非合作博弈论基础

2024-2025 学年短学期 数据要素市场

吴一航

yhwu_is@zju.edu.cn

浙江大学计算机科学与技术学院

2024 年 7 月 4 日

单人决策问题

在之前的讨论中，我们已经见到了很多个人最大化自身效用的单人决策问题，这些问题可以被描述为

$$\max_{x} \ u(x)$$
$$\text{s.t. } x \in X$$

- 例如我们遇到的消费者最大化效用、数据卖家最大化利润问题等
- 然而在现实生活中，很多问题是多人参与的，每个人的决策会影响到其他人的效用，每个人的效用受他人的决策影响

$$\max_{x_i} \ u(x_i, x_{-i})?$$

其中 $x_{-i} = (x_1, x_2, \cdots, x_{i-1}, x_{i+1}, \cdots, x_n)$，即 x_{-i} 表示除了 i 之外的其他人的决策，这在博弈论中是常用的记号

一个例子：垄断与寡头

假设你是一个书店老板，你的书店是某个小镇上唯一的书店，每本书的成本为 20 元．你知道每个顾客只要书的价格不超过 200 元就会买，那么你的决策是什么？[1]

[1] https://www.haifeng-xu.com/cmsc35401win24/slides/lec1-intro.pdf

一个例子：垄断与寡头

假设你是一个书店老板，你的书店是某个小镇上唯一的书店，每本书的成本为 20 元．你知道每个顾客只要书的价格不超过 200 元就会买，那么你的决策是什么？[1]

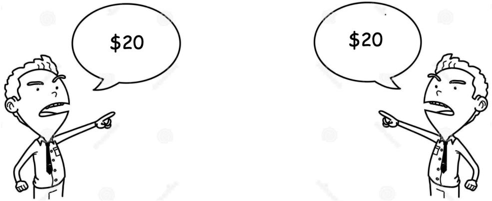
[1] https://www.haifeng-xu.com/cmsc35401win24/slides/lec1-intro.pdf

一个例子：垄断与寡头（Cont’d）

假设你是一个书店老板，但是现在小镇上有另外一个书店，你们每本书的成本均为 20 元，你们都知道每个顾客只要书的价格不超过 200 元就会买，那么你们的决策是什么?

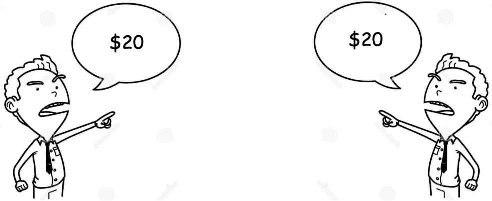
一个例子：垄断与寡头（Cont’d）

假设你是一个书店老板，但是现在小镇上有另外一个书店，你们每本书的成本均为 20 元，你们都知道每个顾客只要书的价格不超过 200 元就会买，那么你们的决策是什么?



一个例子：垄断与寡头（Cont’d）

假设你是一个书店老板，但是现在小镇上有另外一个书店，你们每本书的成本均为 20 元，你们都知道每个顾客只要书的价格不超过 200 元就会买，那么你们的决策是什么？



1. 如果另一个书店的成本是 25 元呢？
2. 如果两个书店勾结呢？

一个游戏

请各位同学不要互相交流，给出一个 0 到 100 之间的实数，提交到学在浙大上，最后所给数字最接近全班给出的数字的平均数的 1/2 的同学获胜，可以获得 2 分总评奖励

游戏的分析：凯恩斯选美博弈

请各位同学不要互相交流，给出一个 0 到 100 之间的实数，提交到学在浙大上，最后所给数字最接近全班给出的数字的平均数的 1/2 的同学获胜，可以获得 2 分总评奖励

游戏的分析：凯恩斯选美博弈

请各位同学不要互相交流，给出一个 0 到 100 之间的实数，提交到学在浙大上，最后所给数字最接近全班给出的数字的平均数的 1/2 的同学获胜，可以获得 2 分总评奖励

如果玩这个游戏的都是非常 “理性” 的人，他们会如何分析这个游戏?

1. 从随机性来看，大家给出的数字应该是均匀分布的，因此均值大约是 50，那么最接近均值的数字是 25
2. 但是大家都很理性，所有人都会认为别人会给出 25，因此此时给出 12.5 是最好的选择
3. 但是如果大家都这么想，那么给出 6.25 是最好的选择，以此类推，这一博弈的结果是所有理性人给出 0

游戏的分析：凯恩斯选美博弈（Cont'd）

这里的分析是否有值得怀疑的点?

1. 所有人都这么理性在现实中是几乎不可能的
2. 有的“聪明人”可能会搅局

游戏的分析：凯恩斯选美博弈（Cont’d）

这里的分析是否有值得怀疑的点?

1. 所有人都这么理性在现实中是几乎不可能的
2. 有的“聪明人”可能会搅局

英国著名经济学家凯恩斯在著作《就业、利息与货币通论》中提出选美博弈:

> 职业投资就像是报纸选美大赛：参赛者必须从 *100* 张照片中选出 *6* 张最美的面孔，最终选择最接近所有参赛者平均偏好的参赛者将会获胜. 所有每位参赛者面临相同的难题：他们都必须选择那些他们认为其他参赛者会最喜欢的面孔，而不是自己认为最美的面孔. 我们不能依据个人判断来选出自己认为的最美的面孔，甚至不是依据什么是最好看的平均观点做出选择，我们已经达到了第三层次，将智力投入到预测平均观点将会如何选择最美的面孔上. 并且我相信有些人已经做到了第四，第五或者更高层次.

游戏的分析：凯恩斯选美博弈（Cont’d）

这里的分析是否有值得怀疑的点?

1. 所有人都这么理性在现实中是几乎不可能的
2. 有的“聪明人”可能会搅局

英国著名经济学家凯恩斯在著作《就业、利息与货币通论》中提出选美博弈:

> 职业投资就像是报纸选美大赛：参赛者必须从 *100* 张照片中选出 *6* 张最美的面孔，最终选择最接近所有参赛者平均偏好的参赛者将会获胜．所有每位参赛者面临相同的难题：他们都必须选择那些他们认为其他参赛者会最喜欢的面孔，而不是自己认为最美的面孔．我们不能依据个人判断来选出自己认为的最美的面孔，甚至不是依据什么是最好看的平均观点做出选择，我们已经达到了第三层次，将智力投入到预测平均观点将会如何选择最美的面孔上．并且我相信有些人已经做到了第四，第五或者更高层次．

所以你在游戏中站在了第几层?

游戏的分析：凯恩斯选美博弈（Cont’d）

1. 凯恩斯认为这种思想可以解释股票的价格，股票的价格并不基于其资产的价值，甚至不是基于其他投资者对资产价值的看法，而是基于投资者所认为的其他投资者所持有的关于资产价值的一般观点，甚至以及更高层次的估计

游戏的分析：凯恩斯选美博弈（Cont’d）

1. 凯恩斯认为这种思想可以解释股票的价格，股票的价格并不基于其资产的价值，甚至不是基于其他投资者对资产价值的看法，而是基于投资者所认为的其他投资者所持有的关于资产价值的一般观点，甚至以及更高层次的估计
2. 2011 年，美国公共广播电台的 Planet Money 节目进行了如下实验：实验者首先观看分别有关猫，懒猴和北极熊的三个动物视频，再被分为两组．第一组需要回答：你认为最可爱的动物是哪一种；第二组需要回答：你认为其他实验者认为最可爱的动物是哪一种．实验结果表明，第一组回答猫的人数最多，占 50%；而第二组回答猫的人数占到了总数的 76%
   - 这一实验结果与凯恩斯的理论一致：每个人估计其他人平均的偏好与每个人自身的真实偏好的分布是不一致的

经典博弈：囚徒困境

一个犯罪团伙的两名成员 1 和 2 被捕，他们在两个独立的房间里接受审问，他们之间无法通信

	不承认	承认
不承认	(-1,-1)	(-15,0)
承认	(0,-15)	(-5,-5)

他们会如何选择?

- 这一问题的理性分析后面马上介绍
- 在分析之前我们先从前面讨论的问题中提取出博弈的基本要素

博弈的基本要素

定义 (博弈的要素)

1. 多个人进行决策
2. 博弈的行为相互影响

总而言之，是指多个人的交互决策

博弈的基本要素

定义 (博弈的要素)

1. 多个人进行决策
2. 博弈的行为相互影响

总而言之，是指多个人的交互决策

如何表示博弈?

