软件学报 ISSN 1000-9825, CODEN RUXUEW
Journal of Software,2018,29(11):3484−3499 [doi: 10.13328/j.cnki.jos.005292]

E-mail: jos@iscas.ac.cn
http://www.jos.org.cn
Tel: +86-10-62562563

基于严凹函数的粗糙集不确定性度量*

黄国顺, 文 翰

(佛山科学技术学院 数学与大数据学院,广东 佛山 528000)

通讯作者: 黄国顺, E-mail:fshgs_72@163.com

摘 要: 通过语义分析,提出了一种拓展的粗糙集不确定性度量公理化定义;将香农熵函数推广到严凹函数,提出了一类以条件概率为自变量、基于严凹函数的粗糙集不确定性度量公式,它是严凹函数值的加权平均.在此基础上,得到一系列粗糙集不确定性度量方法.从严凹函数视角讨论了基于模糊熵的不确定性度量方法,发现现有多种能够用于度量粗糙集不确定性的模糊熵函数都是所提出方法的特殊情形.比较了粗糙度、改进粗糙度和所提出方法的区别和联系,最后设计了一些算例,比较了各种方法的异同,验证了基于严凹函数的粗糙集不确定性度量与粗糙集不确定性语义是一致的.

关键词: 不确定性度量;严凹函数;模糊熵;粗糙度

中图法分类号: TP18

中文引用格式: 黄国顺,文翰.基于严凹函数的粗糙集不确定性度量.软件学报,2018,29(11):3484−3499. http://www.jos.org.cn/1000-9825/5292.htm

英文引用格式: Huang GS, Wen H. Uncertainty measures of rough set based on strictly concave functions. Ruan Jian Xue Bao/Journal of Software, 2018,29(11):3484−3499 (in Chinese). http://www.jos.org.cn/1000-9825/5292.htm

Uncertainty Measures of Rough Set Based on Strictly Concave Functions

HUANG Guo-Shun, WEN Han

(School of Mathmatics and Big Data, Foshan University, Foshan 528000, China)

Abstract: Based on the semantic analysis, a general axiomatic definition of uncertainty measure for rough set is proposed. By extending the Shannon entropy function to strictly concave function, a class of uncertainty measures based on strictly concave function are put forward. They are weighted average of strictly concave function, whose variable is a conditional probability. It follows that a series of measuring methods are developed. The measuring methods based on fuzzy entropy are discussed under the view of strictly concave function. It is proved that they are the special cases of the method proposed in this paper. The difference and relationship among roughness measure, modified rough measure and the uncertainty measure based on strictly concave function are discussed. Finally, some examples are designed to compare the methods discussed in this paper. It is found that the proposed uncertainty measures based on strictly concave function are consistent with the semantics of uncertainty of rough set.

Key words: uncertainty measure; strictly concave function; fuzzy entropy; roughness measure

粗糙集不确定性度量是数据挖掘、粒计算、粗糙集等领域的重要研究内容,已被广泛应用于属性约简、粒度计算和三支决策等领域[1−3].

整体而言,目前主要从 3 个方面对粗糙集的不确定性度量展开研究,即基于纯粗糙集的方法、信息熵及其

* 基金项目: 广东省自然科学基金(2015A030313636); 广东省教育厅普通高校特色创新类项目(2014KTSCX152)

Foundation item: Natural Science Foundation of Guangdong Province (2015A030313636); Project of Department of Education of Guangdong Province (2014KTSCX152)

收稿时间: 2015-04-03; 修改时间: 2015-11-21, 2016-05-25; 采用时间: 2017-04-10; jos 在线出版时间: 2018-04-16

CNKI 网络优先出版: 2018-04-16 10:37:40, http://kns.cnki.net/kcms/detail/11.2560.TP.20180416.1037.001.html

变种公式和模糊熵方法.Pawlak[4]基于纯粗糙集的方法提出粗糙度来度量粗糙集的不确定性问题,由于该方法只关注正域和边界域而忽视负域信息的影响,导致有时边界域分离出无关的知识颗粒时其值不变,这不符合我们的直觉.Beaubouef 等人和梁吉业等人[5,6]通过将粗糙度与知识粒度做积来度量粗糙集不确定性问题.徐宝文等人和王向阳等人[7,8]各自提出类似方法,但这类方法仍会产生一些问题,即只要划分变细,就会导致不确定性度量严格变小,即使与待刻画集合完全无关的负域细分,也会导致计算结果严格变小,这是不合理的.滕书华等人[9]基于一般二元关系提出一种新的不确定性度量集成方法.其次是基于信息熵及其变形公式的度量方法.针对完备信息系统,Duntsch 等人[10]从信息熵角度研究了粗糙集规则预测的不确定性度量问题, Wierman[11]基于信息熵研究了粗糙集的知识粒度度量问题.国内学者苗夺谦、王国胤等人较早涉足这方面的研究[12,13],李健等人[14]讨论了完备信息系统信息熵形式的不确定性度量,魏巍等人研究了几种不确定性度量之间的差异[15].针对不完备信息系统,多位学者给出了基于信息熵及其变形公式的不确定性度量方法,如梁吉业等人给出的信息熵和粗糙熵方法[16,17],Bianucci 等人和朱萍等人提出的 co-entropy[18,19],钱宇华等人提出的混合熵[20]等.代建华等人将信息熵方法推广到集值信息系统和区间值决策系统的不确定性度量研究中[21,22].Chakrabary 等学者从模糊性视角讨论了粗糙集的不确定性度量问题,借助模糊熵来刻画粗糙集的不确定性[23],梁吉业等人提出一种新的模糊熵度量方法[24],王国胤等人研究了在不同知识粒度下的粗糙集不确定性度量问题[25],提出了一种基于信息熵的修正模糊度度量方法,本质上是 Luca 和 Termin 对数型模糊熵的变形[26].魏巍等人对基于模糊熵的粗糙集不确定性度量方法展开系统研究,指出,并不是所有的模糊熵方法都适用于刻画粗糙集不确定性问题,同时给出了 2 个判定准则[27].但这类方法的缺陷也非常明显:首先,模糊熵本身存在诸多限制,例如隶属函数要求在 0.5 处达到最大值,关于 0.5 对称等;其次,并不是每一个模糊熵方法都能用来度量粗糙集的不确定性问题,这恰好说明模糊熵与粗糙集不确定性在语义上的差异性,至于有些模糊熵能用于刻画粗糙集的不确定性,那只能说明两者之间有交集.因此,必须跳出模糊熵的思维模式,去寻找更为广义的适用于粗糙集不确定性度量的方法.

近年来,胡军等人[28]提出了一种公理化形式的粗糙集不确定性度量定义,给出若干不确定性度量约束准则,但其中非负性约束条件过于宽松.黄国顺等人[29]对粗糙集不确定性度量进行语义分析,提出一类基于条件概率的粗糙集不确定性度量方法,证明了条件信息熵可由基于香农熵函数的不确定性度量公式诱导而来,从而揭示出条件信息熵的不确定性本质.

本文针对完备信息系统展开讨论,受文献[30]的启发,先将香农熵函数推广到严凹函数,提出一类基于严凹函数的粗糙集不确定性度量方法,理论上证明了现有多种粗糙集不确定度量方法,包括基于香农熵函数的信息熵和各种模糊熵函数,都是该方法的特殊情形.然后讨论了粗糙度、改进粗糙度和基于严凹函数不确定性度量方法之间的区别和联系.计算实例表明,本文建议的方法对边界域划分变细最敏感和合理.最后设计了一些算例,比较了各种方法的异同,验证了本文提出的粗糙集不确定性度量与粗糙集的不确定性语义是一致的.

1 相关基本概念

本节先回顾粗糙集相关基本知识和几种不确定性度量方法.

信息系统是一个四元组,记作 $I=\langle U,V,f,A\rangle$,其中,U 是一组对象的非空有限集合,称为论域;A 为有限属性集,$V=\bigcup_{a\in A}V_a$, V_a 为属性 a 的值域;$f:U\times A\to V$ 是信息函数.对 U 上的任意属性集 $P\subseteq A$ 定义不可分辨关系 $ind(P)=\{(x,y)\in U^2|\forall a\in P,f(x,a)=f(y,a)\}$,关系 $ind(P)$构成 U 的一个划分,记作 $U/ind(P)$,简记为 U/P.U/P 中的包含对象 x 的元素$[x]_P=\{y|\forall a\in P,f(x,a)=f(y,a)\}$称为等价类.若 $A=C\cup D$,C 为有限的条件属性集,D 为有限的决策属性集,$C\cap D=\varnothing$, $V=\bigcup_{a\in C}V_a$,$f:U\times(C\cup D)\to V$,则称为决策信息系统,记作 DIS.

给定信息系统 $IS=\langle U,V,f,A\rangle$,假设 $P\subseteq A,X\subseteq U$,称 $\underline{apr}_P(X)=\{x\in U:[x]_P\subseteq X\}$和$\overline{apr}_P(X)=\{x\in U:[x]_P\cap X\neq\varnothing\}$分别为 X 关于 P 的下近似集和上近似集.

显然,粗糙集 X 的上、下近似集将论域 U 分割成 3 个两两互不相交的区域,即正域、边界域和负域,其中,正域 $POS_P(X)=\underline{apr}_P(X)$,边界域 $BND_P(X)=\overline{apr}_P(X)-\underline{apr}_P(X)$,负域 $NEG_P(X)=U-\overline{apr}_P(X)$.

定义 1[6]. 设 $IS=\langle U,V,f,A\rangle$,$P,Q\subseteq A$,我们称 Q 粗于 P(或 P 细于 Q)当且仅当对任意的非空 $X_i\in U/P$,存在 $Q_j\in U/Q$,使得 $X_i\subseteq Q_j$,记作 $U/P\preceq U/Q$.如果 $U/P\preceq U/Q$,且存在某个 $X_{i_0}\in U/P,Q_{j_0}\in U/Q$,使得 $X_{i_0}\subset Q_{j_0}$,则称 Q 严格粗于 P(或 P 严格细于 Q),记作 $U/P\prec U/Q$.

特别地,如果 P 为恒等关系,即 $U/P=\{\{x\}|x\in U\}$,此时划分粒度最细;如果 P 为全域关系,即 $U/P=\{U\}$,此时划分粒度最粗.

对$\forall x\in U$,若$[x]_P=[x]_Q$,则称划分 U/P 和 U/Q 相等,记作 $U/P=U/Q$,简记为 $P\equiv Q$.

Wierman[11]在讨论知识粒度不确定性度量时,提出一种大小同构的概念,即对$\forall P,Q\subseteq A$,若存在双射 $h:U/P\to U/Q$,使得对任意的 $X_i\in U/P$,有$|X_i|=|h(X_i)|$,则称划分 U/P 和 U/Q 大小同构,记作 $U/P\cong U/Q$,简记作 $P\cong Q$.

