文章编号:1005-0523(2018)04-0083-06

最优近似粗糙集的矩阵计算

罗来鹏,刘二根,范自柱

(华东交通大学理学院,江西 南昌 330013)

摘要:在粗糙集理论与应用研究中,由近似空间导出粗糙集的上、下近似集是它的核心基础。但是这种刻画只是给出一个范围,为了更精确刻画目标概念,推广粗糙集的应用,针对粗糙集的最优近似集的计算问题,提出一种基于矩阵的计算方法,为此引入布尔矩阵、关系矩阵、集合矩阵、集合相似度矩阵等定义,通过对目标概念与其近似集之间相似度讨论,得到它们之间相似度变化规律,由此建立粗糙集最优近似集的矩阵新判断与算法。最后通过实例对上述结论作进一步说明。

关键词:粗糙集;最优近似集;相似度;矩阵

中图分类号:TP18 **文献标志码**:A

粗糙集是由Pawlak教授在1982提出的不确定性处理数学理论[1],由于不需要先验知识,一直以来在模式识别、机器学习、数据挖掘等领域倍受关注[2-3]。目标概念在近似空间上的表示与计算是粗糙理论与应用中最基本的也是最重要的内容之一。由于目标概念在近似空间上的不确定性,通常使用它的上、下近似集对其近似表示。这种表示只是给出目标概念近似集的一个区间范围。那么在上、下近似集之间是否存在一个集合,它不仅可以用基本知识基表示,而且与目标概念相似度最高?使张清华教授在文献[4]中将粗糙集转化为模糊集方法进行描述,然后针对粗糙集边界域中元素的隶属度的不同,利用截集的方式来构造目标概念 的近似集,讨论了随着知识空间中知识粒度的变化、粗糙集的近似集与目标概念的近似度的变化规律,找到一个相对较好的近似集 $R_{0.5}(X)$。文献[5-6]就最优近似集的问题作了进一步完善,得到一些新的理论成果。但关于如何求近似空间上一个目标概念的最优近似集未给出具体计算方法。

矩阵作为数学中的重要理论工具在信息科学中得到广泛应用。自从Guan J W等人提出粗糙集的矩阵计算[7]以来,在用矩阵来研究粗糙集理论与应用取得了很多理论成果。文献[8]提出基于矩阵的就变精度粗糙集模型以及属性约简;文献[9]则提出用矩阵来表示粒与粒计算;文献[10]提出用矩阵来研究粗糙集近似集的动态变化;文献[11]用矩阵方法研究不协调决策表的属性约简等等。从上述研究可以看出用矩阵方法来研究粗糙集具有直观、易于在计算机上实现等优点。本论文从矩阵的角度提出粗糙集最优近似集新的判定与计算方法。

1 集合矩阵与运算

定义1[7-8] 设矩阵 $\boldsymbol{D}=(d_{ij})_{m\times n}$,且 $d_{ij}=0$ 或1,则称矩阵 $\boldsymbol{D}$ 为布尔矩阵。

定义2[7-8] 设布尔矩阵 $\boldsymbol{A}=(a_{ij})_{m\times n}$,$\boldsymbol{B}=(b_{ij})_{m\times n}$,$\boldsymbol{C}=(c_{ij})_{m\times n}$,若对于 $1\leqslant i\leqslant m$,$1\leqslant j\leqslant n$ 有 $c_{ij}=a_{ij}\vee b_{ij}$,则记 $\boldsymbol{C}=\boldsymbol{A}\vee\boldsymbol{B}$;若 $c_{ij}=a_{ij}\wedge b_{ij}$,则记 $\boldsymbol{C}=\boldsymbol{A}\wedge\boldsymbol{B}$;若 $a_{ij}\leqslant b_{ij}$ 则 $\boldsymbol{A}\leqslant\boldsymbol{B}$。其中∨,∧分别为取小,取大。

定义3[7-8] 设 U 是元素有序排列的有限论域,即 $U=\{x_1,x_2,\cdots,x_n\}$,对于任意 $E\subseteq U$,可定义一个布尔矩阵

收稿日期:2018-01-15
基金项目:国家自然科学基金项目(61472138)
作者简介:罗来鹏(1973—),男,副教授,主要研究方向为粗糙集与数据挖掘。

与其对应，记为：$\boldsymbol{D}_E=[e_i]_{n\times1}$（下文都是这种约定），其中 $e_i=\begin{cases}1 & x_i\in E\\0 & x_i\notin E\end{cases}$，矩阵 $\boldsymbol{D}_E$ 称为集合 E 在论域 U 上的集合矩阵。

