# 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 11 日

# 从非合作博弈到合作博弈

在相关均衡中我们已经了解到，有些时候博弈的参与者可能需要通过仲裁人的协调来得到更好的结果．我们考虑将外部的仲裁人内部化，通过博弈参与者之间的谈判来达成协议．我们期望博弈的参与者能够通过**有效合作（cooperate efficiently）**通过谈判协调到一个具有更高福利的均衡．

# 从非合作博弈到合作博弈

在相关均衡中我们已经了解到，有些时候博弈的参与者可能需要通过仲裁人的协调来得到更好的结果．我们考虑将外部的仲裁人内部化，通过博弈参与者之间的谈判来达成协议．我们期望博弈的参与者能够通过**有效合作（cooperate efficiently）**通过谈判协调到一个具有更高福利的均衡．

**公平假设（equity hypothesis）**：博弈的参与者具有平等机会参与的有效谈判结果，应该与一个知道博弈的共同知识的公正仲裁人所做的推荐一致．

1. 如果仲裁人推荐的不是一个有效合作的结果，那么博弈的参与者应该能够通过谈判达成一个更好的结果，因此不会愿意接受推荐；
2. 如果公正仲裁人的推荐的好处显而易见，那么谈判也自然应当趋向于这一结果．

# 分摊美元博弈

**例 (分摊美元（Divide the Dollars）博弈)**

有两个参与者，每个人都可以在 0 到 100 美元之间提出他所需要的某个钱数．如果两个人提出的钱数之和不超过 100 美元，那么他们将按照他们提出的钱数来分配这 100 美元．如果他们提出的钱数之和超过 100 美元，那么他们都只能得到 0 美元．

# 分摊美元博弈

**例 (分摊美元（Divide the Dollars）博弈)**

有两个参与者，每个人都可以在 0 到 100 美元之间提出他所需要的某个钱数．如果两个人提出的钱数之和不超过 100 美元，那么他们将按照他们提出的钱数来分配这 100 美元．如果他们提出的钱数之和超过 100 美元，那么他们都只能得到 0 美元．

这一博弈的纳什均衡数量是无限的：任意的 $(x, 100-x)$ 都是纳什均衡，其中 $x \in [0, 100]$；$(100, 100)$ 也是纳什均衡，还有很多混合策略纳什均衡．问题是：这些纳什均衡中哪一个是最好的呢？

# 分摊美元博弈

**例 (分摊美元（Divide the Dollars）博弈)**

有两个参与者，每个人都可以在 0 到 100 美元之间提出他所需要的某个钱数．如果两个人提出的钱数之和不超过 100 美元，那么他们将按照他们提出的钱数来分配这 100 美元．如果他们提出的钱数之和超过 100 美元，那么他们都只能得到 0 美元．

这一博弈的纳什均衡数量是无限的：任意的 $(x, 100 - x)$ 都是纳什均衡，其中 $x \in [0, 100]$；$(100, 100)$ 也是纳什均衡，还有很多混合策略纳什均衡．问题是：这些纳什均衡中哪一个是最好的呢?

显然我们认为，此时最好能有一个公正的仲裁人来协调这一博弈，推荐策略 $(50, 50)$，即每个人都提出 50 美元．并且我们也相信，如果是两个具有平等能力的谈判人，他们也会达成这一结果，因为这一结果符合我们对**公平（equity）**和**效率（efficiency）**的直觉．

# 两人讨价还价博弈

## 定义 (两人讨价还价博弈)

一个两人讨价还价博弈是有序二元组 $(S,d)$，其中

- $S \subseteq \mathbb{R}^2$ 是一个非空紧致的凸集，称为备选方案集；
- $d=(d_1,d_2)\in S$ 称为**分歧点（disagreement point）**，即两个参与者的底线，例如博弈的最大最小值，或者保证金；
- 存在满足 $x>d$[a] 的备选方案 $x=(x_1,x_2)\in S$.

我们用 $\mathcal{F}$ 表示所有两人讨价还价博弈的集合.

[a] 即向量的每个元素都大于

# 两人讨价还价博弈

**定义 (两人讨价还价博弈)**

一个两人讨价还价博弈是有序二元组 $(S, d)$，其中

- $S \subseteq \mathbb{R}^2$ 是一个非空紧致的凸集，称为备选方案集；
- $d = (d_1, d_2) \in S$ 称为**分歧点（disagreement point）**，即两个参与者的底线，例如博弈的最大最小值，或者保证金；
- 存在满足 $x > d$[a] 的备选方案 $x = (x_1, x_2) \in S$.

我们用 $\mathcal{F}$ 表示所有两人讨价还价博弈的集合.

[a] 即向量的每个元素都大于

讨价还价的解概念 $\varphi$ 是一个映射 $\varphi : \mathcal{F} \to S$，对于任意的博弈 $(S, d) \in \mathcal{F}$，$\varphi(S, d)$ 是一个策略，称为**讨价还价解（bargaining solution）**．事实上我们可以将 $\varphi(S, d)$ 视为一个对于博弈 $(S, d)$ 的公正仲裁人的推荐策略.

# 两人讨价还价博弈

## 定义 (两人讨价还价博弈)

一个两人讨价还价博弈是有序二元组 $(S, d)$，其中

- $S \subseteq \mathbb{R}^2$ 是一个非空紧致的凸集，称为备选方案集；
- $d = (d_1, d_2) \in S$ 称为**分歧点（disagreement point）**，即两个参与者的底线，例如博弈的最大最小值，或者保证金；
- 存在满足 $x > d$[a] 的备选方案 $x = (x_1, x_2) \in S$.

我们用 $\mathcal{F}$ 表示所有两人讨价还价博弈的集合.

[a]即向量的每个元素都大于

讨价还价的解概念 $\varphi$ 是一个映射 $\varphi : \mathcal{F} \to S$，对于任意的博弈 $(S, d) \in \mathcal{F}$，$\varphi(S, d)$ 是一个策略，称为**讨价还价解（bargaining solution）**. 事实上我们可以将 $\varphi(S, d)$ 视为一个对于博弈 $(S, d)$ 的公正仲裁人的推荐策略.

我们希望讨价还价解能够满足一些基本性质，用于衡量为什么仲裁人会推荐，或者为什么谈判可以达成这一结果.

# 纳什讨价还价公理

## 纳什讨价还价解公理

1. 强有效性：对于任意的 $x \in S$，若 $x \geqslant \varphi(S,d)$，则 $\varphi(S,d) = x$；
2. 个人理性：$\varphi(S,d) \geqslant d$；
3. 协变性：对任意的 $\lambda_1, \lambda_2 > 0$，以及 $\gamma_1, \gamma_2$，若

$$G = \{(\lambda_1 x_1 + \gamma_1, \lambda_2 x_2 + \gamma_2) \mid (x_1, x_2) \in S\}$$
$$w = (\lambda_1 d_1 + \gamma_1, \lambda_2 d_2 + \gamma_2)$$

   则 $\varphi(G,w) = (\lambda_1 \varphi_1(S,d)_1 + \gamma_1, \lambda_2 \varphi_2(S,d)_2 + \gamma_2)$；
4. 无关备选方案独立性：对任一子集 $T \subseteq S$，若 $\varphi(S,d) \in T$，则 $\varphi(T,d) = \varphi(S,d)$；
5. 对称性：若 $d_1 = d_2$，且若 $(x_1,x_2) \in S$，则 $(x_2,x_1) \in S$，则 $\varphi(S,d)_1 = \varphi(S,d)_2$ .

# 纳什讨价还价解

1. 强有效性：表明纳什讨价还价解是帕累托最优的，因为不可能有两个同时更好的备选方案；
2. 个人理性：表明纳什讨价还价解是一个合理的解，因为它至少能够满足每个参与者的底线；
3. 协变性：如果讨价还价博弈经过了整体线性变换，那么纳什讨价还价解也应当经过相应的线性变换；
4. 无关备选方案独立性：剔除那些不会被选取的可行的选择对象，不影响讨价还价解；
5. 对称性：两个参与人的地位完全一致，则应当得到相同的结果.

# 纳什讨价还价解

1. 强有效性：表明纳什讨价还价解是帕累托最优的，因为不可能有两个同时更好的备选方案；
2. 个人理性：表明纳什讨价还价解是一个合理的解，因为它至少能够满足每个参与者的底线；
3. 协变性：如果讨价还价博弈经过了整体线性变换，那么纳什讨价还价解也应当经过相应的线性变换；
4. 无关备选方案独立性：剔除那些不会被选取的可行的选择对象，不影响讨价还价解；
5. 对称性：两个参与人的地位完全一致，则应当得到相同的结果.