为了度量用已知划分知识刻画未知集合的完全程度,Pawlak[4]给出了近似精度和粗糙度的概念.其中,X 关于划分 U/P 的近似精度定义为

$$\alpha_P(X)=\frac{|\underline{apr}_P(X)|}{|\overline{apr}_P(X)|}\tag{1}$$

相应的粗糙度定义为

$$\rho_P(X)=1-\alpha_P(X)=\frac{|BND_P(X)|}{|\overline{apr}_P(X)|}\tag{2}$$

然而,粗糙度有时在边界域分离出负域中的知识颗粒时具有二义性,即有时会保持不变,有时会严格减小.举例如下.

例 1:设 $U=\{x_1,x_2,x_3,x_4,x_5,x_6\}$,$X_1=\{x_1,x_5\}$,$X_2=\{x_1,x_3\}$,$U/P_1=\{\{x_1,x_2,x_3,x_4\},\{x_5,x_6\}\}$,$U/P_2=\{\{x_1,x_3\},\{x_2,x_4\},\{x_5,x_6\}\}$.

显然,$U/P_2\prec U/P_1$,在划分细分过程中分离出与 $X_i(i=1,2)$无关的颗粒$\{x_2,x_4\}$.对于 X_1,有 $\rho_{P_1}(X_1)=\rho_{P_2}(X_1)=1$.但对于 X_2,却有 $\rho_{P_1}(X_2)>\rho_{P_2}(X_2)$.

为此,Beaubouef 等[5]提出一种粗糙熵度量公式.

定义 2[5]. 设 $IS=\langle U,V,f,A\rangle$,$P\subseteq A$,$U/P=\{X_1,X_2,\dots,X_m\}$,$X\subseteq U$,那么 X 在 U/P 中的粗糙熵定义为

$$E_P(X)=-\rho_P(X)\sum_{i=1}^{m}\frac{|X_i|}{|U|}\log_2\frac{1}{|X_i|}\tag{3}$$

梁吉业等人在文献[2]证明了公式(3)本质上是将粗糙度与知识粒度做积.这样做可以克服粗糙度的某些不足,但同时,公式(3)会带来另外一个问题,即只要划分变细,粗糙熵 $E_P(X)$就会严格变小,即使与待刻画集合完全无关的负域发生细分,它也会严格变小,具体反例见例 2.

例 2:设 $U=\{x_1,x_2,\dots,x_8\}$,$X_3=\{x_1,x_2,x_5\}$,$U/P_3=\{\{x_1,x_2,x_3,x_4\},\{x_5,x_6\},\{x_7,x_8\}\}$,$U/P_4=\{\{x_1,x_2,x_3,x_4\},\{x_5,x_6\},\{x_7\},\{x_8\}\}$,显然,$U/P_4\prec U/P_3$,细分过程中只是将与 X_3 无关的知识颗粒$\{x_7,x_8\}$细分成$\{x_7\}$,$\{x_8\}$,其他没变.

此时,$\rho_{P_3}(X_3)=\rho_{P_4}(X_3)=1$,但 $E_{P_3}(X_3)>E_{P_4}(X_3)$.

苗夺谦等人和王国胤等人提出信息熵和条件信息熵等概念[12,13].

定义 3[13]. 设 $IS=\langle U,V,f,A\rangle$,$P\subseteq A$,记 $U/P=\{X_1,X_2,\dots,X_m\}$,则 P 在论域 U 上的信息熵定义为

$$H(P)=-\sum_{X_i\in U/P}p(X_i)\log_2 p(X_i)\tag{4}$$

其中,$p(X_i)=|X_i|/|U|$.

定义 4[13]. 给定 $DIS=\langle U,V,f,C\cup D\rangle$,设 $P\subseteq C$,$U/P=\{X_1,X_2,\dots,X_m\}$,$U/D=\{D_1,D_2,\dots,D_r\}$,那么 D 相对 P 的条件信息熵定义为

$$H(D|P)=-\sum_{X_i\in U/P}p(X_i)\sum_{D_j\in U/D}p(D_j|X_i)\log_2 p(D_j|X_i)\tag{5}$$

其中,$p(X_i)=|X_i|/|U|$,$p(D_j/X_i)=|X_i\cap D_j|/|X_i|$,$=1,2,\dots,m$,$j=1,2,\dots,r$.

2 基于严凹函数的粗糙集不确定性度量

本节先对粗糙集不确定性度量进行语义分析,给出一种扩展的粗糙集不确定度量公理化定义,然后将香农熵函数推广到严凹函数,提出一类基于严凹函数的粗糙集不确定性度量方法.

根据三支决策的观点,经典 Pawlak 粗糙集可看作是(1,0)三支粗糙集,它不允许任何据真错误和采伪错误发生,条件概率不为 1 或 0 的事件一律延迟决策,是一种“厌恶风险”的决策方式[31],它的不确定性主要来自边界域.因此,下文都是基于边界域讨论粗糙集的不确定性度量问题.

对于一个未知集合或概念 X 在已知划分 U/P 下的不确定性度量,我们认为,它至少应该满足如下几个约束条件.

(1) 非负性:即不确定性度量是关于划分 U/P 和待刻画集合 X 的非负函数.

(2) 相对性:X 在划分 U/P 下的不确定性度量是一个相对概念,其大小不但与划分 U/P 有关,而且与 X 在 U/P 中的分布结构有关.相同的 X 在大小同构的不同划分下可能有不同的不确定性度量值.

(3) 只要 X 在划分 U/P 中是确定集,即可认为它是确定的,此时应该没有不确定性,其不确定性度量值应该为 0,即使 U/P 再进一步细分成 U/Q,X 在划分 U/Q 中仍是确定集,其不确定性度量仍为 0,而不必要求只在恒等关系时才达到最小值 0;反之,只要不确定性度量值为 0,即意味着 X 在划分 U/P 下是确定集.

(4) 与 X 无关的负域即使进一步细分,还是与 X 无关,从而不会影响 X 的不确定性度量值.

(5) 来自边界域的细分一般会导致不确定性变小,如果细分过程中边界域分离出正域或负域中的知识颗粒,它的不确定性值应该严格变小.

根据以上语义分析,我们曾在文献[29]提出如下定义:

定义 5[29]. 设 $IS=\langle U,V,f,A\rangle$,$P\subseteq A$,$X\subseteq U$,$\Pi(U)$是 U 上所有划分的全体集合,$P(U)$是 U 的幂集,若存在 $\Pi(U)\times P(U)$到实数集的映射函数 $ER(X|P):\Pi(U)\times P(U)\rightarrow R^1$,如果满足如下条件,则称它为粗糙集 X 在划分 U/P 下的不确定性度量.

(1) 非负性:$ER(X|P)\geqslant 0$,并且 $ER(X|P)=0$ 当且仅当 $BND_P(X)=\varnothing$.

(2) 不变性:若 $BND_P(X)/P\equiv BND_Q(X)/Q$,那么 $ER(X|P)=ER(X|Q)$.

(3) 单调性:如果 $U/P\prec U/Q$,则 $ER(X|P)\leqslant ER(X|Q)$;如果 $U/P\prec U/Q$,且$|BND_P(X)|<|BND_Q(X)|$,则

$$ER(X|P)<ER(X|Q).$$

定义 5 的非负性强调当且仅当 $BND_P(X)=\varnothing$时,它的不确定性值为 0 且达到最小,从而避免了不具有不确定性但其度量值不为 0 的情形发生;不变性则强调当边界域划分完全相同时,它们具有相同的不确定性度量值,这意味着此时即使正域或负域细分,其不确定性值保持不变;单调性则强调,只有当细分过程中边界域分裂出确定区域(正域或负域)知识颗粒时,其不确定性度量值才会严格变小.上述 3 个约束条件都是合理的.

更进一步地,文献[29]提出两种基于条件概率的粗糙集不确定性度量和相应的知识不确定性度量方法,其中,与条件信息熵相关的度量定义如下.

定义 6[29]. 设 $IS=\langle U,V,f,A\rangle$,$P\subseteq A$,$X\subseteq U$,$U/P=\{X_1,X_2,\ldots,X_m\}$,定义公式:

$$ER_1(X\mid P)=-\sum_{i=1}^{m}\frac{|X_i\cap X|}{|U|}\log_2\frac{|X_i\cap X|}{|X_i|} \tag{6}$$

由它诱导出的知识不确定性度量定义如下.

定义 7[29]. 设 $IS=\langle U,V,f,A\rangle$,$P,Q\subseteq A$,$U/P=\{X_1,X_2,\ldots,X_m\}$,$U/Q=\{Y_1,Y_2,\ldots,Y_n\}$,知识 U/Q 在知识 U/P 中的不确定性度量定义为

$$ER_1(Q\mid P)=-\sum_{j=1}^{n}\sum_{i=1}^{m}\frac{|X_i\cap Y_j|}{|U|}\log_2\frac{|X_i\cap Y_j|}{|X_i|} \tag{7}$$

文献[29]研究结果表明,条件信息熵 $H(Q|P)$就是不确定性度量 $ER_1(Q|P)$,即 $H(Q|P)=ER_1(Q|P)$[29],从而揭示出条件信息熵的不确定性本质.

然而,定义 5 并没有刻画出粗糙集不确定性度量的相对性特点.我们知道,如果两边界域划分相同,那么它们

一定是大小同构的,但边界域大小同构不能保证同一个集合 X 在这些划分下的不确定性度量值相同.仍以 $ER_1(X|P)$为例,具体说明如下.

例 3:假设 $X=\{x_1,x_2,x_3,x_5,x_6,x_8\}$,$U=\{x_1,x_2,\ldots,x_{10}\}$,$U/P_1=\{\{x_1,x_2\},\{x_3,x_4,x_5,x_7\},\{x_6,x_8,x_9,x_{10}\}\}$,$U/P_2=\{\{x_1,x_2\},\{x_3,x_4,x_5,x_6\},\{x_7,x_8,x_9,x_{10}\}\}$.

显然,$BND_{P_1}(X)/P_1$ 与 $BND_{P_2}(X)/P_2$ 是大小同构的,但 $ER_1(X|P_1)\neq ER_1(X|P_2)$.

因此,度量粗糙集的不确定性必须带上它的分布结构信息.为此,我们提出如下分布同构概念.

定义 8. 设 $IS=\langle U,V,f,A\rangle$,对$\forall P,Q\subseteq A,X\subseteq U$,若存在双射 $h:U/P\to U/Q$,使得对任意的 $X_i\in U/P$,有$|X_i|=|h(X_i)|$,且$|X_i\cap X|=|h(X_i)\cap X|$,则称 X 在划分 U/P 和 U/Q 下具有相同的分布,简称为分布同构,记作 $U/P\approx U/Q$.

