

文章编号:1001-7402(2021)06-0087-14

直觉模糊序信息系统中双量化粗糙集模型^{*}

汤悦¹,李文涛¹,徐伟华¹,詹弢²

(1. 西南大学 人工智能学院,重庆 400715;

2. 西南大学 数学与统计学院,重庆 400715)

摘要:研究经典信息系统中单量化粗糙集模型在直觉模糊序信息系统中的表达形式,建立直觉模糊序信息系统下双量化粗糙集模型,给出直觉模糊序信息系统下双量化粗糙集模型求解上近似和下近似的算法,最后通过实际案例分析构造的模型。

关键词:直觉模糊集;序信息系统;双量化;粗糙集

中图分类号:O159 文献标识码:A

1965 年美国加利福尼亚 L.A.Zadeh 教授发表的《模糊集合》一文标志着模糊数学的诞生^[1]。其中一个重要概念是隶属度,它表示一个元素对一个集合的隶属程度,隶属度的取值范围是 $[0,1]$ ^[2]。由于实际生活中的多数问题不仅是模糊的,还可能是直觉模糊的,因此 1986 年保加利亚学者 Atanassov 基于模糊理论提出了直觉模糊的概念^[3]。直觉模糊集合不仅包含隶属度的概念还包含非隶属度的概念。这一概念的引入更加切合实际生活中的模糊性不确定性问题,能够包含更多不确定性信息。结合实际生活中有些问题不仅是直觉模糊而且有序^[4-5],因此本文在直觉模糊的基础上考虑序关系,即直觉模糊序关系^[6-7]。

粗糙集是一种处理不确定性事物的工具^[8],它能够用确定的集合即上下近似逼近不确定的集合。经典粗糙集上下近似的定义条件过于严苛,在实际应用时有很大局限性。于是学者们提出了扩展模型:决策粗糙集模型(DTRS)和程度粗糙集模型(GRS)^[9-12]。这两个模型分别能够将相对量化信息和绝对量化信息运用到上下近似的计算中,且对上下近似的定义也具有一定的灵活性。其中 DTRS 是将等价类属于目标集合的条件概率代替经典粗糙集上下近似中严格的包含和是否有交集的关系;GRS 是基于等价类与目标集合重叠的程度来定义上下近似,通过设定一个相交元素个数的阈值 k 来实现。但是这两个模型在使用过程中只能考虑到一种量化信息,为此 2015 年 Li 和 Xu 在文献^[13]中基于经典粗糙集模型提出了双量化决策粗糙集(Dq-DTRS)的概念,模型将绝对量化信息和相对量化信息结合起来,定义了两个双量化模型:DqI-DTRS 模型和 DqII-DTRS 模型。对比单量化模型,双量化模型能够充分

* 收稿日期:2021-06-18;修订日期:2021-07-24

基金项目:大学生创新创业训练计划项目(rgznxy202002;101202010635586);国家自然科学基金资助项目(61772002;61976245);中央高校基本科研业务费项目(SWU119045);重庆市教育委员会“成渝地区双城经济圈建设”科技创新项目(KJCX2020009);重庆市教育委员会科学技术研究计划项目(KJQN202100205;KJQN202100206;KJQN202003806)

作者简介:汤悦(2002-),女,江西吉安人,本科生,研究方向:粗糙集;李文涛(1987-),男,湖北咸宁人,讲师,博士,研究方向:粗糙集,模糊集,粒计算,认知概念学习;徐伟华(1979-),男,山西浑源人,教授,博士生导师,研究方向:信息科学,模糊数学,人工智能粗糙集,应用数学;詹弢(通讯作者)(1991-),四川雅安人,讲师,博士,研究方向:运筹学,控制理论及其应用。

的利用信息系统的相对量化信息和绝对量化信息^[14]；因此双量化模型比单量化模型具有更好的容错能力。随着双量化粗糙集理论的发展与研究的深入,近几年来已经有不少学者对犹豫模糊^[15]、双论域^[16]、不完备序信息系统^[17]及双论域模糊关系^[18]等环境下的双量化模型做了一系列的研究,但是有关直觉模糊序信息系统中的双量化粗糙集模型的研究还很少,为此有必要对直觉模糊序信息系统下双量化粗糙集进行深入的研究与探讨。

本文的研究动机主要分为以下三点:

(1)双量化模型克服了单量化模型的缺点,充分考虑了信息系统的相对量化信息和绝对量化信息,但多数都是基于等价关系,条件过于严苛。

(2)尽管关于双量化模型为主题的研究非常多,但关于直觉模糊序信息系统双量化粗糙集模型甚少被讨论到,因此本研究可以弥补双量化粗糙集理论的不足。

(3)本研究以直觉模糊序信息系统为出发点,将单量化模型和双量化模型推广至直觉模糊序信息系统,并考虑了模型的性质,丰富了双量化粗糙集理论。

文章第1部分回顾了相关基本知识,第2部分将经典粗糙集下的决策理论粗糙集模型(DTRS)和程度粗糙集模型(GRS)推广到直觉模糊序信息系统中^[19],在直觉模糊序信息系统中的序关系下研究单量化模型和双量化模型的定义。第3部分通过一个有关评委评分的实例来初步说明这个模型,接着第4板块给出了直觉模糊序信息系统求解两个双量化模型上下近似的算法。第5部分使用了四个不同维度的数据集进行数值仿真实验,结合实验结果对模型进行了更加全面深入的分析。最后,第6部分对全文进行总结。

1 预备知识

本节简要回顾直觉模糊序信息系统(简记为 IF)、单量化模型及双量化模型的相关基本知识。

定义 1.1 称 $IF=(Ud, At, F, V)$ 为直觉模糊信息系统,其中 $Ud=\{x_1, x_2, \dots, x_m\}$ 为论域,是一个非空有限集; $At=\{a_1, a_2, \dots, a_m\}$ 为非空属性集; F 为 Ud 和 At 的关系集; V 是属性 At 的值域; 满足 $\forall f \in F, \forall x \in Ud, \forall a \in At$ 都有 $f(x, a) = \langle \tau_a(x), \nu_a(x) \rangle$, 且 $0 \leq \tau_a(x), \nu_a(x) \leq 1, 0 \leq \tau_a(x) + \nu_a(x) \leq 1, \tau_a(x)$ 称为 x 在条件属性 a 下的隶属度, $\nu_a(x)$ 称为 x 在条件属性 a 下的非隶属度。其中定义 $f(a) = \{f(x, a) | a \in At\}$ 为直觉模糊集。

定义 1.2 称 $IF=(Ud, At, F, V)$ 为直觉模糊序信息系统, $\forall a \in At, f \in F, x_i, x_j \in Ud, IF^{\geq}$ 为直觉模糊序信息系统上的序关系满足 $f(x_i, a) \leq f(x_j, a) \Leftrightarrow \tau_a(x_i) \leq \tau_a(x_j)$ 且 $\nu_a(x_i) \geq \nu_a(x_j)$, 那么 x_i 在 x_j 的优势类中; 对于属性集 C 满足 $C \subseteq At$, 记 $x \in Ud$ 在直觉模糊序信息系统下关于属性子集 C 的优势类为 $[x]_C^{IF^{\geq}}$, 简写为 $[x]^{IF^{\geq}}$. 在直觉模糊序信息系统 $IF=(Ud, At, F, V)$ 中, 对于目标集合 $X \subseteq Ud$, 目标集合 X 在直觉模糊序关系 IF^{\geq} 下的近似算子定义为:

$$\begin{aligned} \overline{IF^{\geq}}(X) &= \{x \in Ud \mid X \cap [x]^{IF^{\geq}} \neq \emptyset\} \\ \underline{IF^{\geq}}(X) &= \{x \in Ud \mid [x]^{IF^{\geq}} \subseteq X\} \end{aligned} \quad (1)$$

若 $\overline{IF^{\geq}}(X) \neq \underline{IF^{\geq}}(X)$, 则称 X 在直觉模糊序关系 IF^{\geq} 下是粗糙的。

