

商近似空间中粗糙集的研究

赵焕敏 马利文

(北京邮电大学理学院,北京 100876)

摘要:本文从新的角度研究粗糙集,即基于覆盖近似空间上的一种等价关系,定义了覆盖近似空间上的商近似空间和此商近似空间上的 2 对上下近似算子,并详细研究了这 2 对近似算子的性质.

关键词:覆盖粗糙集;邻域;余邻域;商近似空间

中图分类号:TP18

DOI:10.19789/j.1004-9398.2020.02.002

0 引言

覆盖粗糙集理论作为 Pawlak 经典粗糙集理论的一种推广,在解决智能系统中的许多复杂问题方面得到应用,有丰富的理论价值.越来越多的学者从不同的角度来研究粗糙集理论(如文献[1-11]).在 Pawlak 经典粗糙集理论和其他广义粗糙集理论中,研究的核心内容是空间中粗糙子集的上下近似,所以很多文献中,定义和研究了覆盖近似空间上各种不同的上下近似算子.本文从新的思路出发,基于覆盖近似空间上的一种等价关系,定义了一种新的近似空间:覆盖近似空间的商近似空间.通过给定源空间中元素的邻域和余邻域的概念,定义了所研究商近似空间上的 2 对上下近似算子,并详细介绍其具有的良好性质,结合文献[12]中对商空间性质的保持性的证明,给出一些例子作理解,为进一步解决实际问题提供新的方法和理论支撑.

1 基础概念

本节简单介绍覆盖粗糙集理论中的一些基础概念及相关性质.

定义 1^[10] 设 U 是一个有限集合, \mathcal{C} 是由 U 的非空子集构成的集族且 $U = \bigcup_{C \in \mathcal{C}} C$, 则称集族 \mathcal{C} 是集合 U 的一个覆盖,且称有序对 (U, \mathcal{C}) 是一个覆盖近似空间.

设 U 是全集,对于 U 的每一个子集 X ,用 $-X$ 表示它的余集.

定义 2^[2,9-10] 设 (U, \mathcal{C}) 是一个覆盖近似空间.对于任一 $x \in U$,称

$$N(x) = \bigcap \{C \in \mathcal{C} : x \in C\},$$

$$M(x) = -(\bigcup \{C \in \mathcal{C} : x \notin C\})$$

分别为元素 x 的邻域和余邻域.

明显地,余邻域的定义还可以表示为 $M(x) = \bigcap \{-C : (C \in \mathcal{C}) \wedge (x \notin C)\}$.若对于任一 $C \in \mathcal{C}$,有 $x \in C$,则规定 $M(x) = U$.

定义 3^[3] 设 (U, \mathcal{C}) 是一个覆盖近似空间, $A \subset U$,称

$$N(A) = \bigcup \{N(x) : x \in A\},$$

$$M(A) = \bigcup \{M(x) : x \in A\}$$

分别为集合 A 的邻域和余邻域.

命题 1^[2] 设 (U, \mathcal{C}) 是一个覆盖近似空间, $x, y \in U$,则 $y \in N(x)$ 当且仅当 $x \in M(y)$.

命题 2^[2,10] 设 (U, \mathcal{C}) 是一个覆盖近似空间,则对于每一 $C \in \mathcal{C}$,有

$$C = \bigcup_{x \in C} N(x), \quad -C = \bigcup_{x \in -C} M(x).$$

命题 3^[2] 邻域算子 N 和余邻域算子 M 都满足传递性,即若 $x \in N(y)$ 且 $y \in N(z)$,则 $x \in N(z)$;若 $x \in M(y)$ 且 $y \in M(z)$,则 $x \in M(z)$.

用 $\underline{P}(A)$ 表示集合 A 的下近似, $\overline{P}(A)$ 表示集合 A 的上近似.

