第一课.1 完全信息的静态博弈

王彬，暨南大学经济学系，binwang@jnu.edu.cn

阅读材料：吉本斯 1.1；课后请完成作业 1.3、作业 1.4、作业 1.5。

1. 基本概念

静态（同时行动）博弈研究的是以下一个类型的博弈：

- 参与人同时选择策略；
- 所有参与人选择的策略组合决定各参与人的收益。

此处所讲的同时行动，并不是要求参与人在同一个时点选择策略，如果一个参与人在选择策略的时候，他并不能观察到其他参与人的选择，即为静态（同时行动）博弈；反之，如果某参与人在轮到他选择的时候，他已经能观察到其他参与人的选择，即为动态（依次行动）博弈。

完全信息，是指参与人的收益函数是共同知识（common knowledge）。用更通俗的语言：

- 我知道我在不同的策略组合下的收益。
- 你知道我知道我在不同的策略组合下的收益。
- 我知道你知道我知道…
- 你知道我知道你知道我知道…
- 我知道你知道我知道你知道我知道…
- …

静态博弈一般用策略式表述(strategic-form representation)或者叫标准式表述(normal-form representation)。与之对应的动态博弈用扩展式表述(extensive-form representation)。

静态博弈$G=\{S_1,\cdots,S_n;u_1,\cdots,u_n\}$的策略式表述包括：

- **参与人**$\{1,2,\cdots,n\}$;
- 每个参与人的**策略空间**$S_i, i\in\{1,2,\cdots,n\}$;
- 每个参与人的**收益函数**$u_i(s_1,\cdots s_j,\cdots,s_n), i\in\{1,2\cdots,n\}, s_j\in S_j$，其中$(s_1,\cdots s_j,\cdots,s_n), s_j\in S_j$叫做策略组合。

我们可以通过囚徒困境这个实例来理解静态博弈的要素。

案例 1 囚徒困境：警察抓到两个盗贼嫌疑人，但无实物证据，需要两个人的口供定罪。警察将两人分别关在两个屋子里。如果两人都抗拒，最后无口供，只能按入侵私宅判一年刑罚，如果都坦白，都判三年。如一人抗拒，一人坦白，坦白从宽直接释放，抗拒从严判四年刑罚。

首先我们要判断此博弈是不是完全信息的静态博弈，参与人都知道对方的收益函数，所以是完全信息；两个囚徒分别囚禁在两个屋子，虽然不是同时点作出决策，但一方做决策时，无法看到对方的选择，各自独立作出抉择，所以是静态博弈。符合完全信息的静态博弈的要求。

这个博弈中，**参与人**有两个，囚徒 1 和囚徒 2；每个囚徒有两个策略，所以**策略空间**都是{抗拒,坦白}；因为有两个参与人，每个参与人有两种选择，所以一共有四种**策略组合**(抗拒，抗拒)、(抗拒，坦白)、(坦白，抗拒)、(坦白，坦白)，囚徒 1 在策略组合(抗拒，抗拒)的收益为-1，在策略组合(抗拒，坦白)的收益为-4，在策略组合(坦白，抗拒)的收益为 0，在策略组合(坦白，坦白)的收益为-3，囚徒 1 的**收益函数**也可以用表格（表 1）来表示。

表 1 囚徒 1 的收益函数

策略组合	收益
(抗拒，抗拒)	-1
(抗拒，坦白)	-4
(坦白，抗拒)	0
(坦白，坦白)	-3

囚徒 2 在策略组合(抗拒，抗拒)的收益为-1，在策略组合(抗拒，坦白)的收益为 0，在策略组合(坦白，抗拒)的收益为-4，在策略组合(坦白，坦白)的收益为-3。囚徒 2 的**收益函数**也可以用表格（表 2）来表示。

表 2 囚徒 2 的收益函数

策略组合	收益
(抗拒，抗拒)	-1
(抗拒，坦白)	0
(坦白，抗拒)	-4
(坦白，坦白)	-3

用语言描述会非常长，为了简便，在只有两个参与人的博弈中，我们一般用双矩阵来描述，双矩阵包括了博弈的基本三要素：参与人、策略空间以及收益函数。双矩阵的行表示囚徒 1 的策略抗拒和坦白，列表示囚徒 2 的策略抗拒和坦白，而每一个矩阵位置的内容表示在策略组合下双方的收益，第一个数字是囚徒 1 的收益，第二个位置是囚徒 2 的收益，如第一行第二列的(-4，0)表示囚徒 1 在(抗拒，坦白)这个策略组合下（囚徒 1 选择抗拒，囚徒 2 选择坦白）时的收益为-4，而囚徒 2 在这个策略组合下的收益为 0。