下面依据文献^[9-12]定义经典粗糙集下的单量化模型: 决策粗糙集模型(DTRS)和程度粗糙集模型(GRS)。

定义 1.3 设 $I=(Ud, At, F, D, V)$ 为决策信息系统, D 为决策属性。有条件属性集 $C \subseteq At, R_C$ 为条件属性子集 C 下的等价关系; α, β 为参数且满足 $0 \leq \beta \leq \alpha \leq 1$, 对于 $X \subseteq Ud$ 关于等价关系 R_C 在参数 (α, β) 上的概率上近似和概率下近似定义为:

$$\begin{aligned} \overline{R}_{(\alpha, \beta)}(X) &= \{x \in Ud \mid P(X \mid [x]_{R_C}) > \beta\} \\ \underline{R}_{(\alpha, \beta)}(X) &= \{x \in Ud \mid P(X \mid [x]_{R_C}) \geq \alpha\} \end{aligned} \quad (2)$$

$[x]_R$ 为等价关系 R_C 划分出的等价类, $P(X | [x]_R)$ 为条件概率, 等价于 $P(X | [x]_R) = \frac{|[x]_R \cap X|}{|[x]_R|}$ 。

当 X 关于参数 (α, β) 的概率上近似不等于概率下近似时, 则认为 X 在 (α, β) 此近似下是 R_C 粗糙的, 反之为精确的。此模型称为决策粗糙集模型(DTRS)。

其中目标集 X 的关于参数 (α, β) 概率近似的三个近似区域: 正域、负域及边界域的定义如下:

$$\begin{aligned} pos(X) &= \{x \in Ud \mid P(X | [x]_R) \geq \alpha\} \\ neg(X) &= \{x \in Ud \mid P(X | [x]_R) \leq \beta\} \\ bn(X) &= \{x \in Ud \mid \beta < P(X | [x]_R) < \alpha\} \end{aligned} \quad (3)$$

定义 1.4 设 $k \leq |Ud|$ ($|Ud|$ 为 Ud 的基数), 目标集合 X 在程度 k 下的上下近似定义为:

$$\begin{aligned} \bar{R}_k(X) &= \{x \in Ud \mid |[x]_R \cap X| > k\} \\ \underline{R}_k(X) &= \{x \in Ud \mid |[x]_R| - |[x]_R \cap X| \leq k\} \end{aligned} \quad (4)$$

当集合 X 的上近似不等于下近似时称 X 在程度 k 下是 R 粗糙的, 反之为精确的。称此模型为程度粗糙集模型(GRS)。在程度粗糙集中 $\underline{R}_k(X) \subseteq \bar{R}_k(X)$ 并不一定成立, 其上下近似受 k 取值的影响。

2 直觉模糊序信息系统中双量化粗糙集模型

由于现实生活中存在很多模糊性、不确定性的问题, 并且有些问题既是基于直觉模糊序关系的, 但是涉及直觉模糊序信息系统量化领域的研究还很少, 为此很有必要研究直觉模糊序信息系统下的量化模型。通过定义 1.1 和定义 1.2 我们清楚了直觉模糊序信息系统的概念, 定义 1.3 和定义 1.4 介绍了单量化模型在经典粗糙集中的表现形式。为了研究四个量化模型在直觉模糊序信息系统下的定义形式以及相关性质, 本节依次将四种量化模型引入到直觉模糊序信息系统中, 并给出了每个模型近似算子的定义形式, 同时研究模型的相关性质并进行了推理证明。

2.1 直觉模糊序信息系统中单量化粗糙集模型

本文将直觉模糊序信息系统简称为: IF ; 为方便引入双量化粗糙集模型, 下面首先给出了 IF 中单量化粗糙集模型的定义形式, 以及相关基本性质。

定义 2.1 设 $IF = (Ud, At, F, V)$ 为直觉模糊序信息系统, 目标集合 $X \subseteq Ud, C \subseteq At, IF^{\geq}$ 为条件属性子集 C 下的直觉模糊序关系; k 为自然数且满足 $k \leq |Ud|$ ($|Ud|$ 为 Ud 的基数), X 在程度 k 下的上下近似定义如下:

$$\begin{aligned} \overline{IF}_{(k)}^{\geq}(X) &= \{x \in Ud \mid |[x]^{IF^{\geq}} \cap X| > k\} \\ \underline{IF}_{(k)}^{\geq}(X) &= \{x \in Ud \mid |[x]^{IF^{\geq}}| - |[x]^{IF^{\geq}} \cap X| \leq k\} \end{aligned} \quad (5)$$

当 X 的上近似不等于下近似时称 X 在程度 k 下是 IF^{\geq} 粗糙的, 反之为精确的。称此模型为直觉模糊序信息系统程度粗糙集模型, 简记为 IF-GRS. 在 IF-GRS 中 $\overline{IF}_{(k)}^{\geq}(X) = \underline{IF}_{(k)}^{\geq}(X)$ 不一定成立。

性质 2.1 设 $IF = (Ud, At, F, V)$ 为直觉模糊序信息系统, $C \subseteq At, IF^{\geq}$ 为条件属性子集 C 下的直觉模糊序关系, k 为自然数且满足 $k \leq |Ud|$, 对于目标集合 $X, Y \subseteq Ud$ 有如下性质成立:

$$\begin{aligned} (L_1) \quad \underline{IF}_{(k)}^{\geq}(X) &= \sim \overline{IF}_{(k)}^{\geq}(\sim X), \\ (U_1) \quad \overline{IF}_{(k)}^{\geq}(X) &= \sim \underline{IF}_{(k)}^{\geq}(\sim X). \\ (L_2) \quad \underline{IF}_{(k)}^{\geq}(\emptyset) &= \overline{IF}_{(k)}^{\geq}(\emptyset) = \emptyset, \\ (U_2) \quad \underline{IF}_{(k)}^{\geq}(Ud) &= \overline{IF}_{(k)}^{\geq}(Ud) = Ud. \\ (L_3) \quad \underline{IF}_{(k)}^{\geq}(X \cap Y) &\subseteq \underline{IF}_{(k)}^{\geq}(X) \cap \underline{IF}_{(k)}^{\geq}(Y), \\ (U_3) \quad \overline{IF}_{(k)}^{\geq}(X) \cup \overline{IF}_{(k)}^{\geq}(Y) &\subseteq \overline{IF}_{(k)}^{\geq}(X \cup Y). \\ (L_4) \quad \underline{IF}_{(k)}^{\geq}(X \cup Y) &\subseteq \underline{IF}_{(k)}^{\geq}(X) \cup \underline{IF}_{(k)}^{\geq}(Y), \end{aligned}$$

$$(U_4) \overline{IF}_{(k)}^{\geq}(X) \cap \overline{IF}_{(k)}^{\geq}(Y) \subseteq \overline{IF}_{(k)}^{\geq}(X \cap Y).$$

$$(L_5) X \subseteq Y \Rightarrow \underline{IF}_{(k)}^{\geq}(X) \subseteq \underline{IF}_{(k)}^{\geq}(Y),$$

$$(U_5) X \subseteq Y \Rightarrow \overline{IF}_{(k)}^{\geq}(X) \subseteq \overline{IF}_{(k)}^{\geq}(Y).$$

证明 由于:

$$\begin{aligned} & \overline{IF}_{(k)}^{\geq}(\sim X) \\ &= \{x \in Ud \mid |[x]^{IF^{\geq}} \cap \sim X| > k\} \\ &= \sim \{x \in Ud \mid |[x]^{IF^{\geq}} \cap \sim X| \leq k\} \\ &= \sim \{x \in Ud \mid |[x]^{IF^{\geq}}| - |[x]^{IF^{\geq}} \cap X| \leq k\} \\ &= \sim \underline{IF}_{(k)}^{\geq}(X) \end{aligned}$$

