**21188142: 课程综合实践 II（数据要素交易基础）** **2025-2026 学年短学期**

# HW 1: 博弈论与多臂老虎机算法基础

教师：刘金飞，助教：吴一航 日期: 2025 年 6 月 27 日

## 1.1 占优策略均衡与纳什均衡的关系

证明如下关于占优策略均衡与纳什均衡的关系的结论：

1. 如果每个参与人 $i$ 都有一个占优于其它所有策略的策略 $s_i^*$，那么 $s^* = (s_1^*, \cdots, s_n^*)$ 是纳什均衡；
2. 如果每个参与人 $i$ 都有一个严格占优于其它所有策略的策略 $s_i^*$，那么 $s^* = (s_1^*, \cdots, s_n^*)$ 是博弈的唯一纳什均衡。

## 1.2 $N$ 人古诺竞争

假设在古诺竞争中，一共有 $J$ 家企业。当市场中所有企业总产量为 $q$ 时，市场价格为 $p(q) = a - bq$。且每个企业生产单位产品的成本都是同一个常数 $c$，即企业 $i$ 的产量为 $q_i$ 时该企业的成本为 $c_i(q_i) = c \cdot q_i$。假设 $a > c \geqslant 0$，$b > 0$，

1. 求纳什均衡下所有企业的总产量以及市场价格；
2. 讨论均衡价格随着 $J$ 变化的情况，你有什么启示?
3. 讨论 $J \to \infty$ 的均衡结果，你有什么启示?

## 1.3 公地悲剧

假设有 $I$ 个农场主，每个农场主均有权利在公共草地上放牧奶牛。一头奶牛产奶的数量取决于在草地上放牧的奶牛总量 $N$：当 $N < \overline{N}$ 时，$n_i$ 头奶牛产生的收入为 $n_i \cdot v(N)$；而当 $N \geqslant \overline{N}$ 时，$v(N) \equiv 0$。假设每头奶牛的成本为 $c$，且 $v(0) > c$，$v' < 0$，$v'' < 0$，所有农场主同时决定购买多少奶牛，所有奶牛均会在公共草地上放牧（注：假设奶牛的数量可以是小数，也就是无需考虑取整的问题）。

1. 将上述情形表达为策略式博弈；

2. 求博弈的纳什均衡下所有农场主购买的总奶牛数（可以保留表达式的形式，不用求出具体解）；
3. 求所有农场主效用之和最大（社会最优）情况下的总奶牛数（可以保留表达式的形式，不用求出具体解），并与上一问的结果比较，你能从中得到什么启示?

## 1.4 贝叶斯纳什均衡

考虑如下的不完全信息博弈：

- $I = \{1, 2\}$：1 和 2 分别是行、列参与人
- $T_1 = \{A, B\}, T_2 = \{C\}$：参与人 1 有两个类型，参与人 2 有一个类型
- $p(A, C) = \dfrac{1}{3}, p(B, C) = \dfrac{2}{3}$
- 每个参与人有两个可能的行动，下图所示矩阵给出了两种类型向量下的收益矩阵（左图为 $t = (A, C)$ 时的博弈，右图为 $t = (B, C)$ 时的博弈）：

| | $L$ | $R$ |
|---|---|---|
| $T$ | 2, 0 | 0, 3 |
| $B$ | 0, 4 | 1, 0 |

| | $L$ | $R$ |
|---|---|---|
| $T$ | 0, 3 | 3, 1 |
| $B$ | 2, 0 | 0, 1 |

求解该博弈的所有贝叶斯纳什均衡.

## 1.5 混合策略的不完全信息解释

考虑以下**抓钱博弈 (grab the dollar)**：桌子上放 1 块钱，桌子的两边坐着两个参与人，如果两人同时去抓钱，每人罚款 1 块；如果只有一人去抓，抓的人得到那块钱；如果没有人去抓，谁也得不到什么。因此，每个参与人的策略是决定抓还是不抓。

| | | 参与人 2 | |
|---|---|---|---|
| | | 抓 | 不抓 |
| 参与人 1 | 抓 | −1, −1 | 1, 0 |
| | 不抓 | 0, 1 | 0, 0 |

抓钱博弈描述的是下述现实情况：一个市场上只能有一个企业生存，有两个企业在同时决定是否进入。如果两个企业都选择进入，各亏损 100 万；如果只有一个企业进入，进入者盈利 100 万；如果没有企业进入，每个企业既不亏也不盈。

