v_\alpha)$ 为直觉模糊数(IFN),并用它来表示直觉模糊集中的元素,其中 $\mu_\alpha \in [0, 1], v_\alpha \in [0, 1], \mu_\alpha + v_\alpha \leq 1$ 。

然而,由于客观事物的复杂性和不确定性,用单个数值表示直觉模糊集的隶属度和非隶属度在一定程度上是不准确的,可能会造成信息的损失,此时可以用区间数取代单个数值。Atanassov 和 Gargov 推广了直觉模糊集,提出了区间直觉模糊集的概念,见定义 1.2。

定义 1.2 假设 X 为非空集合,称

$$\bar{A} = \{ \langle x, \tilde{\mu}_{\bar{A}}(x), \tilde{v}_{\bar{A}}(x) \rangle \mid x \in X \} \quad (1.2)$$

为区间直觉模糊集(IVIFS)。其中,隶属度区间 $\tilde{\mu}_{\bar{A}}(x) \subset [0, 1]$,非隶属度区间 $\tilde{v}_{\bar{A}}(x) \subset [0, 1], x \in X, \tilde{\mu}_{\bar{A}}^U(x) + \tilde{v}_{\bar{A}}^U(x) \leq 1$ 。显然,如果 $\tilde{\mu}_{\bar{A}}^L(x) = \tilde{\mu}_{\bar{A}}^U(x)$ 且 $\tilde{v}_{\bar{A}}^L(x) = \tilde{v}_{\bar{A}}^U(x)$,那么区间直觉模糊集退化为直觉模糊集。

根据上述定义,区间直觉模糊集的基本元素为一个有序的区间对,分别表示元素 x 隶属于 \bar{A} 的程度和非隶属于 \bar{A} 的程度。同样,为了方便表述,徐泽水将 $\tilde{\alpha} = (\tilde{\mu}_{\bar{A}}(x), \tilde{v}_{\bar{A}}(x)) = ([a, b], [c, d])$ 称为区间直觉模糊数(IVIFN),并用它表示区间直觉模糊集中的元素。其中隶属度区间 $[a, b] \subset [0, 1]$,非隶属度区间 $[c, d] \subset [0, 1], b + d \leq 1$ 。

1.3.2 语言标度信息

语言决策方法是一种近似技术,通过语言变量的语言值来表示定性方面的决策信息。按照表示方式的不同,语言评估标度主要包括加性语言评估标度和积性语言评估标度。考虑到加性语言评估标度的术语下标涵盖整个实数域,而积性语言评估标度的术语下标仅是正实数域。为了便于后续模型的构建,本书主要介绍加性语言评估标度。

令 $S = \{s_\alpha \mid \alpha = -t, \dots, t\}$ 是一个有限和全序离散评估标度集,其中, t 是一个正整数, s_α 代表了一个语言变量的可能值,并要求:① $s_\alpha < s_\beta$, 如果 $\alpha < \beta$;② 否定运算: $\text{neg}(s_\alpha) = s_{-\alpha}$ 特别地, $\text{neg}(s_0) = s_0$ 。例如,九标度 S 可以表示为: $S = \{s_{-4} = \text{extremelypoor}, s_{-3} = \text{verypoor}, s_{-2} = \text{poor}, s_{-1} = \text{slightlypoor}, s_0 = \text{fair}, s_1 = \text{slightlygood}, s_2 = \text{good},$

$s_3 = \text{verygood}, s_4 = \text{extremelygood}$ }。

要保留所有给定的信息, Xu 将离散评估标度集 S 扩展到连续的评估标度集 $\bar{S} = \{s_\alpha | \alpha \in [-q, q]\}$, 其中, $q (q > t)$ 是一个足够大的正整数。如果 $s_\alpha \in S$, 那么 s_α 被称为一个原始加性语言标度, 否则, s_α 被称为虚拟加性语言标度。在一般情况下, 专家使用原始的语言术语来评估备选方案, 而虚拟的语言术语只会出现在运算中。