显然论域上的任何一个集合都与一集合矩阵一一对应，并且 $\boldsymbol{D}_U=[1,1,\cdots,1]^{\mathrm{T}}$，$\boldsymbol{D}_\Phi=[0,0,\cdots,0]^{\mathrm{T}}=0$。

对于 $F,G\subseteq U$，很容易得到如下性质。

性质 1 1）集合 F 的基数 $|F|=|\boldsymbol{D}_F|=\sum_{i=1}^{n} f_i$（下文 $|\cdot|$ 都表示集合或者对应集合矩阵基数）；

2）$\boldsymbol{D}_{F^c}=\boldsymbol{D}_U-\boldsymbol{D}_F$；

3）$\boldsymbol{D}_{F\cap G}=\boldsymbol{D}_F\wedge\boldsymbol{D}_G$，$\boldsymbol{D}_{F\cup G}=\boldsymbol{D}_F\vee\boldsymbol{D}_G$，$\boldsymbol{D}_{F-G}=\boldsymbol{D}_{F\cap G^c}=\boldsymbol{D}_F\wedge\boldsymbol{D}_G$；

4）$F\subseteq G\Leftrightarrow\boldsymbol{D}_F\leqslant\boldsymbol{D}_G$ 。

2 粗糙集的矩阵表示

根据文献[8]，对于近似空间上粗糙集基于矩阵算子描述如下：设 (U,R) 为近似空间，$n=|U|$，R 为等价关系，则由 R 可导出关系矩阵 $\boldsymbol{D}_R=[r_{ij}]_{n\times n}$，其中 $r_{ij}=\begin{cases}1 & x_iRx_j\\0 & \text{否则}\end{cases}$，以及划分矩阵 $\boldsymbol{D}_{U/R}=\{\boldsymbol{D}_{X_1},\boldsymbol{D}_{X_2},\cdots,\boldsymbol{D}_{X_k}\}$。其中对于任意 $i\neq j$，有 $\boldsymbol{D}_{X_i}\wedge^1\boldsymbol{D}_{X_j}=0$，且 $\bigvee_{i=1}^{k}\boldsymbol{D}_{X_i}=\boldsymbol{D}_U$。

定义 4[8] 令 $\boldsymbol{M}_R=[m_{ij}]_{n\times n}$，其中 $m_{ij}=r_{ij}/\sum_{j=1}^{n} r_{ij}$ 则称 $\boldsymbol{M}_R$ 为近似空间上 (U,R) 的矩阵算子。

对于任意 $\boldsymbol{D}_X\leqslant\boldsymbol{D}_X$，令 $\boldsymbol{M}_R\times\boldsymbol{D}_X=[p_i]_{n\times1}$，则 $\boldsymbol{D}_X$ 在近似空间 (U,R) 上的下、上近似集集合矩阵分别为

$$\underline{\boldsymbol{D}}_R^{(X)}=[q_i]_{n\times1}\ \text{其中}\ q_i=\begin{cases}1 & p_i=1\\0 & \text{其余}\end{cases}$$

$$\overline{\boldsymbol{D}}_R^{(X)}=[q_i]_{n\times1}\ \text{其中}\ q_i=\begin{cases}1 & <p_i<1\\0 & \text{其余}\end{cases}$$

边界域矩阵 $\boldsymbol{D}_N^{(X)}=\overline{\boldsymbol{D}}_R^{(X)}-\underline{\boldsymbol{D}}_R^{(X)}$，负域矩阵为 $\boldsymbol{D}_U-\overline{\boldsymbol{D}}_R^{(X)}$。

3 最优近似粗糙集的矩阵计算

3.1 相似度与近似集矩阵表示

设 $\boldsymbol{D}_A,\boldsymbol{D}_B\leqslant\boldsymbol{D}_U$，则集合 A,B 的相似度矩阵定义为：$S(A,B)=\dfrac{|\boldsymbol{D}_A\wedge\boldsymbol{D}_B|}{|\boldsymbol{D}_A\vee\boldsymbol{D}_B|}$。