则证 L_1, U_1 类似 L_1 可证. L_2 与 U_2 显然. 由于:

$$\begin{aligned} & \underline{IF}_{(k)}^{\geq}(X \cap Y) \\ &= \{x \in Ud \mid |[x]^{IF^{\geq}}| - |[x]^{IF^{\geq}} \cap (X \cap Y)| \leq k\} \\ &\Rightarrow \{x \in Ud \mid |[x]^{IF^{\geq}}| - |[x]^{IF^{\geq}} \cap X| \leq k\} \cap \{x \in Ud \mid |[x]^{IF^{\geq}}| - |[x]^{IF^{\geq}} \cap Y| \leq k\} \\ &= \underline{IF}_{(k)}^{\geq}(X) \cap \underline{IF}_{(k)}^{\geq}(Y) \end{aligned}$$

即性质 L_3 成立. 由于:

$$\begin{aligned} & \overline{IF}_{(k)}^{\geq}(X) \cup \overline{IF}_{(k)}^{\geq}(Y) \\ &= \{x \in Ud \mid |[x]^{IF^{\geq}} \cap X| > k\} \cup \{x \in Ud \mid |[x]^{IF^{\geq}} \cap Y| > k\} \\ &\Rightarrow \{x \in Ud \mid (|[x]^{IF^{\geq}} \cap X| \cup |[x]^{IF^{\geq}} \cap Y|) > k\} \\ &\Rightarrow \{x \in Ud \mid |[x]^{IF^{\geq}} \cap (X \cup Y)| > k\} \\ &= \overline{IF}_{(k)}^{\geq}(X \cup Y) \end{aligned}$$

即性质 U_3 成立. 依据定义有:

$$\begin{aligned} & \underline{IF}_{(k)}^{\geq}(X \cup Y) \\ &= \{x \in Ud \mid |[x]^{IF^{\geq}}| - |[x]^{IF^{\geq}} \cap (X \cup Y)| \leq k\} \\ &= \{x \in Ud \mid |[x]^{IF^{\geq}}| - |[x]^{IF^{\geq}} \cap X| \leq k\} \cap \{x \in Ud \mid |[x]^{IF^{\geq}}| - |[x]^{IF^{\geq}} \cap Y| \leq k\} \\ &\Rightarrow \{x \in Ud \mid |[x]^{IF^{\geq}}| - |[x]^{IF^{\geq}} \cap X| \leq k\} \cup \{x \in Ud \mid |[x]^{IF^{\geq}}| - |[x]^{IF^{\geq}} \cap Y| \leq k\} \\ &= \underline{IF}_{(k)}^{\geq}(X) \cup \underline{IF}_{(k)}^{\geq}(Y) \end{aligned}$$

则证 L_4 , 类似可证 U_4 . 由于: $\underline{IF}_{(k)}^{\geq}(X) = \{x \in Ud \mid |[x]^{IF^{\geq}}| - |[x]^{IF^{\geq}} \cap X| \leq k\}$, $\underline{IF}_{(k)}^{\geq}(Y) = \{x \in Ud \mid |[x]^{IF^{\geq}}| - |[x]^{IF^{\geq}} \cap Y| \leq k\}$, 由 $X \subseteq Y$ 推出 $\underline{IF}_{(k)}^{\geq}(X) \subseteq \underline{IF}_{(k)}^{\geq}(Y)$, 即证 L_5, U_5 类似 L_5 可证.

定义 2.2 设 $IF = (Ud, At, F, V)$ 为直觉模糊序信息系统, $C \subseteq At$, IF^{\geq} 为条件属性子集 C 下的直觉模糊序关系; 参数 α, β 满足 $0 \leq \beta \leq \alpha \leq 1$, IF^{\geq} 下集合 $X \subseteq Ud$ 在参数 (α, β) 上的上下近似算子定义为:

$$\begin{aligned} \overline{IF}_{(\alpha, \beta)}^{\geq}(X) &= \{x \in Ud \mid P(X \mid [x]^{IF^{\geq}}) > \beta\} \\ \underline{IF}_{(\alpha, \beta)}^{\geq}(X) &= \{x \in Ud \mid P(X \mid [x]^{IF^{\geq}}) \geq \alpha\} \end{aligned} \quad (6)$$

其中, $[x]^{IF^{\geq}}$ 为 IF^{\geq} 确定的直觉模糊偏序类, $P(X \mid [x]^{IF^{\geq}})$ 为条件概率等价于 $P(X \mid [x]^{IF^{\geq}}) = \frac{|[x]^{IF^{\geq}} \cap X|}{|[x]^{IF^{\geq}}|}$. 称此模型为直觉模糊序信息系统下的决策粗糙集模型(简记为 IF-DTRS). 当 X 关于

参数 (α, β) 上的概率上近似不等于概率下近似时, 则认为 X 在 (α, β) 概率近似下是 IF^{\geq} 粗糙的. 同时目标集 X 关于参数 (α, β) 概率近似的三个近似区域: 正域(简记为 pos)、负域(简记为 neg)及边界域(简

记为 bn) 的定义如下:

$$\begin{aligned} pos(X) &= \{x \in Ud \mid P(X \mid [x]^{IF^{\geq}}) \geq \alpha\} \\ neg(X) &= \{x \in Ud \mid P(X \mid [x]^{IF^{\geq}}) \leq \beta\} \\ bn(X) &= \{x \in Ud \mid \beta < P(X \mid [x]^{IF^{\geq}}) < \alpha\} \end{aligned} \quad (7)$$

性质 2.2 设 $IF = (Ud, At, F, V)$ 为直觉模糊序信息系统, $C \subseteq At$, IF^{\geq} 为条件属性子集 C 下的直觉模糊序关系, α, β 为参数且满足 $0 \leq \beta \leq \alpha \leq 1$, 对于目标集合 $X, Y \subseteq Ud$, 类似于 IF-GRS 模型, IF-DTRS 有如下性质成立:

$$\begin{aligned} (L_6) \underline{IF}_{(\alpha, \beta)}^{\geq}(X) &= \sim \overline{IF}_{(\alpha, \beta)}^{\geq}(\sim X), \\ (U_6) \overline{IF}_{(\alpha, \beta)}^{\geq}(X) &= \sim \underline{IF}_{(\alpha, \beta)}^{\geq}(\sim X), \\ (L_7) \overline{IF}_{(\alpha, \beta)}^{\geq}(\emptyset) &= \emptyset, \\ (U_7) \underline{IF}_{(\alpha, \beta)}^{\geq}(Ud) &= Ud, \\ (L_8) \underline{IF}_{(\alpha, \beta)}^{\geq}(X \cap Y) &\subseteq \underline{IF}_{(\alpha, \beta)}^{\geq}(X) \cap \underline{IF}_{(\alpha, \beta)}^{\geq}(Y), \\ (U_8) \overline{IF}_{(\alpha, \beta)}^{\geq}(X) \cup \overline{IF}_{(\alpha, \beta)}^{\geq}(Y) &\subseteq \overline{IF}_{(\alpha, \beta)}^{\geq}(X \cup Y), \\ (L_9) \underline{IF}_{(\alpha, \beta)}^{\geq}(X \cup Y) &\subseteq \underline{IF}_{(\alpha, \beta)}^{\geq}(X) \cup \underline{IF}_{(\alpha, \beta)}^{\geq}(Y), \\ (L_9) \overline{IF}_{(\alpha, \beta)}^{\geq}(X) \cap \overline{IF}_{(\alpha, \beta)}^{\geq}(Y) &\subseteq \overline{IF}_{(\alpha, \beta)}^{\geq}(X \cap Y). \end{aligned}$$

证明 由性质 2.1 证明易证。说明: 当 $\alpha = 0$ 时 $\underline{IF}_{(\alpha, \beta)}^{\geq}(\emptyset) \neq \emptyset$; 当 $\beta = 1$ 时 $\overline{IF}_{(\alpha, \beta)}^{\geq}(Ud) \neq Ud$ 。