定义 (博弈的数学表达：基础版)

博弈可以被表达为一个三元组 $G = (I, (S_i)_{i\in I}, (u_i)_{i\in I})$，其中

1. **参与人（player）集合**：$i \in I$
2. 每个参与者可以选择的**策略（strategy）**集合：S_i
3. **效用函数（payoff function）**：$u_i : S_1 \times \cdots \times S_n \to \mathbb{R}$

博弈的基本要素

定义 (博弈的要素)

1. 多个人进行决策
2. 博弈的行为相互影响

总而言之，是指多个人的交互决策

如何表示博弈?

定义 (博弈的数学表达：基础版)

博弈可以被表达为一个三元组 $G = (I, (S_i)_{i\in I}, (u_i)_{i\in I})$，其中

1. **参与人（player）集合**：$i \in I$
2. 每个参与者可以选择的**策略（strategy）**集合：S_i
3. **效用函数（payoff function）**：$u_i : S_1 \times \cdots \times S_n \to \mathbb{R}$

例如，在囚徒困境中，$I = \{1, 2\}$，$S_1 = S_2 = \{承认, 不承认\}$，u_1 和 u_2 可以通过上面的表格给出，如 $u_1(承认, 承认) = -5$

博弈的基本要素（Cont'd）

定义 (博弈的数学表达：基础版)

博弈可以被表达为一个三元组 $G = (I, (S_i)_{i\in I}, (u_i)_{i\in I})$，其中

1. **参与人（player）集合**：$i \in I$
2. 每个参与者可以选择的**策略（strategy）**集合：S_i
3. **效用函数（payoff function）**：$u_i : S_1 \times \cdots \times S_n \to \mathbb{R}$

1. 这种博弈表达称为策略式博弈（strategic game）
2. 更为复杂的博弈（扩展式、不完全信息）还会有其它描述的要素
3. 参与人我们要求是理性且智能的
    - 理性人假设在微观经济学中已经介绍，即参与人会选择最大化自身效用的行动
    - 智能人假设参与人有能力分析博弈的全局
    - 例如在竞争市场中，只要求是理性人，因为只需要最大化自身效用，并不要求对市场的全局有深刻的认识
    - 参与人理性、智能是共同知识：我知道你知道我知道你知道⋯⋯
    - 讨论：为什么我们一般都参与人是理性的情况？

博弈论

总而言之，博弈论（game theory）可以被定义为对智能的理性决策者之间冲突与合作的数学模型研究

- 博弈论为分析那些涉及两个或更多个参与者且其决策会影响相互间福利水平的情况提供了一般性的数学方法
- 近代博弈论始于 Zermelo、Borel、Von Neumann、Morgenstern 等人的工作
- 人类对于如何设计物理系统来控制自然物质已经懂得许多，但对于如何建立社会体制来调节面临冲突的人类行为却做得不够

博弈的解

博弈论的核心：给定一个博弈，关于"将会发生什么，我们能说些什么". 这一问题有至少三种不同可能的解释：

1. 经验的、描述性的解释：在给定的博弈中，参与人如何展开博弈
2. 规范的解释：在给定的博弈中，参与人"应该"如何展开博弈
3. 理论的解释：假定参与人的行为是"合理的"或"理性的"，那么我们能推测出什么

博弈的解

博弈论的核心：给定一个博弈，关于“将会发生什么，我们能说些什么”．这一问题有至少三种不同可能的解释：

1. 经验的、描述性的解释：在给定的博弈中，参与人如何展开博弈
2. 规范的解释：在给定的博弈中，参与人“应该”如何展开博弈
3. 理论的解释：假定参与人的行为是“合理的”或“理性的”，那么我们能推测出什么

第一种解释涉及对参与人实际行为的观察，偏向于心理学与行为经济学的领域；第二种解释适用于仲裁者、立法者等，他们需要根据商定的原则（如公正、效率等）决定博弈的结果；第三种解释则是通过理论方法预期一个博弈的合理结果．

博弈的解

博弈论的核心：给定一个博弈，关于“将会发生什么，我们能说些什么”．这一问题有至少三种不同可能的解释：

1. 经验的、描述性的解释：在给定的博弈中，参与人如何展开博弈
2. 规范的解释：在给定的博弈中，参与人“应该”如何展开博弈
3. 理论的解释：假定参与人的行为是“合理的”或“理性的”，那么我们能推测出什么

第一种解释涉及对参与人实际行为的观察，偏向于心理学与行为经济学的领域；第二种解释适用于仲裁者、立法者等，他们需要根据商定的原则（如公正、效率等）决定博弈的结果；第三种解释则是通过理论方法预期一个博弈的合理结果．

定义 (博弈的解/解概念)

博弈的解或**解概念**（**solution concept**）是对于一个博弈的一种预期结果，通常是一个策略组合，即参与人的行动选择，或收益的分配结果

博弈论无处不在

- 扑克、国际象棋、围棋等游戏
- 广告拍卖、频谱拍卖、价格竞争、讨价还价
- 政策制定、国家治理、选举
- 国家安全、国际关系
- ……

博弈的分类

一、 非合作博弈：参与人之间没有合作，选择行动之后效用是各自的效用，与他人无关

- 分类依据一：是否完全信息，即参与人之间是否互相知道对方的效用函数，是否知道博弈的全局信息
- 分类依据二：静态博弈或动态博弈，即参与人的行动是一次同时完成的，还是序贯进行的
- 在两个互相看不见的房子里进行石头剪刀布，不要求同时完成，但是行动的先后不会影响结果，因此是静态博弈
- 四大类博弈：完全信息静态博弈（如囚徒困境）、完全信息动态博弈（如价格领袖模型）、不完全信息静态博弈（如拍卖）、不完全信息动态博弈（如扑克牌）

博弈的分类

一、 非合作博弈：参与人之间没有合作，选择行动之后效用是各自的效用，与他人无关

- 分类依据一：是否完全信息，即参与人之间是否互相知道对方的效用函数，是否知道博弈的全局信息
- 分类依据二：静态博弈或动态博弈，即参与人的行动是一次同时完成的，还是序贯进行的
- 在两个互相看不见的房子里进行石头剪刀布，不要求同时完成，但是行动的先后不会影响结果，因此是静态博弈
- 四大类博弈：完全信息静态博弈（如囚徒困境）、完全信息动态博弈（如价格领袖模型）、不完全信息静态博弈（如拍卖）、不完全信息动态博弈（如扑克牌）

二、 合作博弈：考虑参与人之间合作后产生的联合效用

- 重点关注如何分配联合效用，有很多的解概念（收益分配方式）
- 博弈规范解释的应用（公平分配）
- 目前广泛应用于数据估值

囚徒困境与占优策略

继续我们囚徒困境的例子，一个犯罪团伙的两名成员 1 和 2 被捕，他们在两个独立的房间里接受审问，他们之间无法通信

	不承认	承认
不承认	(-1,-1)	(-15,0)
承认	(0,-15)	(-5,-5)

理性的参与者会观察到：

1. 参与人 1 发现，无论对方选择承认或不承认，自己选择承认都会比不承认效用更高
2. 参与人 2 发现，无论对方选择承认或不承认，自己选择承认都会比不承认效用更高
3. 这时我们说不承认是一个**严格劣策略（strictly dominated strategy）**，即无论对方选择什么，自己选择这个策略都是最差的

囚徒困境与占优策略（Cont’d）

定义 (严格占优)

给定参与人 i 的策略 s_i，如果他有另一个策略 t_i，使得对于所有 $s_{-i} \in S_{-i}$，$u_i(t_i, s_{-i}) > u_i(s_i, s_{-i})$，则称 s_i 是参与人 i 的一个**严格劣策略（strictly dominated strategy）**. 此时我们称 s_i 被 t_i **严格占优（strictly dominated）**，或者说 t_i **严格占优于（strictly dominates）** s_i.

囚徒困境与占优策略（Cont'd）

定义 (严格占优)

给定参与人 i 的策略 s_i，如果他有另一个策略 t_i，使得对于所有 $s_{-i} \in S_{-i}$，$u_i(t_i, s_{-i}) > u_i(s_i, s_{-i})$，则称 s_i 是参与人 i 的一个**严格劣策略（strictly dominated strategy）**. 此时我们称 s_i 被 t_i **严格占优（strictly dominated）**，或者说 t_i **严格占优于（strictly dominates）** s_i.

1. 博弈论中假定，理性人不会选择严格劣策略，这并不是一个很强的假设，是符合常识的
2. 在博弈论中，参与人是理性的是共同知识
3. 因此囚徒困境中的解是 (承认, 承认)

囚徒困境的进一步讨论

囚徒困境的结果看起来有些不合理?

