Lecture 4:非合作博弈论基础（二）

一、混合策略纳什均衡

1.混合策略的引入

纳什均衡是求解博弈的强大工具。然而很可惜的是，仍然存在相当一部分博弈无法找到纳什均衡，甚至是非常常见的博弈，例如石头剪刀布博弈:

	石头	剪刀	布
石头	0,0	1,-1	−1,1
剪刀	−1,1	0,0	1,-1
布	1,-1	−1,1	0,0

不难验证这个博弈没有纳什均衡——这也符合预期，毕竟石头剪刀布的游戏从来没有一个稳定的策略；

- 考虑简单的情况，例如一个参与人永远出石头，那么另一个人只要观察到这一点，就可以永远出布，这样的情况显然无法构成均衡；
- 所以可以猜想，稳定的策略必定带有随机性，各个参与人要让自己的行为不可捉摸，这就引入了混合策略（mixed strategy）的概念。

2.混合策略

定义：令$G=\left(N,(S_i)_{i\in N},(u_i)_{i\in N}\right)$为一个策略型博弈。一个混合策略(mixed strategy)是S_i上的概率分布，参与人i的混合策略集记为

$$\sum_i=\left\{\sigma_i:S_i\to[0,1]:\sum_{s_i\in S_i}\sigma_i(s_i)=1\right\}$$

其中$\sigma_i(s_i)$表示参与人i在该混合策略下选择策略s_i的概率。

- 因此混合策略就是给每个S_i中的策略（称之为纯策略（pure strategy））一个概率，然后按照这个概率随机选择策略。
- 例如在石头剪刀布博弈中，$\left(\frac{1}{3},\frac{1}{3},\frac{1}{3}\right)$就是一种混合策略，表示每个纯策略（出石头、剪刀和布）被选择的概率都是$\frac{1}{3}$。
- 纯策略是混合策略特例：只有一个策略概率为 1，其余为 0。
- 还有一个记号：对每个参与人i，令$\Delta(S_i)$为S_i上的概率分布集合，即

$$\Delta(S_i)=\left\{p:S_i\to[0,1]:\sum_{s_i\in S_i}p(s_i)=1\right\}$$

  则显然有$\sum_i=\Delta(S_i)$；
- 当S_i是连续策略空间时，求和需要替换为积分，当然本课程不讨论连续策略空间下的混合策略；
- 有混合策略后，博弈中参与人的效用函数也需要做相应的调整，需要适应有混合策略的情况。

3.博弈的混合扩展

定义：令$G=\left(N,(S_i)_{i\in N},(u_i)_{i\in N}\right)$为一个策略型博弈。$G$的混合扩展（mixed extension）是一个博弈

$$\Gamma=\left(N,(\Sigma_i)_{i\in N},(U_i)_{i\in N}\right)$$

其中$\Sigma_i=\Delta(S_i)$是参与人i的混合策略集，他的收益函数$U_i:\Sigma\to\mathbb{R}$将每个混合策略向量$\sigma=(\sigma_1,...,\sigma_n)\in\Sigma_1\times...\times\Sigma_n$映射到一个实数

$$U_i(\sigma)=\mathbb{E}[u_i(\sigma)]=\sum_{s\in S}\prod_{j=1}^{n}\sigma_j(s_j)u_i(s_1,...,s_n)$$

- 这里使用了冯诺伊曼-摩根斯坦恩效用函数：每个纯策略$(s_1,...,s_n)$出现的概率为$\prod_{j=1}^{n}\sigma_j(s_j)$，因此效用的本质是参与人$i$在混合策略向量$\sigma$下的期望收益；
- 这里还蕴含一个假定：每个参与人的行动相互独立。

4.混合策略纳什均衡

类似于纯策略纳什均衡，可以给出混合策略纳什均衡的定义：

定义：给定一个博弈的混合扩展$\Gamma=\left(N,(\Sigma_i)_{i\in N},(U_i)_{i\in N}\right)$，一个混合策略向量$\sigma^*=(\sigma_1^*,...,\sigma_n^*)$是一个混合策略纳什均衡，若对每个参与人$i$，有

$$U_i(\sigma^*)\geq U_i(\sigma_i,\sigma_{-i}^*),\forall\sigma_i\in\Sigma_i$$

例如，两个参与人都选择混合策略$\sigma_1^*=\sigma_2^*=\left(\frac{1}{3},\frac{1}{3},\frac{1}{3}\right)$时，$(\sigma_1^*,\sigma_2^*)$构成混合策略纳什均衡；