2.2 直觉模糊序信息系统中双量化粗糙集模型

定义 2.3 设 $IF = (Ud, At, F, V)$ 为直觉模糊序信息系统, $X \subseteq Ud, C \subseteq At, IF^{\geq}$ 为条件属性子集 C 下的直觉模糊序关系; 在 IF 中 DqI-DTRS 模型^[13] 的近似算子定义为:

$$\begin{aligned} \overline{IF}_{(\alpha, \beta)}^{\geq}(X) &= \{x \in Ud \mid P(X \mid [x]^{IF^{\geq}}) > \beta\} \\ \underline{IF}_{(k)}^{\geq}(X) &= \{x \in Ud \mid |[x]^{IF^{\geq}}| - |[x]^{IF^{\geq}} \cap X| \leq k\} \end{aligned} \quad (8)$$

简记为 IF-DGRSI 模型。

定义 2.4 在 IF 中 DqII-DTRS 模型^[13] 的近似算子定义为:

$$\begin{aligned} \overline{IF}_{(k)}^{\geq}(X) &= \{x \in Ud \mid |[x]^{IF^{\geq}} \cap X| > k\} \\ \underline{IF}_{(\alpha, \beta)}^{\geq}(X) &= \{x \in Ud \mid P(X \mid [x]^{IF^{\geq}}) \geq \alpha\} \end{aligned} \quad (9)$$

简记为 IF-DGRSII 模型。

由于在经典双量化粗糙集模型中, 下近似包含于上近似的关系不再成立, 因此文献[13]提出了正域、负域、上边界域和下边界域的概念, 现推广到直觉模糊序信息系统的双量化粗糙集模型中, 其定义如下:

定义 2.5 在 IF-DGRSI 中正域(pos)、负域(neg)、上边界域($upbn$)和下边界域($lowbn$)定义为:

$$\begin{aligned} pos(X) &= \overline{IF}_{(\alpha, \beta)}^{\geq}(X) \cap \underline{IF}_{(k)}^{\geq}(X) \\ neg(X) &= \sim (\overline{IF}_{(\alpha, \beta)}^{\geq}(X) \cup \underline{IF}_{(k)}^{\geq}(X)) \\ upbn(X) &= \overline{IF}_{(\alpha, \beta)}^{\geq}(X) - \underline{IF}_{(k)}^{\geq}(X) \\ lowbn(X) &= \underline{IF}_{(k)}^{\geq}(X) - \overline{IF}_{(\alpha, \beta)}^{\geq}(X) \end{aligned} \quad (10)$$

IF-DGRSII 中正域(pos)、负域(neg)、上边界域($upbn$)和下边界域($lowbn$)定义类似。当元素 x 属于正域时, 判断为接受决策; 当元素 x 属于负域时, 判断为拒绝决策; 当元素 x 属于上边界时, 判断为待定决策, 但更偏向于拒绝; 当元素 x 属于下边界域时, 判断为待定决策, 但更偏向于接受。

性质 2.3 在 IF-DGRSI 模型和 IF-DGRSII 模型中分别有如下性质成立:

$$(1) \overline{IF}_{(a,\beta)}^{\geq}(X) \subseteq \overline{IF}^{\geq}(X), \underline{IF}_{(k)}^{\geq}(X) \supseteq \underline{IF}^{\geq}(X),$$

$$(2) \overline{IF}_{(k)}^{\geq}(X) \subseteq \overline{IF}^{\geq}(X), \underline{IF}_{(a,\beta)}^{\geq}(X) \supseteq \underline{IF}^{\geq}(X).$$

证明 (1)在 IF-DGRSI 模型下,根据定义 2.3 中模型上下近似的定义可得到如下推导:

$$\begin{aligned} \overline{IF}_{(a,\beta)}^{\geq}(X) &= \{x \in Ud \mid P(X \mid [x]^{IF^{\geq}}) > \beta\} \\ &\Rightarrow \{x \in Ud \mid P(X \mid [x]^{IF^{\geq}}) > 0\} \\ &\Rightarrow \{x \in Ud \mid [x]^{IF^{\geq}} \cap X \neq \emptyset\} = \overline{IF}^{\geq}(X) \\ \underline{IF}_{(k)}^{\geq}(X) &= \{x \in Ud \mid [x]^{IF^{\geq}} \subseteq X\} \\ &\Rightarrow \{x \in Ud \mid |[x]^{IF^{\geq}}| - |[x]^{IF^{\geq}} \cap X| = 0\} \\ &\Rightarrow \{x \in Ud \mid |[x]^{IF^{\geq}}| - |[x]^{IF^{\geq}} \cap X| \leq k\} = \underline{IF}_{(k)}^{\geq}(X) \end{aligned}$$

即证 $\overline{IF}_{(a,\beta)}^{\geq}(X) \subseteq \overline{IF}^{\geq}(X), \underline{IF}_{(k)}^{\geq}(X) \supseteq \underline{IF}^{\geq}(X)$ 。

(2)在 IF-DGRSII 模型中:

$$\begin{aligned} \overline{IF}_{(k)}^{\geq}(X) &= \{x \in Ud \mid |[x]^{IF^{\geq}} \cap X| \geq k\} \\ &\Rightarrow \{x \in Ud \mid |[x]^{IF^{\geq}} \cap X| \neq \emptyset\} = \overline{IF}^{\geq}(X) \\ \underline{IF}_{(a,\beta)}^{\geq}(X) &= \{x \in Ud \mid [x]^{IF^{\geq}} \subseteq X\} = \{x \in Ud \mid P(X \mid [x]^{IF^{\geq}}) = 1\} \\ &\Rightarrow \{x \in Ud \mid P(X \mid [x]^{IF^{\geq}}) \geq \alpha\} = \underline{IF}_{(a,\beta)}^{\geq}(X) \end{aligned}$$

即证 $\overline{IF}_{(k)}^{\geq}(X) \subseteq \overline{IF}^{\geq}(X), \underline{IF}_{(a,\beta)}^{\geq}(X) \supseteq \underline{IF}^{\geq}(X)$ 。

通过性质 2.3 可知直觉模糊序信息系统双量化粗糙集模型是直觉模糊序信息系统下经典粗糙集模型的一个推广。

3 案例分析

结合 IF 系统中提出的四个模型,下文给出了一个直觉模糊序信息系统的实例,通过求解集合 X 在四个模型下的上下近似来解释模型,并研究了模型上下近似间的联系和双量化模型参数取不同值时对 X 的描述能力。

例 3.1 表 1 给出一个关于评委给参赛歌手评分的直觉模糊序信息系统,其中 $IF = (Ud, At \cup \{d\}, F, V), Ud = \{x_1, x_2, \dots, x_8\}, At = \{a_1, a_2, \dots, a_5\}$,依照评判结果目标集合 X 取 $\{x_4, x_5, x_6, x_8\}$ 。

表 1 评委评分直觉模糊信息表

X	a_1 (仪态)	a_2 (音准)	a_3 (难度)	a_4 (介绍)	a_5 (音色)	d(结果)
x_1	$\langle 0.4, 0.5 \rangle$	$\langle 0.3, 0.5 \rangle$	$\langle 0.8, 0.2 \rangle$	$\langle 0.4, 0.5 \rangle$	$\langle 0.7, 0.2 \rangle$	0
x_2	$\langle 0.3, 0.5 \rangle$	$\langle 0.4, 0.5 \rangle$	$\langle 0.6, 0.2 \rangle$	$\langle 0.4, 0.5 \rangle$	$\langle 0.7, 0.3 \rangle$	0
x_3	$\langle 0.3, 0.5 \rangle$	$\langle 0.1, 0.8 \rangle$	$\langle 0.8, 0.1 \rangle$	$\langle 0.4, 0.5 \rangle$	$\langle 0.7, 0.3 \rangle$	0
x_4	$\langle 0.1, 0.8 \rangle$	$\langle 0.1, 0.8 \rangle$	$\langle 0.4, 0.5 \rangle$	$\langle 0.1, 0.8 \rangle$	$\langle 0.8, 0.2 \rangle$	1
x_5	$\langle 0.7, 0.3 \rangle$	$\langle 0.4, 0.5 \rangle$	$\langle 0.9, 0.1 \rangle$	$\langle 0.4, 0.4 \rangle$	$\langle 0.8, 0.1 \rangle$	1
x_6	$\langle 0.3, 0.6 \rangle$	$\langle 0.4, 0.6 \rangle$	$\langle 0.7, 0.2 \rangle$	$\langle 0.5, 0.5 \rangle$	$\langle 0.8, 0.2 \rangle$	1
x_7	$\langle 0.4, 0.5 \rangle$	$\langle 0.4, 0.5 \rangle$	$\langle 0.8, 0.2 \rangle$	$\langle 0.4, 0.5 \rangle$	$\langle 0.8, 0.2 \rangle$	0
x_8	$\langle 0.4, 0.4 \rangle$	$\langle 0.4, 0.5 \rangle$	$\langle 0.9, 0.1 \rangle$	$\langle 0.7, 0.3 \rangle$	$\langle 0.8, 0.2 \rangle$	1