1. 为什么两个人不会选择更好的点?（理性人的假设）
2. 现实中两个人选择合作，原因可能是什么?

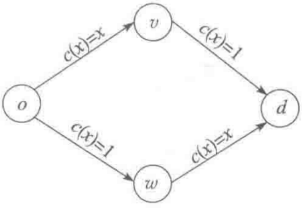
囚徒困境的进一步讨论

囚徒困境的结果看起来有些不合理?

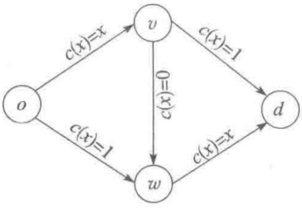
1. 为什么两个人不会选择更好的点?(理性人的假设)
2. 现实中两个人选择合作，原因可能是什么?

囚徒困境本质：出于个人理性的决策无法达到社会最优

1. 囚徒困境的例子：布雷斯悖论、公地悲剧、内卷
2. 解决囚徒困境：强制力、机制设计、长期关系



a）原始网络



b）变更后的网络

重复剔除严格劣策略

我们再来看一个例子，练习如何利用严格占优的解概念来求解博弈，假设参与人 1 代表行参与者，参与人 2 代表列参与者

	L	M	R
T	1, 0	1, 2	0, 1
B	0, 3	0, 1	2, 0

重复剔除严格劣策略

我们再来看一个例子，练习如何利用严格占优的解概念来求解博弈，假设参与人 1 代表行参与者，参与人 2 代表列参与者

	L	M	R
T	1, 0	1, 2	0, 1
B	0, 3	0, 1	2, 0

不难发现，策略 R 被策略 M 严格占优，因此我们可以剔除策略 R，得到新的博弈

	L	M
T	1, 0	1, 2
B	0, 3	0, 1

进一步地，策略 B 被策略 T 严格占优，因此我们可以剔除策略 B，最后 L 被 M 严格占优，因此最终的解是 (T, M)．这一过程被称为**重复剔除严格劣策略（iterated elimination of strictly dominated strategies）**，当然这一操作的背后假定是理性人是共同知识

弱占优

有的博弈没有严格劣策略，例如：

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

尽管没有严格劣策略，但策略 B 确实有特殊之处：和策略 T 相比，策略 B 虽然不能总是给出更好的结果，但至少不会更差，并且在列参与人选择 L 时，策略 B 会给出更好的结果．这种情况下，我们称策略 B **弱占优于（weakly dominates）**策略 T．

定义 (弱占优)

给定参与人 i 的策略 s_i，如果他有另一个策略 t_i，满足如下两个条件：

1. 对于所有 $s_{-i} \in S_{-i}$，$u_i(t_i, s_{-i}) \geqslant u_i(s_i, s_{-i})$
2. 至少存在一个 $s_{-i} \in S_{-i}$，使得 $u_i(t_i, s_{-i}) > u_i(s_i, s_{-i})$

则称 s_i 是参与人 i 的一个**弱劣策略（weakly dominated strategy）**．此时我们称 s_i 被 t_i **弱占优（weakly dominated）**，或者说 t_i **弱占优于（weakly dominates）** s_i．

弱占优（Cont'd）

1. 一般而言，除非强调严格占优，否则我们默认占优是指弱占优

弱占优（Cont’d）

1. 一般而言，除非强调严格占优，否则我们默认占优是指弱占优
2. 理性参与人不会使用（弱）劣策略
   - 可以用于重复剔除劣策略寻找博弈的解
   - 但是比严格劣策略的版本对理性的要求更强
   - 颤抖的手原则：考虑列参与人分别以 x 和 $1-x$ 的概率选择 L 和 R（$0<x<1$），那么行参与人会选择 B，因为 T 的期望效用是 $x+2(1-x)=2-x$，而 B 的期望效用是 2

弱占优（Cont'd）

1. 一般而言，除非强调严格占优，否则我们默认占优是指弱占优
2. 理性参与人不会使用（弱）劣策略
   - 可以用于重复剔除劣策略寻找博弈的解
   - 但是比严格劣策略的版本对理性的要求更强
   - 颤抖的手原则：考虑列参与人分别以 x 和 $1-x$ 的概率选择 L 和 R（$0 < x < 1$），那么行参与人会选择 B，因为 T 的期望效用是 $x + 2(1-x) = 2-x$，而 B 的期望效用是 2
3. 重复剔除劣策略的过程中如果只有严格劣策略，那么结果不依赖于剔除的顺序（自行证明），但是剔除弱劣策略的顺序可能会影响结果

	L	C	R
T	1, 2	2, 3	0, 3
M	2, 2	2, 1	3, 2
B	2, 1	0, 0	1, 0

按顺序 T, R, B, C 和 B, L, C, T 以及 T, C, R 剔除弱劣策略，最终的解不同（自行验证）

纳什均衡

考虑如下的恶策略型二人博弈：

	L	C	R
T	0, 6	6, 0	4, 3
M	6, 0	0, 6	4, 3
B	3, 3	3, 3	5, 5

- 不难发现如下博弈策略之间没有占优关系，于是我们需要思考新的解概念
- 如果一个参与人已知他人使用的策略，那么他参加的博弈实际上就是要选择一个“最佳应对”
- 考虑策略组合 (B, R)，此时每个人都不愿意单独偏离这一组合，因为 B 和 R 各自是对方的最佳应对．也就是说，如果其它参与人确实根据 (B, R) 选择了策略，那么每个人都不愿意单独偏离这一组合，所以这一策略组合是稳定的

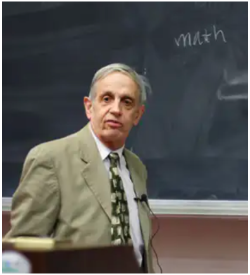
纳什均衡（Cont'd）

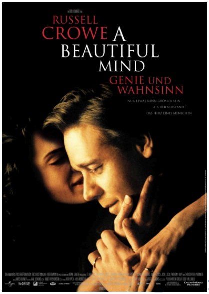
定义 (纳什均衡)

给定一个博弈，一个策略组合 $s^* = (s_1^*, \cdots, s_n^*)$ 是一个**纳什均衡（Nash equilibrium）**，如果对于每个参与人 i，有

$$u_i(s^*) \geqslant u_i(s_i, s_{-i}^*)$$

如果 $u_i(\hat{s}_i, s_{-i}) > u_i(s)$，那么策略 $\hat{s}_i$ 是参与人 i 有利可图的策略偏离，纳什均衡的策略向量不允许存在有利可图的策略偏离





纳什均衡的等价定义

定义 (最佳应对)

令 s_{-i} 为参与人 i 之外的所有参与人的策略组合，参与人 i 的策略 s_i 是 s_{-i} 的一个**最佳应对（best response）**，如果满足

$$u_i(s_i, s_{-i}) = \max_{t_i \in S_i} u_i(t_i, s_{-i})$$

纳什均衡的等价定义

定义 (最佳应对)

令 s_{-i} 为参与人 i 之外的所有参与人的策略组合，参与人 i 的策略 s_i 是 s_{-i} 的一个**最佳应对（best response）**，如果满足

$$u_i(s_i, s_{-i}) = \max_{t_i \in S_i} u_i(t_i, s_{-i})$$

定义 (纳什均衡的等价定义)

如果对于每个参与人 i，s_i^* 是 s_{-i}^* 的一个最佳应对，那么策略组合 $s^* = (s_1^*, \cdots, s_n^*)$ 是一个纳什均衡.

证明留作习题，注意等价需要证明两个方向可以互相推导

例子：协调博弈

协调博弈：博弈双方协调各自的策略对双方都有利

	a	b
A	1, 1	0, 0
B	0, 0	3, 3

显然 (A, a) 和 (B, b) 都是纳什均衡，(A, b) 和 (B, a) 都不是纳什均衡

例子：协调博弈

协调博弈：博弈双方协调各自的策略对双方都有利

	a	b
A	1, 1	0, 0
B	0, 0	3, 3

显然 (A,a) 和 (B,b) 都是纳什均衡，(A,b) 和 (B,a) 都不是纳什均衡

	F	C
F	2, 1	0, 0
C	0, 0	1, 2

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．因此 (F,F) 和 (C,C) 都是纳什均衡

例子：协调博弈

协调博弈：博弈双方协调各自的策略对双方都有利

	a	b
A	1, 1	0, 0
B	0, 0	3, 3

显然 (A,a) 和 (B,b) 都是纳什均衡，(A,b) 和 (B,a) 都不是纳什均衡

	F	C
F	2, 1	0, 0
C	0, 0	1, 2

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．因此 (F,F) 和 (C,C) 都是纳什均衡

- 夫妻对这两个均衡的偏好是什么？有没有办法进一步协调？

例子：协调博弈（Cont'd）

下图展示的也是协调博弈，称为“安全困境”，描述二战后美国和苏联的博弈情形.