特别地,如果两边界域满足定义 8,则称为边界域分布同构,记作 $BND_P(X)/P\approx BND_Q(X)/Q$.当然,如果两边界域划分相同,则它们肯定是边界域分布同构的.为此,可提出定义 5 的如下拓展形式.

定义 9. 设 $IS=\langle U,V,f,A\rangle,P,Q\subseteq A,X\subseteq U$,记 $II(U)$是 U 上所有划分集合,$P(U)$是 U 的幂集,若存在 $II(U)\times P(U)$到实数集的映射函数 $ER(X|P):II(U)\times P(U)\to R^1$ 满足以下条件,那么称 $ER(X|P)$为粗糙集 X 在知识 U/P 下的不确定性度量.

(1) 非负性:$ER(X|P)\geqslant 0$,并且 $ER(X|P)=0$ 当且仅当 $BND_P(X)=\varnothing$.

(2) 不变性:若 $BND_P(X)/P\approx BND_Q(X)/Q$,则 $ER(X|P)=ER(X|Q)$.

(3) 单调性:如果 $U/P\prec U/Q$,则 $ER(X|P)\leqslant ER(X|Q)$;如果 $U/P\prec U/Q$,且$|BND_P(X)|<|BND_Q(X)|$,那么,

$$ER(X|P)<ER(X|Q).$$

与定义 5 相比,定义 9 将不变性条件由 $BND_P(X)/P\equiv BND_Q(X)/Q$ 放宽为 $BND_P(X)/P\approx BND_Q(X)/Q$,其余条件不变.

定理 1. 设 $IS=\langle U,V,f,A\rangle,P\subseteq A,X\subseteq U$,那么 $ER_1(X|P)$是定义 9 下的不确定性度量.

证明:根据定义 6 和定义 9,即知结论成立. □

由于 $ER_1(X|P)$可重写成如下等价形式:

$$ER_1(X\mid P)=-\sum_{i=1}^{m}\frac{|X_i|}{|U|}\times\frac{|X_i\cap X|}{|X_i|}\log_2\frac{|X_i\cap X|}{|X_i|} \tag{6'}$$

它可以看成是以条件概率$|X_i\cap X|/|X_i|$为自变量的香农熵函数值的加权平均.又注意到,香农熵函数是一种特殊的严凹函数,因此自然想到将香农熵函数推广到一般严凹函数,藉此提出一种以条件概率为自变量的基于严凹函数的粗糙集不确定性度量方法.

定义 10[32]. 设 $f(x)$是定义在凸集 I 上的实值函数,若对于任意的 $x,y\in I,x\neq y,0<\lambda<1$,有 $f(\lambda x+(1-\lambda)y)>\lambda f(x)+(1-\lambda)f(y)$,则称 $f(x)$为 I 上的严凹函数.

定理 2. 假设 $f(x)$是$[0,1]$上的非负严凹函数且 $f(0)=f(1)=0$,那么对任意的 $x\in(0,1)$,有 $f(x)>0$.

证明:对任意的 $x\in(0,1)$,存在$\lambda\in(0,1)$,使得 $x=1-\lambda$,从而 $f(x)=f(\lambda\times 0+(1-\lambda)\times 1)>\lambda f(0)+(1-\lambda)f(1)=0$. □

定义 11. 设 $IS=\langle U,V,f,A\rangle,P\subseteq A,X\subseteq U,U/P=\{X_1,X_2,\ldots,X_m\}$,$f(x)$是$[0,1]$上的非负严凹函数且 $f(0)=f(1)=0$,定义公式:

$$ER_f(X\mid P)=\sum_{i=1}^{m}\frac{|X_i|}{|U|}f\left(\frac{|X_i\cap X|}{|X_i|}\right) \tag{8}$$

定理 3. $ER_f(X|P)$是定义 9 下的不确定性度量.

证明:

(1) 非负性.

显然,$ER_f(X|P)\geqslant 0$.若 $ER_f(X|P)=0$,则对任意的 $X_i\in U/P$,$f(|X_i\cap X|/|X_i|)=0$.根据定理 2,一定有 $p(X|X_i)=0$ 或者 $p(X|X_i)=1$.若 $X=\varnothing$,则显然有 $BND_P(X)=\varnothing$;若 $X\neq\varnothing$,由于$\bigcup_{i=1}^{m}X_i=U$, $X\subseteq U$,必存在某个 X_{i_0} 使得 $X_{i_0}\cap X\neq\varnothing$, 从而必有 $p(X\mid X_{i_0})=1$,这意味着只要 $X_{i_0}\cap X\neq\varnothing$ 就有 $X_{i_0}\subseteq X$, 从而 $\overline{apr}_P X=\underline{apr}_P X$, 即 $BND_P(X)=\varnothing$.

反之,若 $BND_P(X)=\varnothing$,则易证 $ER_f(X|P)=0$.

(2) 若 $BND_P(X)/P\approx BND_Q(X)/Q$,那么$|X_i\cap X|/|X_i|=|h(X_i)\cap X|/|h(X_i)|$,显然有 $ER_f(X|P)=ER_f(X|Q)$.

(3) 如果 $U/P\prec U/Q$,则 $ER_f(X|P)\leqslant ER_f(X|Q)$.

假设 $U/P=\{X_1,X_2,\ldots,X_m\}$,$U/Q=\{Y_1,Y_2,\ldots,Y_n\}$,因 U/P 是 U/Q 的一个细分,记 $T[Y_j]=\{X_i|X_i\subseteq Y_j\}$,$j=1,2,\ldots,n$.
设 $T[Y_j]=\{X_{j_1},X_{j_2},\ldots,X_{j_l}\}$,那么,

$$\sum_{k=1}^{l}\frac{p(X_{j_k})}{p(Y_j)}p(X\mid X_{j_k})=\sum_{k=1}^{l}\frac{|X_{j_k}|}{|Y_j|}\frac{|X_{j_k}\cap X|}{|X_{j_k}|}=\frac{|Y_j\cap X|}{|Y_j|}=p(X\mid Y_j).$$

根据 $f(x)$的严凹性及 Jensen 不等式,

$$\begin{aligned}ER_f(X\mid P)&=\sum_{i=1}^{m}\frac{|X_i|}{|U|}f\left(\frac{|X_i\cap X|}{|X_i|}\right)\\&=\sum_{j=1}^{n}p(Y_j)\sum_{k=1}^{l}\frac{p(X_{j_k})}{p(Y_j)}f\left(\frac{|X_{j_k}\cap X|}{|X_{j_k}|}\right)\\&\leqslant\sum_{j=1}^{n}p(Y_j)f\left(\sum_{k=1}^{l}\frac{p(X_{j_k})}{p(Y_j)}p(X\mid X_{j_k})\right)\\&=\sum_{j=1}^{n}p(Y_j)f(X\mid Y_j)=ER_f(X\mid Q),\end{aligned}$$

且等号成立的充要条件是: $p(X\mid X_{j_1})=p(X\mid X_{j_2})=\ldots=p(X\mid X_{j_l})=p(X\mid Y_j)$.

若 $U/P\prec U/Q$ 且$|BND_P(X)|<|BND_Q(X)|$,则存在元素 $x_0\in BND_Q(X)$,使得 $p(X|[x_0]_P)\neq p(X|[x_0]_Q)$,根据前面证明过程,有 $ER_f(X|P)<ER_f(X|Q)$. □

定义 11 中的公式(8)可看作是以条件概率为自变量的严凹函数值的加权平均.显然,满足边界条件 $f(0)=f(1)=0$ 的非负严凹函数很多,由于条件概率的取值在区间[0,1],为讨论方便,下文规定 $x\in[0,1]$,同时规定 $0\log_2 0=0$.

(1) $f_1(x)=-x\log_2 x$

它是经典香农熵函数,当 $x\in(0,1)$时, $f_1''(x)=-\frac{1}{x}\ln 2<0$ 且 $f_1(0)=f_1(1)=0$,所以 $f_1(x)$是满足定义 11 的严凹函数,由它导出的公式 $ER_{f_1}(X\mid P)$ 即为 $ER_1(X|P)$.它是满足定义 9 的不确定性度量,与文献[14]的条件粗糙熵等价.

(2) $f_2(x)=-(1-x)\log_2(1-x)$

由于 $x\in(0,1)$时有 $f_2''(x)=\frac{1}{x-1}\ln 2<0$,易知 $f_2(x)$是满足定义 11 的严凹函数,由它导出的公式:

$$ER_{f_2}(X\mid P)=-\sum_{i=1}^{m}\left(\frac{|X_i|-|X_i\cap X|}{|U|}\right)\log_2\left(\frac{|X_i|-|X_i\cap X|}{|X_i|}\right)\tag{9}$$

是满足定义 9 的不确定性度量.

又因$|X_i|-|X_i\cap X|=|X_i-X|$,所以 $ER_{f_2}(X\mid P)=-\sum_{i=1}^{m}\left(\frac{|X_i-X|}{|U|}\right)\log_2\left(\frac{|X_i-X|}{|X_i|}\right)$.

(3) $f_3(x)=-[x\log_2 x+(1-x)\log_2(1-x)]$

它是由 Luca 和 Termini 提出的对数型模糊熵函数[26].显然,$f_3(x)=f_1(x)+f_2(x)$,根据前面论述易知,当 $x\in(0,1)$时,有 $f_3''(x)<0$.由 $f_3(x)$导出的公式:

$$ER_{f_3}(X\mid P)=-\sum_{i=1}^{m}\frac{|X_i\cap X|}{|U|}\log_2\left(\frac{|X_i\cap X|}{|X_i|}\right)-\sum_{i=1}^{m}\left(\frac{|X_i-X|}{|U|}\right)\log_2\left(\frac{|X_i-X|}{|X_i|}\right)\tag{10}$$

是满足定义 9 的不确定性度量.它与文献[25]中的修正模糊度等价,但又有本质的不同. $ER_{f_3}(X\mid P)$ 只依赖于划分颗粒与待刻画概念的条件概率.

(4) $f_4(x)=xe^{1-x}+(1-x)e^{x}-1$

它是由 Pal 和 Pal 提出的指数型模糊熵函数[33],当 $x\in[0,1]$时, $f_4''(x)=(x-2)e^{1-x}-(1+x)e^{x}<0$,易知它是满足

定义 11 的严凹函数,由它导出的公式:

$$ER_{f_4}(X\mid P)=\sum_{i=1}^{m}\left(\frac{|X_i\cap X|}{|U|}\mathrm{e}^{\frac{|X_i-X|}{|X_i|}}+\frac{|X_i-X|}{|U|}\mathrm{e}^{\frac{|X_i\cap X|}{|X_i|}}\right)-1 \tag{11}$$

是满足定义 9 的不确定性度量.