根据定义 1.2 不难求得此直觉模糊序信息系统下关于条件属性 At 形成的直觉模糊偏序类为:
 $[x_1]^{IF^{\geq}} = \{x_1, x_5, x_7, x_8\}, [x_2]^{IF^{\geq}} = \{x_2, x_5, x_7, x_8\}, [x_3]^{IF^{\geq}} = \{x_3, x_5, x_8\}, [x_4]^{IF^{\geq}} = \{x_4, x_5, x_6, x_7, x_8\}, [x_5]^{IF^{\geq}} = \{x_5\}, [x_6]^{IF^{\geq}} = \{x_6, x_8\}, [x_7]^{IF^{\geq}} = \{x_5, x_7, x_8\}, [x_8]^{IF^{\geq}} = \{x_8\}$ 。

为了直观的体现出每个模型的优劣,这里采用计算粗糙熵的方式来表示每组上下近似对 X 的描述能力。以下采用的是文献[20]中粗糙熵的定义: $\gamma(X) = \rho(X)GK(X)$,粗糙度 $\rho(X) = 1 - |\underline{IF}^{\geq}(X)| / |\overline{IF}^{\geq}(X)|$, $GK(X) = |X|^2 / |Ud|^2$ 为知识粒度, $|*|$ 表示 $*$ 的元素个数。粗糙熵具有不确定性刻画功能,此处使用粗糙熵而非粗糙度是因为粗糙熵在大多数情况下能够有效的克服粗糙度的某些缺陷^[20]。

首先由直觉模糊序信息系统经典粗糙集的概念求得目标集合 X 的上下近似为: $\overline{IF}^{\geq}(X) = \{x_1, x_2, x_3, x_4, x_5, x_6, x_7, x_8\}$, $\underline{IF}^{\geq}(X) = \{x_5, x_6, x_8\}$,求得相应粗糙熵为 0.15625。

基于定义 2.1 取程度 $k=1$ 求出直觉模糊序信息系统下的程度粗糙集模型(IF-GRS)的上下近似如下: $\overline{IF}_{(1)}^{\geq}(X) = \{x_1, x_2, x_3, x_4, x_5, x_6, x_7, x_8\}$, $\underline{IF}_{(1)}^{\geq}(X) = \{x_3, x_4, x_5, x_6, x_7, x_8\}$,粗糙熵为 0.0625。

基于定义 2.3 可求出 IF-DGRSI 模型的上下近似,为了使参数选取更加具有代表性我们使用了如下三组参数进行求解,情况一: $\alpha + \beta = 1$,情况二: $\alpha + \beta > 1$,情况三: $\alpha + \beta < 1$ 。

情况 1:考虑 $k=1, \beta=0.2$: $\overline{IF}_{(0.8,0.2)}^{\geq}(X) = \{x_1, x_2, x_3, x_4, x_5, x_6, x_7, x_8\}$, $\underline{IF}_{(1)}^{\geq}(X) = \{x_3, x_4, x_5, x_6, x_7, x_8\}$,粗糙熵为 0.0625。

情况 2:考虑 $k=1, \beta=0.5$: $\overline{IF}_{(0.7,0.5)}^{\geq}(X) = \{x_3, x_4, x_5, x_6, x_7, x_8\}$, $\underline{IF}_{(1)}^{\geq}(X) = \{x_3, x_4, x_5, x_6, x_7, x_8\}$,粗糙熵为 0。

情况 3:考虑 $k=1, \beta=0.4$: $\overline{IF}_{(0.5,0.4)}^{\geq}(X) = \{x_1, x_2, x_3, x_4, x_5, x_6, x_7, x_8\}$, $\underline{IF}_{(1)}^{\geq}(X) = \{x_3, x_4, x_5, x_6, x_7, x_8\}$,粗糙熵为 0.0625。

基于定义 2.4 并考虑如下三种情况求出 IF-DGRSII 模型的上下近似分别为:

情况 1:考虑 $\alpha=0.8, k=1$: $\overline{IF}_{(1)}^{\geq}(X) = \{x_1, x_2, x_3, x_4, x_5, x_6, x_7, x_8\}$, $\underline{IF}_{(0.8,0.2)}^{\geq}(X) = \{x_4, x_5, x_6, x_7, x_8\}$,粗糙熵为 0.09375。

情况 2:考虑 $\alpha=0.7, k=1$: $\overline{IF}_{(1)}^{\geq}(X) = \{x_1, x_2, x_3, x_4, x_5, x_6, x_7, x_8\}$, $\underline{IF}_{(0.7,0.5)}^{\geq}(X) = \{x_4, x_5, x_6, x_7, x_8\}$,其粗糙熵为 0.09375。

情况 3:考虑 $\alpha=0.5, k=1$: $\overline{IF}_{(1)}^{\geq}(X) = \{x_1, x_2, x_3, x_4, x_5, x_6, x_7, x_8\}$, $\underline{IF}_{(0.5,0.4)}^{\geq}(X) = \{x_1, x_2, x_3, x_4, x_5, x_6, x_7, x_8\}$,粗糙熵为 0。

仔细观察上述上下近似的计算结果可以发现:在 IF-DGRSI 中有: $\overline{IF}_{(0.8,0.2)}^{\geq}(X) \subseteq \overline{IF}^{\geq}(X)$, $\overline{IF}_{(0.7,0.5)}^{\geq}(X) \subseteq \overline{IF}^{\geq}(X)$, $\overline{IF}_{(0.5,0.4)}^{\geq}(X) \subseteq \overline{IF}^{\geq}(X)$; $\underline{IF}_{(1)}^{\geq}(X) \supseteq \underline{IF}^{\geq}(X)$ 。在 IF-DGRSII 中有: $\underline{IF}_{(0.8,0.2)}^{\geq}(X) \supseteq \underline{IF}^{\geq}(X)$, $\underline{IF}_{(0.7,0.5)}^{\geq}(X) \supseteq \underline{IF}^{\geq}(X)$, $\underline{IF}_{(0.5,0.4)}^{\geq}(X) \supseteq \underline{IF}^{\geq}(X)$; $\overline{IF}_{(1)}^{\geq}(X) \subseteq \overline{IF}^{\geq}(X)$ 。即验证了性质 2.3。

通过对比直觉模糊序信息系统中两个双量化模型与经典粗糙集和 GRS 模型的粗糙熵可以发现:IF-DGRSI 模型对目标集合的描述能力更强,其整体粗糙熵小于经典粗糙集和 IF-GRS 模型的粗糙熵,而 IF-DGRSII 模型优于经典粗糙集模型,但次于 IF-GRS 模型。可见 IF-DGRSI 模型更加适用于此问题,因此在模型选择中往往需要结合问题的实际情况来选择最合适的模型。