	不拥有核武器	拥有核武器
拥有核武器	3, 1	2, 2
不拥有核武器	4, 4	1, 3

- 最好的结果是双方都不拥有核武器，因为制造核武器成本高昂，而且核武器容易触发战争，后果严重；稍差一点是自己拥有，对方没有，再次是双方都有，最次是自己没有，对方拥有
- 博弈有两个均衡：都拥有核武器和都不拥有核武器
- 如果双方都不拥有核武器，事实上是在冒险：因为很可能对方在造核武器，这会导致自己处境降到最差；所以双方都会倾向于制造核武器，因为如果对方制造了，那么制造核武器是最佳应对，如果对方没有制造，自己的效用也会从 2 提升到 3
- 安全困境：**更理想、更合理的均衡更危险**

例子：古诺双寡头竞争

例 (古诺双寡头竞争：《关于财富理论的数学原理的研究》)

两家制造商 1 和 2 生产相同的产品，在同一市场中竞争潜在的顾客．两家制造商同时选择产量，需求决定产品的市场价格，市场价格对两家企业而言是相同的．用 q_1 和 q_2 分别表示两家企业的产量，因此 q_1+q_2 是市场的总产量．假设供给为 q_1+q_2，每件产品的价格为 $2-q_1-q_2$．假设两家厂商的单位生产成本分别为正实数 c_1, c_2．这个博弈存在均衡吗?

这是一个两人博弈，每个参与人的策略集合是 $[0,+\infty)$，如果参与人 1 选择策略 q_1，参与人 2 选择策略 q_2，那么参与人 1 的效用（利润）是

$$u_1(q_1,q_2) = (2-q_1-q_2)q_1 - c_1q_1 = q_1(2-q_1-q_2-c_1)$$

参与人 2 的效用是

$$u_2(q_1,q_2) = q_2(2-q_1-q_2-c_2)$$

例子：古诺双寡头竞争（Cont’d）

我们使用最优反应的定义求解纳什均衡，首先求参与人 1 关于 q_2 的最优反应，即求解最大化 $u_1(q_1, q_2)$ 的 q_1：

$$\frac{\partial u_1}{\partial q_1} = 2 - 2q_1 - q_2 - c_1 = 0$$

同理对于参与人 2，我们有

$$\frac{\partial u_2}{\partial q_2} = 2 - q_1 - 2q_2 - c_2 = 0$$

联立方程解得

$$q_1^* = \frac{2 - 2c_1 + c_2}{3}, \quad q_2^* = \frac{2 - 2c_2 + c_1}{3}$$

可以验证这是纳什均衡，并且是唯一的纳什均衡.

纳什均衡的意义

例 (两个将军的协调博弈)

两个将军考虑是否要合作进攻，如果两个将军合作进攻，则各自可以领赏，没有进攻则领不到赏，但如果只有一方进攻，则进攻者会输得很惨，这个博弈会走向何方?

	攻击	不攻击
攻击	(1,1)	(-2,0)
不攻击	(0,-2)	(0,0)

纳什均衡的意义

例 (两个将军的协调博弈)

两个将军考虑是否要合作进攻，如果两个将军合作进攻，则各自可以领赏，没有进攻则领不到赏，但如果只有一方进攻，则进攻者会输得很惨，这个博弈会走向何方?

	攻击	不攻击
攻击	(1,1)	(-2,0)
不攻击	(0,-2)	(0,0)

- 如果全班 75% 的同学都选择攻击，那么攻击成功，参与攻击的获得总评 1 分的加分；如果攻击失败，那么参与攻击的同学会被扣 2 分. 你会选择什么?

纳什均衡的意义

例 (两个将军的协调博弈)

两个将军考虑是否要合作进攻，如果两个将军合作进攻，则各自可以领赏，没有进攻则领不到赏，但如果只有一方进攻，则进攻者会输得很惨，这个博弈会走向何方？

	攻击	不攻击
攻击	(1,1)	(-2,0)
不攻击	(0,-2)	(0,0)

- 如果全班 75% 的同学都选择攻击，那么攻击成功，参与攻击的获得总评 1 分的加分；如果攻击失败，那么参与攻击的同学会被扣 2 分．你会选择什么？
- 现实往往达不到纳什均衡：需要充分的交流或多次博弈达到稳态
- 或者有一个仲裁者建议（相关均衡）：纳什均衡是**稳定的、自我实现的协议**
- 纳什均衡需要精炼
- 均衡与物种演化

安全性：最大最小的概念

纳什均衡带来了稳定性，但有时候稳定性不是唯一的考虑因素，例如：

	L	R
T	2, 1	2, -20
M	3, 0	-10, 1
B	-100, 2	3, 3

显然，这个博弈唯一的均衡是 (B, R)，但行参与人选择 B 其实会很犹豫

- 如果列参与人选了 L 怎么办（不管是因为意外，还是不理性，或是其他原因）
- (B, L) 是灾难性的，因此行参与人可能更偏好 T，虽然只带来 2 的收益，但可以保证避免 -100 的灾难
- 如何保证自己避免最大的损失？选取最差情况最好的策略：最大最小策略

安全性：最大最小的概念（Cont'd）

定义 (最大最小策略)

给定一个博弈，参与人 i 的**最大最小策略（maxmin strategy）**是参与人 i 的一个策略 s_i^*，满足

$$\min_{t_{-i}\in S_{-i}} u_i(s_i^*, t_{-i}) \geqslant \min_{t_{-i}\in S_{-i}} u_i(s_i, t_{-i}), \forall s_i \in S_i$$

即 s_i^* 下的最差情况是最优的，即可以保证自己得到收益

$$\underline{v}_i = \max_{s_i\in S_i} \min_{t_{-i}\in S_{-i}} u_i(s_i, t_{-i})$$

$\underline{v}_i$ 称为参与人 i 的**最大最小值（maxmin value）**，于是最大最小策略等价于

$$u_i(s_i^*, t_{-i}) \geqslant \underline{v}_i, \forall t_{-i} \in S_{-i}$$

有时策略是无限集合，min 和 max 可能不存在，但是可以用下确界和上确界代替

安全性：最大最小的概念（Cont’d）

		Player II		
		L	R	$\min_{s_{\mathrm{II}} \in S_{\mathrm{II}}} u_{\mathrm{I}}(s_{\mathrm{I}}, s_{\mathrm{II}})$
	T	2, 1	2, −20	2
Player I	M	3, 0	−10, 1	−10
	B	−100, 2	3, 3	−100
$\min_{s_{\mathrm{I}} \in S_{\mathrm{I}}} u_{\mathrm{II}}(s_{\mathrm{I}}, s_{\mathrm{II}})$		0	−20	2, 0

三个解概念之间的关系（Cont’d）

定理

1. 参与人 i 占优于其它所有策略的策略是该参与人的最大最小策略，而且这样的一个策略是参与人 i 对其它参与人任何策略向量的最佳应对
2. 一个博弈中，如果每个参与人都有一个占优于其它所有策略的策略，那么占优策略向量是一个均衡点，也是最大最小策略向量
3. 一个博弈中，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 $s^* = (s_1^*, \cdots, s_n^*)$ 是博弈的唯一均衡点，也是唯一的最大最小策略向量
4. 对于每个参与人 i，策略型博弈的每个纳什均衡 s^* 满足 $u_i(s^*) \geq \underline{v}_i$

三个解概念之间的关系（Cont’d）

定理

1. 参与人 i 占优于其它所有策略的策略是该参与人的最大最小策略，而且这样的一个策略是参与人 i 对其它参与人任何策略向量的最佳应对
2. 一个博弈中，如果每个参与人都有一个占优于其它所有策略的策略，那么占优策略向量是一个均衡点，也是最大最小策略向量
3. 一个博弈中，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 $s^*=(s_1^*,\cdots,s_n^*)$ 是博弈的唯一均衡点，也是唯一的最大最小策略向量
4. 对于每个参与人 i，策略型博弈的每个纳什均衡 s^* 满足 $u_i(s^*)\geqslant \underline{v}_i$

