一种基于关系矩阵的决策表正域约简算法

景运革[1,2] 李天瑞[1]

（山西运城学院公共计算机教学部 运城 044000）[1] （西南交通大学信息科学与技术学院 成都 610031）[2]

摘 要 研究了粗糙集属性约简问题，引入等价关系矩阵的诱导矩阵和矩阵的λ-截矩阵等概念来计算决策表的上、下近似集，进而给出基于关系矩阵的决策表正域求解方法，并从理论上证明了该方法的正确性。提出了粗糙集属性核的启发式约简，并用该方法计算最小约简，在属性动态增加时，用矩阵快速更新的方法来改变属性等价关系矩阵，可以快速地计算属性变化后的正域。最后，通过实例分析说明了属性约简的具体操作方法和算法的有效可行性。

关键词 粗糙集，决策表，正域约简，关系矩阵

中图法分类号 TP311 **文献标识码** A

Reduction Algorithm of Positive Domain for Decision Table Based on Relationship Matrix

JING Yun-ge[1,2] LI Tian-rui[1]

(Department of Public Computer Teaching, Yuncheng College, Yuncheng 044000, China)[1]

(School of Information Science and Technology, Southwest Jiaotong University, Chengdu 610031, China)[2]

Abstract This paper discussed the problem of attributes reduction in rough set. We first introduced both induced matrix and λ-cut matrix of equivalence relation matrix to calculate upper and lower approximation of decision tables and then proposed a reduction algorithm of positive domain for decision table based on relationship matrix with the prove of correctness in theory. What's more, a heuristic reduction of attribute core in rough set was proposed to calculate the minimum reduction. With dynamical updating of attributes, we updated attribute equivalence relationship matrix through the method of update of matrix and calculated the positive domain after update of attribute rapidly. The example confirms the feasibility and effectiveness of proposed operation and method of attribute reduction.

Keywords Rough set, Decision tables, Positive domain reduction, Relation matrix

1 引言

粗糙集理论是由波兰学者 Pawlak 于 20 世纪 80 年代初提出的一种新的处理不确定、不精确、不完备知识和信息的数学工具[1]。近年来，粗糙集理论受到了许多研究者的关注，并已成功应用在决策分析、模糊识别、知识发现以及数据挖掘等领域。属性约简[2]是粗糙集理论中重要的研究内容，是在保持属性集的分类或决策能力不变的前提下，在满足决策信息系统决策能力不变的条件下，进行条件属性选择，用产生的新条件属性代替原有条件属性，删除其中不相关或不重要的属性的过程。求解粗糙集的最小属性约简是 NP 问题，主要原因是属性的组合爆炸。所以，减小属性约简的搜索空间，提高约简的效率，对于粗糙集的属性约简具有重要的意义。一个决策信息系统可能存在多个属性约简集合，相对于决策属性集合和条件属性集合的所有约简的交集称为属性核，核中的属性是约简的极限粒度。用核作为计算约简集的起点，可以简化计算约简集，提高属性约简的效率。

矩阵理论是一门具有实用价值的数学理论，它是处理大量有限维空间方式与数量关系的强有力的工具。在经典粗糙集理论中，已给出了许多基于决策表的属性约简算法，如 Skowron[3] 运用区分矩阵研究信息系统和决策表的属性约简问题，Guan 等[4]将信息系统中各属性所形成的不可分辨关系及其性质以及等价关系矩阵的形式进行了系统深入研究，但这些算法只是把矩阵局限于一种表述方式，并没有实质性地参与到计算之中。Yang[5] 从关系矩阵和布尔向量出发，提出矩阵的重量上乘法和重量下乘法的概念，并以此来计算概念的上下近似集，但矩阵重量乘法的结果并非就是近似集的矩阵表示。Liu[6] 提出一种信息系统中概念（论域子集）的布尔列矩阵（布尔列向量）表示法，然后给出粗糙集上近似算子矩阵描述的概念，从矩阵的视角重新考察和研究 Pawlak 粗糙集模型中概念的上下近似集的性质并构建一系列的公理来描述论域子集的上近似算子的特性，但并没有给出直接运用矩阵的运算来计算下近似集的方法。王磊等[7]提出了利用矩阵计算粗糙集上下近似集的方法，通过矩阵的乘法运算和截矩阵运算来计算粗糙集集合的近似集。本文在文献[7]的基础上，给出了利用决策表的布尔列矩阵和条件属性的等价矩阵计算决策表正域的方法，并从理论上证明了该方法的正确性。根据只有基于下近似集不变的粗糙集属性约简才有属性核的存在[9]，提出了基于属性核的启发式约简，即把条件属性对决策属性依赖度最高的属性作为约简的核，在属性动态增加时，用