通过这个实例我们清楚了 IF 系统中双量化模型的求解与使用,同时也初步验证了双量化模型的性质以及双量化模型在实际应用中的优越性。

4 算法设计

在求解知识约简过程中,许多知识约简问题的求解如:近似约简、分配约简及约简^[4]往往需要先求

解目标集合的上下近似,且许多问题都是把决策类作为目标集合进行知识约简的,本节首先给出了算法思想流程图,接着给出了 IF-DGRSI 模型和 IF-DGRSII 模型求解上近似及下近似的算法。

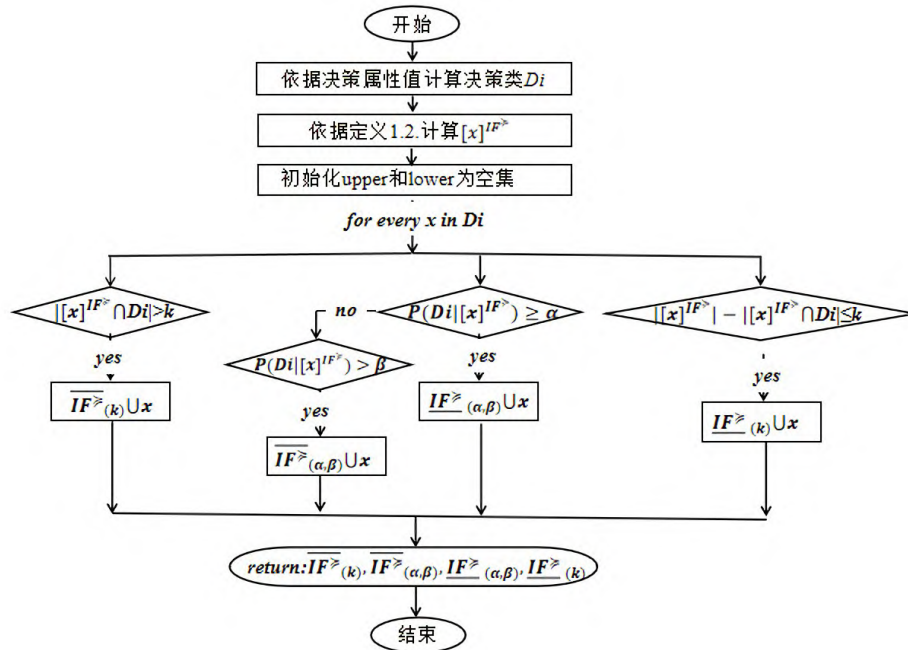


图1 算法思想流程图

算法 4.1 计算 IF-DGRSI 模型的上下近似

输入: 论域 Ud , 属性有 a 个, 样本的个数为 n , 决策属性 D 个数为 m .

输出: 每个决策类 D_i 在 IF-DGRSI 模型下的上近似和下近似。

```

1  begin
2  compute  $Ud/D = \{D_1, D_2, \dots, D_m\}$ ; // 计算决策类
3  for  $x \in Ud$  do
4  compute  $[x]^{IF \geq}$ ; // 依据定义 1.2 计算直觉模糊偏序类
5  end
6  for  $i = 1 : m$  do
7  let  $\overline{IF}_{(a,\beta)}^{\geq}(D_i) = \emptyset$ ;  $\underline{IF}_{(k)}^{\geq}(D_i) = \emptyset$ ;
8  for  $x \in Ud$  do
9  if  $P(D_i | [x]^{IF \geq}) > \beta$  then
10  $\overline{IF}_{(a,\beta)}^{\geq}(D_i) = \overline{IF}_{(a,\beta)}^{\geq}(D_i) \cup \{x\}$ ;
11 end
12 if  $| [x]^{IF \geq} | - | [x]^{IF \geq} \cap D_i | \leq k$  then
13  $\underline{IF}_{(k)}^{\geq}(D_i) = \underline{IF}_{(k)}^{\geq}(D_i) \cup \{x\}$ ;
14 end
15 end
16 end
17 return:  $\overline{IF}_{(a,\beta)}^{\geq}(D_i), \underline{IF}_{(k)}^{\geq}(D_i)$ .
18 end

```


算法 4.2 计算 IF-DGRSII 模型的上下近似

输入:论域 Ud ,属性有 a 个,样本的个数为 n ,决策属性 D 个数为 m .

输出:每个决策类 D_i 在 IF-DGRSII 模型下的上近似和下近似。

```

1  begin
2  compute  $Ud/D = \{D_1, D_2, \dots, D_m\}$ ; //计算决策类
3  for  $x \in Ud$  do
4  compute  $[x]^{IF^{\geq}}$ ; //计算直觉模糊偏序类
5  end
6  for  $i = 1; m$  do
7  let  $\overline{IF}_{(k)}^{\geq}(D_i) = \emptyset; \underline{IF}_{(\alpha, \beta)}^{\geq}(D_i) = \emptyset;$ 
8  for  $x \in Ud$  do
9  if  $|[x]^{IF^{\geq}} \cap D_i| > k$  then
10  $\overline{IF}_{(k)}^{\geq}(D_i) = \overline{IF}_{(\alpha, \beta)}^{\geq}(D_i) \cup \{x\};$ 
11 end
12 if  $P(D_i | [x]^{IF^{\geq}}) \geq \alpha$  then
13  $\underline{IF}_{(\alpha, \beta)}^{\geq}(D_i) = \underline{IF}_{(\alpha, \beta)}^{\geq}(D_i) \cup \{x\};$ 
14 end
15 end
16 end
17 return:  $\overline{IF}_{(k)}^{\geq}(D_i), \underline{IF}_{(\alpha, \beta)}^{\geq}(D_i).$ 
18 end

```

在算法中 k, α 和 β 分别为从键盘输入的参数。下面对算法进时间复杂度分析:设论域中样本个数有 n 个,决策属性包含 m 个取值,条件属性有 r 个。在算法 4.1 中首先计算了决策类,需要遍历每个样本对应决策属性的取值,因此时间复杂度为 $O(n)$,接着有两个平行的 for 循环,总的时间复杂度就为 $O(n)$ 和两个 for 循环相加。其中第一个 for 循环计算直觉模糊偏序类需要两次遍历 Ud ,并且遍历每一个样本在每一个属性下的取值,因此时间复杂度为 $O(n^2 r)$;第二个 for 循环包含一个嵌套 for 循环,则第二个循环的时间复杂度为两个 for 循环的时间复杂度相乘,因此第二个 for 循环的时间复杂度为 $O(mn)$ 。最后算得算法 4.1 的时间复杂度为 $O(n + n^2 r + mn)$,同理,算法 4.2 的时间复杂度为 $O(n + n^2 r + mn)$ 。

5 实验分析

为了从不同的角度体现出直觉模糊序信息系统双量化模型的优越性,本节以四个不同大小的数据集为背景,分别求解了直觉模糊序信息系统下经典粗糙集模型、IF-GRS 模型、IF-DTRS 模型和两个直觉模糊序双量化模型的上下近似;并且利用例 3.1 中的粗糙熵计算公式计算出每个模型结果的粗糙熵值,以此来反应每个模型对目标集合 X 的描述准确性,进而比较出每个模型的优劣。表 2 给出了四个实验数据表的一些基本信息。

表2 数据表基本信息

序号	属性个数	样本个数	目标集合 X 元素个数
A	5	200	100
B	6	440	200
C	9	803	220
D	6	2400	300

本实验利用 python 编程实现,其中 python 版本为 3.6.3。实验数据来源于 UCI 官网,数据类型全为数值型数据。由于本次实验需要直觉模糊数据,对原始数据预处理过程为:先对数据归一化,接着采用 $1-x$ 的形式得到原始数据的非隶属度(x 为原始数据)。实验主要步骤为:首先在不同参数下使用四个不同维度数据表求解出每个模型的上近似元素个数、下近似元素个数;接着通过粗糙熵的计算公式计算出粗糙熵的值。本次实验结合每一个数据集的维度分别对四个数据集随机选取了 100、200、220 和 300 个目标对象构成目标集合 X 。