**定理**

存在唯一的解函数 $\varphi$ 满足五条纳什讨价还价解公理，并且这一解可以最大化如下**纳什积（*Nash product*）**：

$$\varphi(S,d) \in \arg\max_{x\in S, x\geqslant d} (x_1 - d_1)(x_2 - d_2)$$

# 一些评注

事实上，纳什讨价还价公理受到了许多批评：

1. 对协变性的批评：**一个参与人拥有的金钱数量，一般而言会影响他对讨价还价获得更多金钱的态度**
2. 对无关备选方案独立性的批评：如果纳什讨价还价解反应了参与双方妥协的结果，那么**没有被选择的方案可能影响解概念**

当然这并不妨碍这是第一个重要的讨价还价解概念，引出了对合作博弈的探讨．当然纳什讨价还价解还可以通过平均主义、功利主义等角度来理解，感兴趣的同学可以自行查阅相关资料．

# 一些评注

事实上，纳什讨价还价公理受到了许多批评：

1. 对协变性的批评：**一个参与人拥有的金钱数量，一般而言会影响他对讨价还价获得更多金钱的态度**
2. 对无关备选方案独立性的批评：如果纳什讨价还价解反应了参与双方妥协的结果，那么**没有被选择的方案可能影响解概念**

当然这并不妨碍这是第一个重要的讨价还价解概念，引出了对合作博弈的探讨．当然纳什讨价还价解还可以通过平均主义、功利主义等角度来理解，感兴趣的同学可以自行查阅相关资料．

讨价还价有一个著名的推广，即**斯塔尔以及鲁宾斯坦研究的双方轮流报价直到有一个被接受的博弈**：这更类似于我们日常生活中的讨价还价．建模的关键还有一点是贴现率，即参与人讨价还价是有时间成本的．

# 讨价还价的局限

两人讨价还价博弈可以很轻松地推广到 $n$ 人的情形，最终的纳什解最大化的纳什积为

$$\varphi(S,d) \in \arg\max_{x\in S, x\geqslant d} \prod_{i=1}^{n} (x_i - d_i)$$

# 讨价还价的局限

两人讨价还价博弈可以很轻松地推广到 $n$ 人的情形，最终的纳什解最大化的纳什积为

$$\varphi(S,d) \in \arg\max_{x\in S, x\geqslant d} \prod_{i=1}^{n}(x_i - d_i)$$

然而，$n$ 人讨价还价解完全忽视了参与者的一些子集（**联盟（coalition）**）内部的合作关系．我们来看下面两个例子：

**例**

考虑一个三个人参与的博弈，每个人都提出一个**三个参与人的收益配置**，并且三个人的收益之和不能超过 300 美元．考虑以下两种情况：

1. 除非三个人提出相同的配置方案，此时他们将按照他们提出的配置方案来分配，否则他们都得到 0 美元；
2. 除非参与人 1 和 2 提出相同的配置方案，此时他们将按照他们提出的配置方案来分配，否则他们都得到 0 美元．

# 讨价还价的局限（Cont'd）

## 例

考虑一个三个人参与的博弈，每个人都提出一个**三个参与人的收益配置**，并且三个人的收益之和不能超过 300 美元．考虑以下两种情况：

1. 除非三个人提出相同的配置方案，此时他们将按照他们提出的配置方案来分配，否则他们都得到 0 美元；
2. 除非参与人 1 和 2 提出相同的配置方案，此时他们将按照他们提出的配置方案来分配，否则每人都得到 0 美元．

利用讨价还价解的条件可以得出讨价还价解为三个人都提出配置 $(100, 100, 100)$．这对于第一种博弈而言看起来很合理，但对于第二种博弈似乎并不合理

# 讨价还价的局限（Cont'd）

**例**

考虑一个三个人参与的博弈，每个人都提出一个**三个参与人的收益配置**，并且三个人的收益之和不能超过 300 美元．考虑以下两种情况：

1. 除非三个人提出相同的配置方案，此时他们将按照他们提出的配置方案来分配，否则他们都得到 0 美元；
2. 除非参与人 1 和 2 提出相同的配置方案，此时他们将按照他们提出的配置方案来分配，否则每人都得到 0 美元．

利用讨价还价解的条件可以得出讨价还价解为三个人都提出配置 $(100, 100, 100)$．这对于第一种博弈而言看起来很合理，但对于第二种博弈似乎并不合理

- 如果参与人 1 和 2 经过了充分的交流，他们应当同时提出 $(150, 150, 0)$，这对他们两个人而言是更好的结果；

# 讨价还价的局限（Cont'd）

## 例

考虑一个三个人参与的博弈，每个人都提出一个**三个参与人的收益配置**，并且三个人的收益之和不能超过 300 美元．考虑以下两种情况：

1. 除非三个人提出相同的配置方案，此时他们将按照他们提出的配置方案来分配，否则他们都得到 0 美元；
2. 除非参与人 1 和 2 提出相同的配置方案，此时他们将按照他们提出的配置方案来分配，否则每人都得到 0 美元．

利用讨价还价解的条件可以得出讨价还价解为三个人都提出配置 $(100, 100, 100)$．这对于第一种博弈而言看起来很合理，但对于第二种博弈似乎并不合理

- 如果参与人 1 和 2 经过了充分的交流，他们应当同时提出 $(150, 150, 0)$，这对他们两个人而言是更好的结果；
- **有效地谈判（negotiate effectively）**：如果联盟成员的策略中存在一个可行的变化能使他们全部受益，那么将同意实际作出这样一个变化，除非与另外某个同等有效的联盟的协议相矛盾

# 讨价还价的局限（Cont’d）

在第一个例子中，我们没有考虑比参与人全体更小的联盟，第二个例子则引入了 $1, 2$ 使得他们可以有更好的结果，这是 $n$ 人讨价还价解没有考虑的．下面我们看一个更复杂的例子：

**例**

考虑一个三个人参与的博弈，每个人都提出一个**三个参与人的收益配置**，并且三个人的收益之和不能超过 300 美元．如果有任意两个参与者提出相同的配置方案，就按照他们提出的配置方案来分配，否则全体得到 0 美元．

# 讨价还价的局限（Cont’d）

在第一个例子中，我们没有考虑比参与人全体更小的联盟，第二个例子则引入了 1, 2 使得他们可以有更好的结果，这是 $n$ 人讨价还价解没有考虑的．下面我们看一个更复杂的例子：

**例**

考虑一个三个人参与的博弈，每个人都提出一个**三个参与人的收益配置**，并且三个人的收益之和不能超过 300 美元．如果有任意两个参与者提出相同的配置方案，就按照他们提出的配置方案来分配，否则全体得到 0 美元．

显然 $(100, 100, 100)$ 是我们期待的结果．但如果参与人 1 和 2 进行有效谈判，他们应当提出 $(150, 150, 0)$，这时 3 可能和 2 妥协，提出 $(0, 225, 75)$ 以求得到更多的收益．然后 1 可能和 3 妥协，提出 $(113, 0, 187)$．

# 讨价还价的局限（Cont’d）

在第一个例子中，我们没有考虑比参与人全体更小的联盟，第二个例子则引入了 $1,2$ 使得他们可以有更好的结果，这是 $n$ 人讨价还价解没有考虑的．下面我们看一个更复杂的例子：

**例**

考虑一个三个人参与的博弈，每个人都提出一个**三个参与人的收益配置**，并且三个人的收益之和不能超过 300 美元．如果有任意两个参与者提出相同的配置方案，就按照他们提出的配置方案来分配，否则全体得到 0 美元．

显然 $(100, 100, 100)$ 是我们期待的结果．但如果参与人 1 和 2 进行有效谈判，他们应当提出 $(150, 150, 0)$，这时 3 可能和 2 妥协，提出 $(0, 225, 75)$ 以求得到更多的收益．然后 1 可能和 3 妥协，提出 $(113, 0, 187)$．

- 谈判的结果与参与人谈判的顺序和终止条件有关

# 讨价还价的局限（Cont'd）

在第一个例子中，我们没有考虑比参与人全体更小的联盟，第二个例子则引入了 $1, 2$ 使得他们可以有更好的结果，这是 $n$ 人讨价还价解没有考虑的．下面我们看一个更复杂的例子：

**例**

考虑一个三个人参与的博弈，每个人都提出一个**三个参与人的收益配置**，并且三个人的收益之和不能超过 300 美元．如果有任意两个参与者提出相同的配置方案，就按照他们提出的配置方案来分配，否则全体得到 0 美元．

显然 $(100, 100, 100)$ 是我们期待的结果．但如果参与人 1 和 2 进行有效谈判，他们应当提出 $(150, 150, 0)$，这时 3 可能和 2 妥协，提出 $(0, 225, 75)$ 以求得到更多的收益．然后 1 可能和 3 妥协，提出 $(113, 0, 187)$．

- 谈判的结果与参与人谈判的顺序和终止条件有关
- 我们需要一个仲裁人来分析博弈，推荐公平且有效率的解

# 合作博弈的定义

为了研究公平、有效率的解，我们需要形式化地定义具有联盟的博弈：

## 定义 (具有可转移效用的合作博弈)

一个具有**可转移效用**（**transferable utility**）的合作博弈（或称为 TU 博弈）是一个满足如下条件的二元组 $(N; v)$：

1. $N = \{1, 2, \cdots, n\}$ 是一个有限的参与者集合；一个 $N$ 的子集被称为一个**联盟**（**coalition**），全体联盟构成的集合记为 $2^N$；
2. $v : 2^N \to \mathbb{R}$ 称为该博弈的**特征函数**（**characteristic function**），对于任意的 $S \subseteq N$，$v(S)$ 表示联盟 $S$ 的价值（worth），且满足 $v(\varnothing) = 0$.

