

## 分段延迟代价敏感三支决策\*

徐健锋<sup>1,2</sup>, 苗夺谦<sup>1</sup>, 张远健<sup>1</sup>

<sup>1</sup>(同济大学 电子与信息工程学院, 上海 201804)

<sup>2</sup>(南昌大学 软件学院, 江西 南昌 330047)

通信作者: 苗夺谦, E-mail: dqmiao@tongji.edu.cn



**摘要:** 决策粗糙集理论中, 三支决策代价目标函数是典型的单调线性函数. 然而, 在实践经验中经常发现延迟决策的代价与决策概率之间的函数关系往往呈现非单调特性, 决策粗糙集理论的经典代价敏感三支决策模型无法对上述非单调现象进行直接的建模和推理, 导致决策粗糙集理论的应用受到了限制. 为了解决这种具有非单调延迟代价的代价敏感三支决策问题, 提出一种新型分段延迟代价敏感三支决策模型. 该模型定义了具有单调递增和单调递减特性的两组延迟决策损失函数, 并结合经典正负域决策损失函数构造了分段延迟三支决策代价目标函数体系、度量指标和分段决策策略; 然后, 基于条件概率、损失函数及基础度量指标之间关系的 4 种分段延迟代价敏感三支决策分类模式被提了出来, 并且对相应的三支分类阈值进行了推理; 最后, 通过一组典型实例, 验证了分段延迟代价敏感三支决策模型及其三支分类是可行的.

**关键词:** 三支决策; 代价敏感; 目标函数; 分段延迟

**中图法分类号:** TP18

中文引用格式: 徐健锋, 苗夺谦, 张远健. 分段延迟代价敏感三支决策. 软件学报, 2022, 33(10): 3754–3775. <http://www.jos.org.cn/1000-9825/6302.htm>

英文引用格式: Xu JF, Miao DQ, Zhang YJ. Piece-wise Delay Cost-sensitive Three-way Decisions. Ruan Jian Xue Bao/Journal of Software, 2022, 33(10): 3754–3775 (in Chinese). <http://www.jos.org.cn/1000-9825/6302.htm>

## Piece-wise Delay Cost-sensitive Three-way Decisions

XU Jian-Feng<sup>1,2</sup>, MIAO Duo-Qian<sup>1</sup>, ZHANG Yuan-Jian<sup>1</sup>

<sup>1</sup>(College of Electronics and Information Engineering, Tongji University, Shanghai 201804, China)

<sup>2</sup>(School of Software, Nanchang University, Nanchang 330047, China)

**Abstract:** In classic decision-theoretic rough sets (DTRS), the cost objective function of three-way decision is typical monotone linear function. However, it is usually found that the functional relationship between delay decision's cost and probability is non monotonic in practical experience. Hence, the classical cost sensitive three-way decision model in DTRS is not suitable for modeling and reasoning this non monotonic phenomenon. In order to solve the non-monotonic phenomena in the cost sensitive three-way decision problem, a novel piece-wise delay cost sensitive three-way decision model is proposed based on the classical positive/negative domain decision loss functions. The novel model defines two different sets of delay decision loss functions which have the characteristics of monotonous increase and monotonic decrease, and constructs segmented delay three-way decision cost objective function systems, measurement indexes, and segment decision strategies. Then, on the basis of the relationship among conditional probability, loss function and basic metrics, segmented delay cost sensitivity three-way decision model is proposed, and the corresponding three-way classification thresholds reasoning are implemented. Finally, a group of typical analysis examples are used to verify the feasibility of the segmented delay cost sensitivity three-way decision model and its classification.

**Key words:** three-way decisions; cost-sensitive; objective function; segment delay

\* 基金项目: 国家自然科学基金(61763031, 61673301, 61976158); 江西省自然科学基金(20202BAB202018)

收稿时间: 2020-04-04; 修改时间: 2020-09-08; 采用时间: 2020-12-11

三支决策(three-way decisions, 3WD)<sup>[1,2]</sup>理论是在粗糙集理论<sup>[3,4]</sup>研究基础上拓展出的一种新型不确定问题求解思路. 该理论思想来源于人类处理不确定问题时常用的朴素思想, 即具有充分信息的对象往往被快速决策, 而信息不充分的对象往往被延迟决策, 并且等待进一步处理.

三支决策理论目前已成为粒计算与知识发现领域中一个重要的研究方向, 其在理论模型和应用领域获得广泛研究. 三支决策理论模型方面的相关研究包括序贯三支决策<sup>[5]</sup>、三支聚类<sup>[6]</sup>、三支区间集<sup>[7]</sup>和三支模糊集<sup>[8,9]</sup>等, 应用领域的相关研究包括垃圾邮件过滤<sup>[10]</sup>、恶意软件分析<sup>[11]</sup>、人脸识别<sup>[12]</sup>、流计算<sup>[13]</sup>, 以及推荐系统<sup>[14]</sup>等.

决策理论粗糙集是三支决策理论中的重要研究发现之一, 其在经典三支决策模型的基础上引入了贝叶斯决策思想, 提出了一种代价敏感三支决策方法<sup>[15]</sup>.

决策粗糙集理论和应用研究的成果非常丰富, 例如, 文献[16]提出了以时间、金钱为代价的序独立三支决策应用模型, 文献[17]进行了基于代价最小化的属性约简研究, 文献[18]提出了一种多代价结合的混合式代价敏感决策粗糙集模型, 文献[19]研究了基于决策粗糙集模型的动态计算问题.

文献[20]则将经典多粒度及双量化研究与决策粗糙集进行了结合, 提出了多粒度双量化决策粗糙集理论框架. 文献[21]则提出一种代价敏感三支决策与深度神经网络相融合的深度学习模型. 上述研究显示了决策粗糙集丰富的发展活力.

决策粗糙集理论中, 决策损失函数可以看作是由专家给出的不同决策行为的风险系数, 是决策代价计算的基础. 在延迟决策代价目标函数中, 延迟决策的代价与决策条件概率之间存在着单调线性关系, 即: 该目标函数中的延迟决策代价要么随着决策条件概率的增加而单调增加, 要么随着决策条件概率的增加而单调减少. 然而在现实世界中, 当决策条件概率极大和极小时, 该决策被延迟的代价往往高于决策条件概率取值中庸的不确定情况, 例如:

- 1) 在海关的安检决策场景中, 当某旅客被确定为危险人员的概率较大时, 往往会被海关果断扣押或被拒绝入关; 如果某旅客被大概率地确定为安全时, 往往会被海关立即放行. 同样, 当某位旅客被认为可疑人员时, 该旅客往往会被海关延期通关并且被要求接受进一步的审查. 显然, 当旅客为危险人员的概率极高或极低时, 该旅客被延期作进一步盘查的风险显然都较高. 同样, 当某位旅客既无法确定是危险人员也无法确认是安全旅客时, 其被延期通关并且被要求接受进一步安全审查的风险显然低于确定的危险分子或者确认的安全旅客;
- 2) 在大学招录学生的场景中, 当某申请的学生其成绩远远高于录取分数线时往往会被直接录取, 某申请的学生其成绩远远低于录取分数线, 往往也会直接被拒绝. 但是, 当某申请的学生成绩在录取分数线附近时往往会被推迟决定是否被录取. 显然, 对成绩远高于分数线或者远低于分数线的学生被延迟录取的风险都高于录取入学考试成绩在分数线附近的不确定性的考生.

由于经典决策粗糙集理论中采用的延迟决策代价目标函数具有明确的单调特征, 所以经典决策粗糙集无法应用于非单调延迟代价的决策现象. 为了有效解决具有非单调延迟代价现象的代价敏感三支决策难题, 本文提出了一种新型分段延迟代价敏感三支决策模型及其三支分类方法. 该模型针对延迟决策代价的非单调现象, 在经典决策粗糙集理论的基础上, 首次提出了具有单调递增和单调递减特性的两套延迟决策代价目标分段函数, 并且结合经典的正负域决策代价目标函数设计了一种新型分段延迟代价敏感三支决策模型及其三支决策推理. 本文的主要贡献在于拓展了一套新型的复杂代价敏感三支决策理论, 提出了一种非单调延迟三支决策的求解方法.

本文第 1 节主要介绍决策粗糙集的基础理论, 并且根据决策代价目标函数间的几何关系对代价敏感三支决策的阈值及代价关系进行讨论. 第 2 节首先针对决策延迟域代价目标函数的非单调现象, 提出两种具有单调递增和单调递减特性的延迟决策损失函数, 并且结合经典决策粗糙集理论中的正负域决策损失函数构造分段延迟代价敏感三支决策模型, 然后进行相应的代价最优分类推理研究. 第 3 节通过构造 4 组典型的分段延迟代价敏感三支决策模型及概率信息表实例, 以表明本文分段延迟代价敏感三支决策模型及其三支决策推理

的可行性. 第 4 节是对全文的总结及下一步研究计划的展望.

## 1 相关理论研究

### 1.1 三支决策理论

概率粗糙集<sup>[22]</sup>理论中一对确定的集合, 即上近似集与下近似集, 被用来表示具有不确定性的概念. 这两个近似集将论域划分为 3 个互不相交的区域: 正域、负域和边界域.

基于以上粗糙集理论, 加拿大里贾那大学的姚一豫教授将人类处理不确定问题的朴素三支分类思想引入到粗糙集中, 并且进而提出三支决策的概念. 三支决策理论的主要思想是将论域一分为三, 即在经典正负决策域的分类基础上引入延迟区域. 这种将决策或者论域划分为 3 部分的思想, 是人类处理不确定问题思维方式的理论化成果之一.

假设在某个论域空间  $U$  中,  $C \subseteq U$  表示一个目标概念, 条件概率  $\Pr(C|[x]_R) = (C \cap [x]_R) / [x]_R$  表示对象  $x \in U$  相对于  $C \subseteq U$  的隶属度,  $[x]_R \subseteq U$  是等价关系  $R$  的某个等价类.

三支决策的主要概念为: 存在一对阈值  $(\alpha, \beta)$ , 其中,  $0 \leq \beta < \alpha \leq 1$ , 则等价类  $[x]_R$  属于以下 3 种可能的近似分类.

- 1) 正决策域:  $POS_{(\alpha, \beta)}(C) = \{x \in U | \Pr(C|[x]_R) \geq \alpha\}$  或  $POS_{(\alpha, \beta)}(C) = \bigcup \{[x]_R \in U/R | \Pr(C|[x]_R) \geq \alpha\}$ , 其语义为: 如果  $\Pr(C|[x]_R) \geq \alpha$ , 则接受  $x \in C$  和  $[x]_R \in C$  (即正域决策);
- 2) 负决策域:  $NEG_{(\alpha, \beta)}(C) = \{x \in U | \Pr(C|[x]_R) \leq \beta\}$  或  $NEG_{(\alpha, \beta)}(C) = \bigcup \{[x]_R \in U/R | \Pr(C|[x]_R) \leq \beta\}$ , 其语义为: 如果  $\Pr(C|[x]_R) \leq \beta$ , 则拒绝  $x \in C$  (即负域决策);
- 3) 延迟决策域:  $BND_{(\alpha, \beta)}(C) = \{x \in U | \beta < \Pr(C|[x]_R) < \alpha\}$  或  $BND_{(\alpha, \beta)}(C) = \bigcup \{[x]_R \in U/R | \beta < \Pr(C|[x]_R) < \alpha\}$ , 其语义为: 如果  $\beta < \Pr(C|[x]_R) < \alpha$ , 则既不接受也不拒绝  $x \in C$  (即延迟决策).

上述公式显示: 正决策域集合可以由单个阈值  $\alpha$  控制, 负决策域集合可由单个阈值  $\beta$  控制, 延迟决策域集合取决于阈值对  $(\alpha, \beta)$ . 概率三支决策具有以下语义: 当具有确凿的证据和充分的信息时, 我们就会做出接受或拒绝的明确决定; 如果没有足够的信息或有力的证据, 我们宁愿采取更加谨慎的做法, 不做出明确的承诺, 或者延迟决策. 三支决策理论中的延迟决策使我们能够减少错误的接受或者错误的拒绝, 同时, 它也允许我们进一步增加更多的信息或证据, 从而做出更加正确的决定.

Pawlak 粗糙集<sup>[4]</sup>与二支决策是三支决策理论的两种特例<sup>[15]</sup>. 其中, Pawlak 粗糙集模型可视为  $(\alpha, \beta) = (1, 0)$ , 即正决策域与负决策域导出的决策规则都是具有 100% 准确率的一致性决策规则, 所有不一致的决策规则都属于延迟决策域.

二支决策可视为  $(\alpha, \beta) = (\gamma, \gamma)$ , 通常简记为  $\gamma$ . 该模型将所有对象划分为正决策域与负决策域. 此时, 正决策域导出决策规则的条件概率大于阈值  $\gamma$ , 而负决策域中决策规则的条件概率小于阈值  $\gamma$ . 阈值的变化会引起 3 个决策区域大小的变化, 同时获得相对最优三支决策或 3 个区域平衡的决策, 是三支决策的重要研究问题.

### 1.2 决策粗糙集理论

在决策粗糙集理论中, 不同的决策会产生不同的代价损失. 设  $\lambda_{PP}$  和  $\lambda_{PN}$  表示为正类元素  $x \in C$  和负类元素  $x \in C^c$  分别被划分为正域  $POS_{(\alpha, \beta)}(C)$  的损失函数, 其中,  $C^c = U - C$ ;  $\lambda_{NP}$  和  $\lambda_{NN}$  表示为正类元素  $x \in C$  和负类元素  $x \in C^c$  分别被划分为负域  $NEG_{(\alpha, \beta)}(C)$  的损失函数;  $\lambda_{BP}$  和  $\lambda_{BN}$  表示为正域元素  $x \in C$  和负域元素  $x \in C^c$  分别被划分为延迟域  $BND_{(\alpha, \beta)}(C)$  的损失函数.

**定义 1.** 当任一等价类  $[x]_R$  被分别划分至 3 个决策域  $POS_{(\alpha, \beta)}(C)$ 、 $NEG_{(\alpha, \beta)}(C)$ 、 $BND_{(\alpha, \beta)}(C)$  后, 其决策的代价可被描述为正域决策的代价目标函数  $T_P$ 、负域决策的代价目标函数  $T_N$ 、延迟决策的代价目标函数  $T_B$ :

$$\left. \begin{aligned} T_P &= \lambda_{PP} \Pr(C | [x]_R) + \lambda_{PN} \Pr(C^C | [x]_R) \\ T_N &= \lambda_{NP} \Pr(C | [x]_R) + \lambda_{NN} \Pr(C^C | [x]_R) \\ T_B &= \lambda_{BP} \Pr(C | [x]_R) + \lambda_{BN} \Pr(C^C | [x]_R) \end{aligned} \right\} \quad (1)$$

其中,  $C^C$  为  $C$  的补集.

