博弈论第二讲

授课时间: 2021 年 9 月 17 日　授课教师: 张家琳　记录人: 黄耀　董文浩

1 占优策略（Dominant Strategy）

参与人的策略集中，如果存在一个与其他竞争对手可能采取的策略无关的最优选择，则称其为**占优策略 (Dominant Strategy)**，与之相对的其他策略则为**劣势策略**。占优策略严格的数学定义如下：

定义 1 (参与人 i 的占优策略 s_i). 若对 $\forall s_i^{'}, s_{-i}^{'}$, 有 $u_i(s_i, s_{-i}^{'}) \geq u_i(s_i^{'}, s_{-i}^{'})$, 则称 s_i 为参与者 i 的占优策略。

定义 2 (占优策略解 s). 若对 $\forall i$, 有 s_i 是 i 的一个占优策略，则称 s 为一个占优策略解。

占优策略的经典例子有：囚徒困境 (Prisoner's dilemma)、污染博弈 (Pollution game)。

例 1 [**囚徒困境 (Prisoner's dilemma)**] 两个嫌疑犯 A 和 B 在两个房间里分别被警方讯问，他们之间无法相互沟通，每个嫌疑犯都有两个选择：坦白或者沉默。若两人都坦白，则各判刑 8 年；若两人都保持沉默则各判刑 2 年 (因为证据不足); 若一方沉默一方坦白，则沉默方判刑 10 年，坦白方判刑 1 年 (污点证人)。囚徒困境的损失矩阵如表 1 所示。

		A 的选择	
		坦白	沉默
B 的选择	坦白	8，8	1，10
	沉默	10，1	2，2

表 1: 囚徒困境损失矩阵

解 [**分析**] 为了最小化自己的损失，无论对方选择坦白还是沉默，自己选择坦白，都会使得自己判刑减少，因此对 A 和 B 个人来说，坦白是他们的占优策略。当然，这并不是两人的最优解，最优解应该是两人都保持沉默，这样只会被判刑 2 年。

例 2 [**污染博弈 (Pollutiongame)**] 假定有 n 个国家一起进行污染防治，每个国家可以选择控制污或者不控制污染。如表 2 所示，若选择控制污染，仅自身需要承担 3 单位的控制污染成本，其他国家不需要承担额外的污染成本；若选择不控制污染，每个国家都需要额外承担 1 单位的损失。在污染博弈中，每个国家都希望最小化自己的损失。

参与人 i 的策略 s_i	损失
控制污染 ($s_i = 1$)	对参与人 i 造成 3 的损失，对其他参与人不造成损失
不控制污染 ($s_i = 0$)	对每个参与人造成 1 的损失

表 2: 参与人损失定义

解 [**分析**] 根据参与人损失定义, 参与人 i 的损失为:

$$cost_i(s_1, s_2, ..., s_n) = 3s_i + \sum_{j=1}^{n}(1 - s_j) = 2s_i + n - \sum_{j \neq i} s_j$$

为了最小化自己的损失，无论什么其他参与人选择什么策略，参与人 i 都会选择不控制污染 ($s_i = 0$). 因此，对任意的参与人 i, 不控制污染 ($s_i = 0$) 是他的占优策略。

[注意] 当所有参与人都选择自己的占优策略时，每个人的损失都是 n; 一方面，社会效益最优的方案是每个参与人都选择控制污染，这样每个人的损失只有 3。由此可见，选择占优策略未必是最优结果。

2 纳什均衡 (Nash Equilibrium)

在定义纳什均衡之前，我们先给出最优反应的定义。在博弈中，假设其他人所采取的策略是已知的，那么根据其他人的策略而采取能使自己收益做大化的策略，称为**最优反应 (best response)**。

定义 3 (最优反应). *给定 s_{-i}, 若对 $\forall s_i^{'} \in S_i$, 均有 $u_i(s_i, s_{-i}) \geq u_i(s_i^{'}, s_{-i})$，则称 s_i 为参与人 i 相对于 s_{-i} 的最优反应。*

