◎理论研究◎

加权粗糙集的矩阵表示

杨 勇[1],鲁小云[1],李 廉[2]
YANG Yong[1],LU Xiao-yun[1],LI Lian[2]

1.西北师范大学 数学与信息科学学院,兰州 730070
2.兰州大学 信息科学与工程学院,兰州 730000
1.College of Mathematics and Information Science,Northwest Normal University,Lanzhou 730070,China
2.College of Information Science and Engineering,Lanzhou University,Lanzhou 730000,China
E-mail:zt-yang@163.com

YANG Yong,LU Xiao-yun,LI Lian.Matrix presentation of weighted rough sets.Computer Engineering and Applications, 2008,44(18):34-35.

Abstract: The concepts of weighted α-lower product and weighted α-upper product between a relation matrix and a Boolean column vector are presented in weighted approximation space.The results that weighted α-lower and α-upper products are weighted lower and upper approximations,respectively,are proved,thus a new matrix way to study weighted rough sets(VPRSs,Pawlak rough sets) is given.Finally,an algorithm to computer weighted lower and upper approximations is established.
Key words: weighted rough sets;weighted lower and upper approximations;weighted α-lower and upper products

摘 要:在加权近似空间中提出了关系矩阵和布尔列向量加权 α 下乘法和加权 α 上乘法的概念。证明了加权 α 下乘法就是加权下近似,加权 α 上乘法就是加权上近似,从而为加权粗糙集(可变精度粗糙集、经典粗糙集)的计算和研究提供了一种新的矩阵方法。最后给出了计算加权上下近似的算法。
关键词:加权粗糙集;加权上下近似;加权 α 上下乘法
DOI:10.3778/j.issn.1002-8331.2008.18.010 **文章编号:**1002-8331(2008)18-0034-02 **文献标识码:**A **中图分类号:**TP181

粗糙集是一种处理不精确和不确定性知识的数学工具,它采用精确的数学方法分析不精确系统的数据,从中发现隐藏的规律。而在实际应用中,数据集中存在噪音等的干扰,粗糙集理论因限制条件过强导致处理数据的能力大为下降。为此文献[1]提出了具有一定容错性的可变精度粗糙集(VPRS)模型,通过引入一个精度,允许分类的不完全精确性,从而拓宽了分类能力,在现实中被成功应用。但在有些情况下,不同对象的重要程度不同,也就是权重不等,用 VPRS 模型处理时其结果与实际有偏差,这就需要寻找一种扩展的 VPRS 数学模型。最近,马廷淮[2]等提出的既考虑精度又考虑权重的加权粗糙集模型,很好地解决了此类问题,同时该模型为经验知识的总结归纳提供了有益手段。

用矩阵的方法研究粗糙集理论,不仅简单,而且提供了一种新的推理思路。对经典粗糙集的矩阵表示,文献[3-5]已作了详细地研究。本文在上述工作的基础上,将矩阵方法用于加权粗糙集模型,通过引入关系矩阵和布尔列向量的加权重量乘法概念,定义加权 α 下乘法和加权 α 上乘法,证明了加权 α 下乘法和加权 α 上乘法分别是加权下近似和加权上近似,最后给出了求加权粗糙集的矩阵算法。

1 粗糙集理论的基本概念

设 U 是非空有限论域,R 是 U 上的二元等价关系,序对 (U,R) 称为近似空间。对任意的 $X\subseteq U$,X 关于近似空间的一对下近似$\underline{R}X$和上近似定义$\overline{R}X$ 如下:

$$\underline{R}X=\{x\in U|[x]_R\subseteq U\},\ \overline{R}X=\{x\in U|[x]_R\cap X\neq\phi\}$$

其中$[x]_R$是 x 所在的 R 等价类。

Pawlak 粗糙集模型的主要局限性在于它所处理的分类必须是完全正确的或肯定的,对知识的模糊性缺乏相应的处理能力。为此 Ziarko 提出了可变精度粗糙集(VPRS)模型,它在经典粗糙集模型的基础上引入了一个错误分类率,这一方面完善了近似空间的概念,另一方面有利于用粗糙集理论从认为不相关的数据中发现隐藏的事实。