而三支决策整体期望目标可以表示为

$$T(\alpha, \beta) = \sum_{[x]_R \in POS_{(\alpha, \beta)}(C)} T_P \times \Pr([x]_R) + \sum_{[x]_R \in NEG_{(\alpha, \beta)}(C)} T_N \times \Pr([x]_R) + \sum_{[x]_R \in BND_{(\alpha, \beta)}(C)} T_B \times \Pr([x]_R).$$

根据贝叶斯决策理论, 三支决策整体期望目标可以表示为  $\arg \max_{(\alpha, \beta)} T(\alpha, \beta)$ , 其语义是: 寻找到一对阈值  $(\alpha, \beta)$ , 使得  $T(\alpha, \beta)$  的整体代价最小. 由于  $\Pr(C|[x]_R) + \Pr(C^C|[x]_R) = 1$ , 三支决策代价目标函数可改写为如下形式:

$$\left. \begin{aligned} T_N(\Pr(C|[x]_R)) &= (\lambda_{NP} - \lambda_{NN})\Pr(C|[x]_R) + \lambda_{NN} \\ T_P(\Pr(C|[x]_R)) &= (\lambda_{PP} - \lambda_{PN})\Pr(C|[x]_R) + \lambda_{PN} \\ T_B(\Pr(C|[x]_R)) &= (\lambda_{BP} - \lambda_{BN})\Pr(C|[x]_R) + \lambda_{BN} \end{aligned} \right\} \quad (2)$$

经典决策粗糙集理论中通常约定  $(\lambda_{PP} - \lambda_{PN}) < 0$  和  $(\lambda_{NP} - \lambda_{NN}) > 0$ , 其中,  $(\lambda_{PP} - \lambda_{PN}) < 0$  的语义为正域决策代价  $T_P$  随着条件概论  $\Pr(C|[x]_R)$  的增加而减小, 即对象集合  $[x]_R$  被决策为正域  $C$  的代价随着条件概率论  $\Pr(C|[x]_R)$  的增加而减小;  $(\lambda_{NP} - \lambda_{NN}) > 0$  的语义为正域决策代价  $T_N$  随着条件概论  $\Pr(C|[x]_R)$  的增加而增大, 即对象集合  $[x]_R$  被决策为负域  $C^C$  的代价随着条件概率论  $\Pr(C|[x]_R)$  的减少而减小.

根据贝叶斯决策理论, 可以得到如下 3 条决策规则:

$$\left. \begin{aligned} (T_P \leq T_N) \wedge (T_P \leq T_B) &\rightarrow x \in POS_{(\alpha, \beta)}(C) \\ (T_N \leq T_P) \wedge (T_N \leq T_B) &\rightarrow x \in NEG_{(\alpha, \beta)}(C) \\ (T_B \leq T_P) \wedge (T_B \leq T_N) &\rightarrow x \in BND_{(\alpha, \beta)}(C) \end{aligned} \right\} \quad (3)$$

并且, 约定损失函数  $\lambda_{NP} > \lambda_{BP} > \lambda_{PP} > 0$  和  $\lambda_{PN} > \lambda_{BN} > \lambda_{NN} > 0$ , 此时将公式(2)中的代价目标函数  $T_P$ 、 $T_N$ 、 $T_B$  分别代入上述 3 条决策规则公式(3), 可以计算出:

$$\begin{aligned} \alpha &= \frac{\lambda_{PN} - \lambda_{BN}}{(\lambda_{PN} - \lambda_{BN}) + (\lambda_{BP} - \lambda_{PP})} \\ \beta &= \frac{\lambda_{BN} - \lambda_{NN}}{(\lambda_{BN} - \lambda_{NN}) + (\lambda_{NP} - \lambda_{BP})} \end{aligned} \quad (4)$$

显然, 公式(2)中  $T_P$ 、 $T_N$ 、 $T_B$  可以被看作是以  $\Pr(C|[x]_R)$  为变量的线性函数. 其中,  $T_P$  是以  $\lambda_{PP} - \lambda_{PN}$  为斜率  $\lambda_{NP}$  为截距、 $T_N$  是以  $\lambda_{NP} - \lambda_{NN}$  为斜率  $\lambda_{NN}$  为截距、 $T_B$  是以  $\lambda_{BP} - \lambda_{BN}$  为斜率  $\lambda_{BN}$  为截距的几何特征.

如图 1 中三支决策代价目标函数之间的拓扑关系示意图所示: 三支分类的阈值  $\alpha$  是代价目标线性函数  $T_P$  与  $T_B$  的交点横坐标  $\Pr(C|[x]_R)$  的取值, 阈值  $\beta$  是  $T_N$  与  $T_B$  的交点横坐标  $\Pr(C|[x]_R)$  的取值. 当  $1 \geq \alpha > \beta \geq 0$  时, 其三支划分的目标为: 当横坐标  $\Pr(C|[x]_R) \geq \alpha$  时,  $T_P$  的代价目标函数取值相对最低; 当横坐标  $\Pr(C|[x]_R) \leq \beta$  时,  $T_N$  的代价目标函数取值相对最低; 当横坐标  $\beta < \Pr(C|[x]_R) < \alpha$  时,  $T_B$  的代价目标函数取值相对最低.

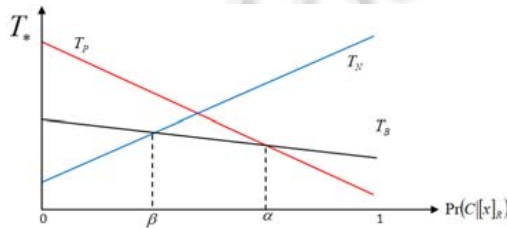


图 1  $T_P$ 、 $T_N$ 、 $T_B$  之间的拓扑关系示意图

注: 该代价目标函数几何示意图中,  $\Pr(C|[x]_R)$  为图中横坐标取值,  $T_*$  为图表纵坐标取值, 其中,  $* \in \{P, B, N\}$ .

根据上述阈值语义, 当  $1 \geq \alpha > \beta \geq 0$  时, 我们可以获得如下三支决策行为:

$$\left. \begin{aligned} \Pr(C|[x]_R) \geq \alpha &\rightarrow x \in POS_{(\alpha, \beta)}(C) \\ \Pr(C|[x]_R) \leq \beta &\rightarrow x \in NEG_{(\alpha, \beta)}(C) \\ \beta < \Pr(C|[x]_R) < \alpha &\rightarrow x \in BND_{(\alpha, \beta)}(C) \end{aligned} \right\} \quad (5)$$

### 1.3 代价目标函数及其阈值分析

基于三支代价目标函数之间的拓扑关系, 可以通过求解  $T_P = T_B$  获得  $\Pr(C|[x]_R) = \frac{\lambda_{PN} - \lambda_{BN}}{(\lambda_{PN} - \lambda_{BN}) + (\lambda_{BP} - \lambda_{PP})}$ , 其取值与经典三支决策理论中的  $\alpha$  取值是一致的, 即  $\alpha = \frac{\lambda_{PN} - \lambda_{BN}}{(\lambda_{PN} - \lambda_{BN}) + (\lambda_{BP} - \lambda_{PP})}$ . 根据  $\alpha$  阈值与代价目标函数

$T_P$ 、 $T_B$  之间的几何语义, 可以获得如下性质.

**性质 1.** 三支决策  $\alpha$  阈值与代价目标函数  $T_P$ 、 $T_B$  之间满足: 当  $\lambda_{PP} - \lambda_{PN} < \lambda_{BP} - \lambda_{BN}$  时, 如果  $\Pr(C|[x]_R) \leq \alpha$ , 则  $T_P \geq T_B$ ; 如果  $\Pr(C|[x]_R) \geq \alpha$ , 则  $T_P \leq T_B$ .

证明: 由公式(2)可知: 代价目标函数  $T_P$  可以看作是以  $\Pr(C|[x]_R)$  为自变量、 $\lambda_{PP} - \lambda_{PN}$  为斜率、 $\lambda_{PN}$  为  $T^*$  轴上的截距的一元线性函数, 代价目标函数  $T_B$  也可以看作是  $\Pr(C|[x]_R)$  为自变量、 $\lambda_{BP} - \lambda_{BN}$  为斜率、 $\lambda_{BN}$  为  $T^*$  轴上的截距的一元线性函数.

由于经典决策粗糙集合理论约定损失函数  $\lambda_{NP} > \lambda_{BP} > \lambda_{PP} > 0$  和  $\lambda_{PN} > \lambda_{BN} > \lambda_{NN} > 0$ , 所以,  $T_P$  和  $T_B$  之间的线性关系如图 2 所示, 存在图 2(a)的  $\lambda_{PP} - \lambda_{PN} < \lambda_{BP} - \lambda_{BN} < 0$  与图 2(b)的  $\lambda_{BP} - \lambda_{BN} > 0 > \lambda_{PP} - \lambda_{PN}$  这两种形态. 显然, 当  $\Pr(C|[x]_R) = \alpha$  时, 代价目标函数  $T_P$  与  $T_B$  相交.

根据初等解析几何公理, 当  $T_P$  和  $T_B$  之间的关系如图 2(a)和图 2(b)所示, 且当  $T_P$  的斜率小于  $T_B$  的斜率时,  $\Pr(C|[x]_R) \geq \alpha \rightarrow T_P \leq T_B$ ,  $\Pr(C|[x]_R) \leq \alpha \rightarrow T_P \geq T_B$ , 故性质 1 成立.  $\square$

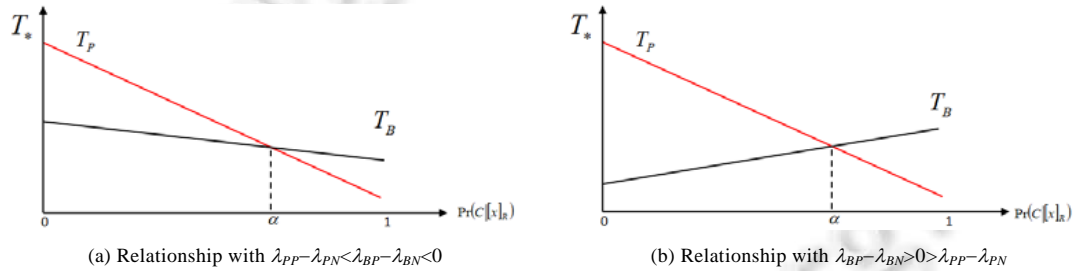


图 2 代价目标函数  $T_P$ 、 $T_B$  之间的拓扑关系图

通过求解  $\Pr(C|[x]_R)$  可获得  $\Pr(C|[x]_R) = \frac{\lambda_{BN} - \lambda_{NN}}{(\lambda_{BN} - \lambda_{NN}) + (\lambda_{NP} - \lambda_{BP})}$ , 其取值与经典三支决策理论中的  $\beta$  取值是一致的, 即  $\beta = \frac{\lambda_{BN} - \lambda_{NN}}{(\lambda_{BN} - \lambda_{NN}) + (\lambda_{NP} - \lambda_{BP})}$ . 根据  $\beta$  阈值与代价目标函数  $T_N$ 、 $T_B$  之间的几何语义, 可以获得如下性质.

**性质 2.** 三支决策  $\beta$  阈值与代价目标函数  $T_N$ 、 $T_B$  之间满足: 当  $\lambda_{NP} - \lambda_{NN} > \lambda_{BP} - \lambda_{BN}$  时, 如果  $\Pr(C|[x]_R) \leq \beta$ , 则  $T_N \leq T_B$ ; 如果  $\Pr(C|[x]_R) \geq \beta$ , 则  $T_N \geq T_B$ .

证明: 由公式(2)可知: 代价目标函数  $T_N$  可以看作是以  $\Pr(C|[x]_R)$  为自变量、 $\lambda_{NP} - \lambda_{NN}$  为斜率、 $\lambda_{NN}$  为  $T^*$  轴上的截距的一元线性函数, 代价目标函数  $T_B$  也可以看作是  $\Pr(C|[x]_R)$  为自变量、 $\lambda_{BP} - \lambda_{BN}$  为斜率、 $\lambda_{BN}$  为  $T^*$  轴上的截距的一元线性函数.

由于经典决策粗糙集合理论约定损失函数  $\lambda_{NP} > \lambda_{BP} > \lambda_{PP} > 0$  和  $\lambda_{PN} > \lambda_{BN} > \lambda_{NN} > 0$ , 所以,  $T_N$  和  $T_B$  之间的线性关系如图 3 所示, 存在图 3(a)的  $\lambda_{NP} - \lambda_{NN} > \lambda_{BP} - \lambda_{BN} > 0$  与图 3(b)的  $\lambda_{NP} - \lambda_{NN} > 0 > \lambda_{BP} - \lambda_{BN}$  这两种形态. 显然, 当  $\Pr$

$(C|[x]_R)=\beta$ 时, 代价目标函数  $T_N$  与  $T_B$  相交. 根据初等解析几何公理, 如图 3 所示, 代价目标函数  $T_N$  的斜率大于  $T_B$  的斜率时,  $\Pr(C|[x]_R) \geq \beta \rightarrow T_N \geq T_B$ ,  $\Pr(C|[x]_R) \leq \beta \rightarrow T_N \leq T_B$ . 故性质 2 成立.  $\square$

通过求解以  $\Pr(C|[x]_R)$  为自变量的代价目标函数  $T_P$ 、 $T_N$  构成的方程组  $T_P=T_N$ , 可获得  $\Pr(C|[x]_R)=\frac{\lambda_{PN}-\lambda_{NN}}{(\lambda_{PN}-\lambda_{NN})+(\lambda_{NP}-\lambda_{PP})}$ . 可以观察发现, 上述  $\Pr(C|[x]_R)$  取值与经典决策粗糙集理论中指定的  $\gamma$  取值是一致的, 即  $\gamma = \frac{\lambda_{PN}-\lambda_{NN}}{(\lambda_{PN}-\lambda_{NN})+(\lambda_{NP}-\lambda_{PP})}$ .

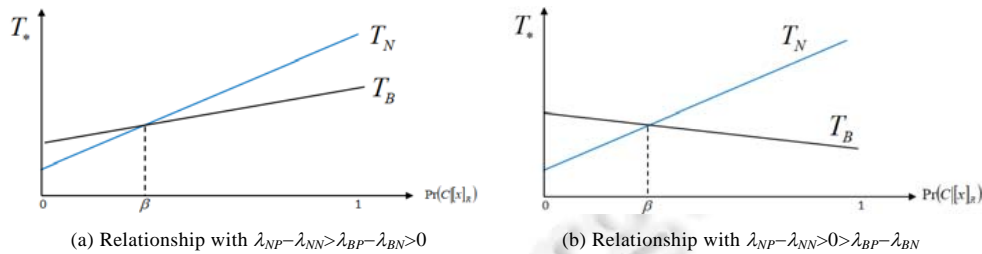


图 3 代价目标函数  $T_N$ 、 $T_B$  之间的拓扑关系图

如图 4 所示: 当  $\lambda_{NP}-\lambda_{NN} > \lambda_{BP}-\lambda_{BN} > \lambda_{PP}-\lambda_{PN}$ , 并且  $0 \leq \alpha < \gamma < \beta \leq 1$  时, 三支决策退化为二支决策. 此时, 当横坐标  $\Pr(C|[x]_R) \geq \gamma$  时,  $T_P$  的代价目标函数取值相对最低; 当横坐标  $\Pr(C|[x]_R) < \gamma$  时,  $T_N$  的代价目标函数取值相对最低.