联盟的价值 $v(S)$ 是一个实数，这实际上蕴含着一个重要的假设：

- 衡量的是联盟 $S$ 的成员可以通过合作获得的可以用一些数值统一衡量的效用，例如金钱，因此可以将总的价值在不同人之间进行分配，于是这样的效用我们也称之为可转移的
- 但声誉、威望等很难被统一衡量，也很难说将一个人的名声转移给另一个人，因此这样的效用称为不可转移的，我们不在接下来的内容中考察，因此后文中的合作博弈均指代 TU 博弈

# 合作博弈的定义（Cont'd）

## 例

我们考虑一个场景，假设有三个人 A，B，C，他们的特点分别如下：

1. A 擅长于发明专利，依靠这一才能年收入达 17 万美元；
2. B 有机敏的商业嗅觉，能准确发掘潜在市场，创建商业咨询公司年收入可达 15 万美元；
3. C 擅长市场营销，开办专门的销售公司年收入可达 18 万美元.

**三个人的才能互补，于是他们考虑合作**：

1. B 可以为 A 提供市场资讯，将 A 的发明专利卖给市场上最有需求的人，这样他们合作每年收入可达 35 万美元；
2. C 利用他的才能销售 A 的发明专利，合作年收入可达 38 万美元；
3. 当然 B 和 C 也可以合作组建一个提供市场咨询和销售一体化的公司，这样每年合作收入可达 36 万美元；
4. A，B，C 合作，A 在 B 的建议下发明最符合市场需求的专利，然后由 C 进行销售，这样他们合作每年收入可达 56 万美元.

# 合作博弈的定义（Cont'd）

用定义形式化这一博弈，此时 $N=\{A,B,C\}$，全体联盟构成的集合

$$2^N=\{\varnothing,\{A\},\{B\},\{C\},\{A,B\},\{B,C\},\{A,C\},\{A,B,C\}\},$$

则该场景可以表示为一个合作博弈 $(N;v)$，其中特征函数 $v$ 定义如下：

$$v(S)=\begin{cases}0 & S=\varnothing\\ 17 & S=\{A\}\\ 15 & S=\{B\}\\ 18 & S=\{C\}\\ 35 & S=\{A,B\}\\ 36 & S=\{B,C\}\\ 38 & S=\{A,C\}\\ 56 & S=\{A,B,C\}\end{cases}$$

通过确定参与者以及特征函数构成的二元组，我们的确可以形式化地完整描述前面例子的所有信息.

# 超可加博弈与凸博弈

## 定义 (超可加博弈)

一个合作博弈 $(N; v)$ 是**超可加（supperadditive）**的，若对于任意的 $S, T \subseteq N$ 且 $S \cap T = \varnothing$，有

$$v(S \cup T) \geqslant v(S) + v(T).$$

- 对这一定义的理解就是，超可加性要求两个不相交的联盟组成更大的联盟时，其价值一定大于等于两个联盟之间不合作的总价值
- 前面的例子，如果 $A$，$B$ 不合作，他们的总收益只有 $17 + 15 = 32$，而如果他们合作，他们的总收益是 35
- 这就是超可加性对应的直观：合作会带来优势互补获得更大收益

# 超可加博弈与凸博弈（Cont’d）

超可加性的定义要求 $S$ 和 $T$ 是不相交的，但我们可以将这一要求放宽：

## 定义 (凸博弈)

一个合作博弈 $(N; v)$ 是**凸（convex）**的，若对于任意的 $S, T \subseteq N$，有

$$v(S \cup T) + v(S \cap T) \geqslant v(S) + v(T).$$

# 超可加博弈与凸博弈（Cont’d）

超可加性的定义要求 $S$ 和 $T$ 是不相交的，但我们可以将这一要求放宽：

**定义 (凸博弈)**

一个合作博弈 $(N; v)$ 是**凸（convex）**的，若对于任意的 $S, T \subseteq N$，有

$$v(S \cup T) + v(S \cap T) \geqslant v(S) + v(T).$$

- 凸博弈一定是超可加的，因为如果我们取的 $S$ 和 $T$ 恰好是不相交的（$S \cap T = \varnothing$），那么上式就变成了超可加博弈的定义

# 超可加博弈与凸博弈（Cont’d）

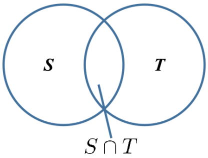
超可加性的定义要求 $S$ 和 $T$ 是不相交的，但我们可以将这一要求放宽：

## 定义 (凸博弈)

一个合作博弈 $(N; v)$ 是**凸（convex）**的，若对于任意的 $S, T \subseteq N$，有

$$v(S \cup T) + v(S \cap T) \geqslant v(S) + v(T).$$

- 凸博弈一定是超可加的，因为如果我们取的 $S$ 和 $T$ 恰好是不相交的（$S \cap T = \varnothing$），那么上式就变成了超可加博弈的定义
- 如果我们将凸博弈定义改写为 $v(S \cup T) \geqslant v(S) + v(T) - v(S \cap T)$，结合下图，根据简单的容斥原理可知，凸性是比超可加性更一般的描述 “合作会带来优势互补获得更大收益” 的性质



# 合作博弈的解概念

解概念会预测博弈的结果，而在合作博弈中，结果就是每个参与者可以从参与联盟得到多少收入，根据因此我们定义合作博弈的解概念如下：

## 定义 (合作博弈的解概念)

一个合作博弈 $(N; v)$（其中 $N = \{1, 2, \cdots, n\}$）的解概念是一个函数 $\varphi$，它将每个博弈 $(N; v)$ 与一个 $\mathbb{R}^n$ 的子集 $\varphi(N; v)$ 联系起来．如果对于任意的博弈 $\varphi(N; v)$ 都是一个单点集，则称这一解概念为**单点解（point solution）**.

# 合作博弈的解概念

解概念会预测博弈的结果，而在合作博弈中，结果就是每个参与者可以从参与联盟得到多少收入，根据因此我们定义合作博弈的解概念如下：

**定义 (合作博弈的解概念)**

一个合作博弈 $(N; v)$（其中 $N = \{1, 2, \cdots, n\}$）的解概念是一个函数 $\varphi$，它将每个博弈 $(N; v)$ 与一个 $\mathbb{R}^n$ 的子集 $\varphi(N; v)$ 联系起来．如果对于任意的博弈 $\varphi(N; v)$ 都是一个单点集，则称这一解概念为**单点解（point solution）**．

更通俗地说，合作博弈的解概念就是一个将每个博弈都映射到一个可行的收入分配集合的函数．这个集合中的每个元素是一个分配向量 $(x_1, x_2, \cdots, x_n)$，其中 $x_i$ 表示参与者 $i$ 在当前分配下可以获得的收入．

# 合作博弈的解概念（Cont’d）

**例**

设 $(N;v)$ 是一个三人参与的合作博弈，其中 $N=\{1,2,3\}$，特征函数 $v$ 定义如下：

$$v(\varnothing)=v(\{1\})=v(\{2\})=v(\{3\})=0,$$
$$v(\{1,2\})=v(\{1,3\})=v(\{2,3\})=v(\{1,2,3\})=1.$$

# 合作博弈的解概念（Cont'd）

## 例

设 $(N;v)$ 是一个三人参与的合作博弈，其中 $N=\{1,2,3\}$，特征函数 $v$ 定义如下：

$$v(\varnothing)=v(\{1\})=v(\{2\})=v(\{3\})=0,$$
$$v(\{1,2\})=v(\{1,3\})=v(\{2,3\})=v(\{1,2,3\})=1.$$

这是我们三人分 300 元的最后一个例子的形式化版本：

1. 如果你是参与者，你会认为两两组成联盟是最好的，结合对称性，你会认为这一博弈的解概念是 $\varphi(N;v)=\left\{\left(\frac{1}{2},\frac{1}{2},0\right),\left(\frac{1}{2},0,\frac{1}{2}\right),\left(0,\frac{1}{2},\frac{1}{2}\right)\right\}$，即你心中的解概念是一个三个元素的集合；
2. 如果你是局外人，你认为三个参与人是“对称的”，因此这一博弈的解概念应当是 $\varphi(N;v)=\left\{\left(\frac{1}{3},\frac{1}{3},\frac{1}{3}\right)\right\}$，只有一个向量组成.

# 核的出发点

我们定义最后的收入分配可以从不同的角度得到很多不同的分配方式，并且他们都是有一定道理的．所以在合作博弈解概念的讨论中，我们将通过引入一些基本原则限制函数 $\varphi$ 的性质，这些原则的不同将带来不同的解概念．

# 核的出发点

我们定义最后的收入分配可以从不同的角度得到很多不同的分配方式，并且他们都是有一定道理的．所以在合作博弈解概念的讨论中，我们将通过引入一些基本原则限制函数 $\varphi$ 的性质，这些原则的不同将带来不同的解概念．

在非合作博弈中，解概念与“均衡”有关，每个参与者在均衡时都没有偏离当前行为的动机．而在合作博弈中，我们可以把不偏离当前行为理解为“不离开联盟”，即**如果一个参与者自己单独离开了联盟，他的收入一定不会增多**．更进一步地，我们在合作博弈中也可能出现一个小的联盟一起离开大的联盟，那么相应的“不偏离”便是每个联盟都不能通过离开大联盟而获得更多的收入，这样可以保证我们在本章开始提到的联盟稳定性．

# 核的定义

## 定义 (核)

一个合作博弈 $(N;v)$（其中 $N=\{1,2,\cdots,n\}$）的**核（core）**是一个解概念 $\varphi$，其中 $\varphi(N;v)$ 由满足以下两个条件的分配向量 $x=(x_1,x_2,\cdots,x_n)$ 组成：

1. **有效率的（efficient）**：$\sum_{i=1}^{n} x_i = v(N)$，即所有参与者分完了整个联盟的全部收入；
2. **联盟理性（coalitionally rational**：对于任意的 $S\subseteq N$，有 $\sum_{i\in S} x_i \geqslant v(S)$，即对于任何联盟而言，他们在大联盟中分配到的收入一定不会比他们自己单独离开大联盟，组成小联盟获得的收入少.