显然 $0\leqslant S(A,B)\leqslant1$，其中 $S(A,B)=0$ 当且仅当 $A\cap B=\Phi$，$S(A,B)=1$ 当且仅当 $A=B$。在本文中，将用相似度来刻画两个集合近似度，也就是说相似度越高，近似度也越高。

根据粗糙集的定义，对于一个目标概念，它的下近似集与其边界域中任何基本知识并都是目标概念的近似集。而下近似集、上近似集只是众多近似集中最小、最大的两个。那么在这些近似集哪个更为近似，下面利用矩阵方法给出新的判定与计算。为了更好推导，根据前面集合矩阵性质很容易得到如下集合矩阵基数性质。

性质 2 1）设 $\boldsymbol{D}_A,\boldsymbol{D}_B\leqslant\boldsymbol{D}_U$，若 $\boldsymbol{D}_A\wedge\boldsymbol{D}_B=0$，则 $|\boldsymbol{D}_A\vee\boldsymbol{D}_B|=|\boldsymbol{D}_A|+|\boldsymbol{D}_B|$；

2）$\boldsymbol{D}_A,\boldsymbol{D}_B,\boldsymbol{D}_C\leqslant\boldsymbol{D}_U$ 且 $\boldsymbol{D}_A\wedge\boldsymbol{D}_B=0$，则 $|(\boldsymbol{D}_A\wedge\boldsymbol{D}_C)\vee(\boldsymbol{D}_B\wedge\boldsymbol{D}_C)|=|(\boldsymbol{D}_A\wedge\boldsymbol{D}_C)|+|(\boldsymbol{D}_B\wedge\boldsymbol{D}_C)|$；

3）$\boldsymbol{D}_A,\boldsymbol{D}_B\leqslant\boldsymbol{D}_U$，$|\boldsymbol{D}_A\vee(\boldsymbol{D}_{B\cap A_C})|=|\boldsymbol{D}_A|+|\boldsymbol{D}_{B\cap A_C}|$。

为了描述需要引入如下记号：设 (U,R) 为近似空间，R 为等价关系，划分矩阵 $\boldsymbol{D}_{U/R}=\{\boldsymbol{D}_{X_1},\boldsymbol{D}_{X_2},\cdots,\boldsymbol{D}_{X_k}\}$ 对于任意 $\boldsymbol{D}_X\leqslant\boldsymbol{D}_U$，令 $\alpha=S(X,\underline{R}(X))=\dfrac{|\boldsymbol{D}_X\wedge\underline{\boldsymbol{D}}_R^{(X)}|}{|\boldsymbol{D}_X\vee\underline{\boldsymbol{D}}_R^{(X)}|}=\dfrac{|\underline{\boldsymbol{D}}_R^{(X)}|}{|\boldsymbol{D}_X|}$，记 $M=\{\boldsymbol{D}_m|\boldsymbol{D}_m\in\boldsymbol{D}_{U/R}\ \text{且}\ \boldsymbol{D}_m\leqslant\boldsymbol{D}_N^{(X)}\}$，则 α 为目标概念 X 与其下近似集的相似度，M 为目标概念 X 边界域中的基本知识基的集合矩阵。对于任意的 $\boldsymbol{D}_k\subseteq M$，$\boldsymbol{D}_k\cup\underline{\boldsymbol{D}}_R^{(X)}$ 都是 X 的近似集矩阵。

3.2 粗糙集最优近似集矩阵判定

定理 1　$\boldsymbol{D}_P\in\boldsymbol{M}$，若 $\frac{|\boldsymbol{D}_P\wedge\boldsymbol{D}_X|}{|\boldsymbol{D}_P\vee\boldsymbol{D}_{X^C}|}\geqslant\alpha$，则 $S(X,\underline{R}(X))\leqslant S(X,P\cup\underline{R}(X))$。