定义 4^[2] 对于覆盖近似空间 (U, \mathcal{C}) 的每一个子集 A ,定义 A 的下近似和上近似分别为:

$$\underline{P}_1(A) = \{x : N(x) \subset A\},$$

$$\overline{P}_1(A) = \{x : N(x) \cap A \neq \emptyset\} = \bigcup_{x \in A} M(x);$$

$$\underline{P}_2(A) = \{x : M(x) \subset A\},$$

$$\overline{P}_2(A) = \{x : M(x) \cap A \neq \emptyset\} = \bigcup_{x \in A} N(x).$$

命题 4 设 (U, \mathcal{C}) 是一个覆盖近似空间, 则对于每一 $C \in \mathcal{C}$, 有

$$\begin{aligned} \underline{P}_1(C) &= C, & \overline{P}_1(-C) &= -C, \\ \underline{P}_2(-C) &= -C, & \overline{P}_2(C) &= C. \end{aligned}$$

2 商近似空间上的粗糙集

本节定义了 1 个覆盖近似空间上的商近似空间和此商近似空间中子集的 2 对上下近似.

定义 5 设 (U, \mathcal{C}) 是 1 个覆盖近似空间, R 是 U 中的 1 个等价关系, 则 $W = U/R$ 是 U 在等价关系 R 下的 1 个商空间. 令 $\mathcal{F} = \{F \subset W: R^{-1}(F) = N(R^{-1}(F))\}$, 即集族 \mathcal{F} 由商空间中原像是源空间中的邻域或邻域的并的集合构成. 容易验证, \mathcal{F} 是 W 的 1 个覆盖, 称 (W, \mathcal{F}) 是源空间 (U, \mathcal{C}) 在等价关系 R 下的商近似空间.

例 1 令 $U = \{x_1, x_2, x_3, x_4, x_5, x_6, x_7, x_8, x_9, x_{10}\}$, 记 $C_1 = \{x_1, x_4\}$, $C_2 = \{x_2, x_3\}$, $C_3 = \{x_1, x_2, x_3\}$, $C_4 = \{x_4, x_5, x_6, x_7\}$, $C_5 = \{x_6, x_7, x_8, x_9\}$, $C_6 = \{x_4, x_5, x_6, x_7, x_{10}\}$, 则 $\mathcal{C} = \{C_1, C_2, C_3, C_4, C_5, C_6\}$ 是 U 的 1 个覆盖. 对于覆盖近似空间 (U, \mathcal{C}) , 若等价关系 R 诱导 U 上的一个划分为 $y_1 = \{x_1\}$, $y_2 = \{x_2, x_3, x_4, x_5\}$, $y_3 = \{x_6, x_7\}$, $y_4 = \{x_8, x_9\}$, $y_5 = \{x_{10}\}$, 则对于 $W = U/R = \{y_1, y_2, y_3, y_4, y_5\}$, 通过对元素邻域的计算, 得到 $F_1 = \{y_1\}$, $F_2 = \{y_2, y_3\}$, $F_3 = \{y_3\}$, $F_4 = \{y_3, y_4\}$, $F_5 = \{y_2, y_3, y_5\}$, 则 $\mathcal{F} = \{F_1, F_2, F_3, F_4, F_5\}$ 是 W 的 1 个覆盖且 (W, \mathcal{F}) 是源空间 (U, \mathcal{C}) 的商近似空间.

定义 6 若 (W, \mathcal{F}) 是源空间 (U, \mathcal{C}) 的商近似空间, 则对于任一 $y \in W$, 称

$$\begin{aligned} n(y) &= \bigcap \{F \in \mathcal{F}: y \in F\}, \\ m(y) &= \bigcap \{-F: (F \in \mathcal{F}) \wedge (y \notin F)\} \end{aligned}$$

分别为元素 y 的商邻域和商余邻域. 对于任一子集 $Y \subset W$, 称

$$\begin{aligned} n(Y) &= \bigcup \{n(y): y \in Y\}, \\ m(Y) &= \bigcup \{m(y): y \in Y\} \end{aligned}$$

分别为集合 Y 的商邻域和商余邻域.

例 2 考虑例 1 的源空间和商近似空间, 取 $Y = \{y_2, y_3\}$, 通过计算得到 $n(Y) = \{y_2, y_3\}$, $m(Y) = \{y_2, y_3, y_4, y_5\}$.

容易验证, 在下列性质方面, 商近似算子 n 、 m 与近似算子 N 、 M 相同.

命题 5 设 (W, \mathcal{F}) 是源空间 (U, \mathcal{C}) 的商近似空间, 则商邻域算子 n 和商余邻域算子 m 都满足传

递性, 且对于任意 $y, z \in W, y \in n(z)$, 当且仅当 $z \in m(y)$.

