Basic Theory of Transferable Utility (TU) Games: Shapley value

I. Overview

- Previously:
  - Nucleolus: set of imputations that minimizes the excess vector in terms of the lexicographic ordering
  - The nucleolus always exists (is nonempty) and is always a singleton.
  - It is defined based on an equity principle that aims to make the excess evenly distributed among all coalitions.
- The solution concept considered here, the Shapley value, is based on a different notion of fairness based on marginal contributions.

II. Shapley Value

- Recall the following concepts, introduced in the section on convex games in the lecture notes on TU games and the core.
- Let $S \subset N$ and $i \notin S$. The **marginal contribution** of i towards $S \subseteq N$ is defined by
$$v(S \cup \{i\}) - v(S).$$
- The Shapley value is a weighted average of the marginal contribution over the coalitions that i can join.
- Define the **marginal contribution of player i with respect to the permutation π** by the following.
$$v(S^{\pi,i} \cup \{i\}) - v(S^{\pi,i})$$
where $S^{\pi,i}$, the set of players that precede i, is defined by
$$S^{\pi,i} = \{j \in N | \pi(j) < \pi(i)\}$$
- The Shapley value is then the average over all the permutations π.

Definition. Let (N, v) be a game. The Shapley value, denoted by $\phi(N, v) = (\phi_i(N, v))_{i \in N}$ is a vector in $\mathcal{R}^n$ given by the following formula.

$$\phi_i(N, v) = \frac{1}{n!} \sum_{\pi \in \Pi} \left(v(S^{\pi,i} \cup \{i\}) - v(S^{\pi,i}) \right) \tag{1}$$

where Π denotes the set of all permutations of N. When there is no confusion, $\phi(N, v)$ is also notated as $\phi(v)$.

- The Shapley value assigns to each player $i \in N$ its expected marginal contribution, where the probability of each ordering is equally likely ($\Rightarrow$ probability of each ordering $= 1/n!$).
- The following equivalent formula is also useful.

Alternative formula:

$$\phi_i(N, v) = \sum_{S \subseteq N \setminus \{i\}} \frac{|S|!(|N| - |S| - 1)!}{|N|!} \left(v(S \cup \{i\}) - v(S) \right) \tag{2}$$

where $|S|$ denotes the number of elements (players) in S. Also, equivalently,

$$\phi_i(N, v) = \sum_{S \subseteq N, S \neq \emptyset} \frac{(|S| - 1)!(|N| - |S|)!}{|N|!} \left(v(S) - v(S \setminus \{i\}) \right) \tag{3}$$

- Generally, $\phi(N, v)$ need not be an imputation, but $\phi(N, v)$ satisfies group rationality. If (N, v) is superadditive, then $\phi(N, v)$ is an imputation.
- Moreover, if (N, v) is convex, then $\phi(N, v) \in \mathcal{C}(N, v)$. This follows from the fact that the imputation $a^\pi \in \mathcal{R}^n$ where for each $i \in N$,

  $$a_i^\pi = v(S^{\pi,i} \cup \{i\}) - v(S^{\pi,i}),$$

  $a^\pi \in \mathcal{C}(N, v)$. Since $\phi(N, v)$ is a convex combination of a^π and the core $\mathcal{C}(N, v)$ is a convex set, $\phi(N, v) \in \mathcal{C}(N, v)$.

III. Example

$$v(S) = \begin{cases} 1 & \text{if } 1 \in S \text{ and } |S| \geq 2 \\ 0 & \text{otherwise} \end{cases}$$

1. Using equation (1): Construct the following table

	Player 1	Player 2	Player 3
$1 \leftarrow 2 \leftarrow 3$			
$1 \leftarrow 3 \leftarrow 2$			
$2 \leftarrow 1 \leftarrow 3$			
$2 \leftarrow 3 \leftarrow 1$			
$3 \leftarrow 1 \leftarrow 2$			
$3 \leftarrow 2 \leftarrow 1$			
Average			

   where, for example, $2 \leftarrow 3 \leftarrow 1$ denotes that $\{2\}$ is the first coalition, followed by 3 joining $\{2\}$ to form $\{2, 3\}$, and then ending in $\{1, 2, 3\}$.