证明　由集合基数性质(2)及不等式性质：$0<\frac{a}{b}\leqslant\frac{c}{d}$，则 $\frac{a}{b}\leqslant\frac{a+c}{b+d}\leqslant\frac{c}{d}$，可得到。

$$\alpha=\frac{|\boldsymbol{D}_X\wedge\underline{\boldsymbol{D}}_R^{(X)}|}{|\boldsymbol{D}_X\vee\underline{\boldsymbol{D}}_R^{(X)}|}\leqslant\frac{|\boldsymbol{D}_X\wedge\underline{\boldsymbol{D}}_R^{(X)}|+|\boldsymbol{D}_P\wedge\boldsymbol{D}_X|}{|\boldsymbol{D}_X\vee\underline{\boldsymbol{D}}_R^{(X)}|+|\boldsymbol{D}_P\wedge\boldsymbol{D}_{X^C}|}=\frac{|\boldsymbol{D}_X\wedge(\boldsymbol{D}_P\vee\underline{\boldsymbol{D}}_R^{(X)})|}{|\boldsymbol{D}_X|+|\boldsymbol{D}_P\wedge\boldsymbol{D}_{X^C}|}=\frac{|\boldsymbol{D}_X\wedge(\boldsymbol{D}_P\vee\underline{\boldsymbol{D}}_R^{(X)})|}{|\boldsymbol{D}_X\vee(\boldsymbol{D}_P\vee\underline{\boldsymbol{D}}_R^{(X)})|}=S(X,P\cup\underline{R}(X))$$

即 $S(X,\underline{R}(X))\leqslant S(X,P\cup\underline{R}(X))$ 并且 $S(X,P\cup\underline{R}(X))\leqslant\frac{|\boldsymbol{D}_P\wedge\boldsymbol{D}_X|}{|\boldsymbol{D}_P\vee\boldsymbol{D}_{X^C}|}$。

性质 3　$\boldsymbol{D}_P\in\boldsymbol{M}$，若 $\frac{|\boldsymbol{D}_P\wedge\boldsymbol{D}_X|}{|\boldsymbol{D}_P\vee\boldsymbol{D}_{X^C}|}<\alpha$，则 $S(X,\underline{R}(X))>S(X,P\cup\underline{R}(X))$。

定理 2　$\boldsymbol{D}_P,\boldsymbol{D}_Q\in\boldsymbol{M}$，若 $\frac{|\boldsymbol{D}_Q\wedge\boldsymbol{D}_X|}{|\boldsymbol{D}_Q\vee\boldsymbol{D}_{X^C}|}\geqslant\frac{|\boldsymbol{D}_P\wedge\boldsymbol{D}_X|}{|\boldsymbol{D}_P\vee\boldsymbol{D}_{X^C}|}\geqslant\alpha$，则 $\alpha\leqslant S(X,P\cup\underline{R}(X))\leqslant S(X,P\cup Q\cup\underline{R}(X))$。

证明　根据定理 1 证明，可以得到

$$S(X,P\cup\underline{R}(X))=\frac{|\boldsymbol{D}_X\wedge(\boldsymbol{D}_P\vee\underline{\boldsymbol{D}}_R^{(X)})|}{|\boldsymbol{D}_X\vee(\boldsymbol{D}_P\vee\underline{\boldsymbol{D}}_R^{(X)})|}\leqslant\frac{|\boldsymbol{D}_X\wedge(\boldsymbol{D}_P\vee\underline{\boldsymbol{D}}_R^{(X)})|+|\boldsymbol{D}_Q\wedge\boldsymbol{D}_X|}{|\boldsymbol{D}_X\vee(\boldsymbol{D}_P\vee\underline{\boldsymbol{D}}_R^{(X)})|+|\boldsymbol{D}_Q\wedge\boldsymbol{D}_{X^C}|}=\frac{|\boldsymbol{D}_X\wedge(\boldsymbol{D}_Q\vee\boldsymbol{D}_P\vee\underline{\boldsymbol{D}}_R^{(X)})|}{|\boldsymbol{D}_X\vee(\boldsymbol{D}_Q\vee\boldsymbol{D}_P\vee\underline{\boldsymbol{D}}_R^{(X)})|}=S(X,\underline{R}(X)\cup P\cup Q)$$

定理 3　若对于任意的 $\boldsymbol{D}_P\in\boldsymbol{M}$ 有 $\frac{|\boldsymbol{D}_P\wedge\boldsymbol{D}_X|}{|\boldsymbol{D}_P\vee\boldsymbol{D}_{X^C}|}\geqslant\alpha$，则 X 在近似空间(U,R)上的最优近似集为 $\overline{R}(X)$。