到稿日期：2013-01-28 返修日期：2013-04-15 本文受国家自然科学基金项目(60873108)资助。

景运革(1970－)，男，博士生，讲师，主要研究方向为粗糙集理论，E-mail：jyg701022@163.com；李天瑞 男，教授，博士生导师。

矩阵更新的方法来改变属性等价关系矩阵，快速计算属性变化后的正域，通过理论分析和实例验证，该算法是有效的。

2 粗糙集与矩阵的相关知识

2.1 粗糙集的基本知识

定义 1 设 $S=(U,A=C\cup D,V,f)$ 是一个决策表，$U=\{x_1,x_2,\cdots,x_n\}$ 表示非空有限的对象集合，也称论域；其中 C 为条件属性集，D 为决策属性集，$V=\bigcup_{a\in C\cup D}V_a$，其中 V_a 是属性 a 的值域，$f:U\times C\cup D\rightarrow V$ 是一个信息函数，即 $a\in C\cup D$，$x\in U$ 有 $f(x,a)\in V_a$。

定义 2 设 U 为非空有限的对象集合，称之为论域。R 是论域 U 上的等价关系。用 U/R 表示 R 的所有等价类的类族(集合)，用 $[x]_R$ 表示在等价关系 R 下包含元素 $x(x\in U)$ 的等价类。对于论域 U 的任意子集 X，其上、下近似集定义如下[8]：

$\underline{R}(X)=\{x\in U|[x]_R\subset X\}$

$\overline{R}(X)=\{x\in U|[x]_R\cap X\neq\Phi\}$

定义 3 设 $S=(U,A=C\cup D,V,f)$ 是一个决策表，令 U/C 和 U/D 为 U 中的等价类，其中 $U/D=\{D_1,D_2,\cdots,D_m\}$，$D_i$ 是 U/D 的任意子集，D 关于 C 的正域记为 $POS_C(D)$，即 $POS_C(D)=\bigcup_{D_i\in U/D}\underline{R}(D_i)$。

定义 4 设 $S=(U,A=C\cup D,V,f)$ 是一个决策表，U 为论域，$C,D\subseteq A$ 为条件属性集和决策属性集，决策属性集 D 与条件属性集 C 的依赖度定义为 $\gamma_C(D)=|POS_C(D)|/|U|$。

2.2 矩阵的相关知识

定义 5(矩阵加法) 如果 $A=(a_{ij})_{m\times n}$，$B=(b_{ij})_{m\times n}$，则有 $A+B=(a_{ij}+b_{ij})_{m\times n}$。

定义 6(矩阵乘法) 如果 $A=(a_{ij})_{m\times n}$，$B=(b_{ij})_{n\times p}$，则 $A\cdot B=(c_{ij})_{m\times p}$，其中 $(c_{ij})_{m\times p}=\sum_{k=1}^{n}a_{ik}\cdot b_{kj}$。

定义 7 同维矩阵相同行相同列位置的元素称为对位元素。

3 基于矩阵的启发式约简

定义 8[9] 决策信息系统 $S=(U,A=C\cup D,V,f)$，U 为论域，C 为条件属性集，D 为决策属性集。D 关于 C 的正域为 $POS_C(D)$，条件属性 C 关于决策属性 D 的约简定义为 C 的一个最小属性子集 $RED(C,D)$，C 关于 D 的所有最小属性集合为 R，C 关于 D 的属性核为 $CORE(C,D)$，D 关于 C 的一个最小属性子集的正域为 $POS_{RED(C,D)}(D)$。且满足：

1) $POS_C(D)=POS_{RED(C,D)}(D)$；

2) 从 $RED(C,D)$ 中去掉任何一个属性，1)不成立；

3) $CORE(C,D)=\cap R$。

基于核的启发式属性约简的算法思想[8]：对一个决策表中所有条件属性集 C 的依赖度大小排序(C 关于 D 的依赖度根据定义四计算)，依赖度最高的条件属性必然存在于核中，再依次将其他条件属性加到依赖度最高的条件属性中，直到 $POS_C(D)=POS_{RED(C,D)}(D)$，算法停止。

3.1 基于矩阵的粗糙集正域计算

通过布尔列矩阵给出等价关系矩阵和等价关系矩阵的诱导矩阵以及诱导矩阵的 λ-截矩阵的表示形式，并以此为基础，求出决策属性 D 的等价关系子集 D_i 的下近似集的布尔列矩阵，给出决策属性 D 关于 C 的正域的计算方法，并从理论上证明该方法的正确性。