命题 6 设 (W, \mathcal{F}) 是源空间 (U, \mathcal{C}) 的商近似空间, 则对于任一 $F \in \mathcal{F}$, 有 $F = n(F)$, $-F = m(-F)$, $R^{-1}(-F) = M(R^{-1}(-F))$.

证明 假设 $F \in \mathcal{F}, n(F) = \bigcup \{n(y): y \in F\}$, 由双包含关系得到 $F = n(F)$, 对于第 2 个等式, 很明显, 左包含 $-F \subseteq \bigcup \{m(y): y \in -F\}$ 成立, 且对于任一 $y \in -F$, 有 $m(y) \subseteq -F$ 成立, 所以 $\bigcup \{m(y): y \in -F\} \subseteq -F$, 故 $-F = \bigcup \{m(y): y \in -F\}$.

对于第 3 个等式, 容易验证, $R^{-1}(-F) \subseteq \bigcup \{M(x): x \in R^{-1}(-F)\}$, 只需要证明 $R^{-1}(-F) \supseteq \bigcup \{M(x): x \in R^{-1}(-F)\}$. 若 $x \in R^{-1}(-F) = -R^{-1}(F)$ 且 $M(x) \not\subseteq R^{-1}(-F)$, 则存在元素 $z \in M(x)$ 且 $z \notin R^{-1}(-F) = -R^{-1}(F)$, 所以 $z \in R^{-1}(F)$ 并 $N(z) \subseteq R^{-1}(F)$. 根据命题 1, 对任意 $x, y \in U, y \in N(x)$ 当且仅当 $x \in M(y)$, 得到 $x \in N(z)$, 所以 $x \in R^{-1}(F)$, 这与 $x \in R^{-1}(-F) = -R^{-1}(F)$ 矛盾. 这证明对任一 $x \in R^{-1}(-F)$, 有 $M(x) \subseteq R^{-1}(-F)$, 故 $R^{-1}(-F) \supseteq \bigcup \{M(x): x \in R^{-1}(-F)\}$.

下面定义商近似空间中子集的上下近似.

定义 7 设 (W, \mathcal{F}) 是源空间 (U, \mathcal{C}) 的商近似空间且 $Y \subset W$, 定义子集 Y 的 2 对下近似和上近似分别为:

$$\begin{aligned} \underline{q}_1(Y) &= \{y \in W: n(y) \subset Y\}, \\ \overline{q}_1(Y) &= \{y \in W: n(y) \cap Y \neq \emptyset\} = m(Y); \\ \underline{q}_2(Y) &= \{y \in W: m(y) \subset Y\}, \\ \overline{q}_2(Y) &= \{y \in W: m(y) \cap Y \neq \emptyset\} = n(Y). \end{aligned}$$

例 3 考虑例 1 的源空间和商近似空间, 取 $Y_1 = \{y_1, y_2\}, Y_2 = \{y_2, y_3, y_4\}$, 通过计算得到:

$$\begin{aligned} \underline{q}_1(Y_1) &= \{y_1\}, & \overline{q}_1(Y_1) &= \{y_1, y_2, y_5\}; \\ \underline{q}_2(Y_1) &= \{y_1\}, & \overline{q}_2(Y_1) &= \{y_1, y_2, y_3\}; \\ \underline{q}_1(Y_2) &= \{y_2, y_3, y_4\}, & \overline{q}_1(Y_2) &= \{y_2, y_3, y_4, y_5\}; \\ \underline{q}_2(Y_2) &= \{y_4\}, & \overline{q}_2(Y_2) &= \{y_2, y_3, y_4\}. \end{aligned}$$

3 商近似空间中近似算子的性质

商近似算子 $\underline{q}_1, \overline{q}_1, \underline{q}_2, \overline{q}_2$ 具有与近似算子 $\underline{P}_1, \overline{P}_1, \underline{P}_2, \overline{P}_2$ 类似的拓扑性质^[2].

命题 7 下近似算子 $\underline{q}_1, \underline{q}_2$ 都是内部算子; 上近似算子 $\overline{q}_1, \overline{q}_2$ 都是闭包算子.

由定义 5 和命题 6, 容易验证命题 8.

