10.5 矩阵博弈的混合策略解

基本信息			
教学主题	矩阵博弈的混合策略解	课时安排	15 分钟
所在章节	第 5 章　指挥决策优化方法/第 4 节　矩阵博弈的混合策略解		

【教学目标】

(1)熟悉矩阵博弈混合策略解的含义。
(2)掌握计算矩阵博弈混合策略解的方法。

【教学重点及难点】

(1)重点：2×2 矩阵博弈混合策略解的计算。
(2)难点：矩阵博弈混合策略解的内涵。

【教学结构设计】

图 10.33 所示为矩阵博弈的混合策略解教学设计框图。

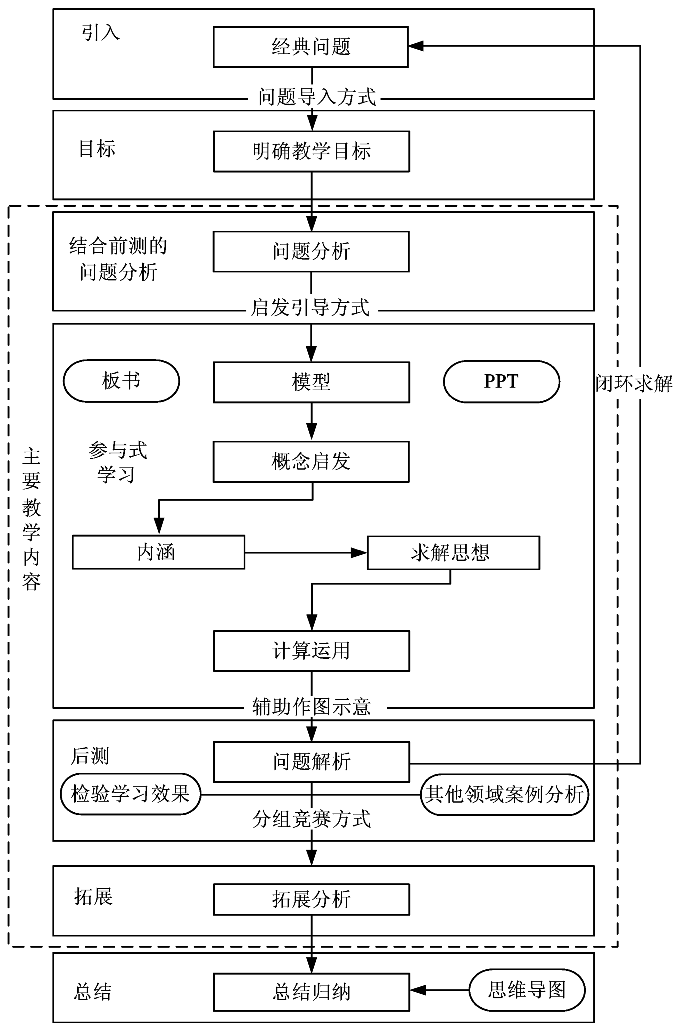

图 10.33　矩阵博弈的混合策略解教学设计框图

教学过程

课堂引入(2 分钟)

(1)问题介绍。

如图 10.34 所示，小偷试图去偷窃有守卫看守的仓库。如果守卫选择睡觉，小偷选择偷，小偷会得逞，收益是 V，守卫收益 $-D$；如果守卫选择睡觉，小偷选择不偷，守卫收益为 S，小偷收益为 0；如果守卫选择不睡觉，小偷选择偷，那么守卫抓住小偷，小偷收益为 $-P$；如果守卫选择不睡觉，小偷选择不偷，收益是均是 0。请问，他们会如何决策？