表3 四个数据表的计算结果

数据表	X 元素个数	平均求解时间	模型	上近似元素个数	下近似元素个数	粗糙熵
A(200×5)	100	6s	IF-经典粗糙集模型	133	78	0.103834
			IF-GRS	$k=2$ 时: 124	111	0.0262096
			IF-DTRS	$\beta=0.2$ 时: 130 $\beta=0.25$ 时: 128 $\beta=0.3$ 时: 110	$\alpha=0.85$ 时: 85 $\alpha=0.75$ 时: 89 $\alpha=0.65$ 时: 94	0.0865385 0.0761719 0.0363636
			IF-DGRSI	$\beta=0.2$ 时: 130 $\beta=0.25$ 时: 128 $\beta=0.3$ 时: 110	$k=2$ 时: 111	0.0365385 0.0332031 -0.0022727
			IF-DGRSII	$k=2$ 时: 124	$\alpha=0.85$ 时: 85 $\alpha=0.75$ 时: 89 $\alpha=0.65$ 时: 94	0.078629 0.0705645 0.0604839
B(440×6)	200	21s	IF-经典粗糙集模型	356	5	0.2037813
			IF-GRS	$k=8$ 时: 206	32	0.1745166
			IF-DTRS	$\beta=0.05$ 时: 267 $\beta=0.06$ 时: 249 $\beta=0.07$ 时: 133	$\alpha=0.97$ 时: 5 $\alpha=0.94$ 时: 5 $\alpha=0.92$ 时: 5	0.2027424 0.2024627 0.1988442
			IF-DGRSI	$\beta=0.05$ 时: 267 $\beta=0.06$ 时: 249 $\beta=0.07$ 时: 133	$k=8$: 32	0.1818491 0.1800591 0.1569005
			IF-DGRSII	$k=8$ 时: 206	$\alpha=0.97$ 时: 5 $\alpha=0.94$ 时: 5 $\alpha=0.92$ 时: 5	0.2015967 0.2015967 0.2015967

C(803 * 9)	220	53s	IF-经典粗糙集模型	803	219	0.0545898
			IF-GRS	$k=15$ 时: 750	293	0.0457372
			IF-DTRS	$\beta=0.15$ 时: 803 $\beta=0.2$ 时: 802 $\beta=0.25$ 时: 800	$\alpha=0.9$ 时: 244 $\alpha=0.8$ 时: 276 $\alpha=0.7$ 时: 315	0.0522529 0.0492295 0.0455057
			IF-DGRSI	$\beta=0.15$ 时: 803 $\beta=0.2$ 时: 802 $\beta=0.25$ 时: 800	$k=15$ 时: 293	0.0476726 0.0476385 0.0475699
			IF-DGRSII	$k=15$ 时: 750	$\alpha=0.9$ 时: 244 $\alpha=0.8$ 时: 276 $\alpha=0.7$ 时: 315	0.0506411 0.0474385 0.0435354
D(2400 * 6)	300	2min56s	IF-经典粗糙集模型	2400	299	0.0136784
			IF-GRS	$k=14$ 时: 2396	764	0.0106427
			IF-DTRS	$\beta=0.2$ 时: 1492 $\beta=0.25$ 时: 1176 $\beta=0.3$ 时: 1003	$\alpha=0.9$ 时: 333 $\alpha=0.75$ 时: 404 $\alpha=0.65$ 时: 473	0.0121377 0.0102572 0.0082565
			IF-DGRSI	$\beta=0.2$ 时: 1492 $\beta=0.25$ 时: 1176 $\beta=0.3$ 时: 1003	$k=14$ 时: 764	0.007624 0.0054741 0.0037232
			IF-DGRSII	$k=14$ 时: 2396	$\alpha=0.9$ 时: 333 $\alpha=0.75$ 时: 404 $\alpha=0.65$ 时: 473	0.0766346 0.0472786 0.0058712

首先基于数据集 A 在求解 IF-DGRSI 模型和 IF-DGRSII 模型的上下近似时考虑了三种情况分别是:情况一: $\alpha + \beta = 1$,情况二: $\alpha + \beta > 1$,情况三: $\alpha + \beta < 1$;参数取值是依据样本大小和目标集合大小而调整的,体现出了直觉模糊序信息系统双量化模型的灵活性。此处 k 取值为 2,是因为在 A 数据集上求出的偏序类大部分元素个数都小于 10 个,因此当 k 取值为 2 时,既能保持上近似相对于经典情况下更小,又能保证下近似相对于经典情况下有更好的容错性。对数据集 B、数据集 C 和数据集 D 参数的选取都类似于上述数据集 A 的选取,此处不再赘述。

依据每个参数值带入对应模型计算的结果如表 3。

从整体结果中可以发现不论是数据集 A、B、C 还是 D 求解的粗糙熵数据中,在 IF-DGRSI 模型或者 IF-DGRSII 模型中总是存在一个粗糙熵最小值(已在表格中用方框标出),即双量化模型在一定程度上能够更好的描述目标集合 X。为了展现更加直观图 2 给出了不同模型和参数下用数据集 A 和 B 求解的粗糙熵数值的变化情况,图 3 给出了数据集 C 和 D 求解的粗糙熵数数值变化图,其中 IF-DTRS(1)、IF-DTRS(2)、IF-DTRS(3) 分别对应于 IF-DTRS 模型取三种不同参数值后得到的子模型,IF-DTRS(1)、IF-DGRSI(2)、IF-DGRSII(3)、IF-DGRSII(1)、IF-DGRSII(2)、IF-DGRSII(3) 类似。

图 2 中可以看出数据集 A 和 B 所得粗糙熵最小值都在 IF-DGRSI 模型的情况一中,图 3 中数据集 C 所求得粗糙熵最小值在 IF-DGRSII 模型的情况三中,数据集 D 所求得粗糙熵最小值在 IF-DGRSI 模

型的情况三中。参考离散型参数得到的结果,可以猜测在参数变化区间内双量化模型中总是存在一组参数值使得该模型下上下近似对目标集合 X 的描述是最精确的,即粗糙熵的值是最小的。也就是说双量化模型中存在一组参数使得该上下近似对目标集合 X 的描述最精确。由于参数变化的范围是连续的,生活中求解出这样的一组参数需要耗费大量的时间和空间;再者若数据集维数很大时,我们求解出这样一组最优参数付出的代价将会是巨大的,因此在实际应用中我们可以采取在连续区间按一定的步长选取参数值的调参手段来得到一个近似最优参数值。