命题 8 设 (W, \mathcal{F}) 是源空间 (U, \mathcal{E}) 的商近似空间, 则对于任一 $F \in \mathcal{F}$, 有:

$$\begin{aligned} q_1(F) &= F, & \bar{q}_1(-F) &= -F; \\ q_2(-F) &= -F, & \bar{q}_2(F) &= F; \\ \underline{P}_1(R^{-1}(F)) &= R^{-1}(F); \\ \bar{P}_1(R^{-1}(-F)) &= R^{-1}(-F), \\ \underline{P}_2(R^{-1}(-F)) &= R^{-1}(-F); \\ \bar{P}_2(R^{-1}(F)) &= R^{-1}(F). \end{aligned}$$

基于此, 命题 9 可以直接得到.

命题 9 设 (W, \mathcal{F}) 是源空间 (U, \mathcal{E}) 的商近似空间, 则对于任一 $F \in \mathcal{F}$, 有:

$$\begin{aligned} \underline{P}_1(R^{-1}(F)) &= R^{-1}(\underline{q}_1(F)) = R^{-1}(F), \\ \bar{P}_1(R^{-1}(-F)) &= R^{-1}(\bar{q}_1(-F)) = R^{-1}(-F); \\ \underline{P}_2(R^{-1}(-F)) &= R^{-1}(\underline{q}_2(-F)) = R^{-1}(-F), \\ \bar{P}_2(R^{-1}(F)) &= R^{-1}(\bar{q}_2(F)) = R^{-1}(F). \end{aligned}$$

下面讨论商邻域和商余邻域在等价关系 R 下逆像的性质.

命题 10 设 (W, \mathcal{F}) 是源空间 (U, \mathcal{E}) 在等价关系 R 下的商近似空间, 则对于任一 $y \in W$, 有:

$$\begin{aligned} (1) R^{-1}(n(y)) &= N(R^{-1}(n(y))); \\ (2) R^{-1}(m(y)) &= M(R^{-1}(m(y))). \end{aligned}$$

证明 根据定义 6 和命题 6, 可以得到:

$$\begin{aligned} (1) R^{-1}(n(y)) &= R^{-1}(\bigcap_{y \in F} F) = \bigcap_{y \in F} R^{-1}(F) = \\ &= \bigcap_{y \in F} \bigcup_{x \in R^{-1}(F)} N(x) = \bigcup_{x \in \bigcap_{y \in F} R^{-1}(F)} N(x) = \\ &= \bigcup_{x \in R^{-1}(\bigcap_{y \in F} F)} N(x) = \bigcup_{x \in R^{-1}(n(y))} N(x) = \\ &= N(R^{-1}(n(y))); \\ (2) R^{-1}(m(y)) &= R^{-1}(\bigcap_{y \in -F} (-F)) = \\ &= \bigcap_{y \in -F} R^{-1}(-F) = \bigcap_{y \in -F} \bigcup_{x \in R^{-1}(-F)} M(x) = \\ &= \bigcup_{x \in \bigcap_{y \in -F} R^{-1}(-F)} M(x) = \bigcup_{x \in R^{-1}(\bigcap_{y \in -F} (-F))} M(x) = \\ &= \bigcup_{x \in R^{-1}(m(y))} M(x) = M(R^{-1}(m(y))). \end{aligned}$$

接下来, 研究近似集在等价关系 R 下像和逆像的性质.

命题 11 设 (W, \mathcal{F}) 是源空间 (U, \mathcal{E}) 在等价关系 R 下的商近似空间, 则有:

(1) 对于任一 $X \subset U$, $R(\bar{P}_1(X)) \subset \bar{q}_1(R(X))$, $R(\bar{P}_2(X)) \subset \bar{q}_2(R(X))$ 成立;

(2) 对于任一 $Y \subset W$, $\bar{P}_1(R^{-1}(Y)) \subset R^{-1}(\bar{q}_1(Y))$, $\bar{P}_2(R^{-1}(Y)) \subset R^{-1}(\bar{q}_2(Y))$, $R^{-1}(\underline{q}_1(Y)) \subset \underline{P}_1(R^{-1}(Y))$, $R^{-1}(\underline{q}_2(Y)) \subset \underline{P}_2(R^{-1}(Y))$ 成立.