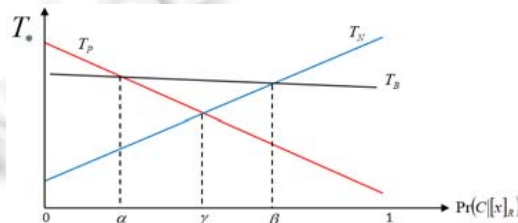


图 4 二支决策情况下,  $T_N$ 、 $T_P$ 、 $T_B$  之间的拓扑关系示意图

因此, 当  $\lambda_{NP}-\lambda_{NN} > \lambda_{BP}-\lambda_{BN} > \lambda_{PP}-\lambda_{PN}$ , 并且  $0 \leq \alpha < \gamma < \beta \leq 1$  时, 我们可以获得如下决策行为:

$$\left. \begin{aligned} \Pr(C|[x]_R) \geq \gamma &\rightarrow x \in POS_{(\alpha, \beta)}(C) \\ \Pr(C|[x]_R) < \gamma &\rightarrow x \in NEG_{(\alpha, \beta)}(C) \end{aligned} \right\} \quad (6)$$

根据  $\gamma$  阈值与代价目标函数  $T_P$ 、 $T_N$  之间的几何语义, 可以获得如下性质.

**性质 3.** 二支决策退化阈值  $\gamma$  与代价目标函数  $T_P$ 、 $T_N$  之间满足关系: 当  $\lambda_{PP}-\lambda_{PN} < \lambda_{NP}-\lambda_{NN}$  时, 如果  $\Pr(C|[x]_R) \leq \gamma$ , 则  $T_P \geq T_N$ ; 如果  $\Pr(C|[x]_R) > \gamma$ , 则  $T_P < T_N$ .

证明: 由公式(2)可知, 代价目标函数  $T_P$  可以看作是以  $\Pr(C|[x]_R)$  为自变量、 $\lambda_{PP}-\lambda_{PN}$  为斜率的一元线性函数, 代价目标函数  $T_N$  可以看作是以  $\Pr(C|[x]_R)$  为自变量、 $\lambda_{NP}-\lambda_{NN}$  为斜率的一元线性函数. 由于经典决策粗糙集理论约定损失函数  $(\lambda_{PP}-\lambda_{PN}) < 0$  和  $(\lambda_{NP}-\lambda_{NN}) > 0$ , 所以, 线性函数  $T_P$  与  $T_N$  之间的拓扑关系如图 5 所示.

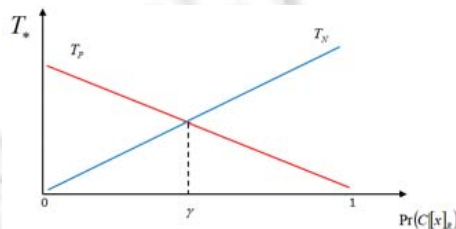


图 5 代价目标函数  $T_P$  与  $T_N$  之间的拓扑关系示意图

根据初等解析几何公理, 当  $T_N$  和  $T_P$  之间的关系如图 5 所示, 代价目标函数  $T_N$  的斜率大于  $T_P$  的斜率时,  $\Pr(C|[x]_R) \geq \gamma \rightarrow T_P \leq T_N$ ,  $\Pr(C|[x]_R) < \gamma \rightarrow T_P > T_N$ . 故性质 3 成立.  $\square$

上述理论研究从三支代价目标函数之间的拓扑关系视角, 对三支决策要素之间的关系及阈值推理进行了系统的讨论. 该研究能够直观地揭示代价敏感三支决策的 3 个主要决策要素: 条件概率、损失函数及决策阈值之间的拓扑关系, 从而为更加复杂的代价敏感三支决策提供了一种新型的研究视角.

## 2 分段延迟代价敏感三支决策

经典决策粗糙集理论中, 3 种决策的代价目标函数都是以条件概率为变量的单调线性函数. 以延迟决策代价目标函数  $T_B = (\lambda_{BP} - \lambda_{BN})\Pr(C|[x]_R) + \lambda_{BN}$  为例, 当  $\lambda_{BP} - \lambda_{BN} < 0$  时, 延迟决策的代价取值将随着条件概率的取值变大而单调变小. 当  $\lambda_{BP} - \lambda_{BN} > 0$  时, 延迟决策的代价取值将随着条件概率的取值变大而单调变大.

然而如图 6 所示, 实际生活经验中, 延迟决策的代价与条件概率取值之间的关系往往呈非单调线性关系.

例如, 在等价类  $[x]_R$  被判断是否属于目标概念  $C$  时, 当条件概率  $\Pr(C|[x]_R)$  趋近于极大值 1 或极小值 0 时,  $[x]_R$  被延迟决策的代价显然高于  $\Pr(C|[x]_R)$  的取值趋向于中间值 1/2 时的代价. 当  $\Pr(C|[x]_R)$  的取值趋向于中间值 1/2 时, 等价类  $[x]_R$  被延迟决策的代价往往趋向于变小.

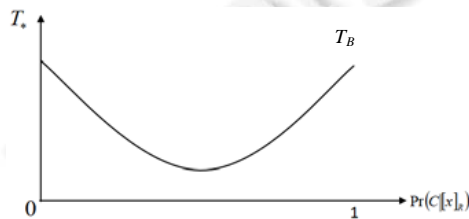


图 6 非单调性  $T_B$  示意图

### 2.1 分段延迟代价敏感三支决策模型

经典决策粗糙集理论中, 三支决策的损失函数通常是专家经验给出的, 不同的专家往往会给出不同的损失函数, 而根据不同的损失函数作用于相同的决策显然也会产生不同的代价. 分段函数具有刻画非单调线性空间的功能. 显然, 通过引入斜率不同并且数量不等的延迟决策代价目标函数  $T_B$ , 并且进行合理的分段集成, 是延迟域代价目标函数非单调现象的一种新型研究思路. 但是, 由于在实际应用中专家提供损失函数往往也是要计算成本的, 引入的延迟决策代价目标分段函数  $T_B$  越多, 损失函数的导入成本也就越高. 同时, 随着越多分段函数  $T_B$  的导入, 分段集成决策推理也会相对更加复杂. 所以, 导入合理数量的分段函数  $T_B$ , 是构造性价比较高的分段延迟代价敏感三支决策模型首要考虑的问题. 因此, 为了尽可能少地增加损失函数的导入数量, 又能够有效地模拟延迟决策代价目标函数非单调语义现象, 本研究采用了具有单调递增和单调递减特性的两套延迟决策代价目标函数作为双分段延迟代价目标函数. 在此基础上, 结合经典三支决策理论中的正负域决策代价目标函数, 构建了一种新型的分段延迟代价敏感三支决策模型 (piece-wise delay cost-sensitive three-way decisions, 简称分段延迟三支决策或者 PWD-3WD). 其中,

- 具有单调递增特性的延迟决策代价目标函数能够表征: 当条件概率  $\Pr(C|[x]_R)$  趋向于极大值时,  $[x]_R$  被延迟决策的代价相对高于  $\Pr(C|[x]_R)$  取值趋向于较小值时的代价;
- 具有单调递减特性的延迟决策代价目标函数能够表征: 当条件概率  $\Pr(C|[x]_R)$  趋近于极小值时,  $[x]_R$  被延迟决策的代价高于  $\Pr(C|[x]_R)$  的取值趋向于相对较大值时的代价.

而当三支决策被要求条件概率  $\Pr(C|[x]_R)$  趋近于极大值 1 和极小值 0 时,  $[x]_R$  被延迟决策的代价都要求高于  $\Pr(C|[x]_R)$  的取值趋向于中位值时的非单调代价现象. 我们可以通过将上述两套延迟决策代价目标函数作为分段函数, 并且设计合理的分段集成策略来实现其非单调函数的模拟.

分段延迟代价敏感三支决策模型采用经典代价敏感三支决策理论中定义的  $\lambda_{PP}$ 、 $\lambda_{PN}$  来表示正类元素  $x \in C$

和负类元素  $x \in C^c$  分别被执行正域决策的损失函数, 采用  $\lambda_{NP}, \lambda_{NN}$  表示为正类元素  $x \in C$  和负类元素  $x \in C^c$  分别被执行负域决策的损失函数.

定义 2. 分段延迟代价敏感三支决策模型中, 正域和负域决策代价目标函数被定义为

$$T_p = (\lambda_{pp} - \lambda_{pn})\Pr(C|[x]_R) + \lambda_{pn},$$
$$T_n = (\lambda_{np} - \lambda_{nn})\Pr(C|[x]_R) + \lambda_{nn},$$

其中约定:  $\lambda_{pp}$  与  $\lambda_{pn}$  之间满足  $(\lambda_{pp} - \lambda_{pn}) < 0$ ,  $\lambda_{np}$  与  $\lambda_{nn}$  之间满足  $(\lambda_{np} - \lambda_{nn}) > 0$ .

在定义 2 给出的分段延迟三支决策中, 正域和负域决策代价目标函数  $T_p$  及  $T_n$  具有如下语义.

- 函数  $T_p$  中, 正类元素  $x \in C$  被划分为正域  $POS_{(\alpha,*)}(C)$  的代价随着  $\Pr(C|[x]_R)$  的增加而单调增加;
- 函数  $T_n$  中, 正类元素  $x \in C$  被划分为正域  $POS_{(\alpha,*)}(C)$  的代价随着  $\Pr(C|[x]_R)$  的增加而单调减少;
- 当  $\Pr(C|[x]_R) = 1$  时, 代价目标函数  $T_p$  中代价值小于代价目标函数  $T_n$  中的代价值;
- 当  $\Pr(C|[x]_R) = 0$  时, 代价目标函数  $T_n$  中的代价值小于代价目标函数  $T_p$  中的代价值.

我们设定正类元素  $x \in C$  和负类元素  $x \in C^c$  分别被执行延迟决策时存在两套延迟损失函数  $\lambda'_{BP}, \lambda'_{BN}$  和  $\lambda''_{BP}, \lambda''_{BN}$ , 其语义为两位专家分别针对正、负类元素被延迟决策提供的两组不同损失评价指标. 基于这两套延迟损失函数, 可以构造具有单调递增和单调递减特性的两个分段延迟决策代价目标函数  $T'_B$  和  $T''_B$ .

定义 3. 在分段延迟代价敏感三支决策模型中, 双分段延迟决策代价目标函数被定义为:

- 正域视角延迟决策代价目标函数  $T'_B = (\lambda'_{BP} - \lambda'_{BN})\Pr(C|[x]_R) + \lambda'_{BN}$ ;
- 负域视角延迟决策代价目标函数  $T''_B = (\lambda''_{BP} - \lambda''_{BN})\Pr(C|[x]_R) + \lambda''_{BN}$ .

其中约定:

- $\lambda'_{BP}, \lambda'_{BN}$  之间满足  $\lambda'_{BP} - \lambda'_{BN} < 0$ ;
- $\lambda''_{BP}, \lambda''_{BN}$  之间满足  $\lambda''_{BP} - \lambda''_{BN} > 0$ ;
- $\lambda'_{BP}, \lambda'_{BN}$  与  $\lambda_{PN}, \lambda_{PP}, \lambda_{NN}, \lambda_{NP}$  之间满足  $\lambda_{PN} > \lambda'_{BN} > \lambda_{NN} > 0$  并且  $\lambda_{NP} > \lambda'_{BP} > \lambda_{PP} > 0$ ;
- $\lambda''_{BP}, \lambda''_{BN}$  与  $\lambda_{PN}, \lambda_{PP}, \lambda_{NN}, \lambda_{NP}$  之间满足  $\lambda_{PN} > \lambda''_{BN} > \lambda_{NN} > 0$  并且  $\lambda_{NP} > \lambda''_{BP} > \lambda_{PP} > 0$ .

上述定义中, 双分段延迟决策代价目标函数的语义为:

- $T'_B$  函数中, 当  $[x]_R \subseteq U$  被划分到  $BND_{(\alpha,\beta)}(C)$  域的代价随着概率  $\Pr(C|[x]_R)$  的增加而单调减少时, 则会随着  $\Pr(C|[x]_R)$  的减少而单调增加, 即  $(\lambda'_{BP} - \lambda'_{BN}) < 0$ ;
- $T''_B$  函数中, 当  $[x]_R \subseteq U$  被划分为  $BND_{(\alpha,\beta)}(C)$  的代价随着概率  $\Pr(C|[x]_R)$  的减少而单调递减时, 则会随着  $\Pr(C|[x]_R)$  的增加而单调递增, 即  $\lambda''_{BP} - \lambda''_{BN} > 0$ .

当采用  $T'_B$  函数作为延迟代价计算依据时: 若  $\Pr(C|[x]_R)$  取值趋近于极大值 1, 则  $[x]_R$  被划分为延迟决策域  $BND_{(\alpha,\beta)}(C)$  的代价大于被划分为  $POS_{(\alpha,\beta)}(C)$  正决策域的代价, 同时小于被划分为负决策域  $NEG_{(\alpha,\beta)}(C)$  的代价, 即  $T_n > T'_B > T_p$ ; 若  $\Pr(C|[x]_R)$  取值趋近于极小值 0, 则  $[x]_R$  被划分为延迟决策域  $BND_{(\alpha,\beta)}(C)$  的代价小于被划分为  $POS_{(\alpha,*)}(C)$  正决策域的代价, 同时大于被划分为负决策域  $NEG_{(\alpha,*)}(C)$  的代价, 即  $T_n < T'_B < T_p$ .

当采用  $T''_B$  函数作为延迟代价计算依据时: 若  $\Pr(C|[x]_R)$  取值趋于极大值 1, 则  $[x]_R$  被划分为延迟决策域  $BND_{(\alpha,\beta)}(C)$  的代价大于被划分为  $POS_{(\alpha,\beta)}(C)$  正决策域的代价, 同时小于被划分为负决策域  $NEG_{(\alpha,\beta)}(C)$  的代价, 即  $T_n > T''_B > T_p$ ; 若  $\Pr(C|[x]_R)$  取值趋近于极小值 0, 则  $[x]_R$  被划分为延迟决策域  $BND_{(\alpha,\beta)}(C)$  的代价小于被划分为  $POS_{(\alpha,\beta)}(C)$  正决策域的代价, 同时大于被划分为负决策域  $NEG_{(\alpha,\beta)}(C)$  的代价, 即  $T_n < T''_B < T_p$ .

由于分段延迟决策代价目标函数  $T'_B$  和  $T''_B$  的定义分别约定了  $(\lambda'_{BP} - \lambda'_{BN}) < 0$  和  $\lambda''_{BP} - \lambda''_{BN} > 0$ , 所有分段延迟决策代价目标函数  $T'_B$  和  $T''_B$  相对于条件概论  $\Pr(C|[x]_R)$  分别具有单调递减和单调递增特性. 通过构造分段延迟决策代价目标函数  $T'_B$  和  $T''_B$  的集成分段策略, 可以构造同时具备单调递减和单调递增特性的延迟决策代价目标函数.

根据分段延迟代价敏感三支决策模型中损失函数的约定, 可以推导获得以下性质.

性质 4. 在分段延迟三支决策模型中, 正域决策的损失函数  $\lambda_{pp}, \lambda_{pn}$  与正、负域视角延迟决策损失函数对  $\lambda'_{BP}, \lambda'_{BN}, \lambda''_{BP}, \lambda''_{BN}$  之间满足  $(\lambda_{pp} - \lambda_{pn}) < (\lambda'_{BP} - \lambda'_{BN}) < 0 < (\lambda''_{BP} - \lambda''_{BN})$ ;



负域决策的损失函数  $\lambda_{NP}, \lambda_{NN}$  与正、负域视角延迟决策损失函数对  $\lambda'_{BP}, \lambda'_{BN}, \lambda''_{BP}, \lambda''_{BN}$  之间满足  $(\lambda_{NP} - \lambda_{NN}) > (\lambda''_{BP} - \lambda''_{BN}) > 0 > (\lambda'_{BP} - \lambda'_{BN})$ .