图 1 囚徒困境

囚徒1 \ 囚徒2	抗拒	坦白
抗拒	(-1,-1)	(-4, 0)
坦白	(0,-4)	(-3,-3)

图注：如果两人都抗拒，最后无口供，都判一年刑罚，如果都坦白，都判三年。如一人抗拒，1 人坦白，坦白从宽直接释放，抗拒从严判四年刑罚。

我们再来看先导课里面提到的三体博弈，按照小说的写法，三体文明首先向地球发送智子以及宇宙舰队水滴，并不是一个静态博弈。为了理解本模型的概念，我们假设人类已经认知到宇宙社会学，知道发射坐标可能威胁并阻止三体文明入侵，双方都知道对方的收益函数并且不能观察到对方的选择。如果地球选择发射坐标，两个文明都归零；如果地球选择和谐共处，若三体选择攻击的话，三体得到两方的资源 20，地球归零，若三体选择和谐共处，则双方都保持原状 10。

分析一下这个博弈，有两个**参与人**地球文明和三体文明；地球文明的**策略空间**是{发射坐标,和谐共处}，而三体文明的**策略空间**是{进攻地球,和谐共处}；因为有两个参与人，每个参与人都有两个策略，所以有 4 种**策略组合**，(发射坐标，进攻地球)、(发射坐标，和谐共处)、(和谐共处，进攻地球)、(和谐共处，和谐共处)，地球文明在这四种策略组合下的**收益**分别是 0、0、0、10，也可以用表格来表示**收益函数**：

表 3 地球文明的收益函数

策略组合	收益
(发射坐标，进攻地球)	0
(发射坐标，和谐共处)	0
(和谐共处，进攻地球)	0
(和谐共处，和谐共处)	10

而三体文明在这四种策略组合下的收益分别是 0、0、20、10，用表格来表示收益函数：

表 4 三体文明的收益函数

策略组合	收益
(发射坐标，进攻地球)	0
(发射坐标，和谐共处)	0
(和谐共处，进攻地球)	20
(和谐共处，和谐共处)	10

用双矩阵来表述就是：

图 2 三体博弈

三体文明

地球文明		进攻地球	和谐共处
	发射坐标	(0, 0)	(0, 0)
	和谐共处	(0, 20)	(10,10)

图注：如果地球选择发射坐标，两方都归零；如果地球选择和平共处，若三体进攻地球，地球归零，而三体得到双方都资源 20，若三体选择和谐共处，则双方保持不变 10。

可以注意到，在静态（同时行动）博弈中，博弈是一次性的(one-shot)，双方选择策略后结束。所以在静态博弈中，策略与行动(action or move)是等价的。而在动态（依次行动）博弈中，策略与行动不等价。

2. 纳什均衡

了解完全信息的静态博弈后，我们感兴趣的是博弈的均衡(equilibrium)或者说结果(outcome)，假设参与人都是**理性**的(rational)，即最大化个人的收益，最优的策略组合是什么？我们先介绍最优的策略组合的特殊情况，**占优策略均衡**，然后再回到一般情况**纳什均衡**。

2.1. 占优策略均衡

我们从完全信息的静态博弈的策略性表述可以看到，参与人的收益是策略组合的函数，其个人收益不仅取决于自己的策略选择，同时也受其他所有参与人策略选择的影响。

在一些特殊的博弈中，无论其他的参与人选择什么策略，参与人选择某一策略都能给参与者带来最大的收益，我们称这种策略为占优策略(dominant strategy)。如果每一参

与人都有占优策略，我们把所有参与人的占优策略的组合称为占优策略均衡。严格来说，

在博弈$G=\{S_1,\cdots,S_n;u_1,\cdots,u_n\}$中，令$s_i^*\in S_i$为参与人$i$策略空间中的一个策略，如果对于**任意**的其他参与人的策略组合$s_{-i}=s_1,\cdots,s_{i-1},s_{i+1},\cdots,s_n$，参与人$i$选择$s_i^*$的收益都大于**此参与人其他任意**策略的收益，即

$$u_i(s_i^*,s_{-i})>u_i(s_i',s_{-i}),\text{对于任意的 } s_{-i} \text{ 以及任意的 } s_i'\neq s_i^*$$