(5) $f_5(x)=x(1-x)$

由于当 $x\in(0,1)$时有 $f_5''(x)=-2<0$,易知 $f_5(x)$是满足定义 11 的严凹函数,由它导出的公式:

$$ER_{f_5}(X\mid P)=\frac{|X|}{|U|}-\sum_{i=1}^{m}\frac{|X_i|}{|U|}\left(\frac{|X_i\cap X|}{|X_i|}\right)^2 \tag{12}$$

是满足定义 9 的不确定性度量,它与文献[24]中的模糊熵等价.

(6) $f_6(x)=\dfrac{x}{1+x}-\dfrac{x}{2}$

显然,当 $x\in[0,1]$时, $f_6''(x)=-2/(1+x)^3<0$,易知 $f_6(x)$是满足定义 11 的严凹函数,由它导出的公式:

$$ER_{f_6}(X\mid P)=\frac{1}{|U|}\left(\sum_{i=1}^{m}\frac{|X_i\|X_i\cap X|}{|X_i|+|X_i\cap X|}-\frac{|X|}{2}\right) \tag{13}$$

是满足定义 9 的不确定性度量.

上述 6 个严凹函数在[0,1]区间上的图像如图 1 所示.显然,它们取值大小不一,其中,$f_6(x)$的取值最小,而 $f_3(x)$取值最大;其次,在这些函数中,只有 $f_3(x),f_4(x),f_5(x)$是模糊熵函数,它们都关于 x=0.5 对称,但 $f_1(x),f_2(x)$和 $f_6(x)$不是模糊熵函数,因为它们都不关于 x=0.5 对称,例如 $f_6(x)$,其图像如图 2 所示.当 $x_0=\sqrt{2}-1$ 时函数值达到最大,而不是在[0,1]区间的中点 x=0.5 处.

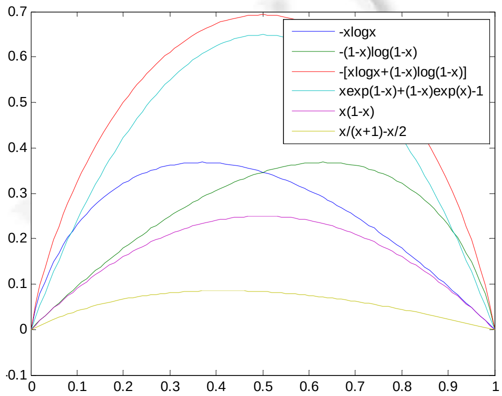

Fig.1 Figure of six strictly concave functions above on interval [0,1]

图 1 上述 6 个严凹函数在区间[0,1]上的图像

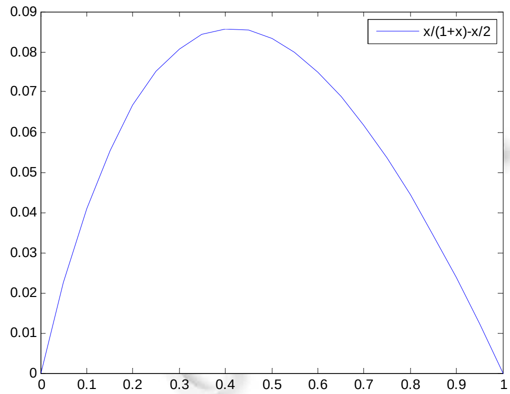

Fig.2 Figure of $y=f_6(x)$ on interval [0,1]

图 2 $y=f_6(x)$在区间[0,1]上的图像

例 4(例 1、例 2 续):利用 $ER_{f_i}(X\mid P)(i=1,2,\ldots,6)$ 分别计算例 1 和例 2,结果见表 1.

Table 1 Calculation results of different strictly concave functions

表 1 不同严凹函数所得计算结果

	$ER_f(X_1\|P_1)$	$ER_f(X_1\|P_2)$	$ER_f(X_2\|P_1)$	$ER_f(X_2\|P_2)$	$ER_f(X_3\|P_3)$	$ER_f(X_3\|P_4)$
f_1	0.500	0.333	0.333	0	0.375	0.375
f_2	0.374	0.333	0.333	0	0.375	0.375
f_3	0.874	0.666	0.666	0	0.75	0.75
f_4	0.544	0.432	0.432	0	0.487	0.487
f_5	0.208	0.167	0.167	0	0.188	0.188
f_6	0.078	0.056	0.056	0	0.0625	0.0625

尽管对于不同的严凹函数,其值大小不一,但表 1 的数据具有如下特点:(1) 只要边界域分离出负域中的知识颗粒,不确定性度量值会严格减小,$ER_f(X_1|P_1)>ER_f(X_1|P_2)$即是这种情况;(2) 当 $BND_{P_j}(X_i)=\varnothing$ 时,相应的不确定性度量达到最小值 0,$ER_f(X_2|P_2)$即是这种情况;(3) 如果只有负域中的知识颗粒发生细分,其他不变,则细分前后不确定性度量值保持不变,$ER_f(X_3|P_3)$和 $ER_f(X_3|P_4)$即是这种情况.

3 基于模糊熵的不确定性度量讨论

文献[27]中系统地讨论了基于模糊熵的粗糙集不确定性度量问题,上节提到的 $f_3(x),f_4(x),f_5(x)$都是模糊熵函数,可用于度量粗糙集的不确定性,本节将从严凹函数的角度出发,进一步考察其他模糊熵是否可用来度量粗糙集的不确定性问题.

设 $R^+=[0,+\infty)$,$F(U)$是论域 U 上的所有模糊集集合,对任意 $\tilde{X}\in F(U)$,记 $\tilde{X}=\mu_{\tilde{X}}(x_1)/x_1+\mu_{\tilde{X}}(x_2)/x_2+\ldots+\mu_{\tilde{X}}(x_n)/x_n$,$\tilde{X}^c$ 表示 $\tilde{X}$ 的补集,即 $\mu_{\tilde{X}^c}(x)=1-\mu_{\tilde{X}}(x)$.

Bhandari 等人引入如下σ-距离[34]:

$$D_L(\tilde{X}_1,\tilde{X}_2)=\frac{1}{2n}\sum_{i=1}^{n}\left((\mu_{\tilde{X}_1}(x)-\mu_{\tilde{X}_2}(x))\ln\frac{1+\mu_{\tilde{X}_1}(x)}{1+\mu_{\tilde{X}_2}(x)}+(\mu_{\tilde{X}_2}(x)-\mu_{\tilde{X}_1}(x))\ln\frac{2-\mu_{\tilde{X}_1}(x)}{2-\mu_{\tilde{X}_2}(x)}\right)\tag{14}$$

范九伦等学者定义了另一种σ-距离[35]:

$$D_E(\tilde{X}_1,\tilde{X}_2)=\frac{1}{2n(1-\mathrm{e}^{-1})}\sum_{i=1}^{n}(2-(1-\mu_{\tilde{X}_1}(x)+\mu_{\tilde{X}_2}(x))\mathrm{e}^{\mu_{\tilde{X}_1}(x)-\mu_{\tilde{X}_2}(x)}-(1-\mu_{\tilde{X}_2}(x)+\mu_{\tilde{X}_1}(x))\mathrm{e}^{\mu_{\tilde{X}_2}(x)-\mu_{\tilde{X}_1}(x)})\tag{15}$$

刘学成在文献[36]中定义了一种基于距离的模糊熵:

$$e_{c_1}^{d}(\tilde{X})=1-d(\tilde{X},\tilde{X}^c),\text{对}\forall\tilde{X}\in F(U)\tag{16}$$

特别地,

$$e_{c_1}^{D_L}(\tilde{X})=\frac{1}{2n}\sum_{i=1}^{n}f_{c_1}^{D_L}(\mu_{\tilde{X}}(x))\tag{17}$$

$$e_{c_1}^{D_E}(\tilde{X})=\frac{1}{2n(1-\mathrm{e}^{-1})}\sum_{i=1}^{n}f_{c_1}^{D_E}(\mu_{\tilde{X}}(x))\tag{18}$$

其中,$f_{c_1}^{D_L}(x)=1-(2x-1)\log_2\dfrac{1+x}{2-x}$,$f_{c_1}^{D_E}(x)=(1-x)\mathrm{e}^{2x-1}+x\mathrm{e}^{1-2x}-\mathrm{e}^{-1}$.

显然,$f_{c_1}^{D_L}(0)=f_{c_1}^{D_L}(1)=0,f_{c_1}^{D_E}(0)=f_{c_1}^{D_E}(1)=0$,当 $x\in[0,1]$时非负,且

$$f_{c_1}''^{D_L}(x)=\ln 2\left[\frac{-6}{(1+x)(2-x)}+\frac{18x^2-18x-9}{(1+x)^2(2-x)^2}\right]<0,\ f_{c_1}''^{D_E}(x)=-4x\mathrm{e}^{2x-1}-(1-x)\mathrm{e}^{1-2x}<0.$$

从而知它们是满足定义 11 的严凹函数.

假设 $P\subseteq A,X\subseteq U,U/P=\{X_1,X_2,\ldots,X_m\}$,对任意 $x\in X_i$,令 $\mu_P^X(x)=|X_i\cap X|/|X_i|$.因此得到由 X 和划分 U/P 导出的模糊集 $\tilde{X}_P=\mu_P^X(x_1)/x_1+\mu_P^X(x_2)/x_2+\ldots+\mu_P^X(x_n)/x_n$.由于对任意 $x,y\in X_i$ 有 $\mu_P^X(x)=\mu_P^X(y)$,因此:

$$e_{c_1}^{D_L}(\tilde{X}_P)=\frac{1}{2}ER_{f_{c_1}^{D_L}}(X|P),e_{c_1}^{D_E}(\tilde{X}_P)=\frac{1}{2(1-\mathrm{e}^{-1})}ER_{f_{c_1}^{D_E}}(X|P).$$

根据定理 3 知,$e_{c_1}^{D_L}(\tilde{X}_P),e_{c_1}^{D_E}(\tilde{X}_P)$是定义 9 下的不确定性度量.