证明: 根据定义 3 中约定了  $\lambda_{PP}, \lambda_{PN}$  与  $\lambda'_{BP}, \lambda'_{BN}$  之间满足  $(\lambda_{PN} > \lambda'_{BN} > 0) \wedge (\lambda'_{BP} > \lambda_{PP} > 0)$ ,

所有  $\lambda_{PN} - \lambda'_{BN} + \lambda'_{BP} - \lambda_{PP} > 0$ , 即  $(\lambda_{PP} - \lambda_{PN}) < (\lambda'_{BP} - \lambda'_{BN})$  成立;

又由于定义 3 约定  $\lambda'_{BP}, \lambda'_{BN}$  之间满足  $\lambda'_{BP} - \lambda'_{BN} < 0$ ;  $\lambda''_{BP}, \lambda''_{BN}$  之间满足  $\lambda''_{BP} - \lambda''_{BN} > 0$ ,

显然,  $(\lambda'_{BP} - \lambda'_{BN}) < (\lambda''_{BP} - \lambda''_{BN})$ .

所以  $(\lambda_{PP} - \lambda_{PN}) < (\lambda'_{BP} - \lambda'_{BN}) < 0 < (\lambda''_{BP} - \lambda''_{BN})$  得证.

同理易证  $(\lambda_{NP} - \lambda_{NN}) > (\lambda''_{BP} - \lambda''_{BN}) > 0 > (\lambda'_{BP} - \lambda'_{BN})$  成立. □

上述性质 4 的语义为: 分段延迟三支决策模型中, 正域决策代价目标函数  $T_P$  的斜率小于正域视角延迟决策代价目标函数  $T'_B$  的斜率; 且由定义 2 和定义 3 可知, 这两个斜率小于 0. 负域决策代价目标函数  $T_N$  的斜率大于正域视角延迟决策代价目标函数  $T'_B$  的斜率; 且由定义 2 和定义 3 可知, 这两个斜率大于 0.

由于决策要素的度量是不确定知识推理的基础, 为了对分段延迟代价敏感三支决策模型进行有效的推理, 我们定义了分段延迟三支决策的 6 种基础度量指标.

**定义 4.** 分段延迟三支决策相关的 6 种基础度量指标分别是: 双正域决策度量指标  $\alpha', \alpha''$ ; 双负域决策度量指标  $\beta', \beta''$ ; 退化决策度量指标  $\gamma$ ; 双延迟决策代价目标函数交叉度量指标  $\mu$ . 其中,

- 正域决策度量指标  $\alpha'$  可以通过求解  $T_P = T'_B$  获得, 其几何语义为代价目标函数  $T_P$  与  $T'_B$  之间交点的

$$\Pr(C|[x]_R) \text{ 的取值, 即 } T_P = T'_B \rightarrow \Pr(C|[x]_R) = \frac{\lambda_{PN} - \lambda'_{BN}}{(\lambda_{PN} - \lambda'_{BN}) + (\lambda'_{BP} - \lambda_{PP})}. \text{ 根据上述 } \Pr(C|[x]_R) \text{ 的取值, 本}$$

$$\text{文设 } \alpha' = \frac{\lambda_{PN} - \lambda'_{BN}}{(\lambda_{PN} - \lambda'_{BN}) + (\lambda'_{BP} - \lambda_{PP})};$$

- 正域决策度量指标  $\alpha''$  可以通过求解  $T_P = T''_B$  获得, 其几何语义为代价目标函数  $T_P$  与  $T''_B$  之间交点的

$$\Pr(C|[x]_R) \text{ 的取值, 即 } T_P = T''_B \rightarrow \Pr(C|[x]_R) = \frac{\lambda_{PN} - \lambda''_{BN}}{(\lambda_{PN} - \lambda''_{BN}) + (\lambda''_{BP} - \lambda_{PP})}. \text{ 根据上述 } \Pr(C|[x]_R) \text{ 的取值, 本}$$

$$\text{文设 } \alpha'' = \frac{\lambda_{PN} - \lambda''_{BN}}{(\lambda_{PN} - \lambda''_{BN}) + (\lambda''_{BP} - \lambda_{PP})};$$

- 负域决策度量指标  $\beta'$  可以通过求解  $T_N = T'_B$  获得, 其几何语义为代价目标函数  $T_N$  与  $T'_B$  之间交点的

$$\Pr(C|[x]_R) \text{ 的取值, 即 } T_N = T'_B \rightarrow \Pr(C|[x]_R) = \frac{\lambda'_{BN} - \lambda_{NN}}{(\lambda'_{BN} - \lambda_{NN}) + (\lambda_{NP} - \lambda'_{BP})}. \text{ 根据上述 } \Pr(C|[x]_R) \text{ 的取值, 本}$$

$$\text{文设 } \beta' = \frac{\lambda'_{BN} - \lambda_{NN}}{(\lambda'_{BN} - \lambda_{NN}) + (\lambda_{NP} - \lambda'_{BP})};$$

- 负域决策度量指标  $\beta''$  可以通过求解  $T_N = T''_B$  获得, 其几何语义为代价目标函数  $T_N$  与  $T''_B$  之间交点的

$$\Pr(C|[x]_R) \text{ 的取值, 即 } T_N = T''_B \rightarrow \Pr(C|[x]_R) = \frac{\lambda''_{BN} - \lambda_{NN}}{(\lambda''_{BN} - \lambda_{NN}) + (\lambda_{NP} - \lambda''_{BP})}. \text{ 根据上述 } \Pr(C|[x]_R) \text{ 的取值, 本}$$

$$\text{文设 } \beta'' = \frac{\lambda''_{BN} - \lambda_{NN}}{(\lambda''_{BN} - \lambda_{NN}) + (\lambda_{NP} - \lambda''_{BP})};$$

- 通过求解  $T_P = T_N$  获得退化决策度量指标  $\gamma = \frac{\lambda_{PN} - \lambda_{NN}}{(\lambda_{PN} - \lambda_{NN}) + (\lambda_{NP} - \lambda_{PP})}$ , 其取值及几何语义仍然与经典

代价敏感三支决策中定义的  $\gamma$  阈值相同;

- 双延迟决策代价目标函数的交叉度量指标  $\mu = \frac{\lambda''_{BN} - \lambda'_{BN}}{(\lambda'_{BP} - \lambda'_{BN}) - (\lambda''_{BP} - \lambda''_{BN})}$  可以通过求解  $T'_B = T''_B$  获得,

其几何语义为代价目标函数  $T'_B$  与  $T''_B$  之间交点的  $\Pr(C|[x]_R)$  的取值.

**性质 5.** 分段延迟三支决策相关的基础度量指标满足:

1.  $0 < \alpha' < 1, 0 < \alpha'' < 1$ ;

2.  $0 < \beta < 1, 0 < \beta' < 1$ .

证明: 由定义 3 可知,  $(\lambda_{PN} - \lambda'_{BN}) > 0$  并且  $(\lambda'_{BP} - \lambda_{PP}) > 0$ ,

所以  $(\lambda_{PN} - \lambda'_{BN}) + (\lambda'_{BP} - \lambda_{PP}) > (\lambda_{PN} - \lambda'_{BN})$ , 即  $0 < \frac{\lambda_{PN} - \lambda'_{BN}}{(\lambda_{PN} - \lambda'_{BN}) + (\lambda'_{BP} - \lambda_{PP})} < 1$ ;

又因为  $\alpha' = \frac{\lambda_{PN} - \lambda'_{BN}}{(\lambda_{PN} - \lambda'_{BN}) + (\lambda'_{BP} - \lambda_{PP})}$ , 所以  $0 < \alpha' < 1$  得证.

同理, 由定义 3 可知,  $\lambda_{PN} > \lambda''_{BN}$  并且  $\lambda''_{BP} > \lambda_{PP}$ ,

所以  $(\lambda_{PN} - \lambda''_{BN}) + (\lambda''_{BP} - \lambda_{PP}) > (\lambda_{PN} - \lambda''_{BN})$ ,

因此,  $0 < \frac{\lambda_{PN} - \lambda''_{BN}}{(\lambda_{PN} - \lambda''_{BN}) + (\lambda''_{BP} - \lambda_{PP})} < 1$  成立;

又因为  $\alpha'' = \frac{\lambda_{PN} - \lambda''_{BN}}{(\lambda_{PN} - \lambda''_{BN}) + (\lambda''_{BP} - \lambda_{PP})}$ , 所以  $0 < \alpha'' < 1$  得证. □

性质 5 中的第 2 点证明从略, 证明过程与性质 5 中的第 1 点相似.

根据性质 5, 分段延迟三支决策模型能够获得两组有效的三支决策阈值. 这两组三支决策阈值的集成是求解前文所述延迟代价目标函数非单调三支决策的基础. 同时, 分段延迟三支决策模型获得的这两组度量指标  $\alpha$ 、 $\beta$  与  $\alpha'$ 、 $\beta'$ , 也看作是两种不同视角下的三支决策.

由于退化决策度量指标  $\gamma$  是由求解  $T_P = T_N$  获得, 则当  $\Pr(C|[x]_R) = \gamma$  时, 正域决策的代价与负域决策的代价相同. 通常, 我们认为此时对该对象进行正、负域决策的不确定程度最大; 而另一方面, 我们也通常认为此时进行延迟决策往往是各方面都相对可以接受的选择. 而当  $\Pr(C|[x]_R) = 0.5$  时, 有可能存在  $T_P > T_N$  或者  $T_P < T_N$ , 所以我们通常无法确定此时进行延迟决策是否各方面都可以接受. 同样,  $T'_B$  与  $T''_B$  之间交点处  $\Pr(C|[x]_R) = \mu$  的取值也存在这样的问题, 并且还有可能出现  $T'_B$  与  $T''_B$  之间交点处的取值  $\mu$  小于 0 或者大于 1 的特殊情况, 因此而更加无法判断此时进行延迟决策是否有实用价值. 根据上述讨论, 本文选择不确定程度最大的决策度量指标  $\Pr(C|[x]_R) = \gamma$  作为双延迟决策代价目标函数的分段转折点. 基于分段转折点  $\gamma$ , 本文给出如下分段延迟三支决策阈值的集成策略.

**定义 5.** 分段延迟代价敏感三支决策的分段策略.

- 正域视角优先策略: 当  $\Pr(C|[x]_R) < \gamma$  时, 采用正域视角延迟决策代价目标函数  $T'_B$ ; 当  $\Pr(C|[x]_R) \geq \gamma$  时, 采用负域视角延迟决策代价目标函数  $T''_B$ ;
- 负域视角优先策略: 当  $\Pr(C|[x]_R) < \gamma$  时, 采用负域视角延迟决策代价目标函数  $T''_B$ ; 当  $\Pr(C|[x]_R) \geq \gamma$  时, 采用正域视角延迟决策代价目标函数  $T'_B$ .

如何在双分段延迟代价三支决策中选择合适的集成策略, 对分段延迟代价三支决策的应用具有重要的指导意义. 下一节将对双延迟代价三支决策各类情况集成策略的推理分别加以展开讨论.

## 2.2 分段延迟代价敏感三支决策推理

本文构造了双延迟决策代价目标函数, 获得两组分段三支决策作为解决延迟域代价目标函数非单调现象的基础. 从本质上讲, 分段延迟三支决策也可以看作是群体决策的一种表现形式. 然而, 分段延迟代价敏感三支决策模型中的两个延迟决策代价目标函数  $T'_B$ 、 $T''_B$  与正负决策代价目标函数  $T_P$ 、 $T_N$  之间的拓扑关系与分段延迟代价三支决策的推理密切相关.

如图 7 所示: 当横坐标  $\Pr(C|[x]_R) > 0$ 、纵坐标  $T_s > 0$  时, 分段延迟三支决策的关系模式示意图可被正、负域决策代价目标函数  $T_P$ 、 $T_N$  划分为 4 个象限. 而两个延迟决策代价目标函数  $T'_B$ 、 $T''_B$  的交点(即  $T'_B(\mu) = T''_B(\mu)$ ), 可能分别处于上述 4 个象限内.

- 当  $T'_B$ 、 $T''_B$  交点处于象限 1 内时, 该交点的延迟代价取值特点为  $(T'_B(\mu) < T_P(\mu)) \wedge (T'_B(\mu) < T_N(\mu))$ ;
- 当  $T'_B$ 、 $T''_B$  交点处于象限 2 内时, 该交点的延迟代价取值特点为  $(T'_B(\mu) < T_P(\mu)) \wedge (T'_B(\mu) > T_N(\mu))$ ;
- 当  $T'_B$ 、 $T''_B$  交点处于象限 3 内时, 该交点的延迟代价取值特点为  $(T'_B(\mu) > T_P(\mu)) \wedge (T'_B(\mu) < T_N(\mu))$ ;

- 当  $T'_B, T''_B$  交点处于象限 4 内时, 该交点的延迟代价取值特点为  $(T'_B(\mu) > T_p(\mu)) \wedge (T'_B(\mu) > T_N(\mu))$ .
- 如图 7(a)所示, 容易证明: 当  $T'_B, T''_B$  交点处于象限 1 时, 分段延迟三支决策模型的基础度量指标  $\gamma$  与  $\alpha, \beta$  和  $\alpha', \beta'$  之间满足  $\beta < \gamma < \alpha$  和  $\beta' < \gamma < \alpha'$ ;
- 如图 7(b)所示, 容易证明: 当  $T'_B, T''_B$  交点处于象限 2 时, 分段延迟三支决策模型的基础度量指标  $\gamma$  与  $\alpha, \beta$  和  $\alpha', \beta'$  之间满足  $\beta < \gamma < \alpha$  和  $\beta' > \gamma > \alpha'$ ;
- 如图 7(c)所示, 容易证明: 当  $T'_B, T''_B$  交点处于象限 3 时, 分段延迟三支决策模型的基础度量指标  $\gamma$  与  $\alpha, \beta$  和  $\alpha', \beta'$  之间满足  $\beta > \gamma > \alpha$  和  $\beta' < \gamma < \alpha'$ ;
- 如图 7(d)所示, 容易证明: 当  $T'_B, T''_B$  交点处于象限 4 时, 分段延迟三支决策模型的基础度量指标  $\gamma$  与  $\alpha, \beta$  和  $\alpha', \beta'$  之间满足  $\beta > \gamma > \alpha$  和  $\beta' > \gamma > \alpha'$ .