   Now, for each row, enter the marginal contributions of each player with respect to the permutation.

   Finally, calculate the average for each column. These entries give the Shapley value.

2. Using equation (2): (explained using the white board)

IV. Axiomatic Characterizations of the Shapley Value

- Fix the set of players N. Let $\mathcal{V}$ be set of functions $v : 2^N \rightarrow \mathcal{R}$ such that (N, v) is a TU game.

- A single-valued solution is a function $\varphi : \mathcal{V} \rightarrow \mathcal{R}^n$.

- The four axioms used in the characterization of the Shapley value are the following.

Pareto Optimality (PO): For each $v \in \mathcal{V}$, $\sum_{i \in N} \varphi_i(v) = v(N)$.

- A player $i \in N$ is said to be a **null player** in the game (N, v) if for all $S \subseteq N \setminus \{i\}$, $v(S \cup \{i\}) = v(S)$.

Null Player Property (N): For each $v \in \mathcal{V}$ and for all null players i in the game (N, v), $\varphi_i(v) = 0$.

- Two players i and j are said to be **symmetric** (or **interchangeable**) in the game (N, v) if for all $S \subseteq N \setminus \{i, j\}$, $v(S \cup \{i\}) = v(S \cup \{j\})$.

Equal Treatment Property (ETP): For each $v \in \mathcal{V}$ and for any pair of players i, j that are interchangeable in the game (N, v), $\varphi_i(v) = \varphi_j(v)$.

- For two games (N, v) and (N, w), define the sum $(N, v + w)$ or simply $v + w$ by

$$(v + w)(S) = v(S) + w(S) \ \forall S \subseteq N$$

- While the previous three properties all were defined within one game, the following property involves values at two games.

Additivity (ADD): For each $v, w \in \mathcal{V}$,

$$\varphi(v + w) = \varphi(v) + \varphi(w)$$

- The Shapley value is characterized as the unique single-valued solution that satisfies the four axioms.

Theorem 1. The Shapley value is the unique single-valued solution on $\mathcal{V}$ that satisfies (PO), (N), (ETP), and (ADD).

- ϕ satisfying (PO), (N), and (ADD) can be checked easily. Checking (ETP) is rather tricky but straightforward. It remains to be shown that for any solution ψ satisfying (PO), (N), (ETP), and (ADD), $\psi = \phi$.

IV. Unanimity Games

- A class of games that plays a key role in the proof: T-unanimity games as defined below.

- A T-unanimity game is a game (N, v_T) such that

$$v_T(S) = \begin{cases} 1 & \text{if } T \subseteq S \\ 0 & \text{otherwise} \end{cases}$$

and define the α-multiple of the game by $(N, \alpha v_T)$ where

$$\alpha v_T(S) = \begin{cases} \alpha & \text{if } T \subseteq S \\ 0 & \text{otherwise} \end{cases}$$

- If ψ is a solution that satisfies (PO), (ETP), (N), and (ADD), by (PO), (ETP), (N),

$$\psi_i(\alpha v_T) = \begin{cases} \frac{\alpha}{|T|} & \text{if } i \in T \\ 0 & \text{otherwise} \end{cases}$$

- Also, since the Shapley value ϕ also satisfies the same axioms, for these games $\phi(\alpha v_T) = \psi(\alpha v_T)$.

V. The Space of Games and Sketch of the Proof

- First, the set of all n-player TU games $\mathcal{V}$ can be viewed as a subspace of $\mathcal{R}^{2^n-1}$.
  - Subspace – $v, w \in \mathcal{V}$ and $\alpha, \beta \in \mathcal{R} \Rightarrow \alpha v + \beta w \in \mathcal{V}$
- Consider the collection of unanimity games $(v_T)_{T \subseteq N, T \neq \emptyset}$. It can be shown that for each $v \in \mathcal{V}$, there exists a unique collection of real numbers $(\alpha_T)_{T \subseteq N, T \neq \emptyset}$ such that

$$v = \sum_{T \subseteq N, T \neq \emptyset} \alpha_T v_T. \tag{4}$$

In other words, $(v_T)_{T \subseteq N, T \neq \emptyset}$ is a **basis** of $\mathcal{V}$.