# 核的定义

## 定义 (核)

一个合作博弈 $(N;v)$（其中 $N=\{1,2,\cdots,n\}$）的**核**（**core**）是一个解概念 $\varphi$，其中 $\varphi(N;v)$ 由满足以下两个条件的分配向量 $x=(x_1,x_2,\cdots,x_n)$ 组成:

1. **有效率的**（**efficient**）: $\sum\limits_{i=1}^{n} x_i = v(N)$，即所有参与者分完了整个联盟的全部收入;
2. **联盟理性**（**coalitionally rational**: 对于任意的 $S\subseteq N$，有 $\sum\limits_{i\in S} x_i \geqslant v(S)$，即对于任何联盟而言，他们在大联盟中分配到的收入一定不会比他们自己单独离开大联盟，组成小联盟获得的收入少.

1. 有效率是非常自然的要求，因为理性的参与者肯定希望收入多多益善，所以要分完所有的收入才合理;
2. 联盟理性则是联盟稳定性的要求，如果一个联盟中的参与者可以通过离开大联盟而获得更多的收入，那么他们就会离开，联盟就不稳定，这与我们定义核的初衷相悖.

# 核的计算

## 例

考虑三个参与人的合作博弈，$N=\{1,2,3\}$，特征函数 $v$ 定义如下：

$$v(\varnothing)=v(\{1\})=v(\{2\})=v(\{3\})=0,$$
$$v(\{1,2\})=v(\{2,3\})=1,\ \ v(\{1,3\})=2,\ \ v(\{1,2,3\})=3.$$

直接设出在核中的向量为 $x=(x_1,x_2,x_3)$，并根据核的定义，这一向量应当满足有效率的：

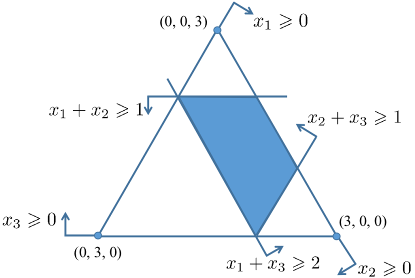
$$x_1+x_2+x_3=v(\{1,2,3\})=3.$$

以及联盟理性，即：

$$x_1+x_2\geqslant v(\{1,2\})=1, x_2+x_3\geqslant v(\{2,3\})=1, x_1+x_3\geqslant v(\{1,3\})=2,$$
$$x_1\geqslant v(\{1\})=0, x_2\geqslant v(\{2\})=0, x_3\geqslant v(\{3\})=0.$$

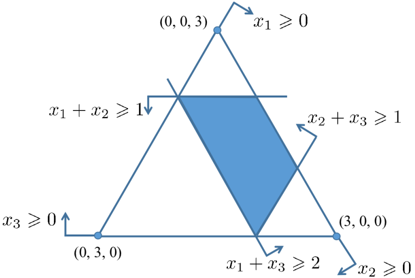
# 核的计算（Cont'd）

接下来的关键问题在于如何求出符合上述等式与不等式的所有向量．事实上，有效率性已经将 $x$ 限制到 $\mathbb{R}^3$ 中的一个平面，然后我们只需将这些不等式约束在这一平面上处理即可．



# 核的计算（Cont'd）

接下来的关键问题在于如何求出符合上述等式与不等式的所有向量．事实上，有效率性已经将 $x$ 限制到 $\mathbb{R}^3$ 中的一个平面，然后我们只需将这些不等式约束在这一平面上处理即可．




## 定理

一个合作博弈 $(N; v)$ 的核是一个凸集．

事实上，如果读者熟悉凸集的性质，我们会发现根据核的定义，核就是一些半空间的交集，因此核一定是凸集．

# 核可能是空集

## 例 (核为空集的博弈)

考虑三个参与人的合作博弈，$N = \{1, 2, 3\}$，特征函数 $v$ 定义如下：

$$v(\varnothing) = v(\{1\}) = v(\{2\}) = v(\{3\}) = 0,$$
$$v(\{1,2\}) = v(\{2,3\}) = v(\{1,3\}) = v(\{1,2,3\}) = 1.$$

设核中向量为 $x = (x_1, x_2, x_3)$，基于下面这一简单的观察发现核为空集：

1. 根据有效率的，我们有
$$x_1 + x_2 + x_3 = v(\{1,2,3\}) = 1. \tag{1}$$
2. 但根据联盟理性，由于 $v(\{1,2\}) = v(\{2,3\}) = v(\{1,3\}) = 1$，因此 $x_1 + x_2 \geqslant 1, x_2 + x_3 \geqslant 1, x_1 + x_3 \geqslant 1$，于是我们有
$$2x_1 + 2x_2 + 2x_3 \geqslant 3. \tag{2}$$

显然 1与 2是矛盾的，因此这一博弈的核为空集.

# 核的讨论

- 一个博弈的核可能为空集，也就是我们无法找到一个没有任何小联盟有动机离开大联盟的分配方式
- 除了可能为空集，前面的例子中求出的核也可能是一个很大的集合，得到核后我们具体应该挑选哪个分配向量呢？这也是核的一个缺陷
- 我们更加希望解概念是一个单点解，那么给定任何一个合作博弈，我们都可以直接计算出一种确定的分配方式，这便是我们下一节引入沙普利值的动机之一，当然我们要因此损失部分的稳定性
- **Bondareva-Shapley 定理**给出了一个核不是空集的充要条件．由于这一定理的陈述需要用到我们没有定义的**平衡（balanced）**等性质，需要较长篇幅来叙述，所以我们这里不给出具体的定理

# Shapley 值的引入

在讨论沙普利值前，我们再讨论一个经典的例子，它能告诉我们除了追求唯一解外，引入沙普利值的第二个很重要的动机：

## 例 (手套博弈)

我们仍然考虑三个参与人的合作博弈，其中 $N = \{1, 2, 3\}$，特征函数 $v$ 定义如下：

$$v(\varnothing) = v(\{1\}) = v(\{2\}) = v(\{3\}) = v(\{1,2\}) = 0,$$
$$v(\{1,3\}) = v(\{2,3\}) = v(\{1,2,3\}) = 1.$$

我们完全可以仿照前面例子的方法画出图示解决本题，这里我们不再重复过程，直接给出结论：核中只有一个向量 $(0, 0, 1)$，即只有当参与者 3 获得全部收入时，这一分配才是有效率的并且联盟理性的.

# Shapley 值的引入（Cont’d）

事实上，上述例子所示的博弈我们称之为**手套博弈**，因为它契合了下面这一故事情景：参与者 1 和 2 各有一只左手手套，而参与者 3 拥有一只右手手套，他们希望通过合作将手套配对，于是当他们单独行动或者只有 1 和 2 组成联盟时效用为 0，其余情况下它们能成功配对一对手套，因此效用为 1 .

# Shapley 值的引入（Cont’d）

事实上，上述例子所示的博弈我们称之为**手套博弈**，因为它契合了下面这一故事情景：参与者 1 和 2 各有一只左手手套，而参与者 3 拥有一只右手手套，他们希望通过合作将手套配对，于是当他们单独行动或者只有 1 和 2 组成联盟时效用为 0，其余情况下它们能成功配对一对手套，因此效用为 1 .

1. 很自然的问题：既然 1 和 2 也为获得 1 单位的效用值贡献了左手手套，那么为什么这一博弈最终的分配只能让 3 获得全部的收益呢

# Shapley 值的引入（Cont'd）

事实上，上述例子所示的博弈我们称之为**手套博弈**，因为它契合了下面这一故事情景：参与者 1 和 2 各有一只左手手套，而参与者 3 拥有一只右手手套，他们希望通过合作将手套配对，于是当他们单独行动或者只有 1 和 2 组成联盟时效用为 0，其余情况下它们能成功配对一对手套，因此效用为 1 .

1. 很自然的问题：既然 1 和 2 也为获得 1 单位的效用值贡献了左手手套，那么为什么这一博弈最终的分配只能让 3 获得全部的收益呢
2. 从核的定义，或者维持联盟的稳定性的角度来看这一结果确实是唯一结果，但它却和我们心中的 "公平性" 相悖．我们希望收入分配是公平的，那么提供了左手手套的 1 和 2 也应当获得一些收入，而不是全部由 3 获得

# Shapley 值的引入（Cont’d）

事实上，上述例子所示的博弈我们称之为**手套博弈**，因为它契合了下面这一故事情景：参与者 1 和 2 各有一只左手手套，而参与者 3 拥有一只右手手手套，他们希望通过合作将手套配对，于是当他们单独行动或者只有 1 和 2 组成联盟时效用为 0，其余情况下它们能成功配对一对手套，因此效用为 1 .