以下证明最后一条，其余定理的证明留作习题．直接利用纳什均衡策略是最优反应即可证明：

$$u_i(s^*)=\max_{s_i\in S_i} u_i(s_i,s_{-i}^*)\geqslant \max_{s_i\in S_i}\min_{t_{-i}\in S_{-i}} u_i(s_i,t_{-i})=\underline{v}_i$$

引入混合策略

考虑如下的石头剪刀布博弈:

	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

- 显然，这个博弈没有纳什均衡
- 那么每个参与人的最优决策是什么呢？我们可以考虑一个参与人是永远出石头，那么另一个人只要观察到这一点，就可以永远出布，因此这样的策略是不合理的（又例如点球大战门将的扑救方向也不能一直保持一致）
- 所以我们可以猜想，最优的策略是随机的，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

混合策略

定义 (混合策略)

令 $G=(I,(S_i)_{i\in I},(u_i)_{i\in I})$ 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

$$\Sigma_i=\{\sigma_i:S_i\to[0,1]\mid\sum_{s_i\in S_i}\sigma_i(s_i)=1\}$$

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．

注意我们还有一个记号，对于每个参与人 i，令 $\Delta(S_i)$ 为 S_i 上的概率分布集合，即

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

因此我们有 $\Sigma_i=\Delta(S_i)$．通俗而言，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．注意，纯策略也可以看作是混合策略的特例，即只有一个策略的概率为 1，其余为 0．

混合策略（Cont'd）

定义 (博弈的混合扩展)

令 $G=(I,(S_i)_{i\in I},(u_i)_{i\in I})$ 为一个策略型博弈．G 的**混合扩展（mixed extension）**是一个博弈

$$\Gamma=(I,(\Sigma_i)_{i\in I},(U_i)_{i\in I})$$

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 $\sigma=(\sigma_1,\cdots,\sigma_n)\in\Sigma_1\times\cdots\times\Sigma_n$ 映射到一个实数

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

- 这里我们用的是冯诺伊曼-摩根斯坦恩博弈的收益函数，求解的是参与人 i 在混合策略向量 σ 下的期望收益
- 这里蕴含着一个假定，就是每个参与人的行动是相互独立的

混合策略纳什均衡

类似于纯策略纳什均衡，我们可以给出混合策略纳什均衡的定义．给定一个博弈的混合扩展 $\Gamma = (I, (\Sigma_i)_{i\in I}, (U_i)_{i\in I})$，一个混合策略向量 σ^* 是一个**混合策略纳什均衡**，如果对于每个参与人 i，有

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

混合策略纳什均衡

类似于纯策略纳什均衡，我们可以给出混合策略纳什均衡的定义．给定一个博弈的混合扩展 $\Gamma = (I, (\Sigma_i)_{i\in I}, (U_i)_{i\in I})$，一个混合策略向量 σ^* 是一个**混合策略纳什均衡**，如果对于每个参与人 i，有

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

定理 (混合策略纳什均衡等价条件)

令 $G = (I, (S_i)_{i\in I}, (u_i)_{i\in I})$ 为一个策略型博弈，Γ 为 G 的混合扩展．一个混合策略向量 σ^* 是 Γ 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

$$U_i(\sigma^*) \geqslant U_i(s_i, \sigma^*_{-i})$$

混合策略纳什均衡

类似于纯策略纳什均衡，我们可以给出混合策略纳什均衡的定义．给定一个博弈的混合扩展 $\Gamma = (I, (\Sigma_i)_{i\in I}, (U_i)_{i\in I})$，一个混合策略向量 σ^* 是一个**混合策略纳什均衡**，如果对于每个参与人 i，有

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

定理 (混合策略纳什均衡等价条件)

令 $G = (I, (S_i)_{i\in I}, (u_i)_{i\in I})$ 为一个策略型博弈，Γ 为 G 的混合扩展．一个混合策略向量 σ^* 是 Γ 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

$$U_i(\sigma^*) \geqslant U_i(s_i, \sigma^*_{-i})$$

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，

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

纳什定理

从前面的石头剪刀布的例子我们知道，并非所有博弈都有纳什均衡．但是纳什定理告诉我们，每个有限的策略型博弈都有至少一个纳什均衡．

定理 (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．

这一定理的证明超出本门课程范围（需要使用布劳威尔不动点定理、角谷不动点定理等），感兴趣的同学可以参考相关书目．

混合策略纳什均衡的求解

性别大战：

	F	C
F	2, 1	0, 0
C	0, 0	1, 2

最自然、最直观的方式还是从最优反应函数的角度出发，但是这里我们需要考虑的是混合策略．对于参与人 1（行参与人）的每个混合策略 $[x(F), (1-x)(C)]$，参与人 2（列参与人）的最优反应集合为

$$\begin{aligned} br_2(x) &= \arg\max_{y\in[0,1]} u_2(x,y) \\ &= \{y \in [0,1] \mid u_2(x,y) \geqslant u_2(x,z), \forall z \in [0,1]\} \end{aligned}$$

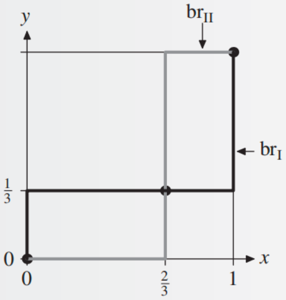
同理对于参与人 2 的每个混合策略 $[y(F), (1-y)(C)]$，参与人 1 的最优反应集合为

$$\begin{aligned} br_1(y) &= \arg\max_{x\in[0,1]} u_1(x,y) \\ &= \{x \in [0,1] \mid u_1(x,y) \geqslant u_1(z,y), \forall z \in [0,1]\} \end{aligned}$$

混合策略纳什均衡的求解（Cont’d）

不难解得

$$br_2(x) = \begin{cases} \{0\}, & x \in [0, 2/3) \\ [0,1], & x \in \{2/3\} \\ \{1\}, & x \in (2/3, 1] \end{cases} \qquad br_1(y) = \begin{cases} \{0\}, & y \in [0, 1/3) \\ [0,1], & y \in \{1/3\} \\ \{1\}, & y \in (/3, 1] \end{cases}$$



三个交点：$(x^*, y^*) = (0, 0)$，$(x^*, y^*) = (2/3, 1/3)$，$(x^*, y^*) = (1, 1)$，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，混合策略纳什均衡对应的收益为 $(2/3, 2/3)$.

混合策略纳什均衡的求解（Cont’d）

定理 (无差异原则)

令 σ^* 为一个混合策略纳什均衡，令 s_i 和 s_i' 为参与人 i 的两个纯策略，如果 $\sigma_i^*(s_i) > 0$，那么 $u_i(s_i, \sigma_{-i}^*) = u_i(s_i', \sigma_{-i}^*)$.

这个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.

混合策略纳什均衡的求解（Cont’d）

定理 (无差异原则)

令 σ^* 为一个混合策略纳什均衡，令 s_i 和 s_i' 为参与人 i 的两个纯策略，如果 $\sigma_i^*(s_i) > 0$，那么 $u_i(s_i, \sigma_{-i}^*) = u_i(s_i', \sigma_{-i}^*)$.

这个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.

关于混合策略纳什均衡的求解，有一种指数时间的通法，需要一定的线性规划基础，因此不在此介绍．我们想要知道的问题是，混合策略纳什均衡的计算看起来非常复杂，是否存在多项式时间的算法?

混合策略纳什均衡的求解（Cont'd）

定理 (无差异原则)

令 σ^* 为一个混合策略纳什均衡，令 s_i 和 s_i' 为参与人 i 的两个纯策略，如果 $\sigma_i^*(s_i) > 0$，那么 $u_i(s_i, \sigma_{-i}^*) = u_i(s_i', \sigma_{-i}^*)$.

这个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.

关于混合策略纳什均衡的求解，有一种指数时间的通法，需要一定的线性规划基础，因此不在此介绍. 我们想要知道的问题是，混合策略纳什均衡的计算看起来非常复杂，是否存在多项式时间的算法?

定理 (陈汐、邓小铁（2009）)

双人博弈纳什均衡的计算是 PPAD 完全问题.