- 可以尝试根据定义验证这一结果；
- 然而一旦开始验证就会发现上述定义不适合于验证这一结果：因为需要对任意的混合策略σ_i都进行验证，展开后的表达式非常复杂。

因此引入一个更为方便的等价条件方便判断：

混合策略纳什均衡等价条件：令$G=\left(N,(S_i)_{i\in N},(u_i)_{i\in N}\right)$为一个策略型博弈，$\Gamma$为$G$的混合扩展。一个混合策略向量$\sigma^*$是$\Gamma$的混合策略纳什均衡，当且仅当对于每个参与人$i$和每一个纯策略$s_i\in S_i$，有

$$U_i(\sigma^*)\geq U_i(s_i,\sigma_{-i}^*)$$

证明: 正向推导只需要注意到纯策略也是特殊的混合策略即可。反过来，对于参与人i的每个混合策略σ_i，

$$U_i(\sigma_i,\sigma_{-i}^*)=\sum_{s_i\in S_i}\sigma_i(s_i)U_i(s_i,\sigma_{-i}^*)\leq\sum_{s_i\in S_i}\sigma_i(s_i)U_i(\sigma^*)=U_i(\sigma^*)$$

5.混合策略纳什均衡计算：最优反应

考虑如下性别大战：一对夫妻要安排他们周末的活动，可选择的活动有看足球赛（F）和听音乐会（C）。丈夫更喜欢看足球赛，而妻子更喜欢听音乐会。如果他们选择的活动不同，那么他们都不会高兴，如果他们选择的活动相同，那么他们都会高兴，只是高兴程度略有不同：

妻子

丈夫		F	C
	F	2, 1	0, 0
	C	0, 0	1, 2

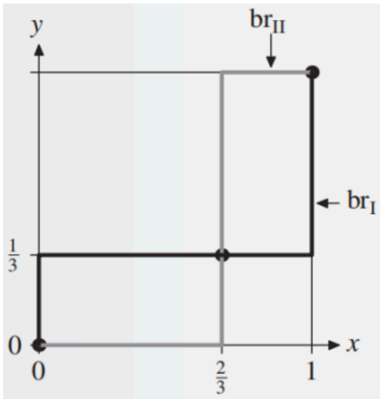
显然 (F,F) 和 (C,C) 是纯策略纳什均衡，但是否存在非纯策略纳什均衡的混合策略纳什均衡呢？

首先展示如何使用最优反应法计算混合策略纳什均衡。记丈夫的混合策略为$(x,1-x)$（表示以x的概率选择F，$1-x$的概率选择C)，妻子的混合策略为$(y,1-y)$。对于丈夫的每个混合策略$(x,1-x)$，妻子的最优反应集合为

$$\begin{aligned}\mathrm{br}_2(x) &= \arg\max_{y\in[0,1]} u_2(x,y)\\ &= \{y\in[0,1]: u_2(x,y)\geq u_2(x,z), \forall z\in[0,1]\}\end{aligned}$$

而$u_2(x,y)=xy\cdot 1+(1-x)(1-y)\cdot 2=2-2x-2y+3xy$，将$x$视为定 值，对$y$求导得到$3x-2$，因此可以得到最优反应集合为（丈夫同理）：

$$\mathrm{br}_2(x)=\begin{cases}\{0\} & x\in[0,\frac{2}{3})\\ [0,1] & x\in\left\{\frac{2}{3}\right\}\\ \{1\} & x\in\left(\frac{2}{3},1\right]\end{cases}, \mathrm{br}_1(y)=\begin{cases}\{0\} & y\in[0,\frac{1}{3})\\ [0,1] & y\in\left\{\frac{1}{3}\right\}\\ \{1\} & y\in\left(\frac{1}{3},1\right]\end{cases}$$