证明　设 $\boldsymbol{M}=\{\boldsymbol{D}_{Q1},\boldsymbol{D}_{Q2},\cdots,\boldsymbol{D}_{Qm}\}$，不妨设 $\frac{|\boldsymbol{D}_{Q1}\wedge\boldsymbol{D}_X|}{|\boldsymbol{D}_{Q1}\vee\boldsymbol{D}_{X^C}|}\leqslant\frac{|\boldsymbol{D}_{Q2}\wedge\boldsymbol{D}_X|}{|\boldsymbol{D}_{Q2}\vee\boldsymbol{D}_{X^C}|}\leqslant\cdots\leqslant\frac{|\boldsymbol{D}_{Qm}\wedge\boldsymbol{D}_X|}{|\boldsymbol{D}_{Qm}\vee\boldsymbol{D}_{X^C}|}$，根据定理 2，得到

$$S(X,\underline{R}(X))\leqslant S(X,Q_1\cup\underline{R}(X))\leqslant\cdots\leqslant S(X,\underline{R}(X))\cup Q_1\cup Q_2\cup\cdots Q_m)=S(X,\overline{R}(X))$$

即 X 在近似空间(U,R)上的最优近似集为 $\overline{R}(X)$。类似得到：

性质 4　若对于任意的 $\boldsymbol{D}_P\in M$，有 $\frac{|\boldsymbol{D}_P\wedge\boldsymbol{D}_X|}{|\boldsymbol{D}_P\vee\boldsymbol{D}_{X^C}|}<\alpha$，则 X 在近似空间(U,R)上的最优近似集为 $\underline{R}(X)$。

定理 4　令 $N=\{\boldsymbol{D}_i|\frac{|\boldsymbol{D}_i\wedge\boldsymbol{D}_X|}{|\boldsymbol{D}_i\wedge\boldsymbol{D}_{X^C}|}\geqslant\alpha,\boldsymbol{D}_i\in M\}$，则 X 在近似空间(U,R)上的最优近似集为 $\underline{R}(X)\cup W^*$。其中 W^* 为集合矩阵 $W=\bigvee_{\boldsymbol{D}_i\in N}\boldsymbol{D}_K$ 对应的集合。

根据上述结果可以看出，目标概念在近似空间上最优近似集在不同条件下，结果不一样，这样也就是说，在属性约简上也应该根据具体决策系统来选择属性约简方法，而不是任意一种约简方法都可行。这为建立基于数据驱动的属性约简算法提供了一种思路。

4 粗糙集最优近似集的矩阵算法

设(U,R)为近似空间，$X\subseteq U$，X 最优近似集算法描述如下：

步骤 1　计算划分矩阵 $\boldsymbol{D}_{U/R}=\{\boldsymbol{D}_{X1},\boldsymbol{D}_{X2},\cdots,\boldsymbol{D}_{Xk}\}$，$\underline{\boldsymbol{D}}_R^{(X)}$，$\boldsymbol{D}_N^{(X)}=\overline{\boldsymbol{D}}_R^{(X)}-\underline{\boldsymbol{D}}_R^{(X)}$ $\alpha=\frac{|\underline{\boldsymbol{D}}_R^{(X)}|}{|\boldsymbol{D}_X|}$。

步骤 2　计算：$M=\{\boldsymbol{D}_m|\boldsymbol{D}_m\in\boldsymbol{D}_{U/R}$ 且 $\boldsymbol{D}_m\leqslant\boldsymbol{D}_N^{(X)}\}$。

步骤 3　计算：$N=\{\boldsymbol{D}_i|\frac{|\boldsymbol{D}_i\wedge\boldsymbol{D}_X|}{|\boldsymbol{D}_i\wedge\boldsymbol{D}_{X^C}|}\geqslant\alpha,\boldsymbol{D}_i\in M\}$。

步骤 4　判断与求解：① 若 $N=\Phi$，X 最优近似集为下近似集 $\underline{R}(X)$；② 若 $M=N$，则 X 最优近似集为上近似集 $\overline{R}(X)$；③ $N\subset\Phi$，则X 最优近似集为 $\underline{R}(X)\cup W^*$。