1. 我们不在此解释 PPAD 完全的含义，只需知道目前是没有多项式时间算法可以计算一般的两人博弈的混合策略纳什均衡
2. 纯策略纳什均衡：PLS 完全性，局部搜索复杂度
3. 稳定匹配

二人零和博弈

定义 (零和博弈)

如果对于每个策略向量 (s_1, s_2)，有 $u_1(s_1, s_2) + u_2(s_1, s_2) = 0$，那么这个博弈是一个（双人）**零和博弈（zero-sum game）**

事实上，一个双人博弈（不一定是零和的）又叫作双矩阵博弈．一个双矩阵博弈可以描述为两个 $m \times n$ 的收益矩阵 A 和 B，其中一个是行参与者的，一个是列参与者的．在零和博弈中，$B = -A$，因此零和博弈只需要一个矩阵就可以表达．

二人零和博弈

定义 (零和博弈)

如果对于每个策略向量 (s_1, s_2)，有 $u_1(s_1, s_2) + u_2(s_1, s_2) = 0$，那么这个博弈是一个（双人）**零和博弈（zero-sum game）**

事实上，一个双人博弈（不一定是零和的）又叫作双矩阵博弈．一个双矩阵博弈可以描述为两个 $m \times n$ 的收益矩阵 A 和 B，其中一个是行参与者的，一个是列参与者的．在零和博弈中，$B = -A$，因此零和博弈只需要一个矩阵就可以表达．

石头剪刀布博弈的零和博弈表示：

	石头	剪刀	布
石头	0	1	-1
剪刀	-1	0	1
布	1	-1	0

- 表格中的每个元素是行参与者的收益，列参与者的收益是这个数的相反数

二人零和博弈（Cont’d）

设收益矩阵为 $A=(a_{ij})_{m\times n}$，当行参与者的混合策略为 $\boldsymbol{x}=(x_1,\cdots,x_m)$，列参与者的混合策略为 $\boldsymbol{y}=(y_1,\cdots,y_n)$ 时，行参与者的期望收益是（列参与人是相反数）

$$u_1(\boldsymbol{x},\boldsymbol{y})=\sum_{i=1}^{m}\sum_{j=1}^{n}a_{ij}x_iy_j=\boldsymbol{x}^{\mathrm{T}}A\boldsymbol{y}$$

二人零和博弈（Cont'd）

设收益矩阵为 $A=(a_{ij})_{m\times n}$，当行参与者的混合策略为 $\boldsymbol{x}=(x_1,\cdots,x_m)$，列参与者的混合策略为 $\boldsymbol{y}=(y_1,\cdots,y_n)$ 时，行参与者的期望收益是（列参与人是相反数）

$$u_1(\boldsymbol{x},\boldsymbol{y})=\sum_{i=1}^{m}\sum_{j=1}^{n}a_{ij}x_iy_j=\boldsymbol{x}^{\mathrm{T}}A\boldsymbol{y}$$

前面我们已经介绍了两种重要的解概念，即纳什均衡（代表稳定性）和最大最小值（代表安全性）．对于零和博弈，我们有一个非常重要的定理，它说明在零和博弈中，纳什均衡和最大最小值是等价的，因此可以同时实现稳定性和安全性．接下来的核心就是要推导出这一定理．

首先我们看参与人 2 的最大最小值和参与人 1 的最大最小值：

$$\underline{v}_1=\max_{\boldsymbol{x}}\min_{\boldsymbol{y}}\boldsymbol{x}^{\mathrm{T}}A\boldsymbol{y}$$

$$\underline{v}_2=\max_{\boldsymbol{y}}\min_{\boldsymbol{x}}\boldsymbol{x}^{\mathrm{T}}(-A)\boldsymbol{y}=-\min_{\boldsymbol{y}}\max_{\boldsymbol{x}}\boldsymbol{x}^{\mathrm{T}}A\boldsymbol{y}$$

二人零和博弈（Cont’d）

故我们定义二人零和博弈的

- 最大最小值为 $\underline{v} = \max_{\boldsymbol{x}} \min_{\boldsymbol{y}} \boldsymbol{x}^{\mathrm{T}} A \boldsymbol{y}$，取得最大最小值的 $\boldsymbol{x}$ 称为行参与者的最大最小策略
- 最小最大值为 $\bar{v} = \min_{\boldsymbol{y}} \max_{\boldsymbol{x}} \boldsymbol{x}^{\mathrm{T}} A \boldsymbol{y}$，取得最小最大值的 $\boldsymbol{y}$ 称为列参与者的最小最大策略

二人零和博弈（Cont’d）

故我们定义二人零和博弈的

- 最大最小值为 $\underline{v} = \max_{\boldsymbol{x}} \min_{\boldsymbol{y}} \boldsymbol{x}^{\mathrm{T}} A \boldsymbol{y}$，取得最大最小值的 $\boldsymbol{x}$ 称为行参与者的最大最小策略
- 最小最大值为 $\overline{v} = \min_{\boldsymbol{y}} \max_{\boldsymbol{x}} \boldsymbol{x}^{\mathrm{T}} A \boldsymbol{y}$，取得最小最大值的 $\boldsymbol{y}$ 称为列参与者的最小最大策略

定理

$\underline{v} \leqslant \overline{v}$.

二人零和博弈（Cont'd）

故我们定义二人零和博弈的

- 最大最小值为 $\underline{v} = \max_{\boldsymbol{x}} \min_{\boldsymbol{y}} \boldsymbol{x}^{\mathrm{T}} A \boldsymbol{y}$，取得最大最小值的 $\boldsymbol{x}$ 称为行参与者的最大最小策略
- 最小最大值为 $\overline{v} = \min_{\boldsymbol{y}} \max_{\boldsymbol{x}} \boldsymbol{x}^{\mathrm{T}} A \boldsymbol{y}$，取得最小最大值的 $\boldsymbol{y}$ 称为列参与者的最小最大策略

定理

$\underline{v} \leqslant \overline{v}$.

证明：令 $y^* = \arg\min_{\boldsymbol{y}} \max_{\boldsymbol{x}} \boldsymbol{x}^{\mathrm{T}} A \boldsymbol{y}$，则

$$\underline{v} = \max_{\boldsymbol{x}} \min_{\boldsymbol{y}} \boldsymbol{x}^{\mathrm{T}} A \boldsymbol{y} \leqslant \max_{\boldsymbol{x}} \boldsymbol{x}^{\mathrm{T}} A \boldsymbol{y}^* = \min_{\boldsymbol{y}} \max_{\boldsymbol{x}} \boldsymbol{x}^{\mathrm{T}} A \boldsymbol{y} = \overline{v}$$

二人零和博弈（Cont’d）

定理

对于一个二人零和博弈，$\underline{v} = \overline{v}$.

定理证明我们留到在线学习一讲，还有一种简单的方法需要用到线性规划对偶，这里不展开．我们称 $\underline{v} = \overline{v}$ 为二人零和博弈的**值**（**value**）.

定理

对于一个二人零和博弈，(x^*, y^*) 是一个纳什均衡当且仅当 $\boldsymbol{x}^*$ 是行参与者的最大最小策略，$\boldsymbol{y}^*$ 是列参与者的最小最大策略.

证明：从纳什均衡出发，我们有

$$
\begin{aligned}
(\boldsymbol{x}^*)^{\mathrm{T}} A \boldsymbol{y}^* &= \max_{\boldsymbol{x}} \boldsymbol{x}^{\mathrm{T}} A \boldsymbol{y}^* \geqslant \min_{\boldsymbol{y}} \max_{\boldsymbol{x}} \boldsymbol{x}^{\mathrm{T}} A \boldsymbol{y} \\
&= \max_{\boldsymbol{x}} \min_{\boldsymbol{y}} \boldsymbol{x}^{\mathrm{T}} A \boldsymbol{y} \geqslant \min_{\boldsymbol{y}} (\boldsymbol{x}^*)^{\mathrm{T}} A \boldsymbol{y} = (\boldsymbol{x}^*)^{\mathrm{T}} A \boldsymbol{y}^*
\end{aligned}
$$

故两个不等号必须取等号，即 $\boldsymbol{x}^*$ 是行参与者的最大最小策略，$\boldsymbol{y}^*$ 是列参与者的最小最大策略.

二人零和博弈（Cont'd）

定理

对于一个二人零和博弈，(x^*, y^*) 是一个纳什均衡当且仅当 $\boldsymbol{x}^*$ 是行参与者的最大最小策略，$\boldsymbol{y}^*$ 是列参与者的最小最大策略.