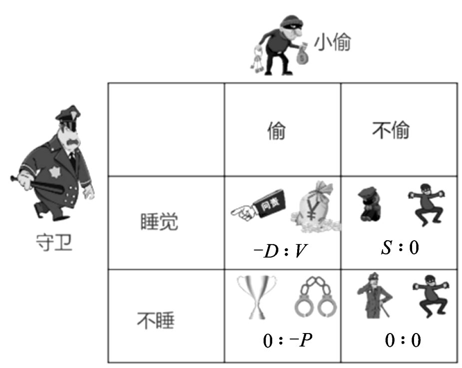

图 10.34　守卫与小偷案例

思考：这是一种什么类型的博弈问题？利用之前给大家讲过的纯策略解求取的划线法，能否找到该博弈问题的纯策略解呢？

(2)引出本讲教学目标。

熟悉矩阵博弈混合策略解的含义，掌握混合策略解的计算方法。

教学手段：问题导入式。

教学设计：通过经典问题，引入博弈模型。首先让学生分析这是一种什么类型的博弈问题。让学生利用上节课讲过的矩阵博弈纯策略解的求解方法，分析该案例的博弈模型，发现没有纯策略解，进而引出本节课学习的内容。

学情反馈：通过提问让学生自己思考分析，可以反映出学生对上节课内容的掌握程度，从而确定需不需要进行集体复习回顾。

教学内容讲解(12 分钟)

(1)混合策略解的含义。

思考：若二人博弈没有纯策略解，则说明了什么问题？以猜拳游戏为例。

提示：猜拳游戏不存在纯策略的 Nash 均衡，即双方都选不出双方可接受、能保持稳定状态的纯策略。为了不让对方猜出自己所选的策略，都只好采用随机的方法来选择策略。

设 n 人有限博弈 $\Gamma=(N,\{S_i\},\{u_i\})$，局中人 i 的策略集：$S_i=\{s_1^{(i)}, s_2^{(i)}, \cdots, s_{m_i}^{(i)}\}$，$i=1,2,\cdots,n$。若局中人 i 以概率 $x_k^{(i)}$ 选择策略，则称 $x_i=(x_1^{(i)}, x_2^{(i)}, \cdots, x_{m_i}^{(i)})$ 为局中人 i 的一个混合策略，其中 $x_k^{(i)} \geqslant 0(k=1,2,\cdots,m_i)$，$\sum_{k=1}^{m_i} x_k^{(i)}=1$。

思考：混合策略解与纯策略解的区别与联系是什么？

教学手段：
(1)游戏启发式。
(2)回顾对比式。

教学设计：
(1)游戏启发：对学生来说，上节课讲过的纯策略解比较容易理解。为了让学生更好地接受、理解混合策略解这个概念，我们以大家熟常玩的猜拳为例，首先分析不存在纯策略解的原因，进而引出混合策略解。
对于不存在 Nash 均衡的博弈，每个局中人都必须对自己选择的策略保密，否则必定会吃亏。

提示：纯策略解是混合策略解的特例(见图 10.35)。局中人 i 的第 k 个纯策略等价于第 k 个分量为 1，其余分量为 0 的混合策略。

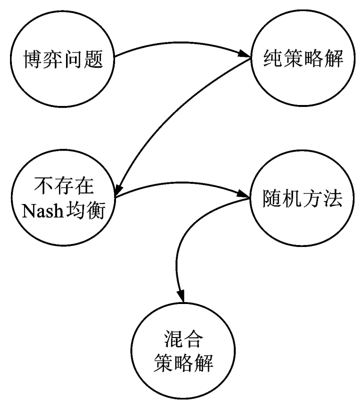

图 10.35　混合策略解与纯策略解的联系

混合策略集的定义：

设有矩阵博弈 $G=\{S_1, S_2; A, B\}$，其中，$S_1=\{\alpha_1, \quad , \alpha_m\}$，$S_2=\{\beta_1, \quad , \beta_n\}$，记

$$S_1^* = \{x \in E^m \mid x_i \geqslant 0,\ i=1, 2, \cdots, m,\ \sum_{i=1}^{m} x_i = 1\}$$

$$S_2^* = \{y \in E^n \mid y_i \geqslant 0,\ i=1, 2, \cdots, n,\ \sum_{i=1}^{n} y_i = 1\}$$

S_1^*，S_2^* 分别称为局中人 I 和局中人 II 的混合策略集，$x\in S_1^*$，$y\in S_2^*$ 分别称为局中人 I 和局中人 II 的混合策略。局中人 i 的混合策略就是策略集 S_i 上的一个概率分布，称(x, y)为一个混合局势。