3.1.1 基于矩阵的粗糙集正域计算方法

定义 9 决策信息系统 $S=(U,A=C\cup D,V,f)$。U 表示非空有限的对象集合，也称论域，且 $U=\{u_1,u_2,\cdots,u_n\}$。C 为条件属性集，决策属性 D 的等价类 $U/D=\{D_1,D_2,\cdots,D_m\}$，$D_i$ 是 U/D 的任意子集。则 D_i 可用布尔列矩阵表示：

$$G(D_i)_{n\times 1}=\begin{bmatrix}x_1\\x_2\\\vdots\\x_n\end{bmatrix},\quad x_i=\begin{cases}1, & u_i\in D_i\\0, & u_i\notin D_i\end{cases}$$

简记 $G(D_i)_{n\times 1}=(x_1,x_2,\cdots,x_n)^{\mathrm{T}}$，$(x_1,x_2,\cdots,x_n)^{\mathrm{T}}$ 表示矩阵$(x_1,x_2,\cdots,x_n)$的转置。U/D 的子集 D_i 和 D_i 对应的 n 维布尔列向量等同看待。

定义 10 设 R 是论域 U 上的等价关系，则其等价关系矩阵 M_R 为：

$$M_R=(m_{i,j})_{n\times n}$$

$$(m_{i,j})_{n\times n}=\begin{cases}1, & u_iRu_j\\0, & u_i\not R u_j\end{cases}$$

定义 11[7] 等价关系矩阵 M_R 所诱导的对角矩阵 $\Lambda_{n\times n}$ 定义为：

$$\Lambda_{n\times n}=\begin{bmatrix}\frac{1}{\sum_{k=1}^{n}m_{1k}} & 0 & \cdots & 0\\ 0 & \frac{1}{\sum_{k=1}^{n}m_{2k}} & \cdots & 0\\ \vdots & \vdots & \vdots & \vdots\\ 0 & 0 & \cdots & \frac{1}{\sum_{k=1}^{n}m_{nk}}\end{bmatrix}$$

引理 1

$$\Lambda_{n\times n}=\begin{bmatrix}\frac{1}{|[u_1]|_R} & 0 & \cdots & 0\\ 0 & \frac{1}{|[u_2]|_R} & \cdots & 0\\ \vdots & \vdots & \vdots & \vdots\\ 0 & 0 & \cdots & \frac{1}{|[u_n]|_R}\end{bmatrix}$$

其中，$|[u_i]_R|$ 表示论域 U 中元素 u_i 所在等价类的元素个数，且 $1\leqslant|[u_i]_R|\leqslant n(1\leqslant i\leqslant n)$。$\Lambda_{m\times n}=diag(\lambda_1,\lambda_2,\cdots,\lambda_n)(\lambda_i=\sum_{j=1}^{n}m_{ij})$，$diag(\lambda_1,\lambda_2,\cdots,\lambda_n)$ 表示对角矩阵。

定义 12[7] 矩阵 $\Lambda_{m\times n}$ 的 λ-截矩阵 $(\Lambda_\lambda)_{m\times n}(0\leqslant\lambda\leqslant 1)$ 可定义为：

$$(a_\lambda)_{ij}=\begin{cases}1, & a_{ij}\geqslant\lambda\\0, & a_{ij}<\lambda\end{cases}$$

定义 13 决策信息系统 $S=(U,A=C\cup D,V,f)$。U 表示非空有限的对象集合，也称论域且 $U=\{u_1,u_2,\cdots,u_n\}$。C 为条件属性集，决策属性 D 的等价类 $U/D=\{D_1,D_2,\cdots,D_m\}$，$D_i$ 是 U/D 的任意子集，则子集 D_i 的下近似集为 $\underline{R}(D_i)$，则 $\underline{R}(D_i)$ 的布尔矩阵表示为 $G(\underline{R}(D_i))$：

$$G(\underline{R}(D_i))=[x_1,x_2,\cdots,x_n]$$

$$x_i=\begin{cases}1, & u_i\in\underline{R}(D_i)\\0, & u_i\notin\underline{R}(D_i)\end{cases}$$

定理 1[7] 设$U=\{u_1,u_2,\cdots,u_n\}$是论域集，R是U上的等价关系，决策属性D的等价类$U/D=\{D_1,D_2,\cdots,D_m\}$，$M_R=(m_{i,j})_{n\times n}$是$R$的等价关系矩阵，$D_i$是$U/D$的任意子集，则子集$D_i$的下近似集的布尔列矩阵为(证明见文献[6])：

$$G(\underline{R}(D_i))=(\Lambda\cdot(M_R\cdot G(D_i)_{n\times 1})_1$$

其中，·表示矩阵的数量积，即$c_{i1}=\sum_{k=1}^{n}m_{ik}x_k$。

$(\Lambda\cdot(M_R\cdot G(D_i)_{n\times 1}))_1$为$(\Lambda\cdot(M_R\cdot G(D_i)_{n\times 1}))$的1-截矩阵。其中关系矩阵$M_R$与论域子集$D_i$的乘积为$n\times 1$阶矩阵，记为$c_{n\times 1}$。