1. 很自然的问题：既然 1 和 2 也为获得 1 单位的效用值贡献了左手手套，那么为什么这一博弈最终的分配只能让 3 获得全部的收益呢
2. 从核的定义，或者维持联盟的稳定性的角度来看这一结果确实是唯一结果，但它却和我们心中的 “公平性” 相悖．我们希望收入分配是公平的，那么提供了左手手套的 1 和 2 也应当获得一些收入，而不是全部由 3 获得
3. 解决方法是通过改变解概念强制要求稳定性的要求，换成其它几条要求并强调公平性，同时，如果能保证对任何博弈的分配方式都是唯一的就更好了．因此，我们接下来将引入沙普利值这一解概念

# Shapley 性质

为了便于形式化地描述下面的性质，我们给出如下记号上的定义：

## 定义

令 $\varphi$ 为一个单点解，即对于任意的合作博弈 $(N; v)$（其中 $N = \{1, 2, \cdots, n\}$），$\varphi(N; v)$ 都是一个单点集，也就是唯一一个 $\mathbb{R}^n$ 中的向量．我们定义 $\varphi_i(N; v)$ 为向量 $\varphi(N; v)$ 中的第 $i$ 个位置的元素，即 $\varphi_i(N; v)$ 表示参与者 $i$ 在博弈 $(N; v)$ 中的分配到的收入．

在这一记号下，我们可以写出以下我们希望沙普利值拥有的性质：

## 定义 (有效率的)

一个解概念 $\varphi$ 是**有效率的（efficiency）**，若对于任意的合作博弈 $(N; v)$（其中 $N = \{1, 2, \cdots, n\}$），有 $\sum_{i=1}^{n} \varphi_i(N; v) = v(N)$．

即所有参与者分完了整个联盟的全部收入，事实上这与核的要求一致，这一条件是理性参与者一定会希望满足的．

# Shapley 性质（Cont’d）

第二个性质是对称性，这一性质是公平性的关键．为了定义对称性，我们需要首先定义两个参与人什么时候是对称的：

**定义 (参与人的对称性)**

令 $(N; v)$ 是一个合作博弈，对于 $i, j \in N$，如果对于任意的 $S \subseteq N \backslash \{i, j\}$，有 $v(S \cup \{i\}) = v(S \cup \{j\})$，则称参与者 $i$ 和 $j$ 是对称的．

- 即对于任何不包含他们的联盟，二者分别加入这一联盟后，联盟增加的效用（边际贡献）是一致的（因为对称性的定义等价于 $v(S \cup \{i\}) - v(S) = v(S \cup \{j\}) - v(S)$）；
- 例如在手套博弈中，1 和 2 是对称的

## Shapley 性质（Cont’d）

第二个性质是对称性，这一性质是公平性的关键．为了定义对称性，我们需要首先定义两个参与人什么时候是对称的：

**定义 (参与人的对称性)**

令 $(N;v)$ 是一个合作博弈，对于 $i,j\in N$，如果对于任意的 $S\subseteq N\backslash\{i,j\}$，有 $v(S\cup\{i\})=v(S\cup\{j\})$，则称参与者 $i$ 和 $j$ 是对称的．

- 即对于任何不包含他们的联盟，二者分别加入这一联盟后，联盟增加的效用（边际贡献）是一致的（因为对称性的定义等价于 $v(S\cup\{i\})-v(S)=v(S\cup\{j\})-v(S)$）；
- 例如在手套博弈中，1 和 2 是对称的

出于公平性的考虑，我们一定希望对称的参与者获得的收入是一致的，这就是对称性的定义：

**定义 (解概念的对称性)**

一个解概念 $\varphi$ 是满足**对称性（symmetry）**的，若对于任意的合作博弈 $(N;v)$，如果 $i,j\in N$ 是对称的，则 $\varphi_i(N;v)=\varphi_j(N;v)$．

# Shapley 性质（Cont'd）

## 定义 (零参与者)

一个解概念 $\varphi$ 是符合**零参与者（null player）**性质的，若对于任意的合作博弈 $(N;v)$，如果 $i \in N$，且对于任意的 $S \subseteq N \backslash \{i\}$，有

$$v(S \cup \{i\}) = v(S),$$

则 $\varphi_i(N;v) = 0$.

- 即如果一个参与人对于任何联盟的加入都不会增加联盟的效用，即他的边际贡献永远为 0，那么他获得的收入应当为 0；
- 这一定义是非常合乎情理的，同时引入这样的性质在数学上也是非常自然的，因为这样的条件属于 "规范性"[1] 的要求，是保证唯一解的重要手段.

[1] 如果读者熟悉行列式的公理化定义，我们要求单位矩阵的行列式为 1，这就是规保证我们能唯一确定出任意行列式值的规范性条件.

# Shapley 性质（Cont'd）

## 定义 (加性)

一个解概念 $\varphi$ 是满足**加性（additivity）**的，若对于任意一组合作博弈 $(N; v)$ 和 $(N; w)$ 有

$$\varphi(N; v + w) = \varphi(N; v) + \varphi(N; w),$$

其中 $(N; v + w)$ 表示对于任意的 $S \subseteq N$，有 $(v + w)(S) = v(S) + w(S)$，即每个联盟的收入是两个博弈的收入之和.

- 这一定义同样是保证唯一解的重要手段，它规定了如何组合一些简单的博弈来得到更复杂的博弈的解
- 这一性质非平凡，因为定义式等号左端加法是联盟特征函数的加法，右端是分配向量的加法，并不是同一种加法[2]
- 一组参与人 $N$，他们可以联合参加一个大项目，或分别参加两个小项目，并且每个联盟 $S$ 在分别参加两个小项目的情况下能获得的总收益与单独参加大项目的收益是一致的，那么每个参与人在大项目中分配的收入应等于在两个小项目中分配的收入之和

[2]在数学上这一性质称为同态性.

# Shapley 值的定义尝试

- 在给定以上四个性质后，我们自然希望构造出一个满足所有性质的沙普利值的计算表达式
- 可以从简单的想法入手，探索如何构建这一表达式．直觉告诉我们沙普利值应当与边际贡献有关，所以我们可以写下如下定义

# Shapley 值的定义尝试

- 在给定以上四个性质后，我们自然希望构造出一个满足所有性质的沙普利值的计算表达式
- 可以从简单的想法入手，探索如何构建这一表达式．直觉告诉我们沙普利值应当与边际贡献有关，所以我们可以写下如下定义

**例**

考虑合作博弈 $(N; v)$，其中 $N = \{1, 2, \cdots, n\}$，我们定义解概念 $\varphi$ 为

$$\varphi_i(N; v) = v(\{1, 2, \cdots, i\}) - v(\{1, 2, \cdots, i-1\}), \quad \forall i \in N.$$

# Shapley 值的定义尝试

- 在给定以上四个性质后，我们自然希望构造出一个满足所有性质的沙普利值的计算表达式
- 可以从简单的想法入手，探索如何构建这一表达式．直觉告诉我们沙普利值应当与边际贡献有关，所以我们可以写下如下定义

**例**

考虑合作博弈 $(N;v)$，其中 $N=\{1,2,\cdots,n\}$，我们定义解概念 $\varphi$ 为

$$\varphi_i(N;v)=v(\{1,2,\cdots,i\})-v(\{1,2,\cdots,i-1\}),\ \forall i\in N.$$

很容易证明这一解概念满足有效率性、零参与者和加性，但不满足对称性．例如考虑 $N=\{1,2,3\}$，特征函数 $v$ 定义如下：

$$v(\varnothing)=v(\{1\})=v(\{2\})=v(\{3\})=v(\{1,2\})=v(\{1,3\})=0,$$
$$v(\{2,3\})=v(\{1,2,3\})=1.$$

不难解得这一解概念下的分配向量为 $\{(0,0,1)\}$，但是我们很容易发现 2 和 3 在这一博弈中是对称的，但是 $\varphi_2(N;v)=0\neq 1=\varphi_3(N;v)$．

# Shapley 值的定义

1. 这一定义能满足除了对称性以外的所有性质，并且我们心里应当有一个强烈的声音告诉我们，这里的不对称性应当来源于我们对这 $n$ 个参与人的排序：不同的排序可能会带来不同的分配结果

# Shapley 值的定义

1. 这一定义能满足除了对称性以外的所有性质，并且我们心里应当有一个强烈的声音告诉我们，这里的不对称性应当来源于我们对这 $n$ 个参与人的排序：不同的排序可能会带来不同的分配结果
2. 那我们为什么不尝试把所有的排序对应的结果取平均呢?

我们记这 $n$ 个参与人的全体排序构成的集合为 $S_n$，根据简单的组合知识可知这一集合中含有 $|S_n| = n!$ 个元素．对于任意排序 $\sigma \in S_n$，定义

$$P_i(\sigma) := \{j \in N : \sigma(j) < \sigma(i)\},$$

即在排序 $\sigma$ 下位于参与人 $i$ 前面的所有参与人的集合．

- 例如若在排序 $\sigma$ 下参与人 $i$ 排在了第一位，那么 $P_i(\sigma) = \varnothing$

基于这一定义，我们便可以将"把所有的排序对应的结果取平均"的想法转化为表达式，就得到了每个参与人的沙普利值的概念．

# Shapley 值的定义（Cont'd）

## 定义 (沙普利值)

令 $(N;v)$ 是一个合作博弈，其中 $N=\{1,2,\cdots,n\}$，则参与人 $i$ 的沙普利值定义为

$$SV_i(N;v)=\frac{1}{n!}\sum_{\sigma\in S_n} v(P_i(\sigma)\cup\{i\})-v(P_i(\sigma)),$$

- 即我们将所有排序的可能性下，基前面失败的定义取了平均；
- 讨论某个具体的博弈时，或其他不引起歧义的情况下，我们也可将参与人 $i$ 的沙普利值记为 $SV_i$，这些 $SV_i$ 按序组成的向量 $SV=(SV_1,SV_2,\cdots,SV_n)$ 即为整个博弈的解概念，即整个博弈的沙普利值.

# Shapley 值的另一定义

前面的定义基于排列中的边际贡献加权平均，我们可以重新整理一下这一定义，得到基于对不同联盟的边际贡献的加权平均定义.