证明：从最大最小值出发，我们有

$$
\begin{aligned}
(\boldsymbol{x}^*)^{\mathrm{T}} A \boldsymbol{y}^* &\geqslant \min_{\boldsymbol{y}} (\boldsymbol{x}^*)^{\mathrm{T}} A \boldsymbol{y} = \max_{\boldsymbol{x}} \min_{\boldsymbol{y}} \boldsymbol{x}^{\mathrm{T}} A \boldsymbol{y} \\
&= \min_{\boldsymbol{y}} \max_{\boldsymbol{x}} \boldsymbol{x}^{\mathrm{T}} A \boldsymbol{y} = \max_{\boldsymbol{x}} \boldsymbol{x}^{\mathrm{T}} A \boldsymbol{y}^* \geqslant \boldsymbol{t}^{\mathrm{T}} A \boldsymbol{y}^*, \forall \boldsymbol{t}
\end{aligned}
$$

因此 $\boldsymbol{x}^*$ 是 $\boldsymbol{y}^*$ 的最优反应，同样的方法也可以证明 $\boldsymbol{y}^*$ 是 $\boldsymbol{x}^*$ 的最优反应，即 $(\boldsymbol{x}^*, \boldsymbol{y}^*)$ 是一个纳什均衡.

二人零和博弈（Cont'd）

定理

对于一个二人零和博弈，(x^*,y^*) 是一个纳什均衡当且仅当 $\boldsymbol{x}^*$ 是行参与者的最大最小策略，$\boldsymbol{y}^*$ 是列参与者的最小最大策略．

证明：从最大最小值出发，我们有

$$
\begin{aligned}
(\boldsymbol{x}^*)^{\mathrm{T}} A \boldsymbol{y}^* &\geqslant \min_{\boldsymbol{y}} (\boldsymbol{x}^*)^{\mathrm{T}} A \boldsymbol{y} = \max_{\boldsymbol{x}} \min_{\boldsymbol{y}} \boldsymbol{x}^{\mathrm{T}} A \boldsymbol{y} \\
&= \min_{\boldsymbol{y}} \max_{\boldsymbol{x}} \boldsymbol{x}^{\mathrm{T}} A \boldsymbol{y} = \max_{\boldsymbol{x}} \boldsymbol{x}^{\mathrm{T}} A \boldsymbol{y}^* \geqslant \boldsymbol{t}^{\mathrm{T}} A \boldsymbol{y}^*, \forall \boldsymbol{t}
\end{aligned}
$$

因此 $\boldsymbol{x}^*$ 是 $\boldsymbol{y}^*$ 的最优反应，同样的方法也可以证明 $\boldsymbol{y}^*$ 是 $\boldsymbol{x}^*$ 的最优反应，即 $(\boldsymbol{x}^*,\boldsymbol{y}^*)$ 是一个纳什均衡．这一定理说明，对于二人零和博弈，

纳什均衡和最大最小值是等价的，因此我们可以通过计算最大最小值来求解纳什均衡，并且零和博弈的解概念（即“值”）的兼具稳定性和安全性．

相关均衡引入：斗鸡博弈

在之前的讨论中，博弈参与人的行为都是独立作出的

- 但是，在有的博弈中（例如协调博弈），我们希望有一个仲裁人站出来，告诉参与人应该如何选择策略
- 在现实中的确存在这样的博弈，例如红绿灯

例 (斗鸡博弈（The game of Chicken）)

两个人开车相向而行，如果两个人都不避让，就会发生车祸．如果两个人都避让，那么就会浪费时间．如果一个人避让，另一个人不避让，那么避让的人会被认为是懦夫．这个博弈的收益矩阵如下：

	避让	不避让
避让	6, 6	2, 7
不避让	7, 2	0, 0

相关均衡引入：斗鸡博弈（Cont’d）

例 (斗鸡博弈（The game of Chicken）)

两个人开车相向而行，如果两个人都不避让，就会发生车祸．如果两个人都避让，那么就会浪费时间．如果一个人避让，另一个人不避让，那么避让的人会被认为是懦夫．这个博弈的收益矩阵如下：

	避让	不避让
避让	6, 6	2, 7
不避让	7, 2	0, 0

这个博弈有三个均衡：

1. （避让，不避让）和（不避让，避让），收益分别为 $(2,7)$ 和 $(7,2)$，这就是红绿灯推荐的均衡，没有人会主动偏离
2. 双方都以 2/3 概率选择避让，1/3 概率选择不避让，收益为 $(14/3, 14/3)$

相关均衡引入：斗鸡博弈（Cont'd）

考虑如下机制：一个仲裁人关于要采取的行动给每个参与者一个建议，但仲裁人不告诉每个参与人其他参与人得到的建议．仲裁人以相同的概率在三个行动向量中选择：

	避让	不避让
避让	1/3	1/3
不避让	1/3	0

尽管参与人不知道别人被推荐了什么策略，但推荐的概率分布是公开的．通过抽签的方法在三个行动向量中选择一个之后，仲裁人向行参与人推荐抽到的行动向量的第一个策略，向列参与人推荐第二个策略，例如抽到了（避让，不避让），那么行参与人被推荐避让，列参与人被推荐不避让．

相关均衡引入：斗鸡博弈（Cont’d）

考虑如下机制：一个仲裁人关于要采取的行动给每个参与者一个建议，但仲裁人不告诉每个参与人其他参与人得到的建议．仲裁人以相同的概率在三个行动向量中选择：

	避让	不避让
避让	1/3	1/3
不避让	1/3	0

尽管参与人不知道别人被推荐了什么策略，但推荐的概率分布是公开的．通过抽签的方法在三个行动向量中选择一个之后，仲裁人向行参与人推荐抽到的行动向量的第一个策略，向列参与人推荐第二个策略，例如抽到了（避让，不避让），那么行参与人被推荐避让，列参与人被推荐不避让．

需要注意的是，如果行参与人收到了避让的建议，那么他认为列参与人收到了避让建议的概率应当是条件概率（其它分析同理）

$$\frac{1/3}{1/3+1/3}=1/2$$

相关均衡引入：斗鸡博弈（Cont'd）

现在证明，两个参与人都无法从单方面偏离仲裁人的建议来获益．正如前面分析的，如果行参与人收到了避让的建议，那么他认为列参与人收到了避让和不避让的建议的概率都是 1/2

1. 如果行参与人遵循建议，那么他的收益是 $1/2 \times 6 + 1/2 \times 2 = 4$
2. 如果不遵循建议，那么他的收益是 $1/2 \times 7 + 1/2 \times 0 = 3.5$

因此，行参与人无法从单方面偏离仲裁人的建议来获益．同理，我们可以验证行列参与人在任何推荐下都无法从单方面偏离仲裁人的建议来获益．

相关均衡引入：斗鸡博弈（Cont’d）

现在证明，两个参与人都无法从单方面偏离仲裁人的建议来获益．正如前面分析的，如果行参与人收到了避让的建议，那么他认为列参与人收到了避让和不避让的建议的概率都是 1/2

1. 如果行参与人遵循建议，那么他的收益是 $1/2 \times 6 + 1/2 \times 2 = 4$
2. 如果不遵循建议，那么他的收益是 $1/2 \times 7 + 1/2 \times 0 = 3.5$

因此，行参与人无法从单方面偏离仲裁人的建议来获益．同理，我们可以验证行列参与人在任何推荐下都无法从单方面偏离仲裁人的建议来获益．

因此这一推荐产生了一个均衡（接下来我们会定义），这一均衡下的期望收益是

$$1/3(6,6) + 1/3(2,7) + 1/3(7,2) = (5,5)$$

这位于斗鸡博弈的三个纳什均衡收益的凸包之外（即不能被线性表示），二者的期望收益之和大于三个纳什均衡的期望收益之和，因此对于参与人而言，这一均衡是更好的．

相关均衡的定义

令 $G=(I,(S_i)_{i\in I},(u_i)_{i\in I})$ 为一个策略型博弈，对于集合 $S=S_1\times\cdots\times S_n$ 上的每个概率分布 p，定义博弈 $\Gamma^*(p)$ 如下：

- 仲裁人根据概率分布 p 在 S 上选择一个策略向量
- 对每个参与人 i，仲裁人向参与人 i 推荐 s_i，但不告知其他参与人的推荐 s_{-i}
- 参与人 i 选择一个策略 s_i'，并不一定是推荐的策略
- 参与人 i 的收益是 $u_i(s_1',\cdots,s_n')$

相关均衡的定义

令 $G = (I, (S_i)_{i\in I}, (u_i)_{i\in I})$ 为一个策略型博弈，对于集合 $S = S_1 \times \cdots \times S_n$ 上的每个概率分布 p，定义博弈 $\Gamma^*(p)$ 如下：