定理 2 决策信息系统$S=(U,A=C\cup D,V,f)$。U表示非空有限的对象集合，也称论域。C为条件属性集，决策属性D的等价类$U/D=\{D_1,D_2\cdots D_n\}$。$D$关于$C$的正域为$POS_C(D)=\bigcup_{i=1}^{n}\underline{R}(D_i)$，则$G(\bigcup_{i=1}^{n}\underline{R}(D_i))=\sum_{i=1}^{n}G(\underline{R}(D_i))$。

证明：设$D_1\in U/D,D_2\in U/D,\cdots D_n\in U/D$，则$D_1\cap D_2\cap\cdots\cap D_n=\Phi$，$\underline{R}(D_i)\subseteq D_i(1\leqslant i\leqslant n)$，则$\underline{R}(D_1)\cap\underline{R}(D_2)\cdots\cap\underline{R}(D_n)=\Phi$，当$n=2$时，$\forall\,\underline{R}(D_1)=\{x_{i1},x_{i2},\cdots,x_{in}\}$，$\underline{R}(D_2)=\{y_{i1},y_{i2},\cdots,y_{in}\}$，$\bigcup_{i=1}^{2}\underline{R}(D_i)=\underline{R}(D_1)\cup\underline{R}(D_2)=\{z_{i1},z_{i2},\cdots,z_{in}\}$

i)若设$x_{ik}\in\underline{R}(D_1)$，$x_{ik}\notin\underline{R}(D_2)$，$x_{ik}\in(\underline{R}(D_1)\cup\underline{R}(D_2))$，则$x_{ik}=1,y_{ik}=0,z_{ik}=1,z_{ik}=x_{ik}+y_{ik}$。

ii)设$x_{ik}\notin\underline{R}(D_1)$，$x_{ik}\in\underline{R}(D_2)$，$x_{ik}\in(\underline{R}(D_1)\cup\underline{R}(D_2))$，则$x_{ik}=0,y_{ik}=1,z_{ik}=1,z_{ik}=x_{ik}+y_{ik}$。

iii)若设$x_{ik}\notin\underline{R}(D_1)$，$x_{ik}\notin\underline{R}(D_2)$，$x_{ik}\notin(\underline{R}(D_1)\cup\underline{R}(D_2))$，则$x_{ik}=0$，$y_{ik}=0$，$z_{ik}=0$，$z_{ik}=x_{ik}+y_{ik}$，则$G(\bigcup_{i=1}^{2}\underline{R}(D_i))=\sum_{i=1}^{2}G(\underline{R}(D_i))$

假设：$G(\bigcup_{i=1}^{n}\underline{R}(D_i))=\sum_{i=1}^{n}G(\underline{R}(D_i))$ 则：

$$G(\bigcup_{i=1}^{n+1}\underline{R}(D_i))=G(\bigcup_{i=1}^{n}\underline{R}(D_i)\cup\underline{R}(D_{n+1}))=G(\sum_{i=1}^{n}G(\underline{R}(D_i)))+G(\underline{R}(D_{n+1}))=G(\sum_{i=1}^{n+1}G(\underline{R}(D_i)))。$$

证毕。

3.1.2 计算决策表的正域的算法描述

输入：等价关系 R、决策属性 D 的等价类 $U/D=\{D_1,D_2,\cdots,D_n\}$。

输出：D 关于 C 的正域 $POS_C(D)$。

Step 1 根据等价关系 R 构建等价关系矩阵 $M_R=(m_{i,j})_{n\times n}$；

Step 2 构建由 $M_R=(m_{i,j})_{n\times n}$所诱导出的对角矩阵 $\Lambda_{n\times n}$；

Step 3 构建等价类 $U/D=\{D_1,D_2,\cdots,D_n\}$的任意子集 D_i 的布尔列矩阵；

Step 4 计算等价类 $U/D=\{D_1,D_2,\cdots,D_n\}$的任意子集 D_i 的下近似集合矩阵；

$G(\underline{R}(D_1))=(\Lambda\cdot(M_R\cdot G(D_1)_{n\times 1})_1$，

$G(\underline{R}(D_2))=(\Lambda\cdot(M_R\cdot G(D_2)_{n\times 1}))_1$，

…

$G(\underline{R}(D_n))=(\Lambda\cdot(M_R\cdot G(D_n)_{n\times 1}))_1$；

Step 5 计算 D 关于 C 的正域 $POS_C(D)$则：

$$G(POS_C(D))=G(\bigcup_{i=1}^{n}\underline{R}(D_i))=\sum_{i=1}^{n}G(\underline{R}(D_i))$$