三个交点：$(x^*,y^*)=(0,0),(x^*,y^*)=\left(\frac{2}{3},\frac{1}{3}\right),(x^*,y^*)=(1,1)$，第 1 个 和第 3 个是纯策略纳什均衡，第 2 个是混合策略纳什均衡。

- 求混合策略纳什均衡下双方的对应的收益，你能从中得到什么启示？
- 从上面的图形能看出混合策略纳什均衡具有什么特点？

6.无差异原则

从上述例子中可以看出，混合策略纳什均衡下双方选择策略F和C的效用是相等的，这一结论可以一般化：

无差异原则：令σ^*为一个混合策略纳什均衡，s_i和s_i'为参与人i的两个纯策略，若$\sigma_i(s_i),\sigma_i^*(s_i')>0$，则$U_i(s_i,\sigma_{-i}^*)=U_i(s_i',\sigma_{-i}^*)$。

定理成立的原因很简单：如果$U_i(s_i,\sigma_{-i}^*)>U_i(s_i',\sigma_{-i}^*)$，那么参与人$i$应该增加$s_i$的概率，这样可以提高自己的收益。

- 被赋予正概率的集合称为混合策略的支撑集合；
- 问题：被严格占优的策略有可能属于混合策略的支撑集合吗；
- 问题：为什么混合策略支撑集的策略无差异，不能只选择其中一个行动或任意选取概率分布？

7.混合策略纳什均衡计算：无差异原则

接下来使用无差异原则计算性别大战的混合策略纳什均衡。使用无差异原 则时首先需要先找到纯策略纳什均衡，否则后续计算可能会忽略。纯策略 纳什均衡显然是(F, F)和(C, C)。

考虑丈夫的混合策略$\sigma_1 = (x, 1-x)$和妻子的混合策略$\sigma_2 = (y, 1-y)$，且 $0 < x < 1, 0 < y < 1$（称为完全混合的均衡）。根据无差异原则必有丈夫选择F和C的效用相等：

$$U_1(F, \sigma_2) = 2y = 1 - y = U_1(C, \sigma_2)$$

解得$y = \frac{1}{3}$，同理可以解得$x = \frac{2}{3}$。因此用无差异原则可以更简便地得到 混合策略纳什均衡。

注意，无差异原则只是取得混合策略纳什均衡的必要条件，并非充分条件，因此求出结果后需要验证。然而本例无需检验，因为本例只有两个策略，两个策略的效用都一致，不存在其他策略得到更高的效用。

8.混合策略纳什均衡的存在性与计算复杂性

尽管并非所有博弈都有纳什均衡，但是下面的纳什定理告诉我们，每个有限的策略型博弈都有至少一个混合策略纳什均衡：

纳什定理：每一个策略型博弈G，如果参与人的个数有限，每个参与人的纯策略数目有限，那么G至少有一个混合策略纳什均衡。

关于混合策略纳什均衡的计算，根据定义可以转化为线性可行性问题，有指数时间的求解方式（实验要求实现），自然的问题是，是否存在多项式时间的通用解法？答案是，不知道是否存在：

定理：双人博弈纳什均衡的计算是 PPAD 完全问题。

二、完全信息动态博弈

1.基本概念

整体博弈被表达为了一棵树，这一类博弈被称为扩展式博弈(extensive-form game)，其中

- 根节点表示博弈的开始，每个叶节点都标志博弈的一个结束点；
- 每个非叶节点上都需要标注这一步的行动者；
- 每个叶节点上需要标注博弈在这一终点下的参与人效用。

扩展式博弈每个参与人的策略是一个向量，表示其在所有可能行动的节点 上的行动。例如蜈蚣博弈中参与人 1 的策略可能是 $(C, C, S, C, ..., S, C)$； 即使选定某一策略后博弈停止，也要将此后所有节点的策略都定义好。

一个扩展式博弈的子博弈（subgame）由一个节点x和所有该节点的后继 节点组成；实际上就是x为根的子树，记为$\Gamma(x)$。

2.完美信息博弈

如果每个参与人在选择行动时，都知道他位于博弈树的哪个节点上，那么这个博弈就是完美信息博弈（game with perfect information），例如蜈蚣博弈，国际象棋等；