在此基础上, 考虑到决策者对事物属性的描述可能介于两个语言术语之间。令 $\tilde{s} = [s_\alpha, s_\beta]$, 其中 $s_\alpha, s_\beta \in \bar{S}$, s_α 和 s_β 分别是上下限, 那么称 \tilde{s} 是不确定的语言变量。而 \tilde{S} 为基于拓展加性语言评估标度 \bar{S} 的不确定语言变量集。

1.4 本书主要内容

决策模型的构建离不开具体的决策环境, 由于区间直觉模糊信息和语言标度信息在决策信息的表示上具有明显优势, 本书将在这两种不同的决策环境下分别进行决策模型的研究与开发。尽管决策环境不同, 但遇到的决策问题却是类似的。换言之, 无论是在区间直觉模糊环境下, 还是在语言标度环境下需要处理的研究问题主要是三类: ①处理决策者的复杂性; ②处理决策属性的复杂性; ③决策信息的合理集成。以下将针对这三类决策问题分别分析复杂多属性大群体决策模型的逻辑结构。

1. 决策者复杂性

本书拟从决策者聚类、决策者权重确定两个角度探讨决策者的复杂性问题。由于实际工程的复杂性, 某个特定决策问题往往会牵扯进大量的相关干系人, 如何根据干系人给出的决策信息将他们合理聚类, 这对于决策问题的简化具有巨大的重要性。另外, 决策过程中, 各个决策者的重要性往往也是不同的, 比如某个领域的权威专家针对实际问题的理解判断显然优于普通公众, 赋予的权重自然需要调整。

针对决策者聚类问题, 第 2 章将给出复杂多属性大群体决策者的偏二叉树 DEA-DA 循环聚类模型。针对目前复杂大群体决策者的聚类尚无成熟聚类标准的现状, 提出将利益团体与实际决策信息相结合的研究思路。通过构建面向偏二叉树 MIP DEA-DA 循环聚类模型, 实现对决策者进行多样本组聚类, 区分出利益偏好一致的各个群体。

针对决策者定权问题, 第 3 章将给出复杂多属性大群体决策者的双层定权模型。该模型基于已有的聚类模型将决策群体分成若干个聚集, 并根据“多数原则”给出各个聚集的权重。其次, 为了使得聚集内决策个体之间偏好具有较好的一致性, 通过聚集内决策者的评价值方差最小模型, 求解出聚集内决策者的最优权重。

2. 决策属性复杂性

本书将从决策属性相关性和决策属性定权两个角度分析决策属性复杂性问题。

实际工程中,由于工程的复杂性,决策问题往往需要众多的属性来加以刻画。属性数量的增多加剧了属性间的复杂程度,属性之间表现出层次性和相关性。属于同一隐性属性的决策属性存在较大的信息重合,而不同隐性属性间也存在较强相关性。

针对决策属性权重问题,第4章提出基于PLS通径模型的复杂多属性大群体的决策属性权重的确定方法,不仅是对构建出的全面和独立的决策属性体系的应用,而且为基于语言评估标度的复杂多属性大群体决策的信息集结研究奠定基础。

针对决策属性相关性问题,第5章提出基于主成分分析的决策属性间的去相关性方法。该模型旨在将原始决策属性转化为几个相互独立的主成分,在表达原始属性主要信息的前提下,在进一步降低属性空间维度的同时,拟解决在复杂多属性大群体决策问题的属性关联问题。

3. 决策信息合理集成

决策的本质在于针对给出的一组方案,根据提供的决策信息得到最优方案。如何合理地将决策信息集成为一个评判方案优劣的某个定量指标是科学决策的重要环节之一。本书拟在合理确定决策者聚集,确定决策者权重,有效融合决策属性,降低决策属性空间维度,解决决策属性的相关性的基础上给出决策信息的综合集成方法,力图提高方案最终评判指标的科学性和合理性。