Step 6 得出决策表 D 关于 C 的正域，算法完毕。

3.2 等价关系矩阵快速更新方法[10]

给定一个决策信息系统$S=(U,A=C\cup D,V,f)$，属性集$A=\{a_1,a_2,\cdots,a_n\}$，$B\subseteq A$且$B\neq\Phi,a_i\in A-B$，增加单属性集$\{a_i\}$至属性集B后，属性集更新为$B'=B\cup\{a_i\}$。设属性集B,B'以及$\{a_i\}$划分论域U所形成的等价关系矩阵分别为$M_B,M_{B'}$和M_{a_i}(均为$n\times n$阶矩阵，$n=|U|$)，则各个等价关系矩阵的元素表示式如下：

$$(m_B)_{ij}=m_{ij}=\begin{cases}1, & f(u_i,a)=f(u_j,a),\forall a\in B\\0, & f(u_i,a)\neq f(u_j,a),\exists a\in B\end{cases}$$

$$(m_{a_i})_{ij}=\omega_{ij}=\begin{cases}1, & f(u_i,a_i)=f(u_j,a_i)\\0, & f(u_i,a_i)\neq f(u_j,a_i)\end{cases}$$

$$(m_{B\cup\{a_i\}})_{ij}=m_{ij}^{+}=\begin{cases}1, & f(u_i,a)=f(u_j,a),\forall a\in(B\cup\{a_i\})\\0, & f(u_i,a)\neq f(u_j,a),\exists a\in(B\cup\{a_i\})\end{cases}$$

定理 3[11] 设决策信息系统$S=(U,A=C\cup D,V,f)$，属性集$A=\{a_1,a_2,\cdots,a_n\}$，$B\subseteq A$且$B\neq\Phi,a_i\in A-B$，属性集B对划分论域所形成的不可区分的矩阵(即等价关系矩阵)为M_B。则增加$\{a_i\}$至属性集B后，更新后的矩阵$M_{B'}$的元素计算式为：$(m_{ij})^{+}=\begin{cases}m_{ij}, & \omega_{ij}=1\\0, & \omega_{ij}=0\end{cases}$

增加属性a_i后等价关系矩阵元素的变化情况如表 1 所列。

表 1 增加属性 a_i 后等价关系矩阵元素的变化情况

m_{ij}	ω_{ij}	m_{ij}^{+}
1	1	1
1	0	0
0	1	0
0	0	0

例如：$M_B=\begin{bmatrix}1&0\\1&0\end{bmatrix}$，$M_{a_i}=\begin{bmatrix}1&0\\0&1\end{bmatrix}$，$\omega_{11}=1,m_{11}^{+}=m_{11}=1$；$\omega_{12}=0,m_{11}^{+}=0$；当$\omega_{21}=0,m_{21}^{+}=0$；$\omega_{22}=1,m_{22}^{+}=m_{22}=0$，则$M_{B'}=\begin{bmatrix}1&0\\0&0\end{bmatrix}$。

定理 4 设决策信息系统$S=(U,A=C\cup D,V,f)$，$B\subseteq A,B\neq\Phi,a_i\in A-B$，$\Lambda_{m\times n}^{B}=diag(\lambda_1,\lambda_2,\cdots,\lambda_n)(\lambda_i=\sum_{j=1}^{n}m_{ij})$，设$\Lambda_{n\times n}^{B\cup a_i}=diag(\lambda_1^{+},\lambda_2^{+},\cdots,\lambda_n^{+})(\lambda_i^{+}=\sum_{j=1}^{n}m_{ij})$，则有：$\lambda_i^{+}=\lambda_i-\sum_{j=1}^{n}(1-\omega_{i,j})\cdot m_{ij}$。

3.3 决策信息系统属性约简算法的描述

输入：决策信息系统 $S=(U,A=C\cup D,V,f)$，U 为论域，C 为条件属性集，D 为决策属性集。条件属性 C 的等价类 $U/C=\{C_1,C_2,\cdots,C_m\}$，子集 C_i 关于 D 的依赖度为 $k=\gamma_{C_i}(D)=|POS_{C_i}(D)|/|U|$，其中 $POS_{C_i}(D)$是 D 关于 C_i 的正域。

输出：最小属性约简为 RED(C,D)。

Step1 计算 D 关于 C 的正域 $POS_C(D)$；

Step2 以属性依赖度对条件属性 C 中的属性排序，假设顺序为 $\gamma_{C_1}(D)>\gamma_{C_2}(D)>\cdots>\gamma_{C_m}(D)$；

