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 $\tilde{X}=\mu_{\tilde{X}}(x_1)/x_1+\mu_{\tilde{X}}(x_2)/x_2+...+\mu_{\tilde{X}}(x_n)/x_n$, $\tilde{X}^c$ 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, $f_{c_1}^{D_L}(x)=1-(2x-1)\log_2\frac{1+x}{2-x}$, $f_{c_1}^{D_E}(x)=(1-x)\mathrm{e}^{2x-1}+x\mathrm{e}^{1-2x}-\mathrm{e}^{-1}$.

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 $\tilde{X}_P=\mu_P^X(x_1)/x_1+\mu_P^X(x_2)/x_2+...+\mu_P^X(x_n)/x_n$.由于对任意 $x,y\in X_i$ 有 $\mu_P^X(x)=\mu_P^X(y)$,因此:

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

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

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

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

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

当 $x=0$ 时,显然有 $f''^{D_E}_{c_2}(x)<0$,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 $f''^{D_E}_{c_2}(x)<0$.

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

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 $\lambda f^{D_L}_{c_2}(x_1)+(1-\lambda)f^{D_L}_{c_2}(x_2)<f^{D_L}_{c_2}(\lambda x_1+(1-\lambda)x_2)$,从而可知 $f^{D_E}_{c_2}(x)$ 是区间[0,1]上的严凹函数. □

显然, $e^{D_L}_{c_2}(\tilde{X}_P)=ER_{f^{D_L}_{c_2}}(X\mid P),e^{D_E}_{c_2}(\tilde{X}_P)=ER_{f^{DE}_{c_2}}(X\mid P)$,又因 $f^{D_L}_{c_2}(0)=f^{D_L}_{c_2}(1)=0,f^{D_E}_{c_2}(0)=f^{D_E}_{c_2}(1)=0$且$f^{D_E}_{c_2}(x)$在区间[0,1]上非负,根据定理 3 可知,由它们导出的公式 $e^{D_L}_{c_2}(\tilde{X}_P),e^{D_E}_{c_2}(\tilde{X}_P)$ 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]:

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

那么,

$$e^{D_L}_{F}(\tilde{X})=\frac{1}{n}\sum_{i=1}^{n}f^{D_L}_{F}(\mu_{\tilde{X}}(x_i)),e^{D_E}_{F}(\tilde{X})=\frac{1}{n}\sum_{i=1}^{n}f^{D_E}_{F}(\mu_{\tilde{X}}(x_i)),$$

其中, $f^{D_L}_{F}(\mu_{\tilde{X}}(x_i)),f^{D_E}_{F}(\mu_{\tilde{X}}(x_i))$ 分别与 $f^{D_L}_{c_2}(x),f^{D_E}_{c_2}(x)$ 相同,从而知 $e^{D_L}_{F}(\tilde{X}_P),e^{D_E}_{F}(\tilde{X}_P)$ 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,$e^{\alpha,0}(\tilde{X}_P)=\dfrac{1}{n}\sum\limits_{i=1}^{n}f^{\alpha,0}(\mu_{\tilde{X}}(x_i))$,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.

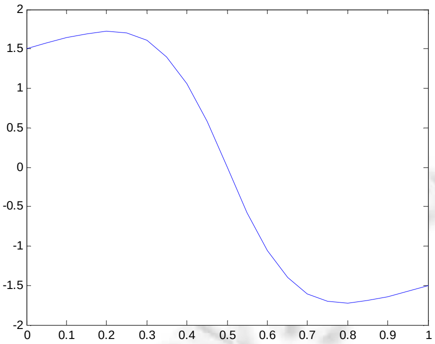

Fig.3 Image of $f'^{\alpha,\beta}(x)$ on [0,1] for α=3, β=0.01

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

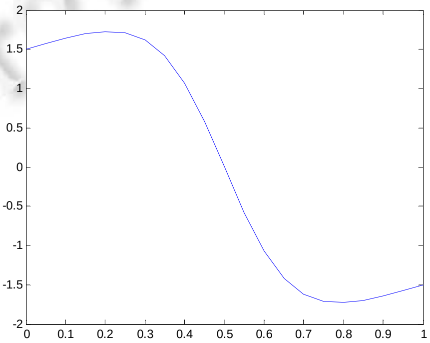

Fig.4 Image of $f'^{\alpha,\beta}(x)$ on [0,1] for α=3, β=0

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

显然,$U/P_i\prec U/P_1(i=2,\dots,6)$.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 $e_{c_1}^{D_L}(\tilde{X})$ 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
$e_{c_1}^{D_L}(\tilde{X}_{1_{P_i}})$	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.