# Shapley 值的另一定义

前面的定义基于排列中的边际贡献加权平均，我们可以重新整理一下这一定义，得到基于对不同联盟的边际贡献的加权平均定义.

假设 $i \in N$ 是一个参与者，$S$ 是某个不包含 $i$ 的联盟，我们考虑有多少种排序 $\sigma$ 恰好可以使得排在 $i$ 的前面就是 $S$，即 $P_i(\sigma) = S$.

- 这时我们相当于将 $n$ 个参与者分成了三个部分，最前面是 $S$，中间是 $i$，最后是 $N\backslash(S \cup \{i\})$
- 这三个部分的参与者的相对顺序是确定的，而这三个部分内的参与者的相对顺序又是可以任意排列的
- 因此这一情况下的排序数目为 $|S|!(n - |S| - 1)!$

# Shapley 值的另一定义

前面的定义基于排列中的边际贡献加权平均，我们可以重新整理一下这一定义，得到基于对不同联盟的边际贡献的加权平均定义.

假设 $i \in N$ 是一个参与者，$S$ 是某个不包含 $i$ 的联盟，我们考虑有多少种排序 $\sigma$ 恰好可以使得排在 $i$ 的前面就是 $S$，即 $P_i(\sigma) = S$.

- 这时我们相当于将 $n$ 个参与者分成了三个部分，最前面是 $S$，中间是 $i$，最后是 $N \backslash (S \cup \{i\})$
- 这三个部分的参与者的相对顺序是确定的，而这三个部分内的参与者的相对顺序又是可以任意排列的
- 因此这一情况下的排序数目为 $|S|!(n - |S| - 1)!$

**定义 (沙普利值等价定义)**

合作博弈 $(N; v)$，$N = \{1, 2, \cdots, n\}$，则参与人 $i$ 的沙普利值定义为

$$SV_i(N; v) = \frac{1}{n!} \sum_{S \subseteq N \backslash \{i\}} |S|!(n - |S| - 1)!(v(S \cup \{i\}) - v(S)).$$

# Shapley 值的性质

**定理**

沙普利值满足有效率的、对称性、零参与者性质和加性的唯一解概念.

- 有效率、零参与者和加性可以通过失败的定义推广而来，凸组合不改变这些性质
- 对称性根据沙普利值等价定义证明是简单的，留作练习
- **载体公理：**若合作博弈 $(N, v)$ 中任一联盟 $R$ 是 $N$ 的一个载体，则 $\sum_{i \in R} SV_i(N; v) = v(R)$
  - 一个联盟 $R$ 是 $(N, v)$ 的一个**载体（carrier）**，若对于任意的 $S \subseteq N$，有 $v(S) = v(S \cap R)$
  - 即不在 $R$ 中的所有参与者其实是**虚设参与者（dummies）**，他们进入任何联盟都不会改变联盟价值
  - 实际上就是有效率和零参与者性质的结合，Shapley 的原版定义如此
- 唯一性的证明比较困难，但很精妙，接下来我们将展开介绍

# Shapley 值的性质（Cont'd）

**定理**

沙普利值满足有效率的、对称性、零参与者性质和加性的唯一解概念.

下面证明唯一性. 观察到 Shapley 值实际上是一个从全体合作博弈构成的集合到 $\mathbb{R}^n$ 的映射，而全体合作博弈是一个 $2^n-1$ 元向量构成的集合（其中 $n=|N|$，因为我们需要规定除了空集外所有 $N$ 的子集的价值）. 因此，我们实际上是要证明满足四条性质的映射

$$\varphi : \mathbb{R}^{2^n-1} \to \mathbb{R}^n$$

是唯一的. 注意出发和到达的空间都是线性空间，并且性质要求这一映射是线性的，所以我们考虑取线性空间的一组基，确定在这组基下的映射值，从而保证这一映射是唯一的.

# Shapley 值的性质（Cont'd）

**定义 (载体博弈)**

令 $T \subseteq N$ 是一个非空联盟，$T$ 上的**载体博弈（carrier game）**是如下合作博弈 $(N, u_T)$，其中对每个联盟 $S \subseteq N$，有

$$u_T(S) = \begin{cases} 1, & \text{若 } T \subseteq S, \\ 0, & \text{其他情况}. \end{cases}$$

# Shapley 值的性质（Cont'd）

## 定义 (载体博弈)

令 $T \subseteq N$ 是一个非空联盟，$T$ 上的**载体博弈**（**carrier game**）是如下合作博弈 $(N, u_T)$，其中对每个联盟 $S \subseteq N$，有

$$u_T(S) = \begin{cases} 1, & \text{若 } T \subseteq S, \\ 0, & \text{其他情况}. \end{cases}$$

我们要证明，所有 $2^n - 1$ 种载体博弈恰好是全体合作博弈的一组基，实际上因为这组向量的个数等于线性空间的维数，因此我们只需要证明线性无关性即可.

# Shapley 值的性质（Cont'd）

## 定义 (载体博弈)

令 $T \subseteq N$ 是一个非空联盟，$T$ 上的**载体博弈**（**carrier game**）是如下合作博弈 $(N, u_T)$，其中对每个联盟 $S \subseteq N$，有

$$u_T(S) = \begin{cases} 1, & 若\ T \subseteq S, \\ 0, & 其他情况. \end{cases}$$

我们要证明，所有 $2^n - 1$ 种载体博弈恰好是全体合作博弈的一组基，实际上因为这组向量的个数等于线性空间的维数，因此我们只需要证明线性无关性即可.

反证法，假设线性相关，则存在一组不全为零的实数 $(\alpha_T)_{T\subseteq N, T\neq\varnothing}$ 使得

$$\sum_{T\subseteq N, T\neq\varnothing} \alpha_T u_T(S) = 0, \forall S \subseteq N.$$

# Shapley 值的性质（Cont'd）

令 $\mathcal{T} = \{T \subseteq N : T \neq \varnothing, \alpha_T \neq 0\}$ 是上页式中非零系数的集联盟集合，由线性相关的假设可知 $\mathcal{T}$ 非空．我们考虑 $\mathcal{T}$ 中的最小联盟 $S_0$，即 $S_0$ 不严格包含任何 $\mathcal{T}$ 中的联盟，则有

$$\sum_{T\subseteq N, T\neq\varnothing} \alpha_T u_T(S_0) = \sum_{T\subset S_0, T\neq\varnothing} \alpha_T u_T(S_0) + \alpha_{S_0} u_{S_0}(S_0) + \sum_{T\not\subseteq S_0} \alpha_T u_T(S_0).$$

$\alpha_T = 0$ 对所有 $T \subset S_0$ 成立，而对每个 $T \not\subseteq S_0$，$u_T(S_0) = 0$，因此上式右端等于 $\alpha_{S_0} u_{S_0}(S_0) = \alpha_{S_0} \neq 0$，和线性相关表达式矛盾，因此线性无关性得证．

# Shapley 值的性质（Cont’d）

令 $\mathcal{T} = \{T \subseteq N : T \neq \varnothing, \alpha_T \neq 0\}$ 是上页式中非零系数的集联盟集合，由线性相关的假设可知 $\mathcal{T}$ 非空．我们考虑 $\mathcal{T}$ 中的最小联盟 $S_0$，即 $S_0$ 不严格包含任何 $\mathcal{T}$ 中的联盟，则有

$$\sum_{T \subseteq N, T \neq \varnothing} \alpha_T u_T(S_0) = \sum_{T \subset S_0, T \neq \varnothing} \alpha_T u_T(S_0) + \alpha_{S_0} u_{S_0}(S_0) + \sum_{T \not\subseteq S_0} \alpha_T u_T(S_0).$$

$\alpha_T = 0$ 对所有 $T \subset S_0$ 成立，而对每个 $T \not\subseteq S_0$，$u_T(S_0) = 0$，因此上式右端等于 $\alpha_{S_0} u_{S_0}(S_0) = \alpha_{S_0} \neq 0$，和线性相关表达式矛盾，因此线性无关性得证．

接下来我们考虑博弈 $(N, u_{T;\alpha})$ 如下：

$$u_{T,\alpha}(S) = \begin{cases} \alpha, & \text{若 } T \subseteq S, \\ 0, & \text{其他情况}. \end{cases}$$

考虑满足有效率、对称性和零参与者性质的解概念在这一博弈下的分配向量，然后我们就可以利用加性得到唯一性．

# Shapley 值的性质（Cont'd）

显然结果是

$$\varphi_i(N; u_{T,\alpha}) = \begin{cases} \dfrac{\alpha}{|T|}, & \text{若 } i \in T, \\ 0, & \text{若 } i \notin T. \end{cases}$$

因为所有的 $i \notin T$ 都是零参与者，而其它 $T$ 中的参与者都是对称的，再结合有效率性可以得到上面的结果.

# Shapley 值的性质（Cont'd）

显然结果是

$$\varphi_i(N; u_{T,\alpha}) = \begin{cases} \dfrac{\alpha}{|T|}, & 若\ i \in T, \\ 0, & 若\ i \notin T. \end{cases}$$

因为所有的 $i \notin T$ 都是零参与者，而其它 $T$ 中的参与者都是对称的，再结合有效率性可以得到上面的结果.