Step3 $RED(C,D)=C_1$；

Step4 for i=1:m

4.1 根据定理 3 计算更新后的 $M_{RED'}$等价关系矩阵。

4.2 根据定理 4 计算更新后的对角矩阵 $\Lambda_{n\times n}^{RED'}=diag(\lambda_1^{+},\lambda_2^{+},\cdots\lambda_n^{+})(\lambda_i^{+}=\sum_{j=1}^{n}m_{ij})$

4.3 根据定理 2 计算更新后 D 关于 $M_{RED'}$的正域 $POS_{RED(C,D)}(D)$

if $POS_C(D)=POS_{RED(C,D)}(D)$

Then $RED(C,D)=RED(C,D)\cup C_i$

End

Step 5　输出最小属性约简 RED(C,D)。

4　算例

决策表如表 2 所列，论域 $U=\{x_1,x_2,x_3,x_4,x_5,x_6\}$，条件属性 $C=\{a,b,c\}$，决策属性 $D=\{d\}$。

表 2　决策表

U	a	b	c	d
x1	2	2	0	1
x2	1	2	0	0
x3	1	2	0	1
x4	0	0	0	0
x5	1	0	1	0
x6	2	0	1	1

根据表 2 和定义 1 求出条件属性等价类和决策属性等价类为：

$U/C=\{(x_1),(x_2,x_3),(x_4),(x_5),(x_6)\}$

$U/a=\{(x_1,x_6),(x_2,x_3,x_5),(x_4)\}$

$U/b=\{(x_1,x_2,x_3),(x_4,x_5,x_6)\}$

$U/c=\{(x_1,x_2,x_3,x_4),(x_5,x_6)\}$

$U/D=\{(x_1,x_3,x_6),(x_2,x_4,x_5)\}$

令 $D_1=\{(x_1,x_3,x_6)\}$ 或 $D_2=\{(x_2,x_4,x_5)\}$。

根据定义 9，决策属性 D 的等价类 U/D 布尔矩阵分别为 $G(D_1)_{n\times1}=(1,0,1,0,0,1)^{\mathrm{T}}$ 和 $G(D_2)_{n\times1}=(0,1,0,1,1,0)^{\mathrm{T}}$。

4.1　计算条件属性关于决策属性 D 的依赖度按降序排列

$$\gamma_a(D)=\frac{2}{6}=\frac{1}{3},\gamma_b(D)=0,\gamma_c(D)=0$$

排序结果：$\gamma_a(D)>\gamma_b(D)=\gamma_c(D)$

4.2　计算决策属性 D 关于条件属性 C 的正域

1. 根据 U/C 可得到条件属性等价关系矩阵和诱导矩阵 $\Lambda_{6\times6}$ 为：

$$M_{6\times6}=\begin{bmatrix}1&0&0&0&0&0\\0&1&1&0&0&0\\0&1&1&0&0&0\\0&0&0&1&0&0\\0&0&0&0&1&0\\0&0&0&0&0&1\end{bmatrix}$$

$$\Lambda_{6\times6}=\begin{bmatrix}1&0&0&0&0&0\\0&0.5&0&0&0&0\\0&0&0.5&0&0&0\\0&0&0&1&0&0\\0&0&0&0&1&0\\0&0&0&0&0&1\end{bmatrix}$$

2. D 关于 C 的正域为：

$$\begin{aligned}G(\underline{R}(D_1))&=(\Lambda\cdot(M_R\cdot G(D_1)_{n\times1})_1\\&=(\Lambda\cdot(M_R\cdot(1,0,1,0,0,1)^{\mathrm{T}})_1\\&=(1,0.5,0.5,0,0,1)^{\mathrm{T}}_1\\&=(1,0,0,0,0,1)^{\mathrm{T}}\end{aligned}$$

$$\begin{aligned}G(\underline{R}(D_2))&=(\Lambda\cdot(M_R\cdot G(D_2)_{n\times1})_1\\&=(\Lambda\cdot(M_R\cdot(0,1,0,1,1,0)^{\mathrm{T}})_1\\&=(0,0.5,0.5,1,1,0)^{\mathrm{T}}_1\\&=(0,0,0,1,1,0)^{\mathrm{T}}\end{aligned}$$