5 实例

设(U,R)为近似空间，U 为有限论域，R 为等价关系，由 R 导出的划分为 $U/R=\{X_1,X_2,X_3,X_4,X_5\}$，其中 $X_1=\{x_1,x_2,x_3\}$，$X_2=\{x_4,x_5\}$，$X_3=\{x_6,x_7,x_8\}$，$X_4=\{x_9\}$，$X_5=\{x_{10}\}$，由此得到矩阵算子与划分矩阵为

$$\boldsymbol{D}_R=\begin{bmatrix} \frac{1}{3} & \frac{1}{3} & \frac{1}{3} & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ \frac{1}{3} & \frac{1}{3} & \frac{1}{3} & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ \frac{1}{3} & \frac{1}{3} & \frac{1}{3} & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & \frac{1}{2} & \frac{1}{2} & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & \frac{1}{2} & \frac{1}{2} & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & \frac{1}{3} & \frac{1}{3} & \frac{1}{3} & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & \frac{1}{3} & \frac{1}{3} & \frac{1}{3} & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & \frac{1}{3} & \frac{1}{3} & \frac{1}{3} & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \end{bmatrix} \boldsymbol{D}_{X_1}=\begin{bmatrix}1\\1\\1\\0\\0\\0\\0\\0\\0\\0\end{bmatrix} \boldsymbol{D}_{X_2}=\begin{bmatrix}0\\0\\0\\1\\1\\0\\0\\0\\0\\0\end{bmatrix} \boldsymbol{D}_{X_3}=\begin{bmatrix}0\\0\\0\\0\\0\\1\\1\\1\\0\\0\end{bmatrix} \boldsymbol{D}_{X_4}=\begin{bmatrix}0\\0\\0\\0\\0\\0\\0\\0\\1\\0\end{bmatrix} \boldsymbol{D}_{X_5}=\begin{bmatrix}0\\0\\0\\0\\0\\0\\0\\0\\0\\1\end{bmatrix}$$

下面从 3 种情形对本文所提最优近似集算法作进一步说明：

1）计算 $\boldsymbol{D}_X=\begin{bmatrix}0\\0\\1\\1\\1\\1\\0\\0\\1\\0\end{bmatrix}$ 的最优近似集，根据前面定义得到：$\underline{\boldsymbol{D}}_R{}^{(X)}=\begin{bmatrix}0\\0\\0\\1\\1\\0\\0\\0\\1\\0\end{bmatrix}$，$\boldsymbol{M}=\begin{bmatrix}1\\1\\1\\0\\0\\1\\1\\1\\0\\0\end{bmatrix},\begin{bmatrix}0\\0\\0\\0\\0\\1\\1\\1\\0\\0\end{bmatrix}=\{\boldsymbol{D}_{X_1},\boldsymbol{D}_{X_2}\}$，$\alpha=\frac{3}{5}$。

$\frac{|\boldsymbol{D}_{X_1}\wedge\boldsymbol{D}_{X_1}|}{|\boldsymbol{D}_{X_1}\wedge\boldsymbol{D}_{X^C}|}=\frac{1}{2}=\frac{|\boldsymbol{D}_{X_1}\wedge\boldsymbol{D}_{X_3}|}{|\boldsymbol{D}_{X_3}\wedge\boldsymbol{D}_{X^C}|}$，所以 X 最优近似集为 $\underline{R}(X)$，事实上这时 $S(X,\overline{R}(X))=\frac{5}{9}$。

2）计算 $\boldsymbol{D}_X=\begin{bmatrix}0\\1\\1\\0\\0\\1\\0\\0\\1\\0\end{bmatrix}$ 的最优近似集，根据定义得到：$\underline{\boldsymbol{D}}_R{}^{(X)}=\begin{bmatrix}0\\0\\0\\0\\0\\0\\0\\0\\1\\0\end{bmatrix}$，$\boldsymbol{M}=\{\boldsymbol{D}_{X_1},\boldsymbol{D}_{X_3}\}$，$\alpha=\frac{1}{4}$。

$\frac{|\boldsymbol{D}_X\wedge\boldsymbol{D}_{X_1}|}{|\boldsymbol{D}_{X_1}\wedge\boldsymbol{D}_{X^C}|}=2$，$\frac{|\boldsymbol{D}_X\wedge\boldsymbol{D}_{X_3}|}{|\boldsymbol{D}_{X_3}\wedge\boldsymbol{D}_{X^C}|}=\frac{1}{2}$，所以 X 最优近似集为 $\overline{R}(X)$，事实上这时 $S(X,\overline{R}(X))=\frac{4}{7}$。