接下来我们终结证明. 根据前面证明的线性无关性，我们知道存在实数 $(\alpha_T)_{T \subseteq N, T \neq \varnothing}$ 使得

$$v(S) = \sum_{T \subseteq N, T \neq \varnothing} \alpha_T u_T(S) = \sum_{T \subseteq N, T \neq \varnothing} u_{T,\alpha_T}(S)$$

利用可加性的要求，我们可以得到

$$\varphi_i(N; v) = \sum_{T \subseteq N, T \neq \varnothing} \varphi_i(N; u_{T,\alpha_T}),$$

这是唯一确定的值，因此我们证明了唯一性.

# Shapley 值计算例题

接下来我们计算手套博弈以沙普利值为解概念得到的分配向量，一方面熟悉沙普利值的计算，另一方面也体现沙普利值蕴含的公平性：

## 例 (手套博弈的沙普利值)

回顾手套博弈中定义的三个参与人的合作博弈，其中 $N = \{1, 2, 3\}$，特征函数 $v$ 定义如下：

$$v(\varnothing) = v(\{1\}) = v(\{2\}) = v(\{3\}) = v(\{1, 2\}) = 0,$$
$$v(\{1, 3\}) = v(\{2, 3\}) = v(\{1, 2, 3\}) = 1.$$

计算这一博弈的沙普利值，并与手套博弈的核进行比较.

# Shapley 值计算例题（Cont'd）

我们使用沙普利值的第一个定义来计算沙普利值，当然读者也可以使用等价定义．我们需要对所有的排序计算每个参与人的边际贡献，然后取平均，我们得到针对参与人 1 和 3 计算结果如下表所示：

| 排序 | 参与人 1 的边际贡献 | 参与人 3 的边际贡献 |
|---|---|---|
| (1,2,3) | $v(\{1\}) - v(\varnothing) = 0$ | $v(\{1,2,3\}) - v(\{1,2\}) = 1$ |
| (1,3,2) | $v(\{1\}) - v(\varnothing) = 0$ | $v(\{1,3\}) - v(\{1\}) = 1$ |
| (2,1,3) | $v(\{1,2\}) - v(\{2\}) = 0$ | $v(\{1,2,3\}) - v(\{1,2\}) = 1$ |
| (2,3,1) | $v(\{1,2,3\}) - v(\{2,3\}) = 0$ | $v(\{2,3\}) - v(\{2\}) = 1$ |
| (3,1,2) | $v(\{1,3\}) - v(\{3\}) = 1$ | $v(\{3\}) - v(\varnothing) = 0$ |
| (3,2,1) | $v(\{1,2,3\}) - v(\{2,3\}) = 0$ | $v(\{3\}) - v(\varnothing) = 0$ |

没有计算参与人 2 是因为 1 和 2 是对称的，因此他们的沙普利值是一致的．

# Shapley 值计算例题（Cont'd）

| 排序 | 参与人 1 的边际贡献 | 参与人 3 的边际贡献 |
|---|---|---|
| (1,2,3) | $v(\{1\}) - v(\varnothing) = 0$ | $v(\{1,2,3\}) - v(\{1,2\}) = 1$ |
| (1,3,2) | $v(\{1\}) - v(\varnothing) = 0$ | $v(\{1,3\}) - v(\{1\}) = 1$ |
| (2,1,3) | $v(\{1,2\}) - v(\{2\}) = 0$ | $v(\{1,2,3\}) - v(\{1,2\}) = 1$ |
| (2,3,1) | $v(\{1,2,3\}) - v(\{2,3\}) = 0$ | $v(\{2,3\}) - v(\{2\}) = 1$ |
| (3,1,2) | $v(\{1,3\}) - v(\{3\}) = 1$ | $v(\{3\}) - v(\varnothing) = 0$ |
| (3,2,1) | $v(\{1,2,3\}) - v(\{2,3\}) = 0$ | $v(\{3\}) - v(\varnothing) = 0$ |

因此我们可以得到

$$SV_1 = SV_2 = \frac{1}{3!}(0+0+0+0+1+0) = \frac{1}{6},$$
$$SV_3 = \frac{1}{3!}(1+1+1+1+0+0) = \frac{2}{3}.$$

故整个博弈的沙普利值为 $SV = (\frac{1}{6}, \frac{1}{6}, \frac{2}{3})$，与核 $(0,0,1)$ 不同，参与人 1 和 2 分配的收入不再是 0，即考虑到了他们也为了配对手套做出的贡献，符合我们对公平性的追求，或者说对“平均主义”的考量．

# 核与沙普利值的关系

尽管手套博弈给了倾向于公平分配的人一个满意的结果，但我们不能否认，因为沙普利值不在核中，三个人共同组成的联盟不稳定

- 例如此时参与人 1 和 3 可以结盟背叛参与人 2，他们商定离开联盟后参与人 2 的那部分收入平分
- 很多博弈中我们无法兼顾公平性和稳定性，但在凸博弈中，我们可以做到这一点，即沙普利值一定在核中

# 核与沙普利值的关系

尽管手套博弈给了倾向于公平分配的人一个满意的结果，但我们不能否认，因为沙普利值不在核中，三个人共同组成的联盟不稳定

- 例如此时参与人 1 和 3 可以结盟背叛参与人 2，他们商定离开联盟后参与人 2 的那部分收入平分
- 很多博弈中我们无法兼顾公平性和稳定性，但在凸博弈中，我们可以做到这一点，即沙普利值一定在核中

**定理**

对于任意的凸博弈 $(N; v)$，沙普利值就在核中.

证明过于技术性，在此省略，大致思想是分两步进行：

1. 第一步我们通过构造出一个符合核的要求的向量证明凸博弈的核一定非空；
2. 第二步我们证明沙普利值能由满足要求的一些向量进行凸组合得到，从而根据核是凸集得到沙普利值就在核中.

# Shapley 值的缺陷

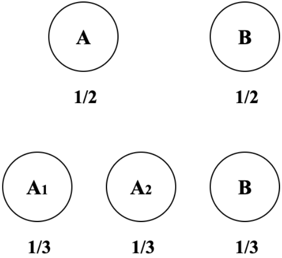
1. 计算困难性：需要计算一个人对 $2^n$ 个联盟的边际贡献，因此需要近似算法；
2. Shapley 值的定义对所有相同大小的联盟都赋予相同的权重，这种对称的处理方式可能不符合实际情况，可以引入**合作结构（cooperation structure）**和**欧文值（Owen value）**等概念来解决这一问题；
3. **Shapley-Shubik 权力指数**：用 Shapley 值衡量人在投票中的重要程度，并以此为依据分团体
   - 例如联合国安理会通过决议，需要五个常任理事国和多少个非常任理事国同意才可以？
   - 对于不同大小的联盟赋予的权重是否合理？
   - 引入**班扎夫值（Banzhaf value）**，详细的讨论见作业．
4. 在数据定价收益分配中的缺陷：对数据的复制、合谋、分裂等敏感
   - 机器学习模型训练等场景下，可能使用了多个数据提供者的数据，Shapley 值可以用来衡量每个数据提供者对模型的贡献，但是如果存在数据的复制、合谋、分裂等行为，Shapley 值可能会带来一些问题

# Shapley 值的缺陷（Cont’d）

在数据定价收益分配中的缺陷：对数据的复制、合谋、分裂等敏感

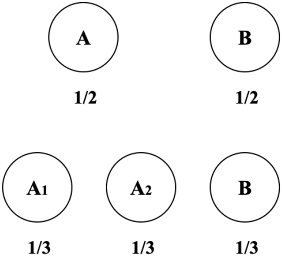
# Shapley 值的缺陷（Cont’d）

在数据定价收益分配中的缺陷：对数据的复制、合谋、分裂等敏感



# Shapley 值的缺陷（Cont’d）

在数据定价收益分配中的缺陷：对数据的复制、合谋、分裂等敏感




1. 数据的复制可以通过查重手段避免
2. 数据的合谋和分裂无法通过技术手段进行检测防御
   - 数据的合谋，即不同的数据提供者将数据合并起来出售以寻求更高的利润（超额收益）
   - 数据的分裂，即一个数据提供者故意将同一组数据拆分为多个部分出售以寻求更高的利润

# Banzhaf 值

## 定义 (2-efficiency)

$\varphi_i(N,v)+\varphi_j(N,v)=\varphi_{p_{ij}}(N',v')$，其中 $\varphi_{p_{ij}}(N',v')$ 代表 $i$ 和 $j$ 合并成一个参与人 $p_{ij}$ 后新的博弈 $(N',v')$ 的解，其中 $N'=N\backslash\{i,j\}\cup\{p_{ij}\}$.

实际上这就是两人合谋不改变分配值的条件（思考）

1. 对称的，一份数据分裂成两份也无法获得超额收益
2. 将 Shapley 值的有效率性替换为这一条后得到了新的四个性质，这四个性质可以保证 Banzhaf 值的唯一性

# Banzhaf 值

## 定义 (2-efficiency)

$\varphi_i(N,v) + \varphi_j(N,v) = \varphi_{p_{ij}}(N',v')$，其中 $\varphi_{p_{ij}}(N',v')$ 代表 $i$ 和 $j$ 合并成一个参与人 $p_{ij}$ 后新的博弈 $(N',v')$ 的解，其中 $N' = N\backslash\{i,j\} \cup \{p_{ij}\}$.