(2)混合策略解的求法。

①基本思想。

按照由简至繁的原则，先考虑两人两策略的 2×2 矩阵博弈问题。

当无纯策略解时，设每个局中人选取各个策略的概率均大于 0，则针对对手策略的最优反应就是基于“无差别原则”——即确定策略的依据是使得对手选取任何策略时的收益是无差别的，从而达到一种均衡，对手无法有倾向性地进行决策。背后的理论仍是源自 Nash 均衡理论。

②2×2 矩阵博弈求解方法。

记双方的混合策略分别为：

$$P_t * = (p^0, p^1),\ P_g * = (q^0, q^1)$$

假设两个赢得矩阵为：

$$A=\begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix},\ B=\begin{bmatrix} b_{11} & b_{12} \\ b_{21} & b_{22} \end{bmatrix}$$

基于“无差别原则”，列方程组求解：

$$p^0 \cdot b_{11}+p^1 \cdot b_{21}=p^0 \cdot b_{12}+p^1 \cdot b_{22}$$
$$p^0+p^1=1$$
$$q^0 \cdot a_{11}+q^1 \cdot a_{12}=q^0 \cdot a_{21}+q^1 \cdot a_{22}$$
$$q^0+q^1=1$$

(2)回顾对比：采用与上次课中的纯策略相关概念对比的方式给学生讲解。让学生理解二者之间的区别与联系。

★课程思政：保守国家秘密无小事。

对于不存在 Nash 均衡的博弈，局中人必须对自己的策略保密，否则可能会造成灾难。保密无小事，小则误事，大则误国。尤其是国家之间的博弈，本来胜券在握，往往会因为一个泄密造成不可挽回的重大损失。

学情反馈：该部分混合策略相对于纯策略的形式发生了较大变化，比较抽象，学生初次接触会有一些难以理解。因此采用案例启发结合与上次课的纯策略的比对分析，引导学生理解这一新的概念。