范九伦和马远良提出了基于正则距离 d 的模糊熵公式[37]:

$$e_{c_2}^{d}(\tilde{X})=1-d(X\cap X^c,U)+d(X\cup X^c,U)\tag{19}$$

因此,

$$e_{c_2}^{D_L}(\tilde{X})=\frac{1}{n}\sum_{i=1}^{n}f_{c_2}^{D_L}(\mu_{\tilde{X}}(x)),e_{c_2}^{D_E}(\tilde{X})=\frac{1}{n}\sum_{i=1}^{n}f_{c_2}^{D_E}(\mu_{\tilde{X}}(x)).$$

其中,

$$f_{c_2}^{D_L}(x)=\begin{cases}1+\dfrac{1}{2}\left(2x-1+\log_2\dfrac{1+x}{2-x}\right), & 0\leqslant x<0.5\\ 1-\dfrac{1}{2}\left(2x-1+\log_2\dfrac{1+x}{2-x}\right), & 0.5\leqslant x\leqslant 1\end{cases},$$

$$f_{c_2}^{D_E}(x)=\begin{cases}1+\dfrac{1}{2(1-\mathrm{e}^{-1})}((2-x)\mathrm{e}^{x-1}+x\mathrm{e}^{1-x}-(x+1)\mathrm{e}^{-x}-(1-x)\mathrm{e}^{x}), & 0\leqslant x<0.5\\ 1-\dfrac{1}{2(1-\mathrm{e}^{-1})}((2-x)\mathrm{e}^{x-1}+x\mathrm{e}^{1-x}-(x+1)\mathrm{e}^{-x}-(1-x)\mathrm{e}^{x}), & 0.5\leqslant x\leqslant 1\end{cases}.$$

推论 1. $f_{c_2}^{D_L}(x)$ 在区间[0,1]上是满足定义 11 的严凹函数.

证明:显然, $f_{c_2}^{D_L}(0)=f_{c_2}^{D_L}(1)=0$, $f_{c_2}^{D_L}(x)$ 是区间[0,1]上的连续函数,且关于 $x=0.5$ 对称.因此,对于任意 $x\in[0,1]$, $f_{c_2}^{D_L}(x)=f_{c_2}^{D_L}(1-x)$,在[0,0.5)上单调递增,在[0.5,1]上单调递减,因此它在 $x=0.5$ 处达到最大.又因:

$$f_{c_2}''^{D_L}(x)=\begin{cases}\dfrac{6x-3}{2(1+x)^2(2-x)^2}, & 0\leqslant x<0.5\\ \dfrac{3-6x}{2(1+x)^2(2-x)^2}, & 0.5\leqslant x\leqslant 1\end{cases},$$

因此,当 $x\in[0,0.5)$或$(0.5,1]$,都有 $f_{c_2}''^{D_L}(x)<0$,从而对任意 x_1,x_2 同时属于[0,0.5)或(0.5,1], $x_1\neq x_2,\lambda\in(0,1)$,都有:

$$\lambda f_{c_2}^{D_L}(x_1)+(1-\lambda)f_{c_2}^{D_L}(x_2)<f_{c_2}^{D_L}(\lambda x_1+(1-\lambda)x_2).$$

所以只需证明:当 x_1,x_2 分布在 $x=0.5$ 的两边时,也有上述结论成立即可.不妨设 $0<x_1<0.5<x_2$,分 3 种情形分别讨论:(1) $x_1+x_2=1$;(2) $x_1+x_2>1$;(3) $x_1+x_2<1$.

- 对于情形(1),因为 $f_{c_2}^{D_L}(x_1)=f_{c_2}^{D_L}(1-x_1)$,所以 $\lambda f_{c_2}^{D_L}(x_1)+(1-\lambda)f_{c_2}^{D_L}(x_2)=f_{c_2}^{D_L}(x_1)<f_{c_2}^{D_L}(0.5)$;
- 对于情形(2),由于 $\lambda f_{c_2}^{D_L}(x_1)+(1-\lambda)f_{c_2}^{D_L}(x_2)=\lambda f_{c_2}^{D_L}(1-x_1)+(1-\lambda)f_{c_2}^{D_L}(x_2)<f_{c_2}^{D_L}(\lambda(1-x_1)+(1-\lambda)x_2)$,又因为$\lambda(1-x_1)+(1-\lambda)x_2>0.5$,若此时有$\lambda x_1+(1-\lambda)x_2>0.5$,根据它在[0.5,1]的单调递减性,有 $f_{c_2}^{D_L}(\lambda(1-x_1)+(1-\lambda)x_2)<f_{c_2}^{D_L}(\lambda x_1+(1-\lambda)x_2)$,从而结论成立;如果$\lambda x_1+(1-\lambda)x_2<0.5$,又因为$\lambda(1-x_1)+(1-\lambda)x_2-0.5>0.5-(\lambda x_1+(1-\lambda)x_2)$,从而 $f_{c_2}^{D_L}(\lambda(1-x_1)+(1-\lambda)x_2)<f_{c_2}^{D_L}(\lambda x_1+(1-\lambda)x_2)$,同样有:
$$\lambda f_{c_2}^{D_L}(x_1)+(1-\lambda)f_{c_2}^{D_L}(x_2)<f_{c_2}^{D_L}(\lambda x_1+(1-\lambda)x_2);$$
- 对于情形(3),与情形(2)证明类似.此处略.

因此,无论何种情形,对任意 $x_1,x_2\in[0,1],x_1\neq x_2,\lambda\in(0,1)$,都有 $\lambda f_{c_2}^{D_L}(x_1)+(1-\lambda)f_{c_2}^{D_L}(x_2)<f_{c_2}^{D_L}(\lambda x_1+(1-\lambda)x_2)$,从而知它是区间[0,1]上的严凹函数. □

为证明 $f_{c_2}^{D_E}(x)$ 也是严凹函数,需要用到如下几个引理:

引理 1[32]. 设 $f(x)$为区间 I 上的严凸函数,则对于 I 中任意的 $a<c<b$,有:

$$\frac{f(b)-f(c)}{b-c}>\frac{f(b)-f(a)}{b-a}>\frac{f(c)-f(a)}{c-a}.$$

引理 2. 设 $f(x)$是 I 上的严凸函数,对 I 中任意 $x_1>x_3>x_4>x_2$,且 $x_1+x_2=x_3+x_4$,那么 $f(x_1)+f(x_2)>f(x_3)+f(x_4)$.

证明:由于 $f(x)$是 I 上的严凸函数,所以对 I 中任意的 $x_1>x_3>x_4>x_2$,根据引理 1,有 $\dfrac{f(x_1)-f(x_3)}{x_1-x_3}>\dfrac{f(x_4)-f(x_2)}{x_4-x_2}$,即 $f(x_1)+f(x_2)>f(x_3)+f(x_4)$. □

推论 2. $f_{c_2}^{D_E}(x)$ 是[0,1]上的严凹函数.

证明:显然, $f_{c_2}^{D_E}(x)$ 是区间[0,1]上的连续函数,且关于 $x=0.5$ 对称,因此对于任意 $x\in[0,1]$, $f_{c_2}^{D_E}(x)=f_{c_2}^{D_E}(1-x)$. 在[0,0.5)上单调递增,在[0.5,1]上单调递减,因此它在 $x=0.5$ 处达到最大.又因

$$f_{c_2}''^{D_E}(x)=\begin{cases}\dfrac{1}{2(1-\mathrm{e}^{-1})}[-x\mathrm{e}^{x-1}-(2-x)\mathrm{e}^{1-x}+(x+1)\mathrm{e}^{x}+(1-x)\mathrm{e}^{-x}],\ 0\leqslant x<0.5\\ \dfrac{-1}{2(1-\mathrm{e}^{-1})}[-x\mathrm{e}^{x-1}-(2-x)\mathrm{e}^{1-x}+(x+1)\mathrm{e}^{x}+(1-x)\mathrm{e}^{-x}],\ 0.5\leqslant x\leqslant 1\end{cases},$$

先讨论 $x\in[0,0.5)$的情形.令 $h(x)=x\mathrm{e}^{x-1}$,那么,

$$f_{c_2}''^{D_E}(x)=\frac{1}{2(1-\mathrm{e}^{-1})}(-h(x)-h(2-x)+h(1+x)+h(1-x)).$$

当 $x=0$ 时,显然有 $f_{c_2}''^{D_E}(x)<0$,当 $0<x<0.5$ 时,因 $2-x>1+x>1-x>x\geqslant 0$,且 $h(x)$是$[0,+\infty)$上的严凸函数,从而 $h(x)+h(2-x)>h(1+x)+h(1-x)$,所以当 $0\leqslant x<0.5$ 时,有 $f_{c_2}''^{D_E}(x)<0$.

类似地,当 $0.5<x\leqslant 1$ 时,有 $f_{c_2}''^{D_E}(x)<0$.

仿照推论 1 的证明过程,类似地有结论:对于任意 $x_1,x_2\in[0,1],x_1\neq x_2,\lambda\in(0,1)$,都有 $\lambda f_{c_2}^{D_L}(x_1)+(1-\lambda)f_{c_2}^{D_L}(x_2)<f_{c_2}^{D_L}(\lambda x_1+(1-\lambda)x_2)$,从而可知 $f_{c_2}^{D_E}(x)$ 是区间[0,1]上的严凹函数. □

显然, $e_{c_2}^{D_L}(\tilde{X}_P)=ER_{f_{c_2}^{D_L}}(X\mid P),e_{c_2}^{D_E}(\tilde{X}_P)=ER_{f_{c_2}^{DE}}(X\mid P)$,又因 $f_{c_2}^{D_L}(0)=f_{c_2}^{D_L}(1)=0,f_{c_2}^{D_E}(0)=f_{c_2}^{D_E}(1)=0$且$f_{c_2}^{D_E}(x)$在区间[0,1]上非负,根据定理 3 可知,由它们导出的公式 $e_{c_2}^{D_L}(\tilde{X}_P),e_{c_2}^{D_E}(\tilde{X}_P)$ 是定义 9 下的不确定性度量.

范九伦等人提出了如下基于距离的模糊熵[38]:

$$e_F^d(\tilde{X})=d(\tilde{X},\tilde{X}_{near})+1-d(\tilde{X},\tilde{X}_{far})\tag{20}$$

那么,

$$e_F^{D_L}(\tilde{X})=\frac{1}{n}\sum_{i=1}^{n}f_F^{D_L}(\mu_{\tilde{X}}(x_i)),e_F^{D_E}(\tilde{X})=\frac{1}{n}\sum_{i=1}^{n}f_F^{D_E}(\mu_{\tilde{X}}(x_i)),$$

其中, $f_F^{D_L}(\mu_{\tilde{X}}(x_i)),f_F^{D_E}(\mu_{\tilde{X}}(x_i))$ 分别与 $f_{c_2}^{D_L}(x),f_{c_2}^{D_E}(x)$ 相同,从而知 $e_F^{D_L}(\tilde{X}_P),e_F^{D_E}(\tilde{X}_P)$ 是定义 9 下的不确定性度量.

在文献[37]中,范九伦和马远良提出了另一种σ-熵,归一化后变成如下形式:

$$e^{\alpha,\beta}(\tilde{X})=\frac{1}{n}\sum_{i=1}^{n}f^{\alpha,\beta}(\mu_{\tilde{X}}(x_i))\tag{21}$$

其中, $f^{\alpha,\beta}(x)=\dfrac{1}{(1-\alpha)\beta}[(x^{\alpha}+(1-x)^{\alpha})^{\beta}-1]$,$\alpha>0,\alpha\neq 1,\beta\neq 0,x\in[0,1]$.