我们就称s_i^*为参与人i的**占优策略**。

在图 1 囚徒困境中，我们可以看到坦白都是两位囚徒的占优策略。对囚徒 1 来说，如果囚徒 2 选择抗拒，若囚徒 1 选择抗拒，收益为-1，若囚徒 1 选择坦白，收益为 0，所以在囚徒 2 选择抗拒时囚徒 1 选择坦白$(0>-1)$；

囚徒1 \ 囚徒2	抗拒
抗拒	(-1,-1)
坦白	(0,-4)

如果囚徒 2 选择坦白，若囚徒 1 选择抗拒，收益为-4，若囚徒 1 选择坦白，收益为-3，所以在囚徒 2 选择坦白时囚徒 1 仍然选择坦白$(-3>-4)$。

囚徒1 \ 囚徒2	坦白
抗拒	(-4, 0)
坦白	(-3,-3)

也就是说无论囚徒 2 选择抗拒还是坦白，囚徒 1 都会选择坦白，所以坦白是囚徒 1 的**占优策略**。

囚徒1 \ 囚徒2	抗拒	坦白
抗拒	(-1,-1)	(-4, 0)

对于囚徒 2，我们可以做类似的分析。对囚徒 2 来说，如果囚徒 1 选择抗拒，若囚徒 2 选择抗拒，收益为-1，若囚徒 2 选择坦白，收益为 0，所以在囚徒 1 选择抗拒时囚徒 2 选择坦白$(0>-1)$；

如果囚徒 1 选择坦白，若囚徒 2 选择抗拒，收益为-4，若囚徒 2 选择坦白，收益为-3，所以在囚徒 1 选择坦白时囚徒 2 仍然选择坦白($-3 > -4$)。

		囚徒 2	
		抗拒	坦白
囚徒 1	坦白	(0,-4)	(-3,-3)

因此，坦白也是囚徒 2 的**占优策略**。

囚徒 1 和囚徒 2 的占优策略组合(坦白，坦白)是囚徒困境博弈唯一的均衡和结果，占优策略均衡。

囚徒困境表明，尽管(抗拒，抗拒)是双方集体最优的策略，但是却并不是一个稳定的均衡，因为每一个囚徒都能通过转换策略到坦白而获得收益的增加；而(坦白，坦白)是个人理性下的均衡解，因为没有囚徒因为偏离这个策略获得收益的增加，反而会被判刑更重。囚徒困境表明，个人理性与集体理性在某些情况下是冲突的，合作之所以不能达到，是因为假设他人选择合作，而我偏离合作可以获得收益增加。

但并不是所有的博弈都能找到占优策略均衡，囚徒困境博弈只是一个非常特殊的例子。考虑以下智猪博弈。

案例 2 智猪博弈：一个猪圈里养着大猪小猪等待吃猪食。若大猪先吃，则大猪吃 7 份，小猪吃 1 份；若小猪先吃，则大猪小猪都吃 4 份；若同时吃，则大猪吃 5 份，小猪吃 3 份。猪食通过猪槽上的两个按钮由大猪小猪控制。若只有大猪按，则小猪先吃；若只有小猪按，则大猪先吃；若同时按，则大猪小猪一起吃，每按一次需要付出 2 单位的成本。

（同时吃，同时吃）	（5，3）	按一次需要付出 2	（按，按）	（3，1）
（先吃，后吃）	（7，1）		（等待，按）	（7，-1）
（后吃，先吃）	（4，4）		（按，等待）	（2，4）
（不吃，不吃）	（0，0）		（等待，等待）	（0，0）

案例 2 以及图 3 智猪博弈中，无论大猪选择什么，小猪等待都能获得最大收益，所以等待是小猪的占优策略，而对于大猪来说，若小猪选择按，大猪选择等待最优，若小猪选择等待，则按是最优选择，因此大猪没有占优策略。