本书的逻辑框架如图1.1所示。

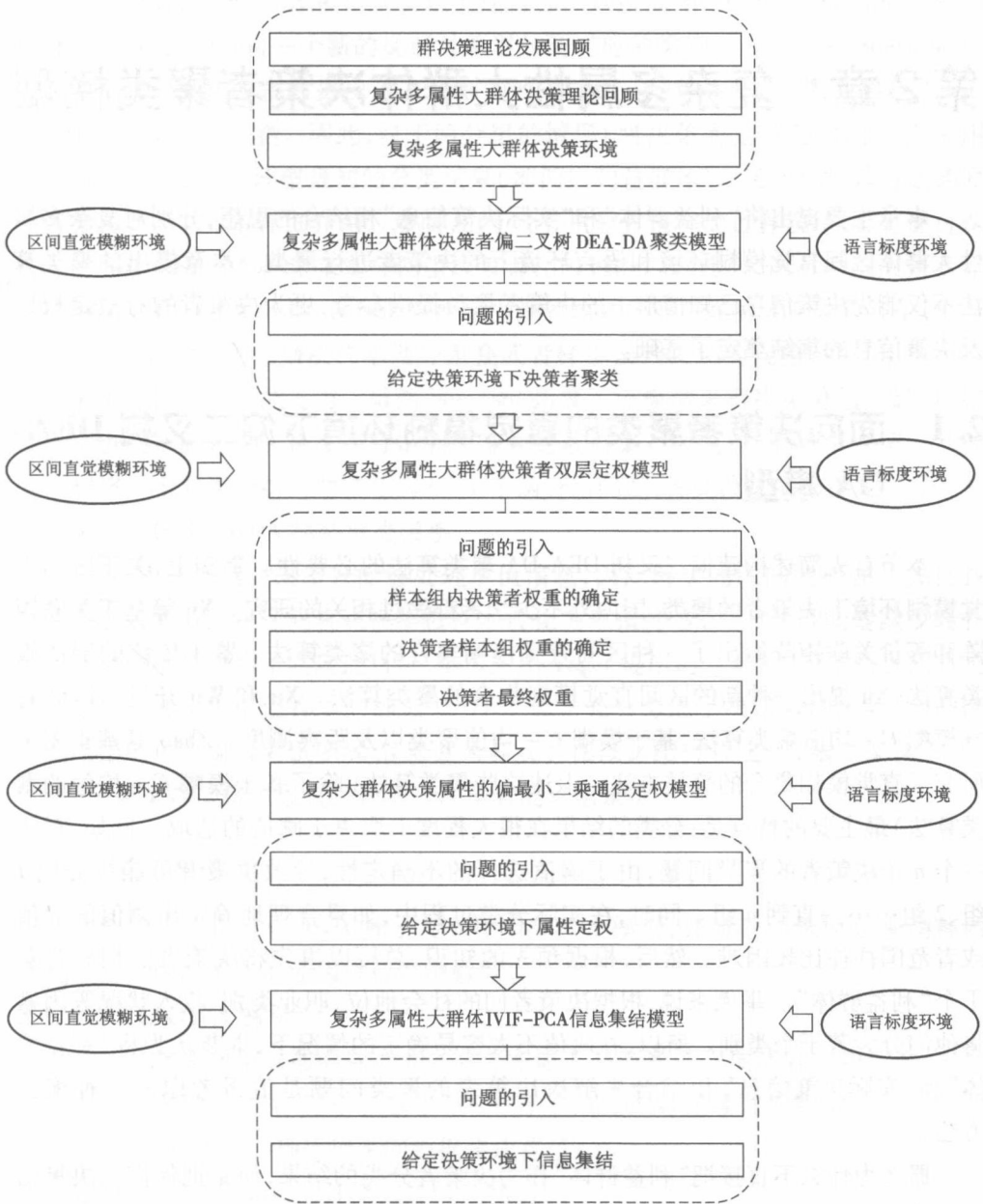


图 1.1 逻辑框架图

第2章 复杂多属性大群体决策者聚类模型

本章主要提出将“利益群体”和“实际决策信息”相结合的思想,分别对复杂多属性大群体区间直觉模糊环境和语言环境下的决策者进行聚类。本章提出的聚类算法不仅能为决策信息已知情形下的决策者聚类提供参考,更为决策者的有效定权以及决策信息的集结奠定了基础。