设 $X,Y\subseteq U$,X 关于 Y 的错误分类率 $c(X,Y)$定义为:

基金项目:国家自然科学基金(the National Natural Science Foundation of China under Grant No.10671158);甘肃省教育厅科研基金(No.0701-16)。
作者简介:杨勇(1967-),男,副教授,博士,主要研究方向:粗糙集理论及应用;鲁小云(1982-),男,硕士研究生,主要研究方向:粗糙集理论;李廉(1952-),男,博士生导师,主要研究方向:代数逻辑。
收稿日期:2007-12-24 **修回日期:**2008-02-27

$$c(X,Y)=\begin{cases}1-\dfrac{|X\cap Y|}{|X|} & |X|\neq 0\\ 0 & |X|=0\end{cases}$$

表示如果把 X 中的元素归类于 Y，则有 $c(X,Y)\times 100\%$ 的元素会归类错误。显然，当 $c(X,Y)=0$ 时，$X\subseteq Y$。

令 $0\leqslant\alpha<0.5$（本文中的 α 都满足这个条件，以后再不单独说明），α 包含关系定义为：$X\overset{\alpha}{\subseteq}Y\Leftrightarrow c(X,Y)\leqslant\alpha$。

设 (U,R) 为近似空间，对任意的 $X\subseteq U$，X 关于近似空间的一对 α 下近似 $\underline{R}_\alpha X$ 和 α 上近似 $\overline{R}_\alpha X$ 定义如下：

$$\underline{R}_\alpha X=\{x\in U|c([x]_R,X)\leqslant\alpha\},\ \overline{R}_\alpha X=\{x\in U|c([x]_R,X)<1-\alpha\}$$

当 $\alpha=0$ 时，α 包含关系退化为经典包含关系，此时 α 上下近似就退化为 Pawlak 粗糙集的上下近似。

在近似空间 (U,R) 中，若对每个元素 $u_i\in U$ 赋予权重 ω_i，则称 (U,R,ω) 为加权近似空间。对任意的 $X\subseteq U$，称 $\mu_X^R(x)=\dfrac{\sum\omega_i,u_i\in X\cap[x]_R}{\sum\omega_i,u_i\in[x]_R}$，$x\in U$ 为 x 相对于 X 在加权近似空间的粗糙隶属度。称集合 $\underline{R}_\alpha^\omega X=\{x\in U|\mu_X^R(x)\geqslant 1-\alpha\}$，$\overline{R}_\alpha^\omega X=\{x\in U|\mu_X^R(x)>\alpha\}$ 分别为 X 关于加权近似空间的加权 α 下近似和加权 α 上近似。

显然，在近似空间 (U,R,ω) 中，若每个对象的权重 $\omega_i=1$，则加权近似空间退化为经典近似空间，加权 α 上下近似就退化为可变精度的 α 上下近似。若 $\omega_i=1$ 且 $\alpha=0$，则加权 α 上下近似就退化为经典粗糙集的上下近似。因此，加权粗糙集是 VPRS 和经典粗糙集的推广，它既继承二者的性质，又增强了数据处理能力，从而在实践中被广泛应用。

2　加权粗糙集上下近似的矩阵表示

设 (U,R,ω) 为近似空间，其中 $U=\{u_1,u_2,\cdots,u_n\}$，则等价关系 R 对应一个 n 阶关系矩阵 $\boldsymbol{M}_R=(m_{ij})_{n\times n}$，$\boldsymbol{M}_R$ 的第 i 行对应的行向量为 $(m_{i1},m_{i2},\cdots,m_{in})$，其中，$m_{ij}=1_{\omega_j}$ 或 0，若 $u_j\in[u_i]_R$，则 $m_{ij}=1_{\omega_j}$，否则 $m_{ij}=0$。记 $Sum(i)=\sum\limits_{m_{ij}=1}\omega_j$，称 $Sum(i)$ 为 $\boldsymbol{M}_R$ 为第 i 行的权重重量。对于任意的 $X\subseteq U$，利用集合特征函数的定义，把 $\boldsymbol{X}$ 可表示成 n 维布尔列向量 $(x_1,x_2,\cdots,x_n)^{\mathrm{T}}$，其中 $x_i=1_{\omega_i}$ 或 0，若 $u_i\in X$，则 $x_i=1_{\omega_i}$，否则 $x_i=0$。