在智猪博弈中，我们并不能找到占优策略均衡。而图 2 三体博弈中，地球文明和三体文明的任何策略都不是占优策略。

图 3 智猪博弈

大猪 \ 小猪	按	等待
按	(3, 1)	(2, 4)
等待	(7, -1)	(0, 0)

作业 1.3: 警察抓到两个盗贼嫌疑人，但无实物证据，需要两个人的口供定罪。警察将两人分别关在两个屋子里。如果两人都抗拒，最后无口供，只能按入侵私宅判R年刑罚，如果都坦白，都判P年。如一人抗拒，一人坦白，坦白从宽判T年，抗拒从严判S年刑罚。其中$R < P$。

(1) 请列出此囚徒困境博弈$G = \{S_1, S_2; u_1, u_2\}$的三要素参与者、策略空间以及收益函数。

(2) 请用双矩阵表示此囚徒困境博弈。

(3) 如果警察想要让坦白成为两位囚徒的占优策略，则R、P、T、S的相互关系是什么？

2.2. 严格劣势策略、重复剔除严格劣势策略

在一些博弈中，无论其他参与人选择何种策略，参与人的某一策略都严格优于另外一策略，我们称这种策略为相对于后一种策略的占优策略，而后一种策略成为前一种策略的劣势策略。用更正式的数学语言表述是：

在博弈$G = \{S_1, \cdots, S_n; u_1, \cdots, u_n\}$中，令$s_i', s_i'' \in S_i$为参与人$i$策略空间中的两个策略，如果对于**任意**的其他参与人的策略组合$s_{-i} = s_1, \cdots, s_{i-1}, s_{i+1}, \cdots, s_n$，参与人$i$选择$s_i'$的收益都大于选择$s_i''$，即

$$u_i(s_i', s_{-i}) > u_i(s_i'', s_{-i}), \text{对于任意的 } s_{-i}$$

我们就称s_i''为参与人i的**严格劣势策略**。如果参与人都是理性的，那其他所有的参与人都知道s_i''是参与人i的严格劣势策略，在寻找均衡的过程中，我们可以从其策略空间中剔除s_i''。对其他的参与人也做同样的动作，依次往复剔除，若$(s_1^*, \cdots, s_n^*)$是重复剔除后的唯一策略组合，我们称之为**重复剔除的占优均衡**。

大猪 \ 小猪	等待
按	(2, 4)
等待	(0, 0)

在智猪博弈中，如果大猪知道小猪是理性的，那大猪知道自己无论选择什么，小猪都会选择等待，因此可以剔除掉小猪的按策略，意识到这点，大猪会在策略空间中，选择按，因此（按，等待）是重复剔除的占优均衡。

然而，重复剔除的占优均衡仍然是特例，我们可以看到无法剔除任何一个策略（如图4），这种情况下，我们不能说任何一个策略组合是最后的均衡和结果。而图 2 三体博弈中，地球文明和三体文明的任何策略都不能被剔除。

图 4

参与人 2

参与人1	L	C	R
T	(0,4)	(4,0)	(5,3)
M	(4,0)	(0,4)	(5,3)
B	(3,5)	(3,5)	(6,6)

作业 1.4:

(1) 请用重复剔除严格劣势策略的方法找到以下博弈的均衡。

参与人 2

参与人1	左	中	右
上	(1,0)	(1,2)	(0,1)
下	(0,3)	(0,1)	(2,0)

(2) 请用重复剔除劣势策略的方法剔除以下博弈的严格劣势策略，做完这个工作后，参与人 1 和 2 还有哪些策略无法剔除。

参与人 2

参与人1	L	C	R
T	(2,0)	(1,1)	(4,2)
M	(3,4)	(1,2)	(2,3)
B	(4,4)	(0,2)	(3,3)

2.3. 纳什均衡

在图 4 中，两个参与人的每一个策略都不能被重复剔除劣势策略的过程给剔除。考虑策略组合(B,R)，如果参与人 1 选择 B，那参与人 2 的最优选择是 R；如果参与人 2 选择 R，那参与人 1 的最优选择是 B。策略组合(B,R)就是一个纳什均衡，给定一个策略

组合中所有其他参与人的策略，所有参与人都不会偏离此组合中的策略，我们就说此策略组合是一个纳什均衡。纳什均衡是策略型稳定的(strategically stable)，也可以说是自我实现的(self-enforcing)。更正式的，