2.1 面向决策者聚类的直觉模糊环境下偏二叉树 DEA-DA 模型

本节首先简述构建偏二叉树 DEA-DA 聚类算法的必要性。事实上,关于区间直觉模糊环境下决策者的聚类,国内外不少学者都做过相关的研究。Xu 等基于关联矩阵和等价关联矩阵给出了一种区间直觉模糊集合的聚类算法。鉴于传统的层次聚类算法, Xu 提出一种新的区间直觉模糊集合的聚类算法。Xu 和 Wu 开发了区间直觉模糊 C -均值聚类算法,基于模糊 C -均值聚类以及距离测度。Zhao 等基于图论研究了直觉模糊集合的聚类方法。上述这些聚类算法(除了基于模糊 C -均值的聚类算法)最主要的特点是:分类的结果在很大程度上取决于阈值的选取。例如,对于一个 n 个决策者的聚类问题,由于阈值选择的不确定性,导致决策者可能被分成 1 组、2 组……一直到 n 组。同时,在实际分类过程中,如果合理地确定出阈值的取值或者范围往往比较困难。然后,根据预先的知识,总可以事先将决策者们划分为若干个“利益群体”。举例来说,根据决策者们的社会地位、职业类别、收入状况等因素将他们分为若干个类别。所以,在阈值不太容易确定的情况下,本节认为将“利益群体”和“实际决策信息”相结合来解决决策者的聚类问题是值得考虑的一种聚类方法。

那么为什么不直接把“利益群体”作为决策者分类的结果,而是把他们与决策信息相结合然后再重新分类呢?下面举例说明这么做的必要性。以消费群体的购物决策为例,对于不同收入水平的消费群体,他们的消费偏好一般是显著不同的,但是即使是位于同一收入水平的消费群体也有可能拥有与该收入水平群体显著不同的消费偏好。所以,一个更为科学的分类方法是:首先按照收入水平将消费群体划分为若干类,然后将每类里面的消费偏好与众不同的消费者挑选出来并划归到消费偏好与之类似的消费类别之中。

本节拟借鉴 DEA-DA 模型来实现上述目标。众所周知,DEA-DA 模型是一种判别分析的方法,可以将一个新的观察对象划分到对应的类别。日本学者 Sueyoshi 对于该模型的研究比较多并提出了一系列的改进模型,其中以 MIP (混合整数规划) DEA-DA 模型最为成熟。因此,对于两分组的情形(将决策者划分为两组),拟采用 MIP DEA-DA 模型来调整最初的分组结果(即两组利益群体)。关于选择该方法的原因将会在 2.1.1.2 节详细介绍。然而,由于计算机软件潜在的计算误差,要想通过执行一次 MIP DEA-DA 模型来划分两组决策者比较困难。因此,采用对两组决策者循环执行 MIP DEA-DA 分析,直到最终两组决策者的划分不再发生改变。事实上,尽管 MIP DEA-DA 模型能够解决样本划分重叠或者样本误判类别的现象,但对于多组样本的聚类分析仍然无能为力。虽然 Sueyoshi 拓展了该模型来解决多分组问题,但扩展模型也只是能处理一种按照特定顺序排列的数据集。因此,本节将引入偏二叉树的思想将一个 h 类分组问题转化为 $h-1$ 个两分组问题,称此改进的 MIP DEA-DA 模型为偏二叉树 DEA-DA 循环聚类模型。

另外,由于提出该模型只能够处理单值信息,因此解决问题的关键在于如何找到一种有效的方法将区间直觉模糊数样本转化为单值样本,并且尽可能地减少转化过后信息的损失。为了实现此目标,本节采用 Yager 提出的 C-OWA 算子将区间直觉模糊数样本转化为单值样本。因为 C-OWA 算子充分考虑了决策者的风险偏好,因而集结出的单个决策者的决策信息更为客观。在该集结过程之后,偏二叉树 DEA-DA 循环聚类模型将可以直接被用于决策者的多组聚类。

2.1.1 预备知识

本部分主要介绍 C-OWA 算子与 MIP DAE-DA 模型,并对利用该模型来调整两个利益群体内决策者的原因进行详细解释。

2.1.1.1 C-OWA 算子

为了集结离散型数据信息,Yager 充分考虑了决策者的态度或者主观偏好,提出了有序加权平均(OWA)算子。基于 OWA 算子,Yager 对连续性型数据的集成算子进行了研究,提出了一种连续区间数据集算子。