则：$G(POS_C(D))=\sum_{i=1}^{2}G(\underline{R}(D_i))=(1,0,0,1,1,1)^{\mathrm{T}}$。

D 关于 C 的正域为：$POS_C(D)=\{x_1,x_3,x_4,x_5\}$。

4.3　计算增加属性集的正域

1. 根据 U/a 可得到条件属性 a 的等价关系矩阵以及待插入的属性集 b 的等价关系矩阵为：

$$M_a=\begin{bmatrix}1&0&0&0&0&1\\0&1&1&0&1&0\\0&1&1&0&1&0\\0&0&0&1&0&0\\0&1&1&0&1&0\\1&0&0&0&0&1\end{bmatrix}$$

$$M_b=\begin{bmatrix}1&1&1&0&0&0\\1&1&1&0&0&0\\1&1&1&0&0&0\\0&0&0&1&1&1\\0&0&0&1&1&1\\0&0&0&1&1&1\end{bmatrix}$$

2. 根据定理 2 可得到$\{a\cup b\}$等价关系矩阵为：

$$M_{a\cup b}=\begin{bmatrix}1&0&0&0&0&0\\0&1&1&0&0&0\\0&1&1&0&0&0\\0&0&0&1&0&0\\0&0&0&0&1&0\\0&0&0&0&0&1\end{bmatrix}$$

3. 动态更新对角矩阵元素。

根据定义 11 可得 a 的对角矩阵的元素：$\Lambda_a=diag(1,3,3,1,3,2)$。

根据定理 4 得$\{a\cup b\}$的对角矩阵的元素：

$$\lambda_1^+=\lambda_1-\sum_{j=1}^{6}(1-\omega_{1,j})\cdot m_{1j}=1-0=1$$

$$\lambda_2^+=\lambda_2-\sum_{j=1}^{6}(1-\omega_{2,j})\cdot m_{2j}=3-1=2$$

$$\lambda_3^+=\lambda_3-\sum_{j=1}^{6}(1-\omega_{3,j})\cdot m_{3j}=3-1=2$$

$$\lambda_4^+=\lambda_4-\sum_{j=1}^{6}(1-\omega_{4,j})\cdot m_{4j}=1-0=1$$

$$\lambda_5^+=\lambda_5-\sum_{j=1}^{6}(1-\omega_{5,j})\cdot m_{5j}=3-2=1$$

$$\lambda_6^+=\lambda_6-\sum_{j=1}^{6}(1-\omega_{6,j})\cdot m_{6j}=2-1=1$$

则，$\Lambda_{\{a\cup b\}}=diag(1,2,2,1,1,1)$。

4. D 关于$\{a\cup b\}$的正域为：

根据引理 1 得 $M_{a\cup b}$ 1-截矩阵的对角矩阵 $\Lambda_{a\cup b}$，D 关于$\{a\cup b\}$的正域 $G(POS_{RED(C,D)}(D))=\sum_{i=1}^{2}G(\underline{R}(D_i))=(1,0,0,1,1,1)^{\mathrm{T}}$，$POS_C(D)=POS_{RED(C,D)}(D)$，所以表 1 的属性约简为$\{a,b\}$。

结束语　本文在矩阵的基本运算和等价关系矩阵及截矩阵计算粗糙集下近似集方法的基础上，提出使用矩阵计算的方法求决策表的正域，利用依赖度的大小，实现了矩阵的启发式约简算法，在属性动态增加时，用矩阵快速更新的方法来改变属性等价关系矩阵，可以快速地计算属性变化后的等价关系矩阵，从而为属性变化后求正域节省了时间。最后通过理论分析和实例表明，本文所提出的基于矩阵的决策表属性约简的算法是可行的。

（下转第 286 页）

对于 BUMP 问题 $f_2(\boldsymbol{X})$，运用本文求解算法进行求解，计算结果如表 4 所列。

表 4 PDO 算法与 BBO 算法的 BUMP 函数测试效果对比

维数 n	目标函数值(PDO 算法)	目标函数值(BBO 算法)	耗时(h)
20	−0.803605881101687	−0.4513385612369127	2.3
30	−0.821789549623916	−0.4781292382376343	2.7
50	−0.832132927828923	−0.5191276301252371	3.4
100	−0.844349528432361	−0.5512423812523893	4.6
1000	−0.843128203982247	−0.6572382712651393	7.3

从表 3 可以看出，对于普通函数优化问题，随着维数的增加，PDO 算法要越来越优于 BBO 算法，而且随着维数的增加，这种优势越来越明显；从表 4 可以看出，对于极难优化的 BUMP 问题，PDO 算法获得最优解的精度要明显高于 BBO 算法。

结束语 本文运用种群动力学理论构造出了可求解大规模优化问题的优化算法，算法中运用种群动力学基本模型提出了竞争、互利、捕食-被食、融合、突变和选择等算子，运用正交拉丁方生成方法提出了试探解的初始化方法，确保了初始试探解具有很好的均衡分散性和整齐可比性。这些方法的运用大幅提升了 PDO 算法的性能，该 PDO 算法具有全局收敛性，且适用于求解高维优化问题。今后需要进一步研究的方向是：