教学手段：

(1)理论分析。

(2)问题解析。

(3)图示分析。

教学设计：

(1)理论分析：给出求解混合策略解的基本思想，基于此给出 2×2 矩阵博弈的求解方法。

(2)问题解析：回顾刚开始引入的小偷与守卫博弈问题。以分组竞赛的形式让学生自己动手计算小偷与守卫的混合策略解。

(3)图示分析：为了加深学生对问题的认识与理解，在公式计算的基础上，采用直观的作图分析的方法，帮助学生理解问题。

③问题解析。

假设守卫选择睡觉的概率是 p^0，选择不睡觉的概率是 p^1，则 $p^0+p^1=1$。

$$\text{收益}=V\times p^0-p\times p^1(\text{偷})$$

$$\text{收益}=0\times p^0-0\times p^1(\text{不偷})$$

$$\begin{cases}V\times p^0-P\times p^1=0\\p^0+p^1=1\end{cases}$$

求解出守卫睡觉的概率为 $p^0=\dfrac{P}{V+P}$，则不睡的概率为 $p^1=\dfrac{V}{V+P}$。

假设小偷选择偷的概率是 q^0，选择不偷的概率是 q^1，则 $q^0+q^1=1$

$$\text{收益}=-D\times q^0+S\times q^1(\text{睡})$$

$$\text{收益}=0\times q^0-0\times q^1(\text{不睡})$$

$$\begin{cases}-D\times q^0+S\times q^1=0\\q^0+q^1=1\end{cases}$$

求解出守卫睡觉的概率为 $q^0=\dfrac{S}{D+S}$，不睡的概率为 $q^0=\dfrac{D}{D+S}$。

为了便于理解，画图演示，如图 10. 36、图 10. 37 所示。

结果分析：由结果可见，加重对守卫的处罚，从长期看，是会降低盗窃发生的概率的。加重对小偷的处罚，只能短期抑制盗窃率。它的主要作用是守卫可以有更多的机会偷懒。

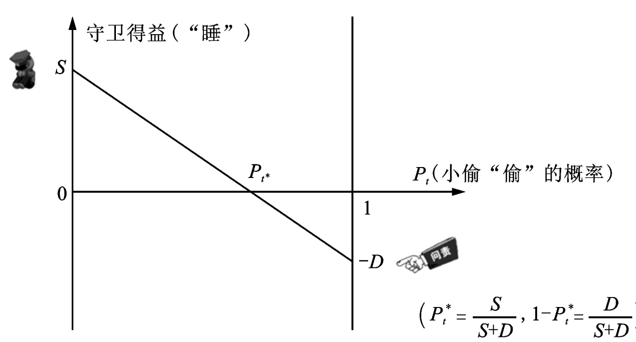

图 10. 36　守卫与小偷案例几何法 1

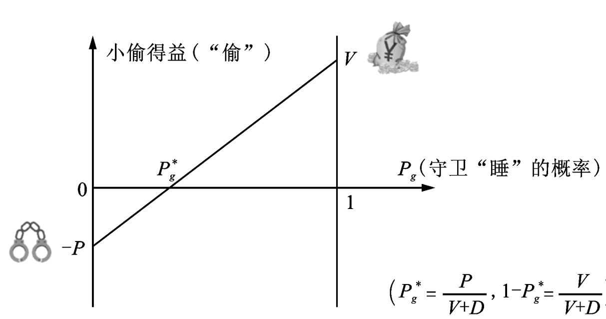

图 10. 37　守卫与小偷案例几何法解析 2

学情反馈：以如何找到混合策略解这一具体问题为驱动，调动起学生学习理论的兴趣，学生急切地期望获得求解方法。只是对背后的理论层面的理解还不够深入，下次课会介绍混合策略解的求解定理，建议学生利用课后查阅相关的理论支撑，做好预习，加深对问题的理解。

横轴表示小偷选择“偷”策略的概率，取值在 0～1。纵轴表示对应于小偷偷窃的不同概率，守卫选择睡这个策略的期望得益。由图 10. 36、图 10. 37 可见，假设小偷选择偷的概率大于 P_t^*，守卫选择睡的期望得益小于 0，因此，他肯定选择不睡，小偷偷则被抓，对小偷来说不可取。小偷偷的概率小于 P_t^*，则守卫睡的期望得益大于 0，天天睡觉也是合算的；即

使小偷提高一些偷的概率，即作案频繁一些，只要不大于 P_t^*，守卫都会选择睡，小偷不用担心被抓。在不被抓的前提下，对小偷而言，偷的概率越大收获越大，因此，他会使得偷的概率趋向于 P_t^*。

类似地，如果守卫睡的概率小于 P_g^*，小偷选择偷的期望得益为负，他自然就不偷了，那守卫就可以放心大胆的睡了。所以在小偷不偷的前提下，守卫自然睡得越多收益越大，所以守卫睡觉的概率会趋向 P_g^*。

下面考虑为了抑制盗窃现象而加重对小偷惩罚时会出现的结果，如图 10. 38 所示。

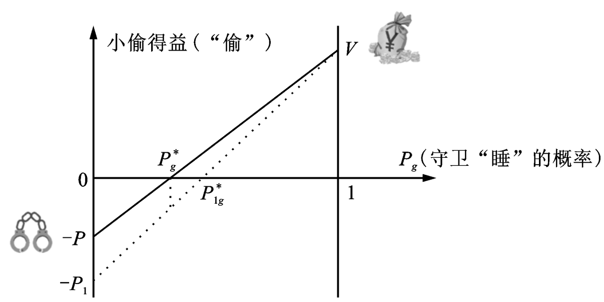

图 10. 38　守卫与小偷案例几何法解析 3

加重对小偷的惩罚会使得 P 增加。如果守卫仍然以 P_g^* 概率睡觉，小偷选择偷的期望收益变为负值，小偷会停止偷。但是长期中，小偷减少偷会使得守卫更多地选择睡，最终守卫会将睡的概率提高到 P_{1g}^*，达到新的均衡。也就是说加重对小偷的惩罚，最多只能抑制短期盗窃发生率，它的主要作用是使得守卫可以更多地偷懒了。