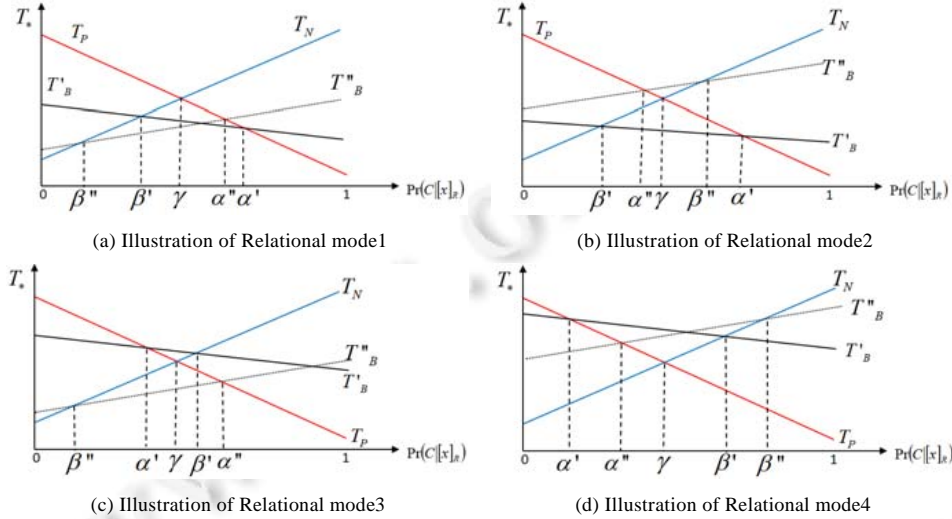


图 7 分段延迟三支决策的关系模式示意图

上述研究发现, 本文分段延迟三支决策模型中定义的基础度量指标  $\gamma, \alpha, \beta, \alpha', \beta'$  之间只可能存在这 4 种典型的拓扑关系, 所以分段延迟代价敏感三支决策模型可被划分为以下 4 种关系模式.

**定义 6.** 分段延迟代价敏感三支决策模型包括的 4 种关系模式.

1. 当分段延迟三支决策基础度量指标之间同时满足  $\beta < \gamma < \alpha$  和  $\beta' < \gamma < \alpha'$  时, 该分段延迟代价敏感三支决策模型被定义为关系模式 1;
2. 当分段延迟三支决策基础度量指标之间同时满足  $\beta < \gamma < \alpha$  和  $\beta' > \gamma > \alpha'$  时, 该分段延迟代价敏感三支决策模型被定义为关系模式 2;
3. 当分段延迟三支决策基础度量指标之间同时满足  $\beta > \gamma > \alpha$  和  $\beta' < \gamma < \alpha'$  时, 该分段延迟代价敏感三支决策模型被定义为关系模式 3;
4. 当分段延迟三支决策基础度量指标之间同时满足  $\beta > \gamma > \alpha$  和  $\beta' > \gamma > \alpha'$  时, 该分段延迟代价敏感三支决策模型被定义为关系模式 4.

根据上述分段延迟代价敏感三支决策的 4 类典型关系模式及分段策略, 可以得到以下 4 种分段延迟三支决策推理.

**定理 1.** 当分段延迟代价敏感三支决策模型属于关系模式 1 时,

- (1) 采用负域视角优先分段策略的 3WD 阈值为  $\alpha = \alpha', \beta = \beta'$ ;
- (2) 采用正域视角优先分段策略的 3WD 阈值为  $\alpha = \alpha', \beta = \beta'$ .

证明: 对于定理 1(1): 当分段延迟代价敏感三支决策模型为关系模式 1 时, 实施负域视角优先分段策略 3WD.

首先证明, 当  $\gamma \leq \Pr(C|[x]_R)$  时的三支决策, 此时需要采用正域视角延迟决策代价目标函数  $T'_B$  与  $T_P$ 、 $T_N$  参与三支决策代价最小化推理.

由于定义 6.1 规定了关系模式 1 中  $\beta < \gamma < \alpha'$  和  $\beta' < \gamma < \alpha'$ , 所以  $\gamma \leq \Pr(C|[x]_R)$  可被进一步划分为  $\gamma \leq \Pr(C|[x]_R) < \alpha'$  和  $\Pr(C|[x]_R) \geq \alpha'$  这两个证明阶段.

(1) 证明  $\gamma \leq \Pr(C|[x]_R) < \alpha'$  时的三支决策.

根据分段延迟三支决策模型性质 4 可知  $(\lambda_{PP} - \lambda_{PN}) < (\lambda'_{BP} - \lambda'_{BN})$ , 此时由性质 1 可以推断: 当  $\Pr(C|[x]_R) < \alpha'$  时,  $T_P > T'_B$ . 又由于关系模式 1 中约定  $\gamma < \alpha'$ , 所以,  $\gamma \leq \Pr(C|[x]_R) < \alpha'$  时,  $T'_B < T_P$ .

由于定义 2 分段延迟代价敏感三支决策模型中规定  $\lambda_{PP} - \lambda_{PN} < 0$  和  $\lambda_{NP} - \lambda_{NN} > 0$ , 所以,  $\lambda_{PP} - \lambda_{PN} < \lambda_{NP} - \lambda_{NN}$ . 由性质 3 可以推断: 当  $\Pr(C|[x]_R) \geq \gamma$  时,  $T_P \leq T_N$ . 同样, 由于关系模式 1 中约定  $\gamma < \alpha'$ , 故此推断: 当  $\gamma \leq \Pr(C|[x]_R) < \alpha'$  时,  $T_P \leq T_N$ .

所以, 根据上述两个结论可以证明:

当  $\gamma \leq \Pr(C|[x]_R) < \alpha'$  时,  $T'_B < T_P \leq T_N$  成立. 即当  $\gamma \leq \Pr(C|[x]_R) < \alpha'$  时,  $x \in BND_{(\alpha, \beta)}(C)$  代价最小.

(2) 证明  $\Pr(C|[x]_R) \geq \alpha'$  时的三支决策.

由分段延迟三支决策模型性质 4 可知  $(\lambda_{PP} - \lambda_{PN}) < (\lambda'_{BP} - \lambda'_{BN})$ , 此时由性质 1 可以推断: 当  $\Pr(C|[x]_R) \geq \alpha'$  时,  $T_P \leq T'_B$ .

由分段延迟三支决策模型性质 4 可知  $(\lambda_{NP} - \lambda_{NN}) > (\lambda'_{BP} - \lambda'_{BN})$ , 此时由性质 2 可以推断: 当  $\Pr(C|[x]_R) > \beta'$  时,  $T_N > T'_B$ .

又由于定义 6.1 分段关系模式 1 中规定  $\beta < \alpha'$ , 所以, 可以推断: 当  $\Pr(C|[x]_R) \geq \alpha'$  时,  $T_N > T'_B$ .

所以, 根据上述两个结论可以证明  $\Pr(C|[x]_R) \geq \alpha' \Rightarrow T_N > T'_B \geq T_P$ .

即当  $\Pr(C|[x]_R) \geq \alpha'$  时,  $x \in POS_{(\alpha, \beta)}(C)$  代价最小.

其次证明当  $\Pr(C|[x]_R) < \gamma$  时的三支决策, 此时需要采用正域视角延迟决策代价目标函数  $T''_B$  与  $T_P$ 、 $T_N$  参与三支决策代价最小化推理.  $\Pr(C|[x]_R) < \gamma$  可被进一步划分为  $\beta' < \Pr(C|[x]_R) < \gamma$  和  $\Pr(C|[x]_R) \leq \beta'$  这两个证明阶段.

(1) 证明  $\beta' < \Pr(C|[x]_R) < \gamma$  时的三支决策.

由于分段延迟代价敏感三支决策模型的性质 4 中规定了  $(\lambda_{NP} - \lambda_{NN}) > (\lambda''_{BP} - \lambda''_{BN})$ , 此时由性质 2 可以推断: 当  $\Pr(C|[x]_R) > \beta'$  时,  $T_N > T''_B$ . 又由于定义 6 约定关系模式 1 中  $\beta' < \gamma$ , 所以, 当  $\gamma > \Pr(C|[x]_R) > \beta'$  时,  $T_N > T''_B$ .

根据定义 2 分段延迟代价敏感三支决策模型中  $\lambda_{PP} - \lambda_{PN} < \lambda_{NP} - \lambda_{NN}$ , 由性质 3 可以推断: 当  $\Pr(C|[x]_R) < \gamma$  时,  $T_P > T_N$ . 又由于定义 6 约定了  $\beta' < \gamma$ , 所以, 当  $\beta' < \Pr(C|[x]_R) < \gamma$  时,  $T_P > T_N$  也成立.

所以, 根据上述两个结论可以证明: 当  $\gamma > \Pr(C|[x]_R) > \beta'$  时,  $T_P > T_N > T''_B$ . 即: 当  $\gamma > \Pr(C|[x]_R) > \beta'$  时,  $x \in BND_{(\alpha, \beta)}(C)$  代价最小.

(2) 证明  $\Pr(C|[x]_R) \leq \beta'$  时的三支决策.

由分段延迟三支决策模型性质 4 可知  $\lambda_{PP} - \lambda_{PN} < \lambda''_{BP} - \lambda''_{BN}$ , 此时由性质 1 可知: 当  $\Pr(C|[x]_R) < \alpha''$  时,  $T_P > T''_B$ . 又由于关系模式 1 中  $\beta' < \alpha''$ , 因此可得  $\Pr(C|[x]_R) \leq \beta'' \rightarrow T_P > T''_B$ .

由于分段延迟代价敏感三支决策模型的性质 4 中规定了  $(\lambda_{NP} - \lambda_{NN}) > (\lambda''_{BP} - \lambda''_{BN})$ , 此时由性质 2 可以推断: 当  $\Pr(C|[x]_R) \leq \beta''$  时,  $T_N \leq T''_B$ .

所以, 根据上述两个结论可以证明: 当  $\Pr(C|[x]_R) \leq \beta''$  时,  $T_P > T''_B \geq T_N$ . 即当  $\Pr(C|[x]_R) \leq \beta''$  时,  $x \in NEG_{(\alpha, \beta)}(C)$  代价最小.

当分段延迟代价敏感三支决策模型属于关系模式 1 时, 由上述证明可获得采用负域视角优先分段策略的三支决策为

$$\begin{aligned} \Pr(C|[x]_R) \geq \alpha' &\rightarrow x \in POS_{(\alpha, \beta)}(C), \\ \alpha' > \Pr(C|[x]_R) > \beta'' &\rightarrow x \in BND_{(\alpha, \beta)}(C), \\ \Pr(C|[x]_R) \leq \beta'' &\rightarrow x \in NEG_{(\alpha, \beta)}(C). \end{aligned}$$

即, 采用负域视角优先分段策略的 3WD 阈值为  $\alpha = \alpha'$ ,  $\beta = \beta''$ , 故定理 1(1)得证.

定理 1(2)的证明过程与定理 1(1)的证明过程类似, 从略. □

**定理 2.** 当分段延迟代价敏感三支决策模型属于关系模式 2 时,

(1) 采用负域视角优先分段策略的 3WD 阈值为  $\alpha=\alpha', \beta=\gamma$ ;

(2) 采用正域视角优先分段策略的 3WD 阈值为  $\alpha=\gamma, \beta=\beta'$ .

证明: 对于定理 2(1): 当分段延迟三支决策模型属于关系模式 2 时, 采用负域视角优先分段策略实施三支决策.

首先证明当  $\gamma < \Pr(C|[x]_R)$  时的三支决策. 此时需要采用正域视角延迟决策代价目标函数  $T'_B$  与  $T_P, T_N$  参与三支决策代价最小化推理.

由于定义 6.2 规定了关系模式 2 中  $\beta < \gamma < \alpha'$  和  $\beta' > \gamma > \alpha'$ , 所以,  $\gamma < \Pr(C|[x]_R)$  时的三支决策可被进一步划分为  $\gamma < \Pr(C|[x]_R) < \alpha'$  和  $\Pr(C|[x]_R) \geq \alpha'$  这两个证明阶段.

(1) 先证  $\gamma < \Pr(C|[x]_R) < \alpha'$  时的三支决策.

由分段延迟三支决策模型性质 4 可知  $(\lambda_{PP} - \lambda_{PN}) < (\lambda'_{BP} - \lambda'_{BN})$ , 此时由性质 1 可以推断: 当  $\Pr(C|[x]_R) < \alpha'$  时,  $T_P > T'_B$ . 又由于关系模式 2 约定  $\gamma < \alpha'$ , 所以, 当  $\gamma < \Pr(C|[x]_R) < \alpha'$  时,  $T_P > T'_B$ .

由于定义 2 分段延迟代价敏感三支决策模型中规定  $\lambda_{PP} - \lambda_{PN} < 0$  和  $\lambda_{NP} - \lambda_{NN} > 0$ , 所以,  $\lambda_{PP} - \lambda_{PN} < \lambda_{NP} - \lambda_{NN}$ . 由性质 3 可以推断: 当  $\Pr(C|[x]_R) > \gamma$  时,  $T_N > T_P$ . 又由于关系模式 2 约定  $\gamma < \alpha'$ , 所以, 当  $\gamma < \Pr(C|[x]_R) < \alpha'$  时,  $T_N > T_P$  也成立.

所以, 根据上述两个结论可以证明: 当  $\gamma < \Pr(C|[x]_R) < \alpha'$  时,  $T'_B < T_P < T_N$ , 即当  $\gamma < \Pr(C|[x]_R) < \alpha'$  时,  $x \in BND_{(\alpha, \beta)}(C)$  代价最小.

(2) 再证  $\Pr(C|[x]_R) \geq \alpha'$  时的三支决策.

由分段延迟三支决策模型性质 4 可知  $(\lambda_{PP} - \lambda_{PN}) < (\lambda'_{BP} - \lambda'_{BN})$ , 由性质 1 可以推断: 当  $\Pr(C|[x]_R) \geq \alpha'$  时,  $T'_B \geq T_P$ .

由分段延迟三支决策模型性质 4 可知  $(\lambda_{NP} - \lambda_{NN}) > (\lambda'_{BP} - \lambda'_{BN})$ , 由性质 2 可得: 当  $\Pr(C|[x]_R) > \beta$  时,  $T_N > T'_B$ . 又由于关系模式 2 中规定  $\beta < \alpha'$ , 因此, 当  $\Pr(C|[x]_R) \geq \alpha'$  时,  $T_N > T'_B$  也成立.

所以, 根据上述两个结论可以证明: 当  $\Pr(C|[x]_R) \geq \alpha'$  时,  $T_N > T'_B \geq T_P$ , 即当  $\Pr(C|[x]_R) > \alpha'$  时,  $x \in POS_{(\alpha, \beta)}(C)$  代价最小.

当  $\Pr(C|[x]_R) \leq \gamma$  时, 需要采用正域视角延迟决策代价目标函数  $T''_B$  与  $T_P, T_N$  参与三支决策代价最小化推理.

$\Pr(C|[x]_R) \leq \gamma$  时的三支决策可被进一步划分为  $\alpha'' < \Pr(C|[x]_R) \leq \gamma$  和  $\Pr(C|[x]_R) \leq \alpha''$  这两个证明阶段.

(1) 先证  $\alpha'' < \Pr(C|[x]_R) \leq \gamma$  时的三支决策.

因为由定义 2 可知, 分段延迟代价敏感三支决策模型中  $\lambda_{PP} - \lambda_{PN} < \lambda_{NP} - \lambda_{NN}$ , 由性质 3 可以推断: 当  $\Pr(C|[x]_R) \leq \gamma$  时,  $T_P \geq T_N$ . 又由于关系模式 2 中规定了  $\alpha'' < \gamma$ , 所以, 当  $\alpha'' < \Pr(C|[x]_R) \leq \gamma$  时,  $T_P \geq T_N$  也成立.