设 $\boldsymbol{M}_R=(m_{ij})_{n\times n}$ 为 n 阶关系矩阵，$\boldsymbol{X}$ 为 n 维布尔列向量，称

$$\boldsymbol{M}_R^\omega\otimes\boldsymbol{X}=(a_1,a_2,\cdots,a_n)^{\mathrm{T}}，其中：a_i=\begin{cases}0 & 若\sum\limits_{j=1}^{n}(m_{ij}\wedge x_j)=0\\ 1_{Ind(i)} & 否则\end{cases},$$

$Ind(i)=\sum\limits_{m_{ij}\wedge x_j=1}\omega_j$ 为 $\boldsymbol{M}_R$ 和 $\boldsymbol{X}$ 的加权重量乘法。

由定义可知，加权重量乘法本质上还是关系矩阵和布尔列向量的取小取大运算，只不过为元素 1 加上满足一定条件的元素权重和的下标。显然，$0<Ind(i)\leqslant Sum(i)$。

例 1　设 (U,R,ω) 为近似空间，其中 $U=\{1_3,2_2,3_2,4_7,5_1,6_1,7_5\}$，$X=\{1,4,5,6\}$，$U/R=\{\{1\},\{2,4\},\{3,5,6\},\{7\}\}$，则：$\boldsymbol{M}_R=$

$$\begin{pmatrix}1_3&0&0&0&0&0&0\\0&1_2&0&1_7&0&0&0\\0&0&1_2&0&1_1&1_1&0\\0&1_2&0&1_7&0&0&0\\0&0&1_2&0&1_1&1_1&0\\0&0&1_2&0&1_1&1_1&0\\0&0&0&0&0&0&1_5\end{pmatrix},\boldsymbol{X}=\begin{pmatrix}1_3\\0\\0\\1_7\\1_1\\1_1\\0\end{pmatrix}，于是\ \boldsymbol{M}_R^\omega\otimes\boldsymbol{X}=\begin{pmatrix}1_3\\1_7\\1_2\\1_7\\1_2\\1_2\\0\end{pmatrix}。$$

在本例中，$Sum(1)=3$，$Sum(2,4)=9$，$Sum(3,5,6)=4$，$Sum(7)=5$，而 $Ind(1)=3$，$Ind(2,4)=7$，$Ind(3,5,6)=2$。

下面定义 $\boldsymbol{M}_R$ 和 $\boldsymbol{X}$ 的加权 α 下乘法 $\boldsymbol{M}_{\underline{R}_\alpha}^\omega\otimes\boldsymbol{X}$ 和加权 α 上乘法 $\boldsymbol{M}_{\overline{R}_\alpha}^\omega\otimes\boldsymbol{X}$：

$$\boldsymbol{M}_{\underline{R}_\alpha}^\omega\otimes\boldsymbol{X}=(b_1,b_2,\cdots,b_n)^{\mathrm{T}}，其中\ b_i=\begin{cases}1，若\dfrac{Ind(i)}{Sum(i)}\geqslant 1-\alpha\\ 0，否则\end{cases};$$

$$\boldsymbol{M}_{\overline{R}_\alpha}^\omega\otimes\boldsymbol{X}=(c_1,c_2,\cdots,c_n)^{\mathrm{T}}，其中\ c_i=\begin{cases}1，若\dfrac{Ind(i)}{Sum(i)}>\alpha\\ 0，否则\end{cases}。$$