(1)如何将两个种群间的竞争、互利、捕食-被食行为扩展为多个种群；

(2)如何深入分析 PDO 算法的动态行为；

(3)如何将更多、更反映种群进化实际情况的种群动力学理论用于改善或构造出新的种群动力学优化算法。

参考文献

[1] 王顺庆，王万雄，徐海根. 生态学稳定性理论与方法[M]. 北京：科学出版社，2004

[2] 陆征一，周义仓. 数学生物学进展[M]. 北京：科学出版社，2006

[3] 李长生. 一类 n 维 Lotka-Volterra 系统稳定平衡点的判别[J]. 潍坊学院学报，2003，3(6)：1-2

[4] 程惠东，孟新柱，王芳. 一类时滞非自治 Lotka-Volterra 扩散生态系统的全局吸引性[J]. 中山大学学报：自然科学版，2008，47(2)：19-22

[5] 周桦，刘佳. 带扩散的具有 Holling-IV 类功能性反应的捕食模型的性质[J]. 南京工业大学学报，2007，29(5)：66-69

[6] 吴兴杰. 具有 Holling IV 功能反应的脉冲食物链模型的动力学性质[J]. 桂林电子科技大学学报，2008，28(03)：248-253

[7] 郑宗剑，林怡平. 一类具有时滞 Holling-IV 型捕食-食饵系统的 Hopf 分支[J]. 云南民族大学学报：自然科学版，2007，16(01)：18-21

[8] 赵延忠. 具有反应扩散的 Leslie 两种群生物竞争模型的稳定性[J]. 青海大学学报，2007，25(5)：62-65

[9] 庞国萍，陈兰荪. 具有脉冲效应和 Holling IV 功能性反应的捕食者食饵系统分析[J]. 南京师大学报：自然科学版，2007，30(2)：1-5

[10] 田灿荣. 一类带时滞竞争模型的周期解[J]. 生物数学学报，2007，22(3)：431-440

[11] Simon D. Biogeography-based Optimization[J]. IEEE Transactions：Evolutionary Computation，2008，12(6)：702-713

[12] 王存睿，王楠楠，段晓东，等. 生物地理学优化算法综述[J]. 计算机科学，2010，37(7)：34-38

[13] 中国现场统计研究会三次设计组. 正交法和三次设计[M]. 北京：科学出版社，1987

[14] Iisufescu M. Finite Markov Processes and Their Applications[M]. Wiley：Chichester，1980

[15] 任子晖，王坚，高岳林. 马尔科夫链的粒子群优化算法全局收敛性分析[J]. 控制理论与应用，2011，28(4)：462-466

[16] 王宜举，修乃华. 非线性优化理论[M]. 北京：科学出版社，2012

[17] 崔志华，曾建潮. 微粒群优化算法[M]. 北京：科学出版社，2011

[18] 陈兰荪，孟新柱，焦建军. 生物动力学[M]. 北京：科学出版社，2009

(上接第 264 页)

参考文献

[1] Pawalk Z. Rough Sets [J]. International Journal of Computer and Information Science，1982，11(5)：341-356

[2] Jensen R，Shen Q. Fuzzy-rough sets assisted attribute selection [J]. IEEE Transactions on Fuzzy Systems，2007，15(1)：73-89

[3] Skowron A，Swiniarski R，Synak P. Approximation Spaces and Information Granulation[C]//Proc of the International Conference on Rough Sets and Current Trends in Computing. Uppsala，Sweden，2004：116-126

[4] Guan J W，Bell D A，Guan Z. Matrix Computation for Information System[J]. Information Sciences，2001，131(1-4)：129-156

[5] 杨勇. 粗糙集的矩阵定义[J]. 计算机工程与应用，2007，43(14)：1-2

[6] Liu Gui-long. The Axiomatization of Rough Set Upper Approximation Operation[J]. Fundamenta Informaticae，2006，69(3)：331-342

[7] 王磊，等. 基于矩阵的粗糙集上下近似的计算方法[J]. 模式识别与人工智能，2011，24(6)：757-762

[8] 刘清. Rough 集及 Rough 推理[M]. 北京：科学出版社，2001：60-70

[9] 陈昊，等. 变精度粗糙集的属性核和最小属性约简算法[J]. 计算机学报，2012，35(5)：1012-1016

[10] Cheng Y. The incremental method for fast computing the rough fuzzy approximations[J]. Data & Knowledge Engineering，2011，70：84-100

[11] Li T R，Ruan D，Gerret W，et al. A rough set based characteristic relation approach for dynamic attribute generalization in data mining[J]. Knowledge-based Systems，2007，20(5)：485-494