当$\beta=1,x\in(0,1)$时,$f''^{\alpha,1}(x)=-\alpha(x^{\alpha-2}+(1-x)^{\alpha-2})<0$,可知 $f^{\alpha,1}(x)$是$(0,1)$上的非负严凹函数,且 $f^{\alpha,1}(0)=f^{\alpha,1}(1)=0$.根据定理 3 可知,由它们导出的公式 $e^{\alpha,1}(\tilde{X}_P)$ 是定义 9 下的不确定性度量.

当$\beta=1/\alpha$时,类似地可证明:当 $x\in(0,1)$时, $f''^{\alpha,\alpha^{-1}}(x)<0,f^{\alpha,\alpha^{-1}}(0)=f^{\alpha,\alpha^{-1}}(1)=0$ 且 $f^{\alpha,\alpha^{-1}}(x)$ 非负,根据定理 3 可知, $e^{\alpha,\alpha^{-1}}(\tilde{X}_P)$ 是定义 9 下的不确定性度量.

然而,并不是对任意的$\beta\neq 0$ 都能保证 $f^{\alpha,\beta}(x)$是严凹函数,由其导出的模糊熵也不可用于度量粗糙集的不确定性,具体反例如下.

例 5:取$\alpha=3,\beta=0.01,f^{3,0.01}(x)=-50((3x^2-3x+1)^{0.01}-1)$,假设 $U=\{x_1,x_2,\dots,x_{20}\},X=\{x_1,x_2\},U/P_1=\{\{x_1,x_2,\dots,x_{10}\},\{x_{11},x_{12},\dots,x_{20}\}\},U/P_2=\{\{x_1,x_2,\dots,x_{20}\}\}$.那么,

$$\tilde{X}_{P_1}=0.2/x_1+0.2/x_2+\dots+0.2/x_{10}+0/x_{11}+\dots+0/x_{20},$$

$$\tilde{X}_{P_2}=0.1/x_1+0.1/x_2+\dots+0.1/x_{20},$$

$$e^{3,0.01}(\tilde{X}_{P_1})=\frac{1}{2}f^{3,0.01}(0.2)=0.162948,$$

$$e^{3,0.01}(\tilde{X}_{P_2})=f^{3,0.01}(0.1)=0.157108.$$

虽然 $U/P_1\prec U/P_2$ 且 $|BND_{P_1}(X)|<|BND_{P_2}(X)|$,但是 $e^{3,0.01}(\tilde{X}_{P_1})>e^{3,0.01}(\tilde{X}_{P_2})$, 违背了定义 9 和文献[27]的定义 2.5(RP2).

当α=3,β=0.01时,模糊熵函数$f^{\alpha,\beta}(x)$的一阶导数$\dfrac{\mathrm{d}f^{\alpha,\beta}(x)}{\mathrm{d}x}=\dfrac{-3}{2}[(3x^2-3x+1)^{\beta-1}(2x-1)]$在[0,1]上的图像如图3所示,虽然满足条件——在$x\in[0,0.5]$时导数大于0,在$x\in[0.5,1]$时导数小于0,但它在区间$\left[0,0.5-\dfrac{\sqrt{9-12\beta}}{18-24\beta}\right]$和$\left[0.5+\dfrac{\sqrt{9-12\beta}}{18-24\beta},1\right]$上是增函数,其原函数的二阶导数会大于0,从而不是严凹函数.

当$\beta\to 0$时,$e^{\alpha,0}(\tilde{X}_P)=\dfrac{1}{n}\sum_{i=1}^{n}f^{\alpha,0}(\mu_{\tilde{X}}(x_i))$,其中,$f^{\alpha,0}(x)=\dfrac{1}{1-\alpha}\ln(x^\alpha+(1-x)^\alpha)$,它也不能保证对任意的$\alpha>0,\alpha\neq 1$是严凹函数,如$\alpha$=3时,$f^{\alpha,0}(x)=-\dfrac{1}{2}\ln(3x^2-3x+1)$,相应的一阶导数$f'^{\alpha,0}(x)=-\dfrac{3}{2}\dfrac{2x-1}{3x^2-3x+1}$在区间$\left[0,\dfrac{3-\sqrt{3}}{6}\right]$和$\left[\dfrac{3+\sqrt{3}}{6},1\right]$上是增函数,从而原函数$f^{\alpha,0}(x)$不是[0,1]区间上的严凹函数,其一阶导数函数的图像如图4所示.

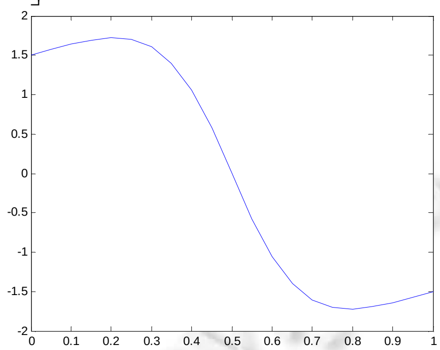

Fig.3 Image of $f'^{\alpha,\beta}(x)$ on [0,1] for α=3, β=0.01

图3 α=3,β=0.01时,$f'^{\alpha,\beta}(x)$在区间[0,1]上的图像

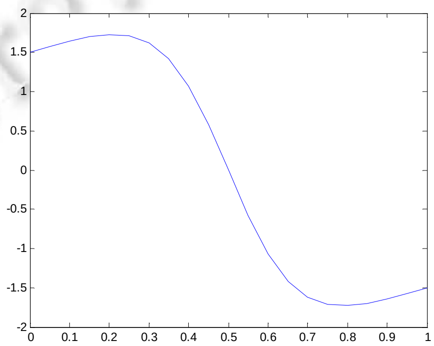

Fig.4 Image of $f'^{\alpha,\beta}(x)$ on [0,1] for α=3, β=0

图4 α=3,β=0时$f'^{\alpha,\beta}(x)$在区间[0,1]上的图像

例6(例5续):取α=3,β=0,那么$e^{3,0}(\tilde{X}_{P_1})=\dfrac{1}{2}f^{3,0.01}(0.2)=0.163482$,$e^{3,0}(\tilde{X}_{P_2})=f^{3,0}(0.1)=0.157355$,虽有$U/P_1\prec U/P_2$且$|BND_{P_1}(X)|<|BND_{P_2}(X)|$,但$e^{3,0}(\tilde{X}_{P_1})>e^{3,0}(\tilde{X}_{P_2})$,违背了定义9和文献[27]的定义2.5(RP2).

根据以上分析,文献[27]中能够用于刻画粗糙集不确定性度量的模糊熵函数都可归结为满足本文定义11的非负严凹函数,其中,文献[27]中的e_{03}^{k},e_{04},e_{05}分别对应着本文的$f_3,f_4,f_5,e_{c_1}^{d_E},e_{c_1}^{d_L},e_{c_6}^{d_E},e_{c_6}^{d_L},e_{nf_4}^{d_E},e_{nf_4}^{d_L}$分别对应着本文的$e_{c_1}^{d_E},e_{c_1}^{d_L},e_{c_2}^{d_E},e_{c_2}^{d_L},e_{F}^{d_E},e_{F}^{d_L}$.由它们导出的公式都是定义9下的不确定性度量.同时发现,文献[27]中的$e_{02}^{\alpha,\beta}$在语义上与粗糙集不确定性度量有时不一致.

4 粗糙度、改进粗糙度和基于严凹函数的不确定性度量之间的关系

由于现有粗糙度的不足,国内外学者纷纷对它进行改进.这节主要讨论粗糙度、改进粗糙度和基于严凹函数不确定性度量之间的关系.

由于粗糙度只依赖于正区域基数在上近似集基数的比例,缺乏对负域信息变化的刻画能力.在例1中,虽然U/P_2没有分离出正区域中的知识颗粒,但我们并不是什么也不知道,仍然得到部分信息,即$\{x_2,x_4\}$一定不在X_1中.Yao提出了如下改进的近似精度和粗糙度公式[39]:

$$\alpha'_P(X)=\frac{|POS_P(X)|+|NEG_P(X)|}{|U|} \tag{22}$$

$$\rho'_P(X)=1-\alpha'_P(X)=\frac{|BND_P(X)|}{|U|} \tag{23}$$

定理 4. 给定 $IS=\langle U,V,f,A\rangle$,假设 $P,Q\subseteq A,X\subseteq U$,如果 $U/P\prec U/Q$,那么,

(1) $\underline{apr}_Q(X)\subseteq\underline{apr}_P(X)\subseteq X\subseteq\overline{apr}_P(X)\subseteq\overline{apr}_Q(X)$;

(2) $POS_Q(X)\subseteq POS_P(X),BND_P(X)\subseteq BND_Q(X)$.

证明:根据上、下近似集以及正域和边界域的定义,即知结论成立. □

定理 5. $\rho'_P(X)$ 是定义 9 下的不确定性度量.

证明:显然,$\rho'_P(X)$ 满足定义 9 的条件(1)和条件(2),下面仅证明单调性.

若 $U/P\prec U/Q$,根据定理 4,易知 $\rho'_P(X)\leqslant\rho'_Q(X)$,如果$|BND_P(X)|<|BND_Q(X)|$,则显然有 $\rho'_P(X)<\rho'_Q(X)$. □

对于$\rho_P(X)$, $\rho'_P(X)$ 和 $ER_f(X|P)$,它们具有如下关系成立.

定理 6. 假设 $U/P\prec U/Q,X\subseteq U$:(1) 如果$\rho_P(X)<\rho_Q(X)$,那么 $\rho'_P(X)<\rho'_Q(X)$;(2) 如果 $\rho'_P(X)<\rho'_Q(X)$,那么,

$$ER_f(X|P)<ER_f(X|Q).$$

证明:

(1) 根据定理 4,由 $U/P\prec U/Q$,易知 $BND_P(X)\subseteq BND_Q(X)$.假设 $\rho'_P(X)<\rho'_Q(X)$ 不成立,那么有$|BND_P(X)|=|BND_Q(X)|$;又因$\rho_P(X)<\rho_Q(X)$,从而$|\overline{apr}_P(X)|>|\overline{apr}_Q(X)|$.与定理 4 矛盾,从而结论(1)成立.

(2) 根据定义 9、公式(23)和定理 3,即有结论成立. □

推论 3. 假设 $U/P\prec U/Q,X\subseteq U$:(1) 若 $ER_f(X|P)=ER_f(X|Q)$,则 $\rho'_P(X)=\rho'_Q(X)$;(2) 若 $\rho'_P(X)=\rho'_Q(X)$,则

$$\rho_P(X)=\rho_Q(X).$$

证明:根据题设和定理 6 即知结论成立. □

根据定理 6 和推论 3,上述 3 种不确定性度量公式中,粗糙度$\rho_P(X)$对知识粒度的划分反应最不敏感,而$ER_f(X|P)$对划分最敏感.