- 仲裁人根据概率分布 p 在 S 上选择一个策略向量
- 对每个参与人 i，仲裁人向参与人 i 推荐 s_i，但不告知其他参与人的推荐 s_{-i}
- 参与人 i 选择一个策略 s_i'，并不一定是推荐的策略
- 参与人 i 的收益是 $u_i(s_1', \cdots, s_n')$

当仲裁人推荐参与人 i 选择 s_i 时，参与人 i 知道其他参与人的推荐是 s_{-i} 的概率为

$$p_i(s_{-i} \mid s_i) = \frac{p(s_i, s_{-i})}{p(s_i)} = \frac{p(s_i, s_{-i})}{\sum_{s_{-i}\in S_{-i}} p(s_i, s_{-i})}$$

相关均衡的定义（Cont'd）

要使得每个参与人接受仲裁人的推荐是一个均衡（即没有人会独自偏离推荐的策略），显然要求为

$$\sum_{s_{-i}\in S_{-i}} p_i(s_{-i}\mid s_i)u_i(s_i,s_{-i}) \geqslant \sum_{s_{-i}\in S_{-i}} p_i(s_{-i}\mid s_i)u_i(s_i',s_{-i}), \forall s_i, s_i' \in S_i$$

可以消去 $p_i(s_{-i}\mid s_i)$ 的分母部分，改写为

$$\sum_{s_{-i}\in S_{-i}} p(s_i,s_{-i})u_i(s_i,s_{-i}) \geqslant \sum_{s_{-i}\in S_{-i}} p(s_i,s_{-i})u_i(s_i',s_{-i}), \forall s_i, s_i' \in S_i$$

相关均衡的定义（Cont'd）

要使得每个参与人接受仲裁人的推荐是一个均衡（即没有人会独自偏离推荐的策略），显然要求为

$$\sum_{s_{-i}\in S_{-i}} p_i(s_{-i} \mid s_i)u_i(s_i, s_{-i}) \geqslant \sum_{s_{-i}\in S_{-i}} p_i(s_{-i} \mid s_i)u_i(s_i', s_{-i}), \forall s_i, s_i' \in S_i$$

可以消去 $p_i(s_{-i} \mid s_i)$ 的分母部分，改写为

$$\sum_{s_{-i}\in S_{-i}} p(s_i, s_{-i})u_i(s_i, s_{-i}) \geqslant \sum_{s_{-i}\in S_{-i}} p(s_i, s_{-i})u_i(s_i', s_{-i}), \forall s_i, s_i' \in S_i$$

如果上式对任意的 i 都满足，我们称满足上述条件的概率分布 p 为博弈 Γ^* 的一个**相关均衡（correlated equilibrium）**.

- 注意：相关均衡是一个概率分布，实际上可以认为是仲裁人的一种推荐策略
- 在相关均衡下，参与人的策略选择不再是独立的，联合分布 p 被直接给出

相关均衡的性质

对于每个混合策略纳什均衡 $\sigma^* = (\sigma_1^*, \cdots, \sigma_n^*)$，我们可以导出一个策略向量集合上的概率分布 p_σ^*：

$$p_\sigma^*(s_1, \cdots, s_n) = \prod_{i=1}^{n} \sigma_i^*(s_i)$$

在纳什均衡 σ^* 下，给定其他参与人的策略向量 σ_{-i}^*，每个参与人以正的概率选择的行动都是给他们带来最大收益的行动（无差异原则），则

$$u_i(s_i, \sigma_{-i}^*) \geqslant u_i(s_i', \sigma_{-i}^*), \forall s_i \in \mathrm{supp}(\sigma_i^*), \forall s_i' \in S_i$$

相关均衡的性质

对于每个混合策略纳什均衡 $\sigma^* = (\sigma_1^*, \cdots, \sigma_n^*)$，我们可以导出一个策略向量集合上的概率分布 p_σ^*：

$$p_\sigma^*(s_1, \cdots, s_n) = \prod_{i=1}^{n} \sigma_i^*(s_i)$$

在纳什均衡 σ^* 下，给定其他参与人的策略向量 σ_{-i}^*，每个参与人以正的概率选择的行动都是给他们带来最大收益的行动（无差异原则），则

$$u_i(s_i, \sigma_{-i}^*) \geqslant u_i(s_i', \sigma_{-i}^*), \forall s_i \in \mathrm{supp}(\sigma_i^*), \forall s_i' \in S_i$$

将这一表达式展开即可得到

定理

对于每个混合策略纳什均衡 σ^*，概率分布 p_{σ^*} 是一个相关均衡.

1. 由此可见，相关均衡是混合策略纳什均衡的一个扩展
2. 直观：纳什均衡本身就产生了一个大家不会偏离的推荐策略

粗糙相关均衡

粗糙相关均衡（coarse correlated equilibrium）是相关均衡的一个扩展，此时参与人只知道自己被推荐策略的概率，而不知道自己被推荐的具体策略，并且不偏离推荐：

$$\sum_{s\in S} p(s)u_i(s) \geqslant \sum_{s\in S} p(s)u_i(s_i', s_{-i}), \forall s_i' \in S_i$$

粗糙相关均衡

粗糙相关均衡（coarse correlated equilibrium）是相关均衡的一个扩展，此时参与人只知道自己被推荐策略的概率，而不知道自己被推荐的具体策略，并且不偏离推荐：

$$\sum_{s\in S} p(s)u_i(s) \geqslant \sum_{s\in S} p(s)u_i(s_i', s_{-i}), \forall s_i' \in S_i$$

可通过相关均衡和粗糙相关均衡之间的关联加深理解：相关均衡为

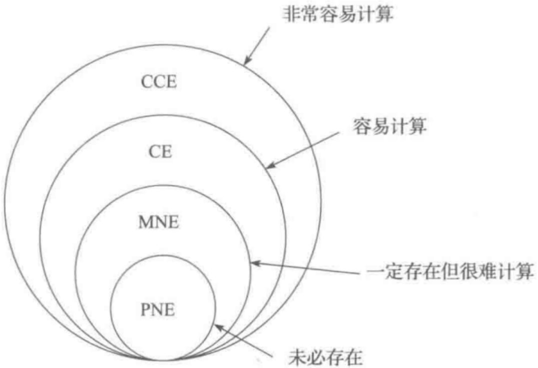
$$\sum_{s_{-i}\in S_{-i}} p(s_i, s_{-i})u_i(s_i, s_{-i}) \geqslant \sum_{s_{-i}\in S_{-i}} p(s_i, s_{-i})u_i(s_i', s_{-i}), \forall s_i, s_i' \in S_i$$

两边对 s_i 求和，得到

$$\sum_{s_i\in S_i}\sum_{s_{-i}\in S_{-i}} p(s_i, s_{-i})u_i(s_i, s_{-i}) \geqslant \sum_{s_i\in S_i}\sum_{s_{-i}\in S_{-i}} p(s_i, s_{-i})u_i(s_i', s_{-i}), \forall s_i' \in S_i$$

整理后就是相关均衡，并且可以理解什么叫“只知道自己被推荐策略的概率，不知道自己被推荐的具体策略”的含义.

均衡之间的关系



均衡概念的层级结构：纯策略纳什均衡(Pure Nash Equilibria，PNE)、混合策略纳什均衡(Mixed Nash Equilibria，MNE)、相关均衡(Correlated Equilibria，CE)和粗糙相关均衡(Coarse Correlated Equilibria，CCE)

综合的例子

考虑有四个智能体的单元自私路由网络，网络很简单，有一个公共起点 o，公共终点 d，边集 $E = \{0, 1, 2, 3, 4, 5\}$ 分别表示 6 条平行的 $o - d$ 边，每条边的代价函数为 $c(x) = x$.

综合的例子

考虑有四个智能体的单元自私路由网络，网络很简单，有一个公共起点 o，公共终点 d，边集 $E = \{0,1,2,3,4,5\}$ 分别表示 6 条平行的 $o-d$ 边，每条边的代价函数为 $c(x) = x$.

1. 纯策略纳什均衡：每个智能体选择不同的边，每人代价 1
2. 混合策略纳什均衡：每个智能体以相同的概率选择其中一条边，每人期望代价 3 / 2
3. 相关均衡：一条边上有两个智能体且另外两条边上分别有一个智能体的局势上的均匀分布，每人期望代价 3 / 2
4. 粗糙相关均衡：和相关均衡类似，但智能体选边结果只能是 $\{0,2,4\}$ 或 $\{1,3,5\}$，那么在两个边集上的均匀分布就是一个粗糙相关均衡，每人期望代价 3 / 2. 注意这不是相关均衡，因为智能体可以通过改变到旁边的边来降低代价到 1