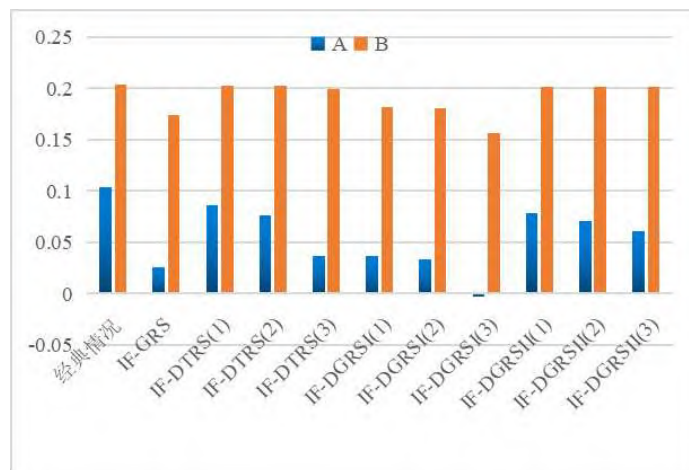


图2 A、B数据集不同参数值下粗糙熵变化情况

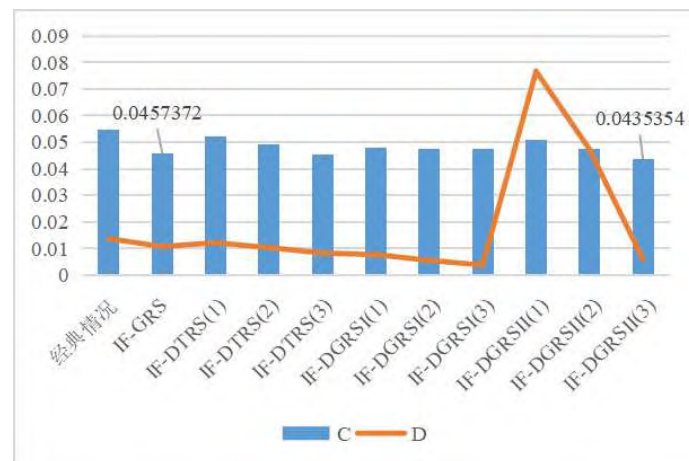


图3 C、D数据集不同参数值下粗糙熵变化情况

接着从具体的数值之间的差值来看,在数据集 A 下模型求解出的粗糙熵极差达到了 0.1056561,数据集 B 和数据集 C 的极差分别为:0.0468808 和 0.0110544,数据集 D 的极差为 0.0099552。这说明粗糙熵极差的大小与数据集的维度之间没有直接的线性关系。从数据 C 和数据集 D 计算的上近似元素个数来看,发现存在几组数据的上近似元素个数是十分接近论域中元素个数的(已在表格中加粗)。通过查看计算得到的偏序类,发现随着数据集的增大,偏序类与偏序类之间相同的元素个数逐渐增多。从序信息系统的角度分析原因,这和数据集的“序”有关,区别于等价关系,偏序关系形成的是覆盖而非划分,偏序类与偏序类间没有明确的界限。假设对象 x 在数据集中有至少 80% 的元素都优于 x ,称 x 为“劣元”。那么数据集中大部分元素都要优于“劣元”,从而“劣元”将会出现在大部分偏序类中,覆盖率增大。与此同时一个“劣元”的偏序类里面可能会包含论域中的全体对象。若在求解上近似时,目标集中含有多个类似于“劣元”的对象,且目标集中包含“劣元”,那么得到的上近似的元素个数就很容易接近于

论域元素的个数。而我们知道,当下近似元素个数不变时,上近似元素个数越大则粗糙度越大,而在 X 和 Ud 确定时知识粒度 $GK(X) = |X|^2 / |Ud|^2$ 为定值,那么粗糙熵的值越大,即模型对目标集合 X 的描述越不精确;当然随着偏序类中元素个数的增大也大大增加了计算量。为了解决此问题则需要将“覆盖”使用某个划分函数^[21]转变为划分,或者采用某些手段降低偏序类之间的覆盖率。虽然已有学者定义了划分函数 $h(x)$ 的概念^[21],不过对于实际问题的解决,仅仅依靠单一的划分函数是不够的,关于划分函数的更多定义形式还有待于接下来进行深入的研究与探讨。

6 结束语

本文在直觉模糊序信息系统中提出两个双量化粗糙集模型的概念,分别是 IF-DGRSI 模型和 IF-DGRSII 模型;研究了每个模型上下近似的求解方法及对应量化模型的性质;接着通过一个关于评委评分的实例求解出每个模型的上下近似,以此来初步解释这四个模型,并通过模型求解的上下近似之间的联系验证了模型的相关性质。为了使结果更加具有说服力,本文还使用了四个不同维度的数据集对模型做了更加全面的数值分析与解释,通过实验发现双量化模型在一定程度上能够更好的描述目标集合。

文中所提出的直觉模糊序信息系统中的四个量化模型进一步丰富了直觉模糊序信息系统的粗糙集理论,同时在数值实验中发现双量化模型对于目标集合 X 最精确描述上下近似算子的存在性。在接下来的研究中,我们可以考虑两个方面:一是研究直觉模糊序信息系统基于双量化模型的不确定性度量和知识约简的启发式算法,以探索出更加高效、便捷的约简方法;二是在不改变序关系的基础上探索出更多能够广泛应用的划分函数,解决“劣元”存在时偏序类间高覆盖率的问题。

参考文献:

- [1] Zadeh L A. Fuzzy sets[J]. Inform and Control, 1965, 8(3): 338-353.
- [2] 谢季坚,刘承平. 模糊数学方法及其应用[M]. 武汉:华中科技大学出版社, 2013: 29-30.
- [3] Atanassov K T. Intuitionistic fuzzy sets[J]. Fuzzy Sets and Systems, 1986, 20(1): 87-96.
- [4] 张文修,梁怡,吴伟志. 信息系统与知识发现[M]. 北京:科学出版社, 2004: 46-57.
- [5] 徐伟华. 序信息系统与粗糙集[M]. 北京:科学出版社, 2013.
- [6] 胡猛,李蒙蒙,徐伟华. 直觉模糊序信息系统下变精度与程度的“逻辑且”粗糙集[J]. 计算机科学, 2017, 44(5): 206-210.
- [7] 桑彬彬,徐伟华. 直觉模糊序信息系统的分配约简[J]. 计算机科学, 2017, 44(6A): 75-79.
- [8] Pawlak Z. Rough sets[J]. International Journal of Computer and Information Sciences, 1982, 11(5): 341-356.
- [9] 余建航,徐伟华. 序信息系统下变精度与程度的“逻辑与”粗糙集[J]. 模糊系统与数学, 2015, 29(4): 147-149.
- [10] Liang D, Liu D, Kobina A. Three-way group decisions with decision-theoretic rough sets[J]. Information Sciences, 2016(345): 46-64.
- [11] 姚晟,吴照玉,陈菊,王维. 基于决策理论粗糙集的一种新属性约简方法[J]. 微电子学与计算机, 2019, 36(5): 76-79.
- [12] 赵天娜,米据生,解滨,梁美社. 基于多伴随直觉模糊粗糙集的三支决策[J]. 南京大学学报, 2017, 53(6): 1082-1086.
- [13] Li W, Xu W. Double-quantitative decision-theoretic rough set[J]. Information Sciences, 2015(316): 54-67.
- [14] 李文涛. 基于粒计算的决策模型与方法研究[D]. 重庆:重庆理工大学, 2015: 5-6, 42-44.
- [15] 杨倩. 犹豫模糊环境下双量化决策模型研究[D]. 重庆:重庆理工大学, 2019: 39-57.
- [16] 孙秉珍,胡晓元. 双论域上量化粗糙集模型及其应用[J]. 计算机工程及其应用, 2017, 53(20): 50-55.
- [17] 林艳丽,刘晓东. 不完备序信息系统下的局部双量化决策粗糙集研究[J]. 山东大学学报, 2020, 55(3): 89-97.
- [18] 李燃,王青海,鲁小云. 双论域上模糊关系的双量化粗糙集研究[J]. 微电子学与计算机, 2019, 36(10): 47-53.
- [19] 温雪俊. 直觉模糊序信息系统的广义优势关系[J]. Journal of Xi'an University(Natural Science Edition), 2016,

19(4):6-8.

[20] 黄国顺,文翰. 基于边界域和知识粒度的粗糙集不确定性度量[J]. 控制与决策,2016,31(6):984-986.

[21] Zhang J,Zhang X,Xu W,Wu Y. Local multigranulation decision-theoretic rough set in ordered information systems [J]. Methodologies and Application,2019,23:13247-13261.

Double-quantitative Rough Set Model in Intuitionistic Fuzzy Ordered Information System

TANG Yue¹,LI Wen-tao¹,XU Wei-hua¹,ZHAN Tao²

(1. College of Artificial Intelligence, Southwest University, Chongqing 400715, China;

2. College of Mathematics and Statistics, Southwest University, Chongqing 400715, China)

Abstract: This paper considers the expression form of single quantization rough set model in intuitionistic fuzzy ordered information system, establishes two double quantization rough set models in intuitionistic fuzzy ordered information system. Moreover, two algorithms for solving the upper approximation and lower approximation of double quantization rough set model in intuitionistic fuzzy ordered information system are presented. Finally, the constructed models are analyzed through an actual example.

Key words: Intuitionistic Fuzzy Set; Ordered Information System; Double-quantitative; Rough Sets