定理 7. 假设 $U/P\prec U/Q,X\subseteq U$.

(1) 如果 $BND_P(X)/P\approx BND_Q(X)/Q$,那么 $ER_f(X|P)=ER_f(X|Q)$,从而$\rho_P(X)=\rho_Q(X)$, $\rho'_P(X)=\rho'_Q(X)$;

(2) 如果$|BND_P(X)|=|BND_Q(X)|$,那么 $\rho'_P(X)=\rho'_Q(X)$,从而$\rho_P(X)=\rho_Q(X)$;

(3) 如果$|BND_P(X)|<|BND_Q(X)|$,那么 $\rho'_P(X)<\rho'_Q(X)$,$ER_f(X|P)<ER_f(X|Q)$;

(4) 若$|BND_P(X)|<|BND_Q(X)|$且$|NEG_P(X)|=|NEG_Q(X)|$,则$\rho_P(X)<\rho_Q(X)$,则 $\rho'_P(X)<\rho'_Q(X)$,$ER_f(X|P)<ER_f(X|Q)$.

证明:(1) 根据定义 9 的不变性准则和推论 3,即有结论成立;(2) 根据公式(23)和推论 3,即知结论成立;(3) 根据公式(23)和定义 9,即知结论成立;(4) 根据粗糙度定义和定理 5,即知结论成立. □

定理 7 的结论(1)表明,如果细分前后两边界域是同分布同构的,上述 3 个公式是一致的,它们在细分前后都保持值不变;结论(2)表明,如果只是保持边界域基数不变,从而正区域和负区域的基数都保持不变,此时 $\rho'_P(X)$ 和$\rho_P(X)$保持不变,而 $ER_f(X|P)$可能不变,也有可能变小,具体反例见例 7;结论(3)表明,如果细化过程中边界域分离出确定区域(正域或负域)中的知识颗粒,此时 $\rho'_P(X)$ 和 $ER_f(X|P)$是一致的,它们都会随划分严格变小,而当边界域分离出负域中的知识颗粒时,粗糙度$\rho_P(X)$有可能变小,也有可能保持不变,具体反例见例 1;结论(4)表明,如果划分细化过程中边界域只分离出正域中的知识颗粒,上述 3 个公式也是一致的,它们都会随着划分的变细而严格变小.

例 7:设 $U=\{x_1,x_2,\ldots,x_8\},X=\{x_1,x_2,x_5\},U/P_1=\{\{x_1,x_2,\ldots,x_6\},\{x_7,x_8\}\},U/P_2=\{\{x_1,x_3,x_5\},\{x_2,x_4,x_6\},\{x_7,x_8\}\},U/P_3=\{\{x_1,x_3,x_4,x_5\},\{x_2,x_6\},\{x_7,x_8\}\}$.

显然,$U/P_i\prec U/P_1(i=2,3)$,$|BND_{P_1}(X)|=|BND_{P_2}(X)|=|BND_{P_3}(X)|$.如果取 $f(x)=f_5(x)$,那么 $ER_f(X|P_1)=ER_f(X|P_3)=3/16$,$ER_f(X|P_2)=2/9$,从而 $ER_f(X|P_2)<ER_f(X|P_1)$.其他满足定义 11 的严凹函数也有类似性质,此处略.

综上所述,粗糙度、改进粗糙度和基于严凹函数的不确定性度量具有一些共同的属性,如定义 9 中的非负性,但它们在某些方面又各不相同,表 2 列出了它们随划分变细时的一些异同.

Table 2 Comparison of different uncertainty measures with refinement

表 2 各种不确定性度量方法随划分变细的变化比较

情形	$\rho_P(X)$	$\rho'_P(X)$	$ER_f(X\|P)$	是否一致
$BND_P(X)=\varnothing$	0	0	0	一致
细分前后边界域分布同构	保持不变	保持不变	保持不变	一致
细分前后边界域基数不变	不变	不变	变小或不变	仅$\rho_P(X)$与 $\rho'_P(X)$一致
边界域只分离出负域中颗粒	变小或不变	严格变小	严格变小	仅 $\rho'_P(X)$ 与 $ER_f(X\|P)$一致
边界域只分离出正域中颗粒	严格变小	严格变小	严格变小	一致

根据表 2,我们有如下结论.

(1) 当边界域为空或者细分前后边界域分布同构或者边界域只分离出正域中的知识颗粒时,上述 3 种度量方法是一致的.

(2) 若细分前后只保留边界域基数不变,此时只有 $\rho'_P(X)$ 与$\rho_P(X)$保持一致,而 $ER_f(X|P)$可能与其不一致.

(3) 如果细分前后边界域只分离出负域中的颗粒,此时 $\rho'_P(X)$ 与 $ER_f(X|P)$保持一致,而$\rho_P(X)$与其不一致.

(4) 当知识划分变细时,如果 $ER_f(X|P)$不变,则其他两个也一定不变;如果$\rho_P(X)$严格变小,则其他两个一定严格变小.

5 数值算例

本节再设计一个计算实例来比较上述3种粗糙集不确定性度量的性能和差异.考虑到上述3种度量方法的语义分歧主要发生在划分变细过程,因此本算例主要针对划分变细展开讨论.

我们知道,知识空间的划分变细包括正域变细、边界域变细和负域细分这 3 类情形.当然,在具体计算实例中,有可能有上述 3 种细分的若干种同时发生.为了看清楚各类细分对粗糙集不确定性方法的影响,我们特意将其设计成每种情况只有一类细分情形发生,即只有正域细分或负域细分或边界域细分.在此基础上衍生出 7 种情况:(1) 只有正域中知识颗粒的细分;(2) 只有负域中知识颗粒的细分;(3) 边界域细分但只分离出正域中的知识颗粒;(4) 边界域为全集时发生细分,但只分离出负域中的知识颗粒;(5) 边界域不等于全集时发生细分,但只分离出负域中的知识颗粒;(6) 边界域细分,但边界域不变且细分前后条件概率不变;(7) 边界域细分,但边界域不变且细分前后条件概率发生改变.

例 8:设 $U=\{x_1,x_2,\ldots,x_{10}\},X_1=\{x_1,x_2,x_5,x_7,x_8\},X_2=\{x_1,x_2,x_5,x_7,x_9\},U/P_1=\{\{x_1,x_2,\ldots,x_6\},\{x_7,x_8\},\{x_9,x_{10}\}\},U/P_2=\{\{x_1,x_2,\ldots,x_6\},\{x_7\},\{x_8\},\{x_9,x_{10}\}\},U/P_3=\{\{x_1,x_2,\ldots,x_6\},\{x_7,x_8\},\{x_9\},\{x_{10}\}\},U/P_4=\{\{x_1,x_2\},\{x_3,\ldots,x_6\},\{x_7,x_8\},\{x_9,x_{10}\}\},U/P_5=\{\{x_1,x_2,x_3,x_5,x_6\},\{x_4\},\{x_7,x_8\},\{x_9,x_{10}\}\},U/P_6=\{\{x_1,x_3,x_4,x_5\},\{x_2,x_6\},\{x_7,x_8\},\{x_9,x_{10}\}\},U/P_7=\{\{x_1,x_3,x_5\},\{x_2,x_4,x_6\},\{x_7,x_8\},\{x_9,x_{10}\}\}$.

显然,$U/P_i\prec U/P_1(i=2,\ldots,6)$.本文除了粗糙度$\rho_P(X)$用到$X_2$外,其余度量都是针对$X_1$展开计算的,它们的计算结果见表 3(其中,严凹函数选 f_1,f_6,模糊熵选 $e^{D_L}_{c_1}(\tilde{X})$ 作代表).

Table 3 Computation results of different uncertainty measures

表 3 不同不确定性度量所得计算结果

	P_1	P_2	P_3	P_4	P_5	P_6	P_7
$\rho_{P_i}(X_1)$	0.75	0.75	0.75	0.50	0.714	0.75	0.75
$\rho'_{P_i}(X_1)$	0.60	0.60	0.60	0.40	0.50	0.60	0.60
$E_{P_i}(X_1)$	1.463	1.313	1.313	0.70	1.115	1.05	1.013
$ER_{f_1}(X_1\|P_i)$	0.30	0.30	0.30	0.20	0.221	0.30	0.275
$ER_{f_6}(X_1\|P_i)$	0.05	0.05	0.05	0.03	0.0375	0.05	0.045
$e^{D_L}_{c_1}(\tilde{X}_{1_{P_i}})$	0.30	0.30	0.30	0.151	0.240	0.30	0.268
$\rho_{P_i}(X_2)$	1.00	0.89	0.89	0.80	1.00	1.00	1.00

表 3 中的计算结果有如下几个特点.

(1) 只要知识划分变细,$E_P(X_1)$就会严格变小.

(2) 当划分变细导致正域或负域中的知识颗粒细分时,除了 $E_P(X_1)$变小外,其他方法都保持不变,U/P_2 和 U/P_3 即是这种情形.

(3) 当边界域分离出正域中的知识颗粒时,上述方法的不确定性值都会严格变小,即使是粗糙度$\rho_P(X)$,在这种情况下也会严格变小,U/P_4 即是这种情形.

(4) 当边界域分离出负域中的知识颗粒时,除了粗糙度$\rho_P(X)$外,其他不确定性度量都会严格变小,U/P_5 即是这种情形.若细分前边界域为整个论域,细分前后粗糙度保持不变;若细分前边界域不等于整个论域时,细分后粗糙度会变小.例如在例 8 中,对于集合 X_1, $BND_{P_1}(X_1)\neq U$, U/P_1 细分成 U/P_5 后, $\rho_{P_5}(X_1)<\rho_{P_1}(X_1)$; 而对于集合 X_2,此时 $BND_{P_1}(X_2)=U$, 同样是 U/P_1 细分后变成 U/P_5,但 $\rho_{P_5}(X_2)=\rho_{P_1}(X_2)$.

(5) 当边界域大小保持不变,同时细分前后条件概率保持不变,除了 $E_P(X_1)$变小外,其余方法所得结果在细分前后都保持不变,U/P_6 即是这样.

(6) 当边界域细分,边界域大小不变且细分前后条件概率有改变,粗糙度和改进粗糙度保持不变,但基于严凹函数的不确定性度量方法所得结果会严格变小,U/P_7 就是这种情形.

上述计算结果表明,

(1) 粗糙度之所以不能较好地刻画出粗糙集不确定性的变化情况,问题出在情形(4),即当细分前边界域为整个论域且边界域分离出负域中的知识颗粒时,细分前后的粗糙度保持不变.这是它不合理之处.

(2) 与粗糙度和改进粗糙度相比,基于严凹函数的不确定性度量方法能更加敏感地反映出划分细分的信息变化,如果细分前后基于严凹函数的不确定性度量值不变,那么粗糙度和改进粗糙度不变.