例 2　在上例中，取 $\alpha=0.4$，由 $\dfrac{Ind(1)}{Sum(1)}=1$，$\dfrac{Ind(2,4)}{Sum(2,4)}=\dfrac{7}{9}$，$\dfrac{Ind(3,5,6)}{Sum(3,5,6)}=\dfrac{1}{2}$ 可得：$\boldsymbol{M}_{\underline{R}_{0.4}}^\omega\otimes\boldsymbol{X}=(1,1,0,1,0,0,0)^{\mathrm{T}}$，$\boldsymbol{M}_{\overline{R}_{0.4}}^\omega\otimes\boldsymbol{X}=(1,1,1,1,1,1,0)^{\mathrm{T}}$。

例 3　在例 1 中，令 $\omega_i=1$，则加权近似空间 (U,R,ω) 退化为近似空间 (U,R)，于是有：$\boldsymbol{M}_R^1\otimes\boldsymbol{X}=(1_1,1_1,1_2,1_1,1_2,1_2,0)^{\mathrm{T}}$。

取 $\alpha=0.4$ 时：

$\boldsymbol{M}_{\underline{R}_{0.4}}^1\otimes\boldsymbol{X}=(1,0,1,0,1,1,0)^{\mathrm{T}}$，$\boldsymbol{M}_{\overline{R}_{0.4}}^1\otimes\boldsymbol{X}=(1,1,1,1,1,1,0)^{\mathrm{T}}$

取 $\alpha=0$ 时：

$\boldsymbol{M}_{\underline{R}_0}^1\otimes\boldsymbol{X}=(1,0,0,0,0,0,0)^{\mathrm{T}}$，$\boldsymbol{M}_{\overline{R}_0}^1\otimes\boldsymbol{X}=(1,1,1,1,1,1,0)^{\mathrm{T}}$

定理　设 (U,R,ω) 为加权近似空间，则对任意的 $X\subseteq U$，有 (1) $\underline{R}_\alpha^\omega X=\boldsymbol{M}_{\underline{R}_\alpha}^\omega\otimes\boldsymbol{X}$；(2) $\overline{R}_\alpha^\omega X=\boldsymbol{M}_{\overline{R}_\alpha}^\omega\otimes\boldsymbol{X}$。

证明　(1) 设 $u_i\in\underline{R}_\alpha^\omega X$，则 $\mu_X^R(u_i)=\dfrac{\sum\omega_j,u_j\in X\cap[u_i]_R}{\sum\omega_j,u_j\in[u_i]_R}\geqslant 1-\alpha$，于是 $\sum\omega_j\neq 0,u_j\in X\cap[u_i]_R$，有 $\boldsymbol{M}_R$ 和 $\boldsymbol{X}$ 的加权重量乘法的第 i 个分量为 1，且其下标为 $Ind(i)=\sum\limits_{m_{ij}\wedge x_j=1}\omega_j=\sum\limits_{u_j\in X\cap[u_i]_R}\omega_j$，而 $Sum(i)=\sum\limits_{m_{ij}=1}\omega_j=\sum\limits_{u_j\in[u_i]_R}\omega_j$ 这样 $\dfrac{Ind(i)}{Sum(i)}\geqslant 1-\alpha$，因此，$\boldsymbol{M}_{\underline{R}_\alpha}^\omega\otimes\boldsymbol{X}$ 的第 i 个分量为 1。反之，若 $\boldsymbol{M}_{\underline{R}_\alpha}^\omega\otimes\boldsymbol{X}$ 的第 i 个分量为 1，重复上述过程，可得 $u_i\in\underline{R}_\alpha^\omega X$。

(2) 同理可证。

例 4　根据上述定理，例 2 中 X 的加权 0.4 上下近似为：$\overline{R}_{0.4}^\omega X=\{1,2,3,4,5,6\}$，$\underline{R}_{0.4}^\omega X=\{1,2,4\}$。例 3 中加权近似空间退化为经典近似空间，加权粗糙集退化为可变精度粗糙集，X 的 0.4 上下近似为：$\overline{R}_{0.4}X=\{1,2,3,4,5,6\}$，$\underline{R}_{0.4}X=\{1,3,5,6\}$；取 $\alpha=$

（下转 82 页）