实际上这就是两人合谋不改变分配值的条件（思考）

1. 对称的，一份数据分裂成两份也无法获得超额收益
2. 将 Shapley 值的有效率性替换为这一条后得到了新的四个性质，这四个性质可以保证 Banzhaf 值的唯一性

## 定义 (Banzhaf 值)

$$B_i(N;v) = \sum_{S \subseteq N\backslash\{i\}} \frac{1}{2^{|N|}} \left(v(S \cup \{i\}) - v(S)\right)$$

和 Shapley 值的区别在于权重的不同，Shapley 值的权重是 $|S|!(n-|S|-1)!/n!$，而 Banzhaf 值的权重是 $2^{|N|}$，即无论联盟大小如何，每个联盟的权重都是一样的

# Semi-value

更一般的，我们有 Semi-value 的概念，它是 Shapley 值和 Banzhaf 值的推广，是考虑 Shapley 性质除有效率性外的其他三个性质的解概念：

$$\varphi_i(N,v) = \sum_{S \subseteq N \setminus \{i\}} p_s^n [v(S \cup \{i\}) - v(S)],$$

其中 $p_s^n$ 是一个只与 $n$ 和 $s$ 相关的权重，满足

$$p_s^n = \int_0^1 t^s (1-t)^{n-s-1} \mathrm{d}\xi(t).$$

其中 $\xi(t)$ 是一个概率分布函数.

# 其它解概念

为了描述其它解概念，我们首先引入一些记号和定义．对于 $(N, v)$ 的任意两个分配结果 $x$ 和 $y$，我们引入记号和定义：

1. $x >_S y$：当且仅当对于任意的 $i \in S$，有 $x_i > y_i$，即 $S$ 中的所有成员都喜欢 $x$ 严格胜于 $y$；
2. $x \geqslant_S y$：当且仅当对于任意的 $i \in S$，有 $x_i \geqslant y_i$；
3. 对任一联盟 $S$ 和任一分配 $x$，我们称 $e(S, x) = v(S) - \sum_{i \in S} x_i$ 为联盟 $S$ 在 $x$ 下的**余额（excess）**．
   - 注意：$e(S, x) \geqslant 0$ 的充要条件是联盟 $S$ 能独自达成分配 $x$．

# 稳定集

**稳定集（stable sets）**是第一个被研究的合作博弈解概念，由冯 · 诺伊曼和摩根斯坦在 1944 年提出. 任何一个分配向量构成的集合 $Z$ 构成一个稳定集需要满足以下两个条件:

1. 若分配 $y \notin Z$，则存在一个分配 $x \in Z$ 和联盟 $S$，使得 $e(S,x) \geqslant 0$ 且 $x >_S y$;
2. 对任意的两个分配 $x, y \in Z$ 以及任意的联盟 $S$，若 $x >_S y$，则 $e(S,x) < 0$.

# 稳定集

**稳定集（stable sets）** 是第一个被研究的合作博弈解概念，由冯 · 诺伊曼和摩根斯坦在 1944 年提出. 任何一个分配向量构成的集合 $Z$ 构成一个稳定集需要满足以下两个条件:

1. 若分配 $y \notin Z$，则存在一个分配 $x \in Z$ 和联盟 $S$，使得 $e(S,x) \geqslant 0$ 且 $x >_S y$;
2. 对任意的两个分配 $x, y \in Z$ 以及任意的联盟 $S$，若 $x >_S y$，则 $e(S,x) < 0$.

第一个条件的含义是：如果有一个分配 $y$ 不在稳定集中，那么一定是因为某个联盟 $S$ 否决了 $y$，因为它有更好的选择 $x$. 第二个条件的含义是：如果联盟 $S$ 对稳定集中有两个分配有严格的喜好差异，那一定是因为好的那个分配自己无法实现.

- 稳定集可能是空集，在核为空集时稳定集也可能不是空集

# 讨价还价集

**讨价还价集（bargaining set）** 是由 Aumann 和 Maschler 在 1964 年提出的解概念．讨价还价集背后的思想是：如果一个参与者害怕另一个参与者对他的异议提出一个反异议，那么他就不会提出异议．

1. 参与人 $i$ 对另一个参与人 $j$ 和一个分配 $x$ 的**异议（objection）**是有序对 $(y, S)$，其中 $y$ 是一个分配，$S$ 是一个联盟，满足

$$i \in S, j \notin S, e(S, y) = 0, y >_S x.$$

即 $i$ 所在的（$j$ 不在的）联盟 $S$ 能达成严格优于 $x$ 的分配 $y$;

2. 在 $i$ 对 $j$ 的异议 $(y, S)$ 提出后，$j$ 对 $i$ 的一个**反异议（counter objection）**是一个有序对 $(z, T)$，其中 $z$ 是一个分配，$T$ 是一个联盟，满足

$$j \in T, i \notin T, T \cap S \neq \varnothing, e(T, z) = 0, z \geqslant_T x, z \geqslant_{T\cap S} y.$$

此时 $j$ 可以形成一个联盟 $T$，能使得 $T$ 和 $S$ 重合的部分达成一个比异议好的分配 $z$，并且能拉拢 $T\backslash S$ 中所有人恢复到至少与 $x$ 相等的分配（因为异议 $y$ 可能使 $T\backslash S$ 的利益受损）．

# 核仁

考虑联盟的余额 $e(S,x)=v(S)-\sum_{i\in S}x_i$，事实上 $e(S,x)$ 衡量了 $S$ 中成员对于 $x$ 的不满意程度

1. 余额是正的，说明 $S$ 能够独立达成一个比 $x$ 更好的分配；
2. 余额越小，不满意程度越低，反之越高.

# 核仁

考虑联盟的余额 $e(S,x)=v(S)-\sum_{i\in S}x_i$，事实上 $e(S,x)$ 衡量了 $S$ 中成员对于 $x$ 的不满意程度

1. 余额是正的，说明 $S$ 能够独立达成一个比 $x$ 更好的分配；
2. 余额越小，不满意程度越低，反之越高.

给定分配 $x$，计算所有联盟在 $x$ 的余额，然后从左到右按递减顺序排列：

$$\theta(x)=(e(S_1,x),e(S_2,x),\cdots,e(S_{2^n},x)).$$

简写为 $\theta(x)=(\theta_1(x),\theta_2(x),\cdots,\theta_{2^n}(x))$.

# 核仁

考虑联盟的余额 $e(S,x)=v(S)-\sum_{i\in S}x_i$，事实上 $e(S,x)$ 衡量了 $S$ 中成员对于 $x$ 的不满意程度

1. 余额是正的，说明 $S$ 能够独立达成一个比 $x$ 更好的分配;
2. 余额越小，不满意程度越低，反之越高.

给定分配 $x$，计算所有联盟在 $x$ 的余额，然后从左到右按递减顺序排列:

$$\theta(x)=(e(S_1,x),e(S_2,x),\cdots,e(S_{2^n},x)).$$

简写为 $\theta(x)=(\theta_1(x),\theta_2(x),\cdots,\theta_{2^n}(x))$.

核仁是配置集的子集，因此我们要首先定义配置集. 合作博弈的配置集 $X(N,v)=\{x\in\mathbb{R}^n:\sum_{i\in N}x_i=v(N),x_i\geqslant v(i),\forall i\in N\}$.

- 显然，配置集就是全体满足有效率和个人理性的分配向量的集合

# 核仁（Cont’d）

## 定义 (核仁)

合作博弈 $(N, v)$ 的**核仁（nucleolus）**是配置集 $X(N, v)$ 中字典序极小的分配向量，即

$$\mathcal{N}(N, v) = \{x \in X(N, v) : \theta(x) \preceq_L \theta(y), \forall y \in X(N, v)\}.$$

字典序 $\preceq$ 的含义是：对于两个向量 $a = (a_1, a_2, \cdots, a_n)$ 和 $b = (b_1, b_2, \cdots, b_n)$，要么 $a = b$，要么存在一个 $k$ 使得 $a_i = b_i, \forall i < k$ 且 $a_k < b_k$，则 $a \preceq_L b$.

# 核仁（Cont'd）

## 定义 (核仁)

合作博弈 $(N, v)$ 的**核仁（nucleolus）**是配置集 $X(N, v)$ 中字典序极小的分配向量，即

$$\mathcal{N}(N, v) = \{x \in X(N, v) : \theta(x) \preceq_L \theta(y), \forall y \in X(N, v)\}.$$

字典序 $\preceq$ 的含义是：对于两个向量 $a = (a_1, a_2, \cdots, a_n)$ 和 $b = (b_1, b_2, \cdots, b_n)$，要么 $a = b$，要么存在一个 $k$ 使得 $a_i = b_i, \forall i < k$ 且 $a_k < b_k$，则 $a \preceq_L b$.

## 定理

核仁 $\mathcal{N}(N, v)$ 包含且仅包含唯一一个分配向量.

# 合作博弈解概念的分类

就一个简单的二人博弈而言，其中

$$N = \{1, 2\}, v(\varnothing) = 0, v(\{1\}) = v(\{2\}) = 0, v(\{1, 2\}) = 1.$$

总结而言，前面定义的所有合作博弈解概念可以大致分为两个类别：

1. **无从反对（unobjectionable）**解概念：只从防止联盟异议的准则来导出的解概念，那么这个博弈的解应该是所有满足要求的分配组成的集合 $\{(\alpha, 1-\alpha) \mid 0 \leqslant \alpha \leqslant 1\}$.
   - 核、讨价还价集、稳定集
   - 包括了参与人能接受的全部分配，而没有形成联盟来要求重新配置
2. **公平（equitable）**解概念：解概念还基于对局中人之间公平的某种考虑（同样有效率），那么这个博弈的解应该是单点配置 $(1/2, 1/2)$.
   - 沙普利值、核仁
   - 有仲裁人的指导方针，根据公平和效率决定焦点均衡