定义 2.1 设 $[a, b]$ 为任意的连续区间,且

$$F_Q([a, b], f) = \int_0^1 \frac{dQ(y)}{dy} H^{-[1]}(y) dy \quad (2.1)$$

其中, f 是定义在 $[a, b]$ 上的非负函数,对应于区间内元素的重要性。 $Q: [0, 1] \rightarrow [0, 1]$ 表示 BUM 函数,该函数具有如下几点性质:① $Q(0) = 0$; ② $Q(1) = 1$; ③ 如果 $x > y$, 那么 $Q(x) \geq Q(y)$ 。另外, $H^{-[1]}(y) (y \in [0, 1])$ 是函数 $H: y = H(x) = \int_x^b g(z) dz$

($x \in [a, b]$)的逆函数。 $g(z) = f(z) / \int_a^b f(x) dx$ 是 f 的标准化式。那么, F_Q 被称为连续区间的 OWA 算子, 即 C-OWA 算子。

特别地, 如果 $f(z) = K$ (K 为常数) 且 $z \in [a, b]$, 那么 $[a, b]$ 中的所有元素有相同的权重。此时, 式(2.1)可以简化为

$$F_Q([a, b]) = \int_0^1 \frac{dQ(y)}{dy} [b - y(b-a)] dy \quad (2.2)$$

通过对式(2.2)的展开运算, Yager 给出 F_Q 的一般式:

$$F_Q([a, b]) = a + (b-a) \int_0^1 Q(y) dy \quad (2.3)$$

令 $\lambda = \int_0^1 Q(y) dy$, 那么式(2.3)可以表示为

$$F_Q([a, b]) = (1-\lambda)a + \lambda b \quad (2.4)$$

$F_Q([a, b])$ 是基于区间端点的线性组合, λ 被称为 BUM 函数 Q 的态度特点且 $\lambda \in [0, 1]$ 。若 λ 越接近 0, 表明 DM 为消极型决策者, 倾向于把区间集成为一个趋近区间下限的较小数, 其对应的 Q 为乐观型 BUM 函数。反之, λ 越接近 1 表明 DM 越是乐观, 倾向于把区间集成为一个趋近区间上限的较大数, 其对应的 Q 为悲观型 BUM 函数。若 $\lambda = 1/2$, 表明 DM 保持中立, 其对应的 Q 为中立型 BUM 函数。

2.1.1.2 MIP DEA-DA 模型

DEA-DA 是由 Sueyoshi 提出的用于判别分析的非参数方法。该方法同时具备了 DEA 和 DA 的特点, 从而使得 DA 方法更有灵活性。对于两个样本组 G_1 和 G_2 , 该方法基本的思想是采用两阶段(阶段 1 和阶段 2)线性规划模型实现有效的样本分组。阶段 1 通过两个超平面识别出两个样本组的重叠区域, 并且将重叠区域以外的样本划分到 G_1 或者 G_2 之中。阶段 2 通过另外两个超平面将重叠区域的样本划分到 G_1 或者 G_2 之中。DEA-DA 的基本思想框架如图 2.1 所示。

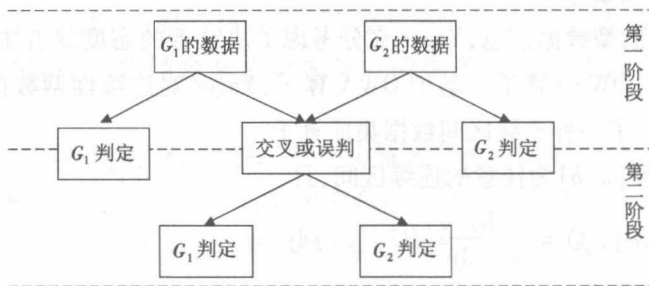


图 2.1 DEA-DA 模型的基本思路框架

注:①重叠区域是指该区域同时包含了 G_1 中的样本和 G_2 中的样本;②误判表明