证明 (1) 因为 $R(X) \subset m(R(X)) = \bar{q}_1(R(X))$, 所以 $X \subset R^{-1}(m(R(X))) = R^{-1}(\bigcup_{y \in R(X)} m(y)) = \bigcup_{y \in R(X)} R^{-1}(m(y)) = \bigcup_{y \in R(X)} (\bigcup_{x \in R^{-1}(m(y))} M(x))$, 则对于任一 $z \in X$, 等式右边存在一个 $M(x)$ 使得 $z \in M(x)$, 所以 $M(z) \subset M(x)$, 从而得到 $M(X) \subset \bigcup_{y \in R(X)} (\bigcup_{x \in R^{-1}(m(y))} M(x)) = R^{-1}(m(R(X)))$, 即 $\bar{P}_1(X) \subset R^{-1}(\bar{q}_1(R(X)))$, 故 $R(\bar{P}_1(X)) \subset \bar{q}_1(R(X))$.

同理, 可以证得对于任一 $X \subset U$, $R(\bar{P}_2(X)) \subset \bar{q}_2(R(X))$.

(2) 根据 (1), 有 $R(\bar{P}_1(R^{-1}(Y))) \subset \bar{q}_1(R(R^{-1}(Y))) = \bar{q}_1(Y)$, 所以 $\bar{P}_1(R^{-1}(Y)) \subset R^{-1}(\bar{q}_1(Y))$.

根据上面的关系式, 可以得到 $R^{-1}(\underline{q}_1(Y)) = R^{-1}(-\bar{q}_1(-Y)) = -R^{-1}(\bar{q}_1(-Y)) \subset -\bar{P}_1(R^{-1}(-Y)) = -\bar{P}_1(-R^{-1}(Y)) = \bar{P}_1(R^{-1}(Y))$.

同样的方法可以证得 $\bar{P}_2(R^{-1}(Y)) \subset R^{-1}(\bar{q}_2(Y))$, 且 $R^{-1}(\bar{q}_2(Y)) \subset \bar{P}_2(R^{-1}(Y))$.

举 1 个例子说明上述性质中的等号不成立.

例 4 对于例 1 中的源空间和商近似空间, 令 $X = \{x_2, x_3\}$, 则 $R(X) = \{y_2\}$, $\bar{P}_1(X) = \{x_2, x_3\}$, $\bar{P}_2(X) = \{x_2, x_3\}$, 通过计算可以得到

$$\begin{aligned} R(\bar{P}_1(X)) &= \{y_2\} \subset \{y_2, y_5\} = \bar{q}_1(R(X)), \\ R(\bar{P}_2(X)) &= \{y_2\} \subset \{y_2, y_3\} = \bar{q}_2(R(X)). \end{aligned}$$

令 $Y_1 = \{y_3\}$, $Y_2 = \{y_5\}$, $Y_3 = Y_4 = \{y_2\}$, 则

$$\begin{aligned} \bar{q}_1(Y_1) &= \{y_2, y_3, y_4, y_5\}, \bar{q}_2(Y_2) = \{y_2, y_3, y_5\}, \\ \bar{q}_1(Y_3) &= \bar{q}_1(Y_4) = \emptyset, \end{aligned}$$

$$R^{-1}(Y_1) = \{x_6, x_7\}, R^{-1}(Y_2) = \{x_{10}\},$$

$$R^{-1}(Y_3) = R^{-1}(Y_4) = \{x_2, x_3, x_4, x_5\}.$$

且下列各式成立:

$$\begin{aligned} \bar{P}_1(R^{-1}(Y_1)) &= \{x_5, x_6, x_7, x_8, x_9, x_{10}\} \subset \\ &= \{x_2, x_3, x_4, x_5, x_6, x_7, x_8, x_9, x_{10}\} = R^{-1}(\bar{q}_1(Y_1)), \\ \bar{P}_2(R^{-1}(Y_2)) &= \{x_4, x_5, x_6, x_7, x_{10}\} \subset \\ &= \{x_2, x_3, x_4, x_5, x_6, x_7, x_{10}\} = R^{-1}(\bar{q}_2(Y_2)), \\ R^{-1}(\bar{q}_1(Y_3)) &= \emptyset \subset \{x_2, x_3, x_4\} = \bar{P}_1(R^{-1}(Y_3)), \\ R^{-1}(\bar{q}_2(Y_4)) &= \emptyset \subset \{x_2, x_3\} = \bar{P}_2(R^{-1}(Y_4)). \end{aligned}$$

4 结论和讨论

在大数据分析和应用中需对数据集进行分类, 把所有类作为新的元素进行研究, 这便是商空间.