由于分段延迟代价敏感三支决策模型的性质 4 中规定了  $(\lambda_{PP} - \lambda_{PN}) < (\lambda''_{BP} - \lambda''_{BN})$ , 由性质 2 可以推断: 当  $\Pr(C|[x]_R) \geq \alpha''$  时,  $T_P \leq T''_B$ . 由于关系模式 2 中规定了  $\alpha'' < \gamma$ , 所以, 当  $\alpha'' < \Pr(C|[x]_R) \leq \gamma$  时,  $T_P \leq T''_B$ .

所以, 根据上述两个结论可以证明: 当  $\alpha'' < \Pr(C|[x]_R) \leq \gamma$  时,  $T''_B \geq T_P \geq T_N$  成立, 即当  $\alpha'' < \Pr(C|[x]_R) \leq \gamma$  时,  $x \in NEG_{(\alpha, \beta)}(C)$  代价最小.

(2) 再证  $\Pr(C|[x]_R) \leq \alpha''$  时的三支决策.

因为分段延迟代价敏感三支决策模型的性质 4 中规定了  $(\lambda_{NP} - \lambda_{NN}) > (\lambda''_{BP} - \lambda''_{BN})$ , 此时由性质 2 可以推断: 当  $\Pr(C|[x]_R) \leq \beta'$  时,  $T_N \leq T''_B$ . 由于定义 6 分段关系模式 2 中规定了  $\alpha'' < \gamma < \beta'$ , 所以, 当  $\Pr(C|[x]_R) \leq \alpha''$  时,  $T_N > T''_B$ .

同样, 因为分段延迟代价敏感三支决策模型的性质 4 中规定了  $(\lambda_{PP} - \lambda_{PN}) < (\lambda''_{BP} - \lambda''_{BN})$ , 由性质 2 可以推断: 当  $\Pr(C|[x]_R) \leq \alpha''$  时,  $T_P \geq T''_B$ .

所以, 根据上述两个结论可以证明: 当  $\Pr(C|[x]_R) \leq \alpha''$  时,  $T_N < T''_B \leq T'_P$ , 即当  $\Pr(C|[x]_R) < \alpha''$  时,  $x \in NEG_{(\alpha, \beta)}(C)$  代价最小.

当分段延迟代价敏感三支决策模型为关系模式 2 时, 由上述证明可得, 采用负域视角优先分段策略的

3WD 为

$$\begin{aligned} \Pr(C|[x]_R) \geq \alpha' &\rightarrow x \in POS_{(\alpha,\beta)}(C), \\ \alpha' > \Pr(C|[x]_R) > \gamma &\rightarrow x \in BND_{(\alpha,\beta)}(C), \\ \Pr(C|[x]_R) \leq \gamma &\rightarrow x \in NEG_{(\alpha,\beta)}(C). \end{aligned}$$

即, 采用负域视角优先分段策略的 3WD 阈值为  $\alpha=\alpha', \beta=\gamma$ , 故定理 2(1)得证.

定理 2(2)的证明过程与定理 2(1)的证明过程类似, 从略.  $\square$

**定理 3.** 当分段延迟代价敏感三支决策模型属于关系模式 3 时,

- (1) 采用负域视角优先分段策略的 3WD 阈值为  $\alpha=\gamma, \beta=\beta'$ ;
- (2) 采用正域视角优先分段策略的 3WD 阈值为  $\alpha=\alpha', \beta=\gamma$ .

证明: 定理 3(1)的证明方法与定理 2(2)的证明类似, 定理 3(2)的证明方法与定理 2(1)的证明类似.  $\square$

**定理 4.** 当分段延迟代价敏感三支决策模型属于关系模式 4 时,

- (1) 采用负域视角优先分段策略的 3WD 退化为二分类, 阈值为  $\alpha=\gamma, \beta=\gamma$ ;
- (2) 采用正域视角优先分段策略的 3WD 退化为二分类, 阈值为  $\alpha=\gamma, \beta=\gamma$ .

证明: 假设某分段延迟三支决策为模式 4, 并且采用负域视角优先分段策略实施 3WD.

首先证明: 当  $\gamma \leq \Pr(C|[x]_R)$  时, 需要采用正域视角延迟决策代价目标函数  $T'_B$  与  $T_P, T_N$  参与三支决策代价最小化推理.

由于定义 6 规定了关系模式 4 满足  $\beta > \gamma > \alpha'$  和  $\beta' < \gamma < \alpha'$ , 所以,  $\gamma \leq \Pr(C|[x]_R)$  可被进一步划分为  $\gamma < \Pr(C|[x]_R) \leq \beta$  和  $\Pr(C|[x]_R) > \beta$  这两个证明阶段.

(1) 先证  $\gamma < \Pr(C|[x]_R) \leq \beta$  时的最优三支决策.

由于分段延迟代价敏感三支决策模型的性质 4 中约定  $(\lambda_{NP} - \lambda_{NN}) > (\lambda'_{BP} - \lambda'_{BN})$ , 因此, 由性质 1 可以推断: 当  $\Pr(C|[x]_R) \leq \beta$  时,  $T_N \leq T'_B$ . 又由于关系模式 4 满足  $\gamma < \beta$ , 所以, 当  $\gamma < \Pr(C|[x]_R) \leq \beta$  时,  $T'_B \geq T_N$  成立.

由于定义 2 分段延迟代价敏感三支决策模型中规定了  $(\lambda_{PP} - \lambda_{PN}) < 0$  和  $(\lambda_{NP} - \lambda_{NN}) > 0$ , 所以,  $\lambda_{PP} - \lambda_{PN} < \lambda_{NP} - \lambda_{NN}$ . 此时由性质 3 可以推断: 当  $\Pr(C|[x]_R) > \gamma$  时,  $T_P < T_N$ . 又由于定义 6 约定了关系模式 4 中  $\gamma < \beta$ , 所以, 当  $\gamma < \Pr(C|[x]_R) \leq \beta$  时,  $T_P < T_N$  也成立.

根据上述两个结论可以推断: 当  $\gamma < \Pr(C|[x]_R) \leq \beta$  时,  $T'_B \geq T_N > T_P$ , 即当  $\gamma < \Pr(C|[x]_R) \leq \beta$  时,  $x \in POS_{(\alpha,\beta)}(C)$  代价最小.

(2) 再证  $\Pr(C|[x]_R) > \beta$  时的最优三支决策.

由于分段延迟代价敏感三支决策模型的性质 4 中规定了  $(\lambda_{PP} - \lambda_{PN}) < (\lambda'_{BP} - \lambda'_{BN})$ , 由性质 1 可以推断: 当  $\Pr(C|[x]_R) > \alpha'$  时,  $T_P < T'_B$ . 又由于  $\alpha' < \gamma < \beta$ , 所以, 当  $\Pr(C|[x]_R) > \beta$  时,  $T_P < T'_B$ .

由分段延迟代价敏感三支决策模型的性质 4 中可得  $(\lambda_{NP} - \lambda_{NN}) > (\lambda'_{BP} - \lambda'_{BN})$ , 由性质 2 可以推断: 当  $\Pr(C|[x]_R) > \beta$  时,  $T_N > T'_B$ .

根据上述两个结论可以推断: 当  $\Pr(C|[x]_R) > \beta$  时,  $T_N > T'_B > T_P$ . 即当  $\Pr(C|[x]_R) > \beta$  时,  $x \in POS_{(\alpha,\beta)}(C)$  代价最小.

其次证明当  $\Pr(C|[x]_R) < \gamma$  时, 需要采用负域视角延迟决策代价目标函数  $T''_B$  与  $T_P, T_N$  参与三支决策代价最小化推理.

$\Pr(C|[x]_R) < \gamma$  可被进一步划分为  $\alpha'' < \Pr(C|[x]_R) \leq \gamma$  和  $\Pr(C|[x]_R) \leq \alpha''$  这两个证明阶段.

(1) 证明  $\alpha'' < \Pr(C|[x]_R) \leq \gamma$  时的最优三支决策.

由分段延迟代价敏感三支决策模型的性质 4 中可知  $(\lambda_{NP} - \lambda_{NN}) > (\lambda_{PP} - \lambda_{PN})$ , 因此, 由性质 3 可以推断: 当  $\Pr(C|[x]_R) < \gamma$  时,  $T_N < T_P$ . 又由于关系模式 4 中规定  $\alpha'' < \gamma$ , 所以, 当  $\alpha'' < \Pr(C|[x]_R) \leq \gamma$  时,  $T_N < T_P$ .

由分段延迟代价敏感三支决策模型的性质 4 中可知  $(\lambda''_{BP} - \lambda''_{BN}) > (\lambda_{PP} - \lambda_{PN})$ , 因此, 由性质 1 可以推断: 当  $\Pr(C|[x]_R) > \alpha''$  时,  $T''_B > T_P$ . 又由于关系模式 4 约定  $\alpha'' < \gamma$ , 所以, 当  $\alpha'' < \Pr(C|[x]_R) \leq \gamma$  时,  $T''_B > T_P$ .

所以, 根据上述两个结论可以证明: 当  $\alpha'' < \Pr(C|[x]_R) \leq \gamma$  时,  $T''_B > T_P > T_N$  成立.

即当  $\alpha'' < \Pr(C|[x]_R) \leq \gamma$  时,  $x \in NEG_{(\alpha, \beta)}(C)$  代价最小.

(2) 证明  $\Pr(C|[x]_R) \leq \alpha''$  时的最优三支决策.

由分段延迟代价敏感三支决策模型的性质 4 中可知  $(\lambda_{NP} - \lambda_{NN}) > (\lambda_{BP}'' - \lambda_{BN}'')$ , 由性质 2 可以推断: 当  $\Pr(C|[x]_R) < \beta''$  时,  $T_N < T_B''$ . 又由于关系模式 4 约定  $\alpha' < \beta'$ , 所以, 当  $\Pr(C|[x]_R) \leq \alpha''$  时,  $T_N < T_B''$  也成立.

由于分段延迟代价敏感三支决策模型的性质 4 中规定了  $(\lambda_{PP} - \lambda_{PN}) < (\lambda_{BP}'' - \lambda_{BN}'')$ , 由性质 1 可以推断: 当  $\Pr(C|[x]_R) \leq \alpha''$  时,  $T_P \geq T_B''$ .

所以, 根据上述两个结论可以推断: 当  $\Pr(C|[x]_R) \leq \alpha''$  时,  $T_P \geq T_B'' > T_N$ , 即当  $\Pr(C|[x]_R) < \alpha''$  时,  $x \in NEG_{(\alpha, \beta)}(C)$  代价最小.

当分段延迟代价敏感三支决策模型属于关系模式 4 时, 由上述证明可得, 采用负域视角优先分段策略的 3WD 为

$$\begin{aligned} \Pr(C|[x]_R) \geq \gamma &\rightarrow x \in POS_{(\alpha, \beta)}(C), \\ \Pr(C|[x]_R) < \gamma &\rightarrow x \in NEG_{(\alpha, \beta)}(C). \end{aligned}$$

即, 采用负域视角优先分段策略的 3WD 阈值为  $\alpha=\gamma, \beta=\gamma$ , 故定理 4(1)得证.

定理 4(2)的证明过程与定理 4(1)的证明过程类似, 从略. □

由定理 4 可以得出结论: 当分段延迟代价敏感三支决策的关系为模式 4 时, 不论采用何种分段策略的 3WD 都退化为二分类, 阈值为  $\alpha=\gamma, \beta=\gamma$ . 所以, 分段延迟代价敏感三支决策的关系为模式 4 时的 3WD 阈值退化为二支决策阈值, 记为  $\gamma$ .

### 3 分段延迟代价敏感三支决策的应用实例

上述基于分段延迟代价敏感三支决策的阈值求解方法, 可通过一个安检系统实例进行说明. 如表 1 所示, 本实例给定了各安检对象等价类  $[x_i]$  归属于安全概念  $C$  的概率信息表. 该安检系统实例数据由 16 个安检对象等价类组成, 其中,  $[x_i]$  表示第  $i$  个等价类,  $\Pr([x_i])$  表示样本属于  $[x_i]$  的概率,  $\Pr(C|[x_i])$  表示等价类  $[x_i]$  是否属于安全概念  $C$  的条件概率. 基于上述概率信息表, 采用分段延迟三支决策方法判断等价类  $[x_i]$  是否属于安全概念  $C$  的三支决策. 当某个等价类  $[x_i]$  将被决策为安全概念  $C$  的正决策域  $POS_{(\alpha, \beta)}(C)$  时, 对象  $[x_i]$  将被评定为安全对象; 当某个等价类  $[x_i]$  将被决策为安全概念  $C$  的延迟决策域  $BND_{(\alpha, \beta)}(C)$  时, 对象  $[x_i]$  将被要求延期接受进一步的安全审查; 当某个等价类  $[x_i]$  将被决策为安全概念  $C$  的负决策域  $NEG_{(\alpha, \beta)}(C)$  时, 对象  $[x_i]$  将被评定为危险对象.

表 1 概率信息表<sup>[23]</sup>

	$[x_1]$	$[x_2]$	$[x_3]$	$[x_4]$	$[x_5]$	$[x_6]$	$[x_7]$	$[x_8]$
$\Pr([x_i])$	0.093	0.088	0.093	0.089	0.069	0.046	0.019	0.015
$\Pr(C [x_i])$	1.0	0.978	0.95	0.91	0.89	0.81	0.72	0.61
	$[x_9]$	$[x_{10}]$	$[x_{11}]$	$[x_{12}]$	$[x_{13}]$	$[x_{14}]$	$[x_{15}]$	$[x_{16}]$
$\Pr([x_i])$	0.016	0.02	0.059	0.04	0.087	0.075	0.098	0.093
$\Pr(C [x_i])$	0.42	0.38	0.32	0.29	0.2	0.176	0.1	0

根据分段延迟代价敏感三支决策理论, 本实例首先给出满足定义 2 要求的正域决策损失函数  $\lambda_{PP}=0$  和  $\lambda_{NP}=0$ 、负域决策损失函数  $\lambda_{NP}=2$  和  $\lambda_{NN}=0$ . 根据上述损失函数可以获得的正域决策代价目标函数和负域决策代价目标函数为

$$\begin{cases} T_P = (0-2)\Pr(C|[x]_R) + 2 \\ T_N = (2-0)\Pr(C|[x]_R) + 0 \end{cases} \quad (7)$$

然后, 同样根据分段延迟代价敏感三支决策理论, 本实例依据定义 3 选取了 4 组不同的分段延迟代价敏感损失函数值, 见表 2, 分别用分段延迟损失函数 1-4 表示. 这 4 组分段延迟代价敏感损失函数值的选择目的是为了分别与公式(7)组合出分段延迟代价敏感三支决策模型的 4 种关系模式, 以便对本文定义的 4 种关系模式情况下分段延迟代价敏感三支决策推理进行验证与分析.