在博弈$G=\{S_1,\cdots,S_n;u_1,\cdots,u_n\}$中，给定其他参与人的策略组合$s_{-i}^*=s_1^*,\cdots,s_{i-1}^*,s_{i+1}^*,\cdots,s_n^*$，$s_i^*$为参与人$i$的最优选择，即

$$u_i(s_i^*,s_{-i}^*)\geq u_i(s_i,s_{-i}^*),\text{对于任意的 } s_i$$

我们就称$s^*=(s_1^*,\cdots,s_n^*)$为博弈G的纳什均衡。

其中s_i^*最大化$u_i(s_i,s_{-i}^*)$的值，我们也称s_i^*为其他参与人策略组合s_{-i}^*的最优反应。

在双矩阵中，我们通过找最优反应来找到纳什均衡。在图 4 中，我们先找给定参与人 2 的策略时参与人 1 的最优反应。若参与人 2 选择 L，参与人 1 的最优反应为 M，我们就在策略组合(M,L)的收益组合中参与人 1 的收益下划横线，表示给定参与人 2 选择 L 时，参与人 1 的选择为 M。依次类推，我们可以知道给定某个参与人的某个选择，另一个参与人的最优反应是什么。当一个支付组合的两个数字下都有横线时，此组合为纳什均衡，因为它符合纳什均衡的定义，给定策略组合中其他参与人的选择，本参与人的最优反应是此策略组合中的策略。如图 5，(B,R)策略组合中的收益组合(6,6)下都有横线，给定参与人的选择 B，参与人的最优反应是 R；给定参与人 2 的选择 R，参与人的最优反应是 B。

图 5

参与人 2

参与人 1	L	C	R
T	(0,$\underline{4}$)	($\underline{4}$,0)	(5,3)
M	($\underline{4}$,0)	(0,$\underline{4}$)	(5,3)
B	(3,5)	(3,5)	($\underline{6}$,$\underline{6}$)

同样，我们也可以寻找图 2 三体博弈的纳什均衡。当三体文明选择进攻地球时，无论地球选择发射坐标或者选择和谐共处，地球都将毁灭归零，所以两个策略都是地球文明对三体选择进攻的最优反应；当三体文明选择和谐共处时，地球文明的最优反应是和谐共处；同理，当地球文明选择发射坐标时，无论三体文明选择进攻地球还是和谐共处，三体文明都将毁灭归零，因此两个策略都是三体文明面对地球选择发射坐标的最优反应，当地球选择和谐共处时，进攻地球是三体文明的最优反应。

由此，我们可以看到三体博弈中有两个纳什均衡(发射坐标,进攻地球)、(和谐共处,进攻地球)。因此，我们可以看到，(和谐共处，和谐共处)不是稳定解，尽管三体因为罗辑的威胁暂时收手，但三体文明在地球选择和谐共处的情况下，永远都有动力进攻地球。因此小说的结局也变成了最差占优策略均衡、严格剔除劣势策略与纳什均衡的关系

占优策略均衡是纳什均衡；如果严格剔除了是策略后，保留下来的策略组合唯一，则此唯一的策略组合为纳什均衡。

的结果(发射坐标，进攻地球)。

图 6 三体博弈

三体文明

地球文明

	进攻地球	和谐共处
发射坐标	(0, 0)	(0, 0)
和谐共处	(0, 20)	(10,10)

图注：如果地球选择发射坐标，两方都归零；如果地球选择和平共处，若三体进攻地球，地球归零，而三体得到双方都资源 20，若三体选择和谐共处，则双方保持不变 10。

作业 1.5:

(1) 请找到以下博弈的纳什均衡。

参与人 2

参与人 1

	左	中	右
上	(1,0)	(1,2)	(0,1)
下	(0,3)	(0,1)	(2,0)

(2) 请找到以下博弈的纳什均衡。

参与人 2

参与人 1

	L	C	R
T	(2,0)	(1,1)	(4,2)
M	(3,4)	(1,2)	(2,3)
B	(4,4)	(0,2)	(3,3)

2.4. 占优策略均衡、重复剔除的占优均衡与纳什均衡

占优策略均衡、重复剔除的占优均衡是纳什均衡的特殊情况，前两种均衡都是纳什均衡。