(3) 基于严凹函数的不确定性度量方法,无论是基于香农熵函数还是基于模糊熵函数的度量公式,都能较好地刻画出粗糙集的不确定性变化信息,语义上也与粗糙集的不确定性保持一致.

6 结束语

本文结合粗糙集不确定性度量的语义分析,给出一种拓展的粗糙集不确定性度量公理化定义,将香农熵函数推广到一般严凹函数,提出一类以条件概率为自变量,基于严凹函数的粗糙集不确定性度量方法,研究结果表明,基于严凹函数的不确定性度量方法不但能够有效地刻画粗糙集不确定性度量问题,而且它能够将现有的基于香农熵和模糊熵的不确定性度量方法统一到本文提出的理论框架下,这将有助于我们看清粗糙集不确定性度量问题的数学本质,同时为不确定性度量研究提供新思路和新视角.然后,我们研究了粗糙度、改进粗糙度和本文提出的基于严凹函数的不确定性度量之间的关系,发现粗糙度对划分变细反应最不敏感,其次是改进粗糙度,基于严凹函数的不确定性度量能够较好地反映出划分变细的信息变化情况.而 Beaubouef 提出的粗糙熵方法则反应过度,即只要划分细分,就会导致粗糙熵严格变小,即使与待刻画集合完全无关的负域的细分,也会导致计算结果严格变小,这是它不合理的地方.最后设计了一组计算实例,发现了粗糙度不能有效刻画粗糙集不确定性问题的根源(即当细分前边界域为整个论域且边界域分离出负域中的知识颗粒时,细分前后粗糙度保持不变),验证了基于严凹函数的不确定性度量方法与粗糙集不确定性在语义上的一致性.下一步我们将研究粗糙集不确定性度量与属性约简、信息粒度之间的理论联系,为基于不确定性度量研究决策信息系统高效算法提供理论基础.同时,基于严凹函数研究更复杂信息系统,如不完备系统、集值信息系统和区间值决策系统等的不确定性度量问题.

致谢 感谢加拿大里贾纳大学(University of Regina)姚一豫(Yiyu Yao)教授的鼓励和帮助.感谢审稿专家的改进意见.

References:

[1] Sun L, Xu JC, Tian Y. Feature selection using rough entropy-based uncertainty measures in incomplete decision systems. Knowledge-Based Systems, 2012,36(1):206−216.

[2] Liang JY, Qian YH. Information granules and entropy theory in information systems. Science in China Series F: Information Sciences, 2008,51(10):1427−1444.

[3] Deng XF, Yao YY. A multifaceted analysis of probabilistic three-way decisions. Fundamenta Informaticae, 2014,132(3):291−313.

[4] Pawlak Z. Rough sets. Int'l Journal of Computer & Information Sciences, 1982,11(5):341−356.

[5] Beaubouef T, Petry FE, Arora G. Information-Theoretic measures of uncertainty for rough sets and rough relational databases. Information Sciences, 1998,109(1-4):185−195.

[6] Liang JY, Wang JH, Qian YH. A new measure of uncertainty based on knowledge granulation for rough sets. Information Sciences, 2009,179:458−470.

[7] Xu BW, Zhou YM, Lu HM. An improved accuracy measure for rough sets. Journal of Computer and System Sciences, 2005,71: 163−173.

[8] Wang XY, Cai N, Yang J, Liu XJ. A new method for measuring uncertainty in rough sets. Journal of Shanghai Jiaotong University, 2006,40(7):1130−1134 (in Chinese with English abstract).

[9] Teng SH, Lu M, Yang AF, Zhang J, Zhuang ZW. A weighted uncertainty measure of rough sets based on general binary relation. Chinese Journal of Computers, 2014,37(3):649−665 (in Chinese with English abstract).

[10] Düntsch I, Gediga G. Uncertainty measures of rough set prediction. Artificial intelligence, 1998,106(1):109−137.

[11] Wierman MJ. Measuring uncertainty in rough set theory. Int'l Journal of General System, 1999,28(4-5):283−297.

[12] Miao DQ, Wang J. On the relationships between information entropy and roughness of knowledge in rough set. Pattern Recognition and Artificial Intelligence, 1998,11(3):34−40 (in Chinese with English abstract).

[13] Wang GY, Yu H, Yang DH. Decision table reduction based on conditional information entropy. Chinese Journal of Computers, 2002,25(7):759−766 (in Chinese with English abstract).

[14] Li J, Shi KQ. Uncertainty measurement of rough sets based on conditional entropy. Systems Engineering and Electronics, 2008, 30(3):473−476 (in Chinese with English abstract).

[15] Wei W, Wei Q, Wang F. Comparative study of uncertainty measure in rough set. Journal of Nanjing University (Natural Sciences), 2015,51(4):714−722 (in Chinese with English abstract).

[16] Liang JY, Shi ZZ. The information entropy, rough entropy and knowledge granulation in rough set theory. Int'l Journal of Uncertainty, Fuzziness and Knowledge-Based Systems, 2004,19(1):37−46.

[17] Liang JY, Shi ZZ, Wierman MJ. Information entropy, rough entropy and knowledge granulation in incomplete information systems. Int'l Journal of General Systems, 2006,35(6):641−654.

[18] Bianucci D, Cattaneo G, Ciucci D. Entropies and co-entropies of coverings with application to incomplete information systems. Fundamenta Informatiace, 2007,75(1-4):77−105.

[19] Zhu P, Wen QY. Entropy and co-entropy of a covering approximation space. Int'l Journal of Approximate Reasoning, 2012,53(4): 528−540.

[20] Qian YH, Liang JY. Combination entropy and combination granulation in rough set theory. Int'l Journal of Uncertainty, Fuzziness and Knowledge-Based Systems, 2008,16(2):179−193.

[21] Dai JH, Wang WT, Xu Q, *et al*. Uncertainty measurement for interval-valued decision systems based on extended conditional entropy. Knowledge-Based Systems, 2012,27(1):443−450.

[22] Dai JH, Tian HW. Entropy measures and granularity measures for set-valued information systems. Information Sciences, 2013, 240(1):72−82.

[23] Chakrabary K, Biswas R, Nanda S. Fuzziness in rough sets. Fuzzy Sets and Systems, 2000,110(2):247−25.

[24] Liang JY, Dang CY, Chin KS, *et al*. A new method for measuring uncertainty and fuzziness in rough set theory. Int'l Journal of General Systems, 2002,31(4):331−342.

[25] Wang GY, Zhang QH. Uncertainty of rough sets in different knowledge granularities. Chinese Journal of Computers, 2008,31(9): 1588−1598 (in Chinese with English abstract).

[26] De Luca A, Termini S. A definition of a non-probabilistic entropy in the setting of fuzzy sets theory. Information and Control, 1972, 20(4):301−312.

[27] Wei W, Liang JY, Qian YH, *et al*. Can fuzzy entropies be effective measures for evaluating the roughness of a rough set? Information Sciences, 2013,232:143−166.

[28] Hu J, Wang GY. Uncertainty measure rule sets on rough sets. Pattern Recognition and Artificial Intelligence, 2010,23(5):606−615 (in Chinese with English abstract).

[29] Huang GS, Zeng FZ, Wen H. Uncertainty measures of rough set based on conditional possibility. Control and Decision, 2015,30(6): 1099−1105 (in Chinese with English abstract).

[30] Huang GS, Zeng FZ, Chen GY, Wen H. Knowledge granularity and relative granularity based on strictly convex function. Pattern Recognition and Artificial Intelligence, 2013,26(10):897−908 (in Chinese with English abstract).

[31] Yao YY. The superiority of three-way decisions in probabilistic rough set models. Information Sciences, 2011,181(6):1080−1096.

[32] Kuang JC. Applied Inequalities. 4th ed., Ji'nan: Shandong Science and Technology Press, 2010 (in Chinese).

[33] Pal NR, Pal SK. Entropy: A new definition and its applications. IEEE Trans. on Systems, Man and Cybernetics, 1991,21(5): 1260−1270.

[34] Bhandari D, Pal NR. Some new information measures for fuzzy sets. Information Sciences, 1993,67(3):209−228.

[35] Fan J, Xie W. Distance measure and induced fuzzy entropy. Fuzzy Sets and Systems, 1999,104(2):305−314.

[36] Liu XC. Entropy, distance measure and similarity measure of fuzzy sets and their relations. Fuzzy Sets and Systems, 1992,52(3): 305−318.

[37] Fan JL, Ma YL. Some new fuzzy entropy formulas. Fuzzy Sets and Systems, 2002,128(2):277−284.

[38] Fan JL, Ma YL, Xie WX. On some properties of distance measures. Fuzzy Sets and Systems, 2001,117(3):355−361.

[39] Yao YY. A note on definability and approximations. In: Peters JF, *et al*., eds. Proc. of the Transactions on Rough Sets VII. Berlin, Heidelberg: Springer-Verlag, 2007. 274−282.

附中文参考文献:

[8] 王向阳,蔡念,杨杰,刘小军.基于近似精度和条件信息熵的粗糙集不确定性度量方法.上海交通大学学报,2006,40(7):1130−1134.

[9] 腾书华,鲁敏,杨阿峰,张军,庄钊文.基于一般二元关系的粗糙集加权不确定性度量.计算机学报,2014,37(3):649−665.

[12] 苗夺谦,王珏.粗糙集理论中知识粗糙性与信息熵关系的讨论.模式识别与人工智能,1998,11(3):34−40.

[13] 王国胤,于洪,杨大春.基于条件信息熵的决策表约简.计算机学报,2002,25(7):759−766.

[14] 李健,史开泉.基于条件粗糙熵的粗集不确定性度量.系统工程与电子技术,2008,30(3):473−476.

[15] 魏巍,魏琪,王锋.粗糙集的不确定性度量比较研究.南京大学学报(自然科学),2015,51(4):714−722.

[25] 王国胤,张清华.不同知识粒度下粗糙集的不确定性研究.计算机学报,2008,31(9):1588−1598.

[28] 胡军,王国胤.粗糙集的不确定性度量准则.模式识别与人工智能,2010,23(5):606−615.

[29] 黄国顺,曾凡智,文翰.基于条件概率的粗糙集不确定性度量.控制与决策,2015,30(6):1099−1105.

[30] 黄国顺,曾凡智,陈广义,文翰.基于严凸函数的知识粒度与相对粒度.模式识别与人工智能,2013,26(10):897−908.

[32] 匡继昌.常用不等式.第4版,济南:山东科学技术出版社,2010.

黄国顺(1972－),男,江西临川人,博士,教授,CCF 高级会员,主要研究领域为粗糙集,粒计算,不确定性度量.

文翰(1977－),男,博士,讲师,主要研究领域为文本挖掘,机器学习.