如果每个人的策略都是最优反应，那么就会形成一种稳定的局面，这时的博弈结果就是**纳什均衡 (Nash Equilibrium)**

纳什均衡，又称为非合作博弈均衡，它描述了博弈达到稳定状态时各方的策略。纳什均衡分为**纯策略纳什均衡（Pure strategy Nash equilibrium** 和**混合策略纳什均衡 (Mixed strategy Nash equilibrium)**。我们首先定义纯策略纳什均衡。

定义 4 (纯策略纳什均衡 s). *若对 $\forall$ 参与人 i，替代策略 $s_i^{'} \in S_i$, 均有 $u_i(s_i, s_{-i}) \geq u_i(s_i^{'}, s_{-i})$, 则该策略向量 $s \in S$ 被称为纳什均衡。*

在一个策略向量 $s \in S$ 达到纳什均衡时，对所有的参与人 i, s_i 都是他的最优反应。即在纳什均衡下，对任何一位参与人来说，单方面改变自己的策略不会带来任何好处。因此，纳什均衡形成，所有人都愿意保持当前的状态。因此，纳什均衡又是一种**稳定解**。

接着，我们定义混合策略以及混合策略纳什均衡。

定义 5 (混合策略). *参与人 i 的混合策略 p_i 是其策略空间 S_i 上的一个概率分布，我们用 $p = (p_1, p_2, ..., p_n)$ 来表示一个混合策略向量。*

定义 6 (混合策略纳什均衡 p). *若对 $\forall$ 参与人 i, S_i 空间上的替代混合策略 $p_i^{'}$, 均有 $E_{s\sim p}(u_i(s_i, s_{-i})) \geq E_{s_i^{'}\sim p_i^{'}, s_{-i}\sim p_{-i}}(u_i(s_i^{'}, s_{-i}))$, 则称该混合策略 p 达到纳什均衡。*

接下来的定理是博弈论最核心的定理之一。

定理 7 (纳什定理). *任何有限玩家、有限策略的博弈都存在混合策略纳什均衡* [1]。

证明 令 $G =< N, (A_i, (u_i) >$ 是一博弈，我们用向量 $(p_1, ..., p_{m_i})$ 的集合来刻画参与人 i 的混合战略集合 $\Delta(A_i)$, 其中 m_i 为有限集合 A_i 的元素个数，p_k 是参与人 i 使用第 k 个纯策略的概率，自然应有对所有 $k, p_k \geq 0$ 且 $\sum_{k=1}^{m_i} p_k = 1$。

在这种刻画方式下，我们容易看到这个向量集合作为行动集合是 {非空的，凸的，紧的} 也就是说，$\Delta(A_i)$ 是欧氏空间的一个非空紧凸子集。由于函数 U_i 是线性的，则它自然是 {连续的，在 $\Delta(A_i)$ 上拟凹的 }

参见纯策略纳什均衡的存在性定理证明，可在 $X_{j\in N}\Delta(A_i)$ 上构造符合条件的最优反应函数 B, 以应用角谷不动点定理，得到：

一定存在不动点 α^*, 使 $\alpha^* \in B(\alpha^*)$. 任何这样的不动点都是这个博弈中的混合策略纳什均衡。 □

接着，我们举几个有关纳什均衡的例子。

例 3 [**公地悲剧 (Tragedy of the commons)**]

背景设定	n 个参与人公用带宽为 1 的信道
参与人 i 的策略	要占用的带宽 $s_i \in [0,1]$
参与人 i 的收益	若 $\sum_{i=1}^n s_i > 1 : u_i(s_1,...,s_n) = 0$, 若 $\sum_{i=1}^n s_i \le 1 : u_i(s_1,...,s_n) = s_i(1-\sum_{i=1}^n s_j)$
游戏目标	个人最优：$max_{s_i} u_i(s_1,...,s_n)$ 全局最优：$max_s \sum_{i=1}^n u_i(s_1,...,s_n)$