但很多博弈不符合这一条件，例如德州扑克或者斗地主等扑克牌游戏，你不知道其他玩家的手牌。

3.子博弈完美均衡

下面介绍完全信息动态博弈的均衡概念，需要扩展普通的纳什均衡概念。

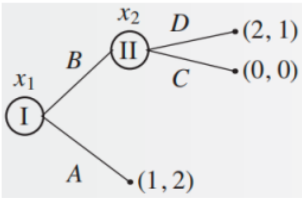
定义：在扩展式博弈Γ中，一个策略向量σ^*是子博弈完美均衡（subgame perfect equilibrium），如果对于博弈的任意子博弈$\Gamma(x)$，局限在那个子博弈的策略向量σ^*是$\Gamma(x)$的纳什均衡：对每个参与人i，每个策略σ_i和子博弈$\Gamma(x)$，

$$u_i(\sigma^* \mid x) \geq u_i(\sigma_i, \sigma_{-i}^* \mid x)$$

这一定义是很直观的，因为如果某个子博弈$\Gamma(x)$上参与人存在有利可图的偏离，那么全局来看这也是一个有利可图的偏离。

当一个博弈存在不止一个均衡时，我们希望基于合理的选择标准选择一些均衡，而剔除另一些均衡，这样的一个选择叫做均衡精炼（equilibrium refinements)。子博弈完美均衡是否是纳什均衡的精炼？换言之，是否存在不是子博弈完美均衡的纳什均衡？

4.子博弈完美均衡的例子



		Player II	
		C	D
Player I	A	1, 2	1, 2
	B	0, 0	2, 1

1. 这一博弈有两个纯策略纳什均衡：(A, C)和(B, D)，参与人 I 更偏好(B, D)，参与人 II 更偏好(A, C)；
2. (A, C)不是子博弈完美均衡，因为在x_2处参与人 II 存在有利可图的偏离：选择D而不是C（因此子博弈完美均衡的确是纳什均衡的精炼）；
3. 在 (A, C)下，I 不会偏离均衡，是因为 II 威胁 I：如果你选择B，我就选择C，然而这个威胁显然是不可置信的，因为如果 I 选择B，那么 II 还是选择D更有利。

5.子博弈完美均衡的充分条件

例子中(A, C)能作为均衡，或者说c这一被d占优的策略可以成为均衡，是因为(A, C)到不了真正要选择C, D 的x_2点。

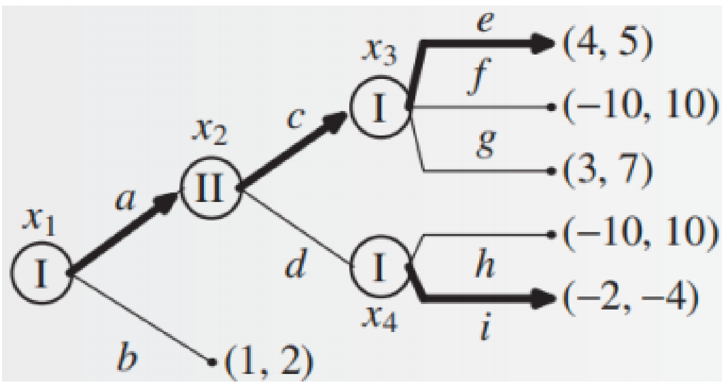
用$P_\sigma(x)$表示当实施策略向量σ时，博弈展开将造访节点x的概率。有如下定理：

定理：令σ^*是扩展式博弈Γ的纳什均衡，如果对所有x都有$P_{\sigma^*} > 0$，那么σ^*是子博弈完美均衡。

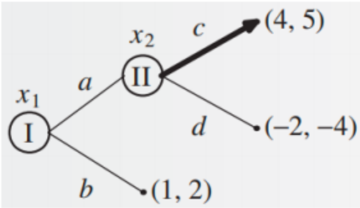
- 定理是显然的，因为如果σ^*不是子博弈完美均衡，那么在某个子博弈$\Gamma(x)$上存在有利可图的偏离，并且这个偏离产生的概率不为 0，因此也可以带来全局的有利可图的偏离；
- 推论：完全混合的纳什均衡是子博弈完美均衡。