不仅减少了研究对象的数量,而且由于每类元素整体作为一个对象,使其共性得以保留,这点在文献[12]中也被详细证明.本文定义了1个覆盖近似空间上的商近似空间,基于邻域和余邻域的概念对此商近似空间作出相关性质的证明和研究.由于邻域和余邻域是近似空间中的点的基本描述元,代表点的性质的基本单位,因而命题10充分说明了商近似空间很好地保留了源空间的邻域与余邻域所代表的性质.这也是商空间理论的初衷,对后面应用新定义的商近似空间模型解决属性约简、决策规则发现等实际问题提供了重要理论基础.事实上,对覆盖近似空间 (U, \mathcal{C}) 的商空间 $W = U/R$,自然想到定义其覆盖为 $\mathcal{S} = \{R(C) :$

$C \in \mathcal{C}\}$,文献[11]通过探究映射 f 所满足的条件,对得到的映射空间做了初步分析.但由这样定义覆盖得到的邻域 $\{n(y) : y \in W\}$ 与余邻域 $\{m(y) : y \in W\}$ 不具有本文定义的商近似空间所具有的命题10.读者可以利用例1的数据自行检验这一点.这便是本文所定义的商近似空间的优越性所在.

命题11表明本文所定义的商近似空间中的上下近似与源空间的上下近似有一种确定的重要的关系,此性质也是上述覆盖 \mathcal{S} 定义法不能具备的.这一点对算子的应用是非常有益的,使得其上下近似的变化能被很好的把握,是放大还是缩小,对结合实例进一步研究粗糙集提供了重要启示.

参 考 文 献

- [1] LIU G L, SAI Y. A comparison of two types of rough sets induced by coverings[J]. International Journal of Approximate Reasoning, 2009, 50 (3) :521 – 528.
- [2] MA L W. On some types of neighborhood-related covering rough sets [J]. International Journal of Approximate Reasoning, 2012, 53(6) :901 – 911.
- [3] MA L W. Some twin approximation operators on covering approximation spaces[J]. International Journal of Approximate Reasoning, 2015, 56:59 – 70.
- [4] POMYKALA J A. Approximation operations in approximation space [J]. Bulletin of the Polish Academy of Science, 1987, 35:653 – 662.
- [5] QIN K Y, GAO Y, PEI Z. On covering rough sets [C]// Yao J, Lingras P, Wu W Z. Rough sets and knowledge technology. Heidelberg:Springer, 2007: 34 – 41.
- [6] YAO Y Y. Relational interpretations of neighborhood operators and rough set approximation operators[J]. Information Sciences, 1998, 111:239 – 259.
- [7] YANG X B, ZHANG M, DOU H L, et al. Neighborhood systems-based rough sets in incomplete information system[J]. Knowledge-Based Systems, 2011, 24:858 – 867.
- [8] YUN Z Q, GE X, BAI X L. Axiomatization and conditions for neighborhoods in a covering to form a partition[J]. Information Sciences, 2011, 181(9) :1735 – 1740.
- [9] ZHU W, WANG F Y. On three types of covering-based rough sets [J]. IEEE Transactions on Knowledge and Data Engineering, 2007, 19 (8) :1131 – 1144.
- [10] ZHU W. Topological approaches to covering rough sets[J]. Information Sciences, 2007, 177 (6) :1499 – 1508.
- [11] WANG C Z, CHEN D G, SUN B Q, et al. Communication between information systems with covering based rough sets [J]. Information Sciences, 2012, 216:17 – 33.
- [12] 张铃,张钺. 问题求解理论及应用[M]. 2版. 北京:清华大学出版社, 2007:18 – 27.

Study of Rough Sets in Quotient Approximation Spaces

ZHAO Huanmin MA Liwen

(School of Science, Beijing University of Posts and Telecommunications, Beijing 100876)

Abstract: From a new point of view, this study defined the quotient approximation space of a covering approximation space with a equivalent relation on it. Then this study defined two pairs of lower and upper approximation operators on it and investigated their properties in detail.

Keywords: covering rough set; neighborhood; complementary neighborhood; quotient approximation space