表 3: 公地悲剧

解 [**分析**] 给定 s_{-i}, 注意到 $u_i(s_1,...,s_n) = s_i(1-\sum_{j\neq i} s_j - s_i)$, 因此当 $s_i = \frac{1}{2}\times(1-\sum_{j\neq i} s_j)$ 时，$u_i(s)$ 取得最大值。解方程组 $s_i = \frac{1}{2}\times(1-\sum_{j\neq i} s_j)$,$\forall i$, 得到方程组的解为 $s_i = \frac{1}{n+1}$。此时参与人 i 的收益为 $u_i(s) = \frac{1}{(n+1)^2}$。注意到这个解是一个纯策略纳什均衡。

另一方面 $\sum_{i=1}^n u_i(s) = \sum_{i=1}^n s_i(1-\sum_{i=1}^n s_j)$。因此，当 $\sum_{i=1}^n s_i = \frac{1}{2}$ 时，有最大全局收益 $\sum_{i=1}^n u_i(s) = \frac{1}{4}$。令 $x_i = \frac{1}{2n}$,$\forall i$, 此时每个参与人的收益为 $u_i(s) = \frac{1}{4n}$。

对比可知，在由政府统一分配带宽时，社会的总效益最高，且玩家的个人收益也高于玩家自行决策所获得的收益。这个例子说明，同占优策略一样，纳什均衡也未必是一个好的解。

例 4 [**猜拳游戏 (Rock-Paper-Scissors)**]

解 [**分析**] 首先，猜拳游戏不存在纯策略纳什均衡，因为无论双方出什么，总有一方会想要改变自己的策略，从而赢得比赛。但是，我们可以证明猜拳游戏中存在一个混合策略纳什均衡，即每个玩家完全随机地出拳，其策略的概率分布为：

	玩家 1	玩家 2
石头	$\frac{1}{3}$	$\frac{1}{3}$
剪刀	$\frac{1}{3}$	$\frac{1}{3}$
布	$\frac{1}{3}$	$\frac{1}{3}$

表 4: 石头剪刀布概率分布

从上表可以看出，当对方出拳的策略完全随机时，己方不论出石头、剪刀还是布，获胜的期望都是 0。因此，由全概率公式得知，任何混合策略获胜都是 0。由此可知，当对方的策略完全随机时，任何混合策略都是一个最优反应（当然，我方最终选择的策略也是完全随机的，因为需要保证相同的论证对方也成立）。所以上述策略是一个纳什均衡。这个例子说明，一个博弈中可能不存在纯策略纳什均衡，但存在混合策略纳什均衡。

下一个例子说明，纳什均衡不一定唯一，一个博弈中可能存在多个纳什均衡。

例 5 [**夫妻购物问题 (Couple Shopping)**]

游戏说明：夫妻两人周末有两种选择：购物与观看球赛，两人需要分别进行决策，决策的收益矩阵如下：

	购物	球赛
购物	6 5	1 1
球赛	1 1	6 5

表 5: 夫妻购物收益矩阵

解 [**分析**] 可以发现，在该问题中，只要夫妻两人做出相同的选择，博弈就能达到纳什均衡，即（购物，购物）和（球赛，球赛）两种方案均为纳什均衡。

3 机制设计 (Mechanism Design)

机制设计就是指站在一个上帝的角度，不是作为参与者，而是作为博弈游戏的设计者，来设计一种游戏机制来达到某种目的，同时研究什么样子的博弈局面是可以通过设计特殊的游戏机制来实现的。其通常需要包括以下要素：

- 第一：需要设计机制
- 第二：设计方需要有一些私有信息和共有信息，私有信息就是每个玩家不公开的信息，比如拍卖时的内心估价，公开信息就是每个玩家给设计者的信息，比如报价。
- 第三：设计者想要达到的目的，比如：希望每个玩家诚实的给出他们的私有信息，设计者获得最大的收益，达到游戏对每个玩家的公平，社会效益的最大化等等
- 第四：需要站在博弈论的视角下
- 第五：是否涉及了金钱