表 2 分段延迟损失函数表  $\lambda'_{BN}$ 

	$\lambda'_{BP}$	$\lambda'_{BN}$	$\lambda''_{BP}$	$\lambda''_{BN}$
分段延迟损失函数 1	0.25	1	1	0.25
分段延迟损失函数 2	0.25	1	1.85	1.6
分段延迟损失函数 3	0.9	1.5	1	0.25
分段延迟损失函数 4	0.9	1.5	1.85	1.6

其中, 分段延迟损失函数 1 构成延期接受进一步安全审查的分段代价目标函数, 包括正域视角延迟决策代价目标函数和负域视角延迟决策代价目标函数为

$$\begin{cases} T'_B = (0.25 - 1)\Pr(C|[x]_R) + 1 \\ T''_B = (1 - 0.25)\Pr(C|[x]_R) + 0.25 \end{cases} \quad (8)$$

其中, 分段延迟损失函数 2 构成延期接受进一步安全审查的分段代价目标函数, 包括正域视角延迟决策代价目标函数和负域视角延迟决策代价目标函数为

$$\begin{cases} T'_B = (0.25 - 1)\Pr(C|[x]_R) + 1 \\ T''_B = (1.85 - 1.6)\Pr(C|[x]_R) + 1.6 \end{cases} \quad (9)$$

其中, 分段延迟损失函数 3 构成延期接受进一步安全审查的分段代价目标函数, 包括正域视角延迟决策代价目标函数和负域视角延迟决策代价目标函数为

$$\begin{cases} T'_B = (0.9 - 1.5)\Pr(C|[x]_R) + 1.5 \\ T''_B = (1 - 0.25)\Pr(C|[x]_R) + 0.25 \end{cases} \quad (10)$$

其中, 分段延迟损失函数 4 构成延期接受进一步安全审查的分段代价目标函数, 包括正域视角延迟决策代价目标函数和负域视角延迟决策代价目标函数为

$$\begin{cases} T'_B = (0.9 - 1.5)\Pr(C|[x]_R) + 1.5 \\ T''_B = (1.85 - 1.6)\Pr(C|[x]_R) + 1.6 \end{cases} \quad (11)$$

根据分段延迟代价敏感三支决策模型的定义, 可将上述公式(8)–公式(11)这 4 类分段延迟决策代价目标函数分别与公式(7)表示的正域决策代价目标函数  $T_P$  及负域决策代价目标函数  $T_N$  组合为 4 种典型分段延迟代价敏感三支决策模型, 记为实例 1–实例 4.

然后, 根据定义 4 求解出每个分段延迟三支决策模型实例相关的基础度量指标, 并且应用定理 1–定理 4 推理得到对应不同分段策略的三支决策阈值. 最后, 根据获得的三支决策阈值对表 1 等价类对象实例实施三支分类决策.

### 3.1 PWD-3WD模型实例1及决策分析

基于公式(7)正域及负域决策代价目标函数  $T_P$ 、 $T_N$  和公式(8)表示的分段延迟代价目标函数  $T'_B$ 、 $T''_B$  构建如下 PWD-3WD 模型实例 1:

$$\begin{cases} T_P = (0 - 2)\Pr(C|[x]_R) + 2 \\ T_N = (2 - 0)\Pr(C|[x]_R) + 0 \\ T'_B = (0.25 - 1)\Pr(C|[x]_R) + 1 \\ T''_B = (1 - 0.25)\Pr(C|[x]_R) + 0.25 \end{cases}$$

根据定义 4 计算上述分段延迟代价敏感三支决策模型实例 1, 可获得如下相关的 5 种基础度量指标:

$$\alpha=0.8, \beta=0.3636, \alpha'=0.6363, \beta'=0.2, \gamma=0.5.$$

本实例分段延迟决策代价目标函数  $T'_B$ 、 $T''_B$  与正、负域决策代价目标函数  $T_P$ 、 $T_N$  以及基础度量指标之间的拓扑关系如图 8 所示, 其基础度量指标满足定义 6.1 的条件  $\beta < \gamma < \alpha$  和  $\beta' < \gamma < \alpha'$ , 所以, 本分段延迟代价敏感三支决策模型属于关系模式 1.



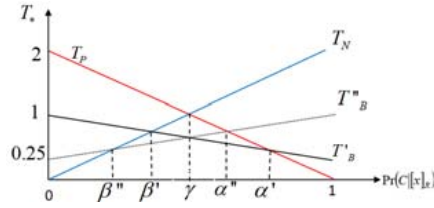


图 8 分段延迟代价敏感三支决策模型 1 示意图

根据定理 1(1), 本实例采用负域视角优先分段策略可以获得三支决策阈值为  $\alpha=0.8, \beta=0.2$ . 图 9(a)展示了本实例采用负域视角优先分段策略的最优三支决策代价关系, 其中, 与阴影部分相衔接的代价目标函数表示当前条件概率情况下选定的决策行为具有最小决策代价, 与阴影部分不衔接的代价目标函数表示其决策代价取值大于与阴影部分衔接的代价目标函数. 所以, 表 1 中满足  $\Pr(C|[x_i]) \geq 0.8$  的对象集合  $[x_i]$  属于三支决策正域, 满足  $\Pr(C|[x_i]) \leq 0.2$  的对象集合  $[x_i]$  属于三支决策负域, 满足  $0.2 < \Pr(C|[x_i]) < 0.8$  的对象集合  $[x_i]$  属于延迟决策域. 其中,  $i \in \{1, 2, 3, \dots, 16\}$ .

根据定理 1(2), 采用正域视角优先分段策略的 3WD 阈值为  $\alpha=0.3636, \beta=0.3636$ . 图 9(b)展示了本实例采用正域视角优先分段策略的最优三支决策代价关系.

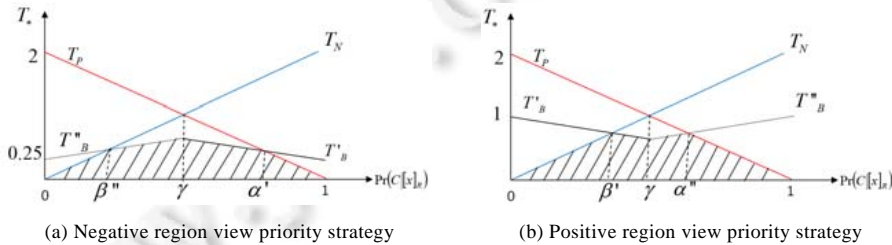


图 9 分段延迟代价敏感三支决策模型 1 的最优代价示意图

使用上述阈值对表 1 中实验对象数据实施三支分类, 得到如下三支决策域分类.

- $POS_{(\alpha,\beta)}(C) = \{[x_1], [x_2], [x_3], [x_4], [x_5], [x_6]\}$  集合中的元素被认定为安全对象;
- $BND_{(\alpha,\beta)}(C) = \{[x_7], [x_8], [x_9], [x_{10}], [x_{11}], [x_{12}]\}$  集合中的元素将被要求延期接受进一步的安全审查;
- $NEG_{(\alpha,\beta)}(C) = \{[x_{13}], [x_{14}], [x_{15}], [x_{16}]\}$  集合中的元素被认定为危险对象.

所以, 表 1 中满足  $\Pr(C|[x_i]) \geq 0.6363$  的对象集合  $[x_i]$  属于安全概念  $C$  的三支决策正域, 满足  $\Pr(C|[x_i]) \leq 0.3636$  的对象集合  $[x_i]$  属于安全概念  $C$  三支决策的负域, 满足  $0.3636 < \Pr(C|[x_i]) < 0.6363$  的对象集合  $[x_i]$  属于安全概念  $C$  的延迟决策域. 根据上述阈值对表 1 中实验对象数据实施三支分类, 得到三支决策域划分为:

- $POS_{(\alpha,\beta)}(C) = \{[x_1], [x_2], [x_3], [x_4], [x_5], [x_6], [x_7]\}$  集合中的元素被认定为安全对象;
- $BND_{(\alpha,\beta)}(C) = \{[x_8], [x_9], [x_{10}]\}$  集合中的元素被要求延期接受进一步的安全审查;
- $NEG_{(\alpha,\beta)}(C) = \{[x_{11}], [x_{12}], [x_{13}], [x_{14}], [x_{15}], [x_{16}]\}$  集合中的元素被认定为危险对象.

图 9(a)和图 9(b)分别展示了不同分段策略的代价最优三支决策示意图, 也体现了本安检系统中采用分段延迟代价敏感三支决策模型 1 中非单调的延迟代价语义.

其中, 图 9(b)采用正域视角优先的分段策略分段延迟代价目标函数语义为:

- 当  $\Pr(C|[x_i]) < \gamma$  时, 分段延迟代价目标函数为  $T'_B = (0.25 - 1)\Pr(C|[x_i]) + 1$ ;
- 当  $\Pr(C|[x_i]) > \gamma$  时, 分段延迟代价目标函数为  $T''_B = (1 - 0.25)\Pr(C|[x_i]) + 0.25$ .

这与安全检测工作中  $\Pr(C|[x_i])$  趋近于中位数时延迟代价最小的经验相符合.

### 3.2 PWD-3WD模型实例2及决策分析

基于公式(7)正域及负域决策代价目标函数  $T_P, T_N$  和公式(9)表示的分段延迟代价目标函数  $T'_B, T''_B$  构建如

下 PWD-3WD 模型实例 2:

$$\begin{cases} T_P = (0-2)\Pr(C|[x]_R) + 2 \\ T_N = (2-0)\Pr(C|[x]_R) + 0 \\ T'_B = (0.25-1)\Pr(C|[x]_R) + 1 \\ T''_B = (1.85-1.6)\Pr(C|[x]_R) + 1.6 \end{cases}$$

根据定义 4 计算上述分段延迟代价敏感三支决策模型实例 2, 可获得如下相关的 5 种基础度量指标:

$$\alpha=0.8, \beta=0.363, \alpha'=0.177, \beta'=0.914, \gamma=0.5.$$

本模型实例分段延迟决策代价目标函数  $T'_B$ 、 $T''_B$  与正、负域决策代价目标函数  $T_P$ 、 $T_N$  以及基础度量指标之间的拓扑关系如图 10 所示, 其基础度量指标满足定义 6.2 的条件  $\beta < \gamma < \alpha$  和  $\beta' > \gamma > \alpha'$ , 所以, 本分段延迟代价敏感三支决策模型属于关系模式 2.

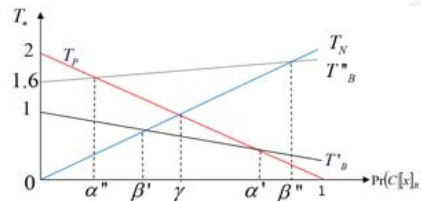


图 10 分段延迟代价敏感三支决策模型 2 示意图

根据定理 2(1), 本实例采用负域视角优先分段策略可以获得三支决策阈值为  $\alpha=0.8$ ,  $\beta=0.5$ . 图 11(a)展示了本实例采用负域视角优先分段策略的最优三支决策代价关系, 其中, 与阴影部分相衔接的代价目标函数表示当前条件概率情况下选定的决策行为具有最小决策代价, 与阴影部分不衔接的代价目标函数表示其决策代价取值大于与阴影部分衔接的代价目标函数. 所以, 表 1 中满足  $\Pr(C|[x_i]) \geq 0.8$  的对象集合  $[x_i]$  属于三支决策正域, 满足  $\Pr(C|[x_i]) \leq 0.5$  的对象集合  $[x_i]$  属于三支决策负域, 满足  $0.5 < \Pr(C|[x_i]) < 0.8$  的对象集合  $[x_i]$  属于延迟决策域, 其中,  $i \in \{1, 2, 3, \dots, 16\}$ .

使用上述阈值对表 1 中实验对象数据实施三支分类, 得到如下三支决策域分类.

- $POS_{(\alpha, \beta)}(C) = \{[x_1], [x_2], [x_3], [x_4], [x_5], [x_6]\}$  集合中的元素被认定为安全对象;
- $BND_{(\alpha, \beta)}(C) = \{[x_7], [x_8]\}$  集合中的元素将被要求延期接受进一步的安全审查;
- $NEG_{(\alpha, \beta)}(C) = \{[x_9], [x_{10}], [x_{11}], [x_{12}], [x_{13}], [x_{14}], [x_{15}], [x_{16}]\}$  集合中的元素被认定为危险对象.

根据定理 2(2), 采用正域视角优先分段策略的 3WD 阈值为  $\alpha=0.5$ ,  $\beta=0.3636$ . 图 11(b)展示了本实例采用正域视角优先分段策略的最优三支决策代价关系, 所以, 表 1 中满足  $\Pr(C|[x_i]) \geq 0.5$  的对象集合  $[x_i]$  属于安全概念 C 的三支决策正域, 满足  $\Pr(C|[x_i]) \leq 0.3636$  的对象集合  $[x_i]$  属于安全概念 C 的三支决策负域, 满足  $0.3636 < \Pr(C|[x_i]) < 0.5$  的对象集合  $[x_i]$  属于安全概念 C 的延迟决策域.

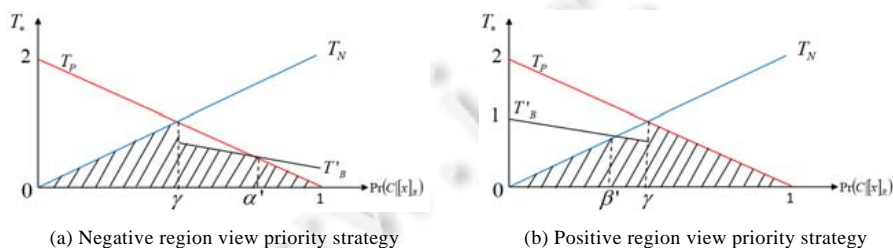


图 11 分段延迟代价敏感三支决策模型 2 的最优代价示意图

根据上述阈值对表 1 中实验对象数据实施三支分类, 得到三支决策域划分为:

- $POS_{(\alpha, \beta)}(C) = \{[x_1], [x_2], [x_3], [x_4], [x_5], [x_6], [x_7], [x_8]\}$  集合中的元素被认定为安全对象;
- $BND_{(\alpha, \beta)}(C) = \{[x_9], [x_{10}]\}$  集合中的元素被要求延期接受进一步的安全审查;

- $NEG_{(\alpha,\beta)}(C) = \{[x_{11}], [x_{12}], [x_{13}], [x_{14}], [x_{15}], [x_{16}]\}$  集合中的元素被认定为危险对象.

图 11(a)和图 11(b)分别展示了不同分段策略的代价最优三支决策示意图,也体现了本安检系统实例中分段延迟代价敏感三支决策的退化现象,其特点是,只有一个分段延迟决策代价目标函数在代价敏感决策中起作用.

### 3.3 PWD-3WD模型实例3及决策分析

基于公式(7)正域及负域决策代价目标函数  $T_P$ 、 $T_N$ 和公式(10)表示的分段延迟代价目标函数  $T'_B$ 、 $T''_B$  构建如下 PWD-3WD 模型实例 3:

$$\begin{cases} T_P = (0-2)\Pr(C|[x]_R) + 2 \\ T_N = (2-0)\Pr(C|[x]_R) + 0 \\ T'_B = (0.9-1.5)\Pr(C|[x]_R) + 1.5 \\ T''_B = (1-0.25)\Pr(C|[x]_R) + 0.25 \end{cases}$$

根据定义 4 计算上述分段延迟代价敏感三支决策模型实例 3,可获得如下相关的 5 种基础度量指标:

$$\alpha'=0.357, \beta'=0.576, \alpha''=0.6363, \beta''=0.2, \gamma=0.5.$$

本模型实例分段延迟决策代价目标函数  $T'_B$ 、 $T''_B$  与正、负域决策代价目标函数  $T_P$ 、 $T_N$  以及基础度量指标之间的拓扑关系如图 12 所示,其基础度量指标满足定义 6.3 的条件  $\alpha' < \gamma < \beta'$  和  $\beta'' < \gamma < \alpha''$ ,所以,本分段延迟代价敏感三支决策模型属于关系模式 3.