如果加重对守卫的处罚意味着 D 增大到 D_1，如图 10. 39 所示。此时，如果小偷选择偷的概率不变，守卫选择睡的期望收益变为负值了，守卫则选择不偷懒睡觉。守卫不睡觉，小偷就只能减少偷的概率，直到 P_t^* 下降到 P_{1t}^*，达到新的策略均衡。

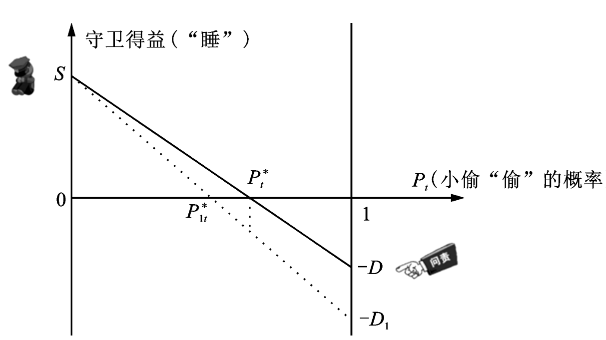

图 10. 39　守卫与小偷案例几何法解析 4

拓展分析：这个小偷与守卫的博弈称为一种“激励的悖论”。因此，在监督与违法之间的博弈中，要减少违法行为，关键是加重对监管部门的惩罚。

课程总结与作业布置(1 分钟)

图 10.40 所示为课程总结思维导图。

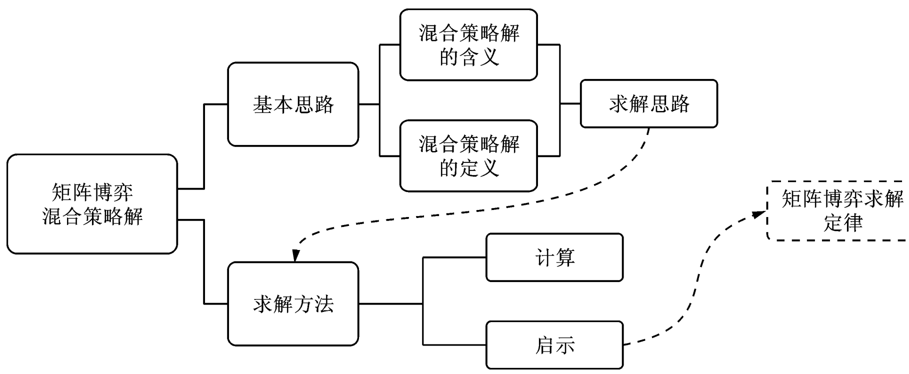

图 10.39　课程总结思维导图

(1)练习题：田忌赛马问题，是否有纯策略解？如果没有，如何利用混合策略解解释？
(2)思考题。
(1)针对军事问题：两个进攻方向时，双方进攻与防御兵力部署问题该如何求解？
(2)多人有限决策问题，如何求解？

教学反思

本次课的内容与之前讲过的内容联系紧密，一方面，要注意学生对之前的知识的掌握情况，适时复习巩固；另一方面，要结合给出的经典问题，将学过的知识融会贯通，尤其要注意强化培养学生量化处理问题的思维逻辑。

本次课知识理论性较强，混合策略解的相关概念比上次课的纯策略解要抽象得多。博弈的纳什均衡理论容易让学生产生畏难情绪，因此，概念方面要与上次课学生已经掌握的概念对比分析，有助于学生的理解；理论方面，要由简单到复杂的思路，针对简单的 2×2 矩阵博弈，简要介绍求解的基本思想，以具体的应用为主线，灵活穿插案例讲解，引起学生的兴趣，抓住学生的注意力，以达到保证教学效果的目的。课堂上可以适当引入游戏环节，比如猜拳游戏，现场演示案例中提到的两个人如何决策的问题，能充分调动学生积极性。让学生明白，生活中、各领域内都有决策问题存在，这是学习驱动力。另外，由本次课的“小偷与守卫”博弈的经典问题，拓展分析监督与违法之间的博弈以及对防止腐败的启示，潜移默化地对学生进行思政教育。