机制设计理论被广泛的应用在各种行业中，最典型的应用就是应用在了拍卖行业，出名的一价拍卖和二价拍卖。

例 6 [**拍卖 (Auction)**]

- 拍卖要在 n 个竞价者 (bidders) 中售卖一个物品，并且拍卖只进行一轮
- 每个竞价者对这个物品有一个心里价值估计 v_i
- 投标人 i 对这个游戏有策略：对这件拍卖品给一个报价 b_i(投标人彼此之间都不知道彼此的报价)
- 拍卖商根据报价 (b_1,b_2,...,b_n) 依据某种机制选择中标人和中标价格 p
- 中标人的收益为 v_i-p，若没有中标，则收益为 0
- 问题就是如何设计一个好的机制，满足一些特殊的要求

解 [**分析**] 最常见的两种拍卖机制就是**一价拍卖** (**first price auction**) 和**二价拍卖** (**second price auction**) 二者分别满足了一些要求，我们分别进行分析

- 在一价拍卖规则下，商品就是由出价最高的那个人获得，并且拍卖成功的那个人需要支付这个最高报价，一价拍卖是大家最容易想到，也最自然的拍卖机制，但是它有许多的性质没有满足，显然，在这种机制下，玩家是没有绝对占优策略的，这样就存在了很多的不确定性，并且这种机制不具有**诚实性**

- 在二价拍卖规则下，商品仍是由出价最高的那个人获得，但是拍卖成功只需要支付第二高的价格，这种设计，就会使每个玩家有占优策略，即报自己的内心估价，下面证明这个结论：
  该结论即是要证对于任意的 i 和任意的 $(b_1,b_2,...,b_n)$, 有 u_i $(v_i,b_{-i}) \geq u_i$ (b_i,b_{-i}), 为方便证明，我们首先假设每人出价各不相同,该假设并不会影响结论的正确性,令 b=max$(b_1,...,b_{i-1},b_{i+1},...b_n)$。

  1. 若 b>v_i, 则若要中标，必须报价比 b 还要高，然后支付 b，但是这样收益为负，因而必须要报价在 0 到 b 之间，这样收益为 0，为占优策略，因而报 v_i 也是占优策略。
  2. 若 b<v_i，则若要中标，只要报价比 b 高即可，而且只需支付 b，获得收益 v_i-b，所以只要报价在 b 到正无穷即可，为占优策略，因而报 v_i 也为占优策略。
  3. 若 b=v_i，则无论是否中标，收益均为 0，所以可以从 0 到正无穷随意报价，则报 v_i 也为占优策略。
     所以，对于每一个玩家来说，占优策略就是真实的报出自己的内心估价，每一个人都会这么报价，在一个机制中，如果所有人都如实的报出了自己的私有信息对每一个人都是占优策略，就称这个机制具有**诚实性**，在一个诚实的机制中，每一个人都会报出自己的私有信息。

所以，由以上讨论可知，二价拍卖相比于一价拍卖确实比较有优势，但是由于拍得者只需要支付第二高的报价，所以相比一价拍卖，可能会降低拍卖主的收益，所以，在实际生活中，这两种方式均有采用。

例 7 [**最适空调温度问题**]

- 一间教室中有 n 个学生，他们对于空调温度的要求并不相同。
- 每个同学 i 有自己的最适温度 T_i 作为私有信息，同时有自己的报告温度 b_i 作为公开信息。
- 决策者需要设计一种机制决定空调最终温度 T
- 对于每一个同学 i，其优化目标为 min $|T - T_i|$
- 若以社会福利最大为目标，则机制需要满足优化目标：min $\sum_{i=1}^{n} |T - T_i|$
- 若以社会公平为目标，则机制需要满足优化目标：min max($|T - T_i|(n \geq i \geq 1)$)

Reference

[1] A Tutorial on the Proof of the Existence of Nash Equilibria by A.X.Jiang, K.Leyton-Brown, University of British Columbia.