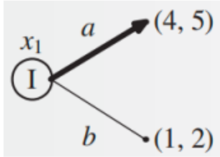
6.逆向归纳法

如何找到完美信息博弈的子博弈完美均衡？直观：要求每个子博弈都是均衡，可以从最小的子博弈出发求解：



从最小的子博弈出发，即$\Gamma(x_3)$和$\Gamma(x_4)$，选择图中加粗的策略（子博弈 的均衡），然后将均衡结果替代子博弈，逐步向上推导到根节点即可（因 此子博弈完美均衡是(ae, c)）





这一方法称为逆向归纳法（backward induction），该方法的应用保证了 每一个子博弈都使用了均衡策略，并且每一步都能做出选择，由此可得：

定理：每个有限完美信息扩展式博弈都至少有一个子博弈完美纯策略均衡。

然而逆向归纳法存在局限性：不是子博弈完美均衡的均衡可能更好：
- 重复囚徒困境有限轮，逆向归纳法会得到两个罪犯在每一轮都选择承认（因此需要新的博弈建模方式描述人们在长期关系中会合作这一事实）；

7.产量领导模型（斯塔克尔伯格模型）
经济学中子博弈完美均衡最基本的应用就是产量领导模型（或称斯塔克尔伯格（Stackelberg）模型），常用于描述有一家厂商处于支配地位或充当自然领导者的行业。例如 IBM 是具有支配地位的行业，通常观察到的其它小企业的行为模式是等待 IBM 宣布新产量然后调整自己的产量决策，此时 IBM 就是斯塔克尔伯格领导者，其它厂商是跟随者。

设市场中有两个厂商：
- 厂商 1 是领导者，选择产量y_1；
- 厂商 2 是跟随者，选择产量y_2；
- 用$p(y_1+y_2)$表示总产量为y_1+y_2时的市场价格；
- $c_1(y_1)$和$c_2(y_2)$表示厂商 1 和 2 在生产y_1和y_2单位商品时的成本。

使用逆向归纳法，第一步应该是求解在厂商 1 的任意策略y_1下厂商 2 的最优反应$y_2=f_2(y_1)$，然后厂商 1 看哪个y_1结合对应的$f_2(y_1)$能实现自己的收益最大化。因此厂商 1 的利润最大化决策可以综合表达为：

$$\max_{y_1} \pi_1(y_1,y_2)=p(y_1+y_2)y_1-c_1(y_1)$$

$$\text{s.t. } y_2=\arg\max_{y_2} \pi_2(y_1,y_2)=p(y_1+y_2)y_2-c_2(y_2)$$

- 这是一个双层优化问题（bi-level optimization problem），即优化的约束条件是另一个优化问题：厂商 1 在做决策时，他知道厂商 2 会根据他的决策做出最优反应；
- 求解过程：先求解厂商 2 的最优反应函数$y_2^*=f_2(y_1)$，然后将其代入厂商 1 的利润函数中，求解厂商 1 的最优产量y_1^*。

例：设总产量为y_1+y_2时的市场价格为$2-y_1-y_2$，并且厂商 1 和 2 的生产一件产品的单位生产成本分别为c_1,c_2，求在该假设下二者的子博弈完美均衡产量。

先写出厂商 1 和 2 的利润函数：

$$\pi_1=(2-y_1-y_2)y_1-c_1y_1$$

$$\pi_2=(2-y_1-y_2)y_2-c_2y_2$$

然后先对给定y_1的情况下求厂商 2 的最优反应，解得

$$y_2=\frac{2-y_1-c_2}{2}$$

然后将y_2代入厂商 1 的利润函数，求解得到最优的

$$y_1^* = \frac{2 + c_2 - 2c_1}{2}$$

最后将y_1代入y_2的表达式，求解得到最优的

$$y_2^* = \frac{2 + 2c_1 - 3c_2}{4}$$

- 上述结果能如何联系到实际？
- 上述求解过程和古诺竞争的区别？
- 事实上古诺竞争和斯塔克尔伯格竞争都是在纳什均衡的概念提出之前就已经被研究了，因此纳什均衡统一了这些博弈背后的思想。

三、不完全信息博弈