3）计算 $\boldsymbol{D}_X=\begin{bmatrix}0\\0\\1\\1\\0\\1\\1\\1\\0\\0\end{bmatrix}$ 的最优近似集，由定义得到：$\boldsymbol{D}_{\underline{R}}^{(X)}=\begin{bmatrix}0\\0\\0\\0\\0\\1\\1\\1\\0\\0\end{bmatrix}$，$\boldsymbol{M}=\{\boldsymbol{D}_{X_1},\boldsymbol{D}_{X_2}\}$，$\alpha=\frac{3}{5}$。

$\frac{|\boldsymbol{D}_X\wedge\boldsymbol{D}_{X_1}|}{|\boldsymbol{D}_{X_1}\wedge\boldsymbol{D}_{X^C}|}=\frac{1}{2}$，$\frac{|\boldsymbol{D}_X\wedge\boldsymbol{D}_{X_2}|}{|\boldsymbol{D}_{X_2}\wedge\boldsymbol{D}_{X^C}|}=1$，所以 X 的最优近似集为 $\underline{R}(X)\cup X_2$，并且 $S(X,\underline{R}(X)\cup X_2)=\frac{2}{3}$，这时的 $S(X,\overline{R}(X))=\frac{5}{8}$。

6 结束语

由于近似空间上基本知识太粗造成概念的不确定性，一直以来目标概念采用上、下近似集进行范围表示。而最优近似概念的提出和最优近似集计算为我们提供一种新视角认识粗糙集度量与表示，为在决策系统边界域的不确定性规则知识的获取提供一种新方法，相比当前变精度粗糙集模型更具有泛化性能。文提出的基于矩阵最优近似集的判定与计算，不仅在理论上证明了计算方法有效，而且简单、直观、易于在计算机上编程实现，有利于建立新属性约简算法，促进粗糙集应用的发展。

参考文献：

[1] PAWLAK Z. Rough sets[J]. Computer and Information Science，1982，11(5)：341-356.

[2] 潘远，王卿. 结合粗糙集约简的纹理特征影像分类[J]. 华东交通大学学报，2016，33(1)：83-88.

[3] 徐鹏，蒋凯，王泽华，等. 基于粗糙集的道路交通事故客观因素显著性分析[J]. 华东交通大学学报，2017，34(6)：66-71.

[4] 张清华，王国胤，肖雨. 粗糙集的近似集[J]. 软件学报，2012，23(7)：1745-1759.

[5] 张清华，薛玉斌，王国胤. 粗糙集的最优近似集[J]. 软件学报，2016，27(2)：295-308.

[6] 张清华，薛玉斌，胡峰，等. 粗糙集近似集不确定性研究[J]. 电子学报，2016，44(7)：1574-1580.

[7] GUAN J W，BELL D A，GUAN Z.Matrix computation for information systems[J]. Information Sciences，2001，131(1/2/3/4)：129-156.

[8l] 罗来鹏，刘二根，曾毅. 粗糙集理论研究的矩阵方法[J]. 系统工程与电子技术，2009，31(4)：859-862.

[9] 王磊，李天瑞. 一种基于矩阵的知识粒度计算方法[J]. 模式识别与人工智能，2013，26(5)：447-453.

[10] 王磊，李天瑞，刘清，等. 对象集变化时近似集动态维护的矩阵方法[J]. 计算机研究与发展，2013，50(9)：1992-2004.

[11] XU W H，LI Y，LIAO X W. Approaches to attribute reductions based on rough set and matrix computation in inconsistent ordered information systems[J]. Knowledge-Based Systems，2012，27(3)：78-91.

Matrix Computation for Optimal Approximation Rough Set

Luo Laipeng, Liu Ergen, Fan Zizhu

(School of Sciences, East China Jiaotong University, Nanchang 330013, China)

Abstract: The upper and lower approximation sets of rough sets derived from approximate space are the core foundation of the rough set theory and its applied research. However, this description only gives a range of rough sets. To depict the target concept and promote rough set application more accurately, this paper, according to the computation problem of the optimal approximate set of rough sets, proposed a calculation method based on matrixes. Then, some concepts such as Boolean matrix, relation matrix, set matrix and similarity degree matrix were introduced. By discussing the similarity degree between the target concept and its approximate set, some relevant rules were obtained. A new judgment and algorithm of optimal approximate set based on matrix for rough set was eventually established. Finally, a practical example was given to further illustrate results above.

Key words: rough set; optimal approximation set; similarity degree; matrix