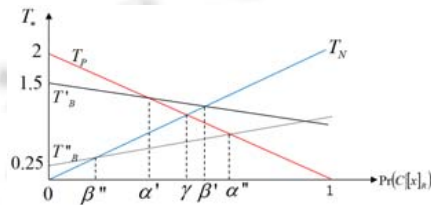


图 12 分段延迟代价敏感三支决策模型 3 示意图

根据定理 3(1),本实例采用负域视角优先分段策略可以获得三支决策阈值为  $\alpha=0.5$ ,  $\beta=0.2$ .图 13(a)展示了本实例采用负域视角优先分段策略的最优三支决策代价关系,其中,与阴影部分相临接的代价目标函数表示当前条件概率情况下选定的决策行为具有最小决策代价,与阴影部分不临接的代价目标函数表示其决策代价取值大于与阴影部分临接的代价目标函数.所以,表 1 中满足  $\Pr(C|[x_i]) \geq 0.5$  的对象集合  $[x_i]$  属于三支决策正域,满足  $\Pr(C|[x_i]) \leq 0.2$  的对象集合  $[x_i]$  属于三支决策负域,满足  $0.2 < \Pr(C|[x_i]) < 0.5$  的对象集合  $[x_i]$  属于延迟决策域.其中,  $i \in \{1, 2, 3, \dots, 16\}$ .

使用上述阈值对表 1 中实验对象数据实施三支分类,得到如下三支决策域分类.

- $POS_{(\alpha,\beta)}(C) = \{[x_1], [x_2], [x_3], [x_4], [x_5], [x_6], [x_7], [x_8]\}$  集合中的元素被认定为安全对象;
- $BND_{(\alpha,\beta)}(C) = \{[x_9], [x_{10}], [x_{11}], [x_{12}]\}$  集合中的元素被要求延期接受进一步的安全审查;
- $NEG_{(\alpha,\beta)}(C) = \{[x_{13}], [x_{14}], [x_{15}], [x_{16}]\}$  集合中的元素被认定为危险对象.

根据定理 3(2),采用正域视角优先分段策略的 3WD 阈值为  $\alpha=0.6363$ ,  $\beta=0.5$ .图 13(b)展示了本实例采用正域视角优先分段策略的最优三支决策代价关系.所以,表 1 中满足  $\Pr(C|[x_i]) \geq 0.6363$  的对象集合  $[x_i]$  属于安全概念 C 的三支决策正域,满足  $\Pr(C|[x_i]) \leq 0.5$  的对象集合  $[x_i]$  属于安全概念 C 的三支决策负域,满足  $0.5 < \Pr(C|[x_i]) < 0.6363$  的对象属于安全概念 C 的延迟决策域.根据上述阈值对表 1 中实验对象数据实施三支分类,得到三支决策域划分为:

- $POS_{(\alpha,\beta)}(C) = \{[x_1], [x_2], [x_3], [x_4], [x_5], [x_6], [x_7]\}$  集合中的元素被认定为安全对象;
- $BND_{(\alpha,\beta)}(C) = \{[x_8]\}$  集合中的元素被要求延期接受进一步的安全审查;
- $NEG_{(\alpha,\beta)}(C) = \{[x_9], [x_{10}], [x_{11}], [x_{12}], [x_{13}], [x_{14}], [x_{15}], [x_{16}]\}$  集合中的元素被认定为危险对象.

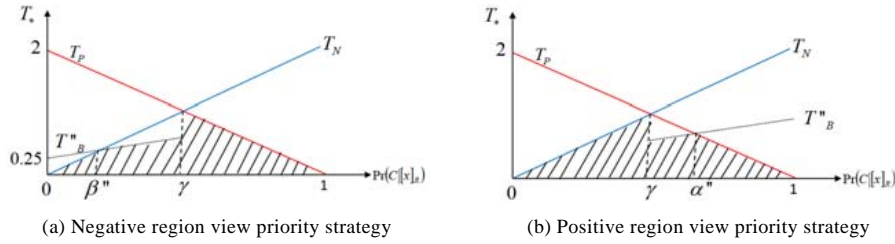


图 13 分段延迟代价敏感三支决策模型 3 的最优代价示意图

图 13(a)和图 13(b)分别展示了不同分段策略的代价最优三支决策示意图,也体现了本安检系统实例 3 分段延迟代价三支决策退化现象,其特点也是只有一个分段延迟决策代价目标函数在代价敏感决策中起作用。

3.4 PWD-3WD模型实例4及决策分析

基于公式(7)正域及负域决策代价目标函数  $T_p$ 、 $T_N$ 和公式(11)表示的分段延迟代价目标函数  $T'_B$ 、 $T''_B$  构建如下 PWD-3WD 模型实例 4:

$$\begin{cases} T_p = (0 - 2)\Pr(C|[x]_R) + 2 \\ T_N = (2 - 0)\Pr(C|[x]_R) + 0 \\ T'_B = (0.9 - 1.5)\Pr(C|[x]_R) + 1.5 \\ T''_B = (1.85 - 1.6)\Pr(C|[x]_R) + 1.6 \end{cases}$$

根据定义 4 计算上述分段延迟代价敏感三支决策模型实例 4,可以获得如下相关的 5 种基础度量指标:

$$\alpha=0.357, \beta=0.576, \alpha'=0.177, \beta'=0.914, \gamma=0.5.$$

本模型实例分段延迟决策代价目标函数  $T'_B$ 、 $T''_B$ 与正、负域决策代价目标函数  $T_p$ 、 $T_N$ 以及基础度量指标之间的拓扑关系如图 14 所示,其基础度量指标满足定义 6.4 的条件满足  $\alpha < \gamma < \beta$ 和  $\alpha' < \gamma < \beta'$ ,所以,本分段延迟代价敏感三支决策模型属于关系模式 4.

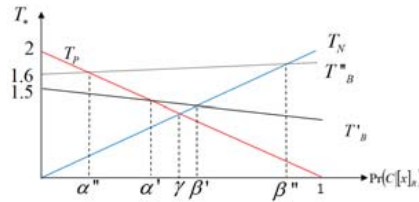


图 14 分段延迟代价敏感三支决策模型 4 示意图

根据定理 4(1)和定理 4(2),可以采用负域视角优先分段策略的三支决策阈值和正域视角优先分段策略的三支决策阈值都是  $\alpha=0.5, \beta=0.5$ . 图 15 展示了本实例采用正或者负域视角优先分段策略的最优三支决策代价关系. 所以,表 1 中满足  $\Pr(C|[x_i]) \geq 0.5$  的对象集合属于安全概念 C 的三支决策正域,满足  $\Pr(C|[x_i]) < 0.5$  的对象属于安全概念 C 的三支决策负域. 根据上述阈值对表 1 中实验对象数据实施三支分类,得到三支决策域划分为:

- $POS_{(\alpha,\beta)}(C) = \{[x_1],[x_2],[x_3],[x_4],[x_5],[x_6],[x_7],[x_8]\}$  集合中的元素被认定为安全对象;
- $NEG_{(\alpha,\beta)}(C) = \{[x_9],[x_{10}],[x_{11}],[x_{12}],[x_{13}],[x_{14}],[x_{15}],[x_{16}]\}$  集合中的元素被认定为危险对象.

图 15 所示为分段延迟代价敏感三支决策模型 4 显示出了二支退化后的决策代价情况.

上述 4 个实例只有实例 1 采用正域视角优先的分段策略分段延迟代价目标函数语义与安全检测工作中  $\Pr(C|[x_i])$ 趋近于中位数时延迟代价最小的经验相符合. 即:当  $\Pr(C|[x_i]) < \gamma$ 时,分段延迟代价目标函数为  $T'_B = (0.25 - 1)\Pr(C|[x]_R) + 1$ ; 当  $\Pr(C|[x_i]) > \gamma$ 时,分段延迟代价目标函数为  $T''_B = (1 - 0.25)\Pr(C|[x]_R) + 0.25$ . 而其余实例都不能表达这种语义. 所以,本实例显示:分段延迟代价敏感三支决策模型的关系模式 1 及正域视角优先的分段策略是唯一能够表达本文开始部分描述的海关安检决策场景以及大学招录学生场景中的非单调延迟决策代价语义.

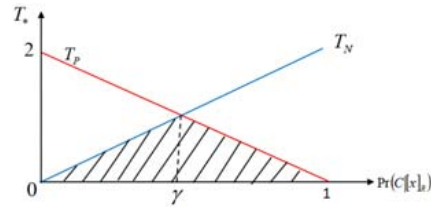


图 15 分段延迟代价敏感三支决策模型 4 的最优代价示意图

#### 4 结束语

本研究首先采用了一种新型的几何视角对代价敏感三支决策理论中各决策代价目标函数及决策阈值之间的拓扑关系进行了分析,并且针对非单调延迟代价现象构造了一种具有两段分段延迟决策代价目标函数的新型分段延迟代价敏感三支决策模型.基于分段延迟代价敏感三支决策代价目标函数之间的逻辑关系推理,拓展了一套分段延迟代价敏感三支分类阈值求解策略.最后,通过一组实验验证了本文提出的分段延迟代价敏感三支决策模型能够有效地求解非单调延迟决策三支分类问题.

本文提出的分段延迟代价敏感三支决策模型采用了双段延迟代价敏感代价目标函数,虽然能够简洁、有效地表达非单调延迟代价敏感现象,但也还存在一定的不足.例如,在现实生活中还可能更加复杂甚至平滑变化的非单调决策代价敏感形态,本文提出的由两段延迟代价敏感代价目标函数构成的分段延迟代价敏感三支决策模型只能近似地表示这种非单调延迟代价敏感状态,无法更加精确地表征这种非单调延迟代价敏感三支决策.另外,在实际应用中,各种复杂的决策代价目标场景都可能遭遇.比如在群体决策中,不同分段的决策者具有不同的复杂风险偏好(如理财投资),甚至有可能有非完全理性的感性偏好.这些不同应用场景中更加复杂的决策代价关系、更多数量的分段延迟代价目标函数的合理应用等问题,也将是本文后续的代价敏感三支决策的重要研究工作.

#### References:

- [1] Yao YY. Three-way decision: An interpretation of rules in rough set theory. In: Peng W, ed. Proc. of the Rough Sets and Knowledge Technology. Gold Coast: Springer Int'l Publishing, 2009. 642–649.
- [2] Yao YY. An outline of a theory of three-way decision. In: Li T, Nguyen HS, Wang G, eds. Proc. of the Rough Sets and Current Trends in Computing. Chengdu: Springer, 2012. 1–17.
- [3] Yao JT, Yao YY, Ziarko W. Probabilistic rough sets: Approximations, decision-makings and applications. Int'l Journal of Approximate Reasoning, 2008, 49(2): 253–254.
- [4] Zhang QH, Xue YB, Wang GY. Optimal approximation sets of rough sets. Ruan Jian Xue Bao/Journal of Software, 2016, 27(2): 295–308 (in Chinese with English abstract). <http://www.jos.org.cn/1000-9825/4854.html> [doi: 10.13328/j.cnki.jos.004854]
- [5] Yang X, Li TR, Fujita H, et al. A unified model of sequential three-way decisions and multilevel incremental processing. Knowledge-based Systems, 2017, 134(1): 172–188.
- [6] Wang PX, Yao YY. CE3: A three-way clustering method based on mathematical morphology. Knowledge-based Systems, 2018, 155(1): 54–65.
- [7] Liang DC, Liu D. Systematic studies on three-way decisions with interval-valued decision-theoretic rough sets. Information Sciences, 2014, 276(C): 186–203.
- [8] Yao YY, Wang S, Deng XF. Constructing shadowed sets and three-way approximations of fuzzy sets. Information Sciences, 2017, 412-413(1): 132–153.
- [9] Liu JB, Zhang LB, Zhou XZ, et al. Three-way decision model under intuitionistic fuzzy information system environment. Journal of Chinese Computer Systems, 2018, 39(6): 1281–1285.
- [10] Fernandes V, Yevseyeva I, Méndez J, et al. A spam filtering multi-objective optimization study covering parsimony maximization and three-way classification. Applied Soft Computing, 2016, 48(1): 111–123.

- [11] Nauman M, Azam N, Yao TT. A three-way decision making approach to malware analysis using probabilistic rough sets. *Information Sciences*, 2016, 374(1): 193–209.
- [12] Li HX, Zhang LB, Huang B, *et al.* Sequential three-way decision and granulation for cost-sensitive face recognition. *Knowledge-based Systems*, 2016, 91(C): 241–251.
- [13] Xu JF, Miao DQ, Zhang YJ, *et al.* A three-way decisions model with probabilistic rough sets for stream computing. *Int'l Journal of Approximate Reasoning*, 2017, 88: 1–22
- [14] Zhang HR, Min F. Three-way recommender systems based on random forests. *Knowledge-based Systems*, 2016, 91(C): 275–286.
- [15] Yao YY. Decision-theoretic rough set models. *Lecture Notes in Computer Science*, 2007, 178(17): 1–12.
- [16] Min F, Liu QH, A hierarchical model for test-cost-sensitive decision systems. *Information Sciences*, 2009, 179(4): 2442–2452
- [17] Yu F, Fan M. Cost-sensitive approximate attribute reduction with three-way decisions. *Int'l Journal of Approximate Reasoning*, 2019, 104: 148–165
- [18] Huaxiong L, Xianzhong Z, Jiabo, Z, *et al.* Cost-sensitive classification based on decision-theoretic rough set model. In: Li T, Nguyen HS, Wang G, eds. *Proc. of the Rough Sets and Current Trends in Computing*. Chengdu: Springer, 2012. 379–388.
- [19] Chen HM, Li TR, Luo C, *et al.* A decision-theoretic rough set approach for dynamic data mining. *IEEE Trans. on Fuzzy Systems*, 2015, 23(6): 1958–1970.
- [20] Xu WH, Guo YT. Generalized multi-granulation double-quantitative decision-theoretic rough set. *Knowledge-based Systems*, 2016, 105: 190–205.
- [21] Li HX, Zhang LB, Zhou XZ, Huang B. Cost-sensitive sequential three-way decision modeling using a deep neural network. *Int'l Journal of Approximate Reasoning*, 2017, 85: 68–78.
- [22] Ziarko W. Probabilistic rough set. In: Yao YY, ed. *Proc. of the Rough Sets, Fuzzy Sets, Data Mining, and Granular Computing*. Regina: Springer, 2005. 283–293.
- [23] Deng XF, Yao YY. A multifaceted analysis of probabilistic three-way decisions. *Fundamenta Informaticae*, 2014, 132(3): 291–313.

#### 附中文参考文献:

- [4] 张清华, 薛玉斌, 王国胤. 粗糙集的最优近似集. *软件学报*, 2016, 27(2): 295–308. <http://www.jos.org.cn/1000-9825/4854.html>  
[doi: 10.13328/j.cnki.jos.004854]



徐健锋(1973—), 男, 博士, 教授, CCF 专业会员, 主要研究领域为粒计算, 粗糙集, 三支决策, 数据挖掘, 机器学习.



张远健(1990—), 男, 博士, 主要研究领域为不确定性分析, 粒计算, 联邦学习.



苗夺谦(1964—), 男, 博士, 教授, 博士生导师, CCF 杰出会员, 主要研究领域为人工智能, 粒计算, 粗糙集, 三支决策, 深度学习, 机器学习.