1. 求抓钱博弈的纯策略纳什均衡；

2. 求抓钱博弈的混合策略纳什均衡；

现在考虑同样的博弈但具有如下不完全信息：如果参与人 $i$ 赢了，他的利润是 $1+\theta_i$（而不是 1）。这里 $\theta_i$ 是参与人的类型，参与人 $i$ 自己知道 $\theta_i$，但另一个参与人不知道。假定 $\theta_i$ 在 $[-\varepsilon, \varepsilon]$ 区间上均匀分布。

参与人 2

| 参与人 1 | | 抓 | 不抓 |
|---|---|---|---|
| | 抓 | −1, −1 | $1+\theta_1, 0$ |
| | 不抓 | $0, 1+\theta_2$ | 0, 0 |

由于两个参与人的情况完全对称，故考虑如下对称贝叶斯纳什均衡（两个人的策略相同）形式：参与人 $i(i=1,2)$ 的策略均为

$$s_i(\theta_i) = \begin{cases} \text{抓}, & \text{如果}\theta_i \geqslant \theta^*, \\ \text{不抓}, & \text{如果}\theta_i < \theta^*. \end{cases}$$

即 $\theta^*$ 是两个参与人抓或不抓的类型分界阈值，其中 $\theta^*$ 是一个待计算确定的参数。

3. 求 $\theta^*$；

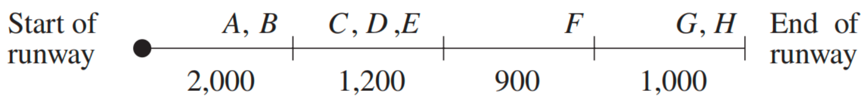
4. 当 $\varepsilon \to 0$ 时，上述贝叶斯纳什均衡会收敛于什么？从中你能得到怎样的启示。

## 1.6 飞机跑道成本分配的沙普利值计算

机场跑道的维护费用通常是向在那个机场降落飞机的航空公司来收取的。但是轻型飞机所需的跑道长度比重型飞机所需的跑道长度短，这就带来了一个问题，如何在拥有不同类型飞机的航空公司之间确定公平的维护费用分摊。

定义一个成本博弈 $(N; c)$（即每个联盟的效用是成本函数 $c$），这里 $N$ 是降落在这个机场上的所有飞机的集合，$c(S)$（对每个联盟 $S$）是能够允许联盟中所有飞机降落的最短跑道的维护费用。如果用沙普利值来确定费用的分摊，**证明：每段跑道的维护费用由使用那段跑道的飞机均摊**。

下图描绘了一个例子，其中标号为 $A, B, C, D, E, F, G$ 和 $H$ 的八架飞机每天都要在这个机场降落。每架飞机所需的跑道的整个长度由图中的区间来表示。例如，飞机 $F$ 需要前三个跑道区间。每个跑道区间的每周维护费用标示在图的下面。例如，$c(A, D, E) = 3200$，$c(A) = 2000$ 和 $c(C, F, G) = 5100$。在这一例子中，$A$ 的沙普利值恰好等于 $2000/8 = 250$，而 $F$ 的沙普利值等于 $2000/8 + 1200/6 + 900/3 = 750$。你的任务是将这一性质推广到一般的情形下给出证明（**提示：使用沙普利值的性质和公式的特点**）。




## 1.7 $\varepsilon$-贪心算法的遗憾分析

令 $\varepsilon_t = t^{-1/3}(K \log t)^{1/3}$，证明：$\varepsilon$-贪心算法的遗憾界为 $O(T^{2/3}(K \log T)^{1/3})$。

提示：整体思路是先考虑求任一时刻 $t+1$ 的期望遗憾 $\mathbb{E}[R_{t+1}]$，然后对这些遗憾求和，具体步骤如下：

1. 对于时刻 $t+1$，注意在前 $t$ 时刻中期望出现 $\sum_{i=1}^{t} \varepsilon_i$ 次探索，则每个臂被选中的平均次数为 $\sum_{i=1}^{t} \varepsilon_i / K$，然后定义事件 $E$ 为

$$|\mu_{t+1}(a) - Q_{t+1}(a)| \leqslant \sqrt{\frac{K \log t}{\sum_{i=1}^{t} \varepsilon_i}},$$

   则接下来的步骤与课上讲的贪心算法分析类似；

2. 证明任一时刻 $t+1$ 的期望遗憾 $\mathbb{E}[R_{t+1}]$ 满足

$$\mathbb{E}[R_{t+1}] \leqslant 3 \left( \frac{1}{t} K \log t \right)^{1/3} + O(t^{-2});$$

   注意其中需要用到 $\varepsilon_t$ 非增（即 $\varepsilon_t \geqslant \varepsilon_{t+1}$）的条件；

3. 将上式从 1 到 $T$ 求和并放缩得到遗憾界。