- To prove that $(v_T)_{T \subseteq N, T \neq \emptyset}$ is a basis, it is sufficient to show that $(v_T)_{T \subseteq N, T \neq \emptyset}$ is linearly independent.
  - A collection of vectors $(x^1, x^2, \cdots, x^k)$ is said to be **linearly independent** if

$$[\alpha_1 x^1 + \alpha_2 x^2 + \cdots \alpha_k x^k = 0] \Rightarrow \alpha_1 = \alpha_2 = \cdots = \alpha_k = 0$$

- Once it is shown that $(v_T)_{T\subseteq N, T\neq\emptyset}$ is a basis, consider $v \in \mathcal{V}$ and coefficients $(\alpha_T)_{T\subseteq N, T\neq\emptyset}$ such that (4) holds. By additivity

$$\psi(v) = \sum_{T\subseteq N, T\neq\emptyset} \psi(\alpha_T v_T) = \sum_{T\subseteq N, T\neq\emptyset} \phi(\alpha_T v_T) = \phi(v).$$

- Thus, the only solution that satisfies (PO), (ETP), (N), and (ADD) on the set of all games $\mathcal{V}$ is the Shapley value ϕ.

VI. Restricting the Class of Games

- Theorem 1 shows that the Shapley value is the unique value on $\mathcal{V}$ satisfying (PO), (ETP), (N), and (ADD).
- When looking at a subclass of games (e.g. superadditive), uniqueness may not hold at first sight.
- However, uniqueness is preserved under a subset $\mathcal{K}$ that is a convex cone that contains all T-unanimity games.
  - A subset of games $\mathcal{K}$ is a convex cone if for any $v, w \in \mathcal{K}$ and $\alpha, \beta \in \mathcal{R}$ with $\alpha \geq 0$ and $\beta \geq 0$:

$$\alpha v + \beta w \in \mathcal{K}$$

- The proof needs to be adjusted, because additivity can only be used for games in $\mathcal{K}$. However, games $\alpha_T v_T$ with $\alpha_T < 0$ may not be in $\mathcal{K}$ even if $v_T \in \mathcal{K}$.
- From the above fact, a parallel result for the class of superadditive games also holds.

VI. Shapley Value and Potential

- Let $\Gamma = \{(N, v) | (N, v) \text{ is a game.}\}$. Let $P : \Gamma \to \mathcal{R}$ be a function. Then, the **marginal contribution** of i with respect to (N, v) according to P, denoted by $D^i P(N, v)$, is defined by the following expression.

$$D^i P(N, v) = \begin{cases} P(N, v) & \text{if } |N| = 1, \\ P(N, v) - P(N \setminus \{i\}, v) & \text{if } |N| \geq 2. \end{cases} \tag{5}$$

where in $(N \setminus \{i\}, v)$, v is restricted to those coalitions $S \subseteq N \setminus \{i\}$, and by convention, $P(\emptyset, v) = 0$.

Definition. A function $P : \Gamma \to \mathcal{R}$ is a **potential function** if

$$\sum_{i \in N} D^i P(N, v) = v(N) \tag{6}$$

- An important property is that there exists at most one potential function.

Lemma. There exists a unique potential function P. Moreover, this P satisfies the following equation.

$$P(N, v) = \frac{v(N) + \sum_{i \in N} P(N \setminus \{i\}, v)}{|N|}. \tag{7}$$

- If a potential P exists, it must satisfy (7) by substituting equation (5) into equation (6).
- The following explicit formula for P establishes existence:

$$P(N, v) = \sum_{S \subseteq N} \frac{(|S| - 1)!(|N| - |S|)!}{|N|!} v(S). \tag{8}$$

- To show (8) holds, use induction on $|N|$ starting with $|N| = 1$ and use equation (7).
- Now, the main result.

Theorem. Let P be a potential function. Then,

$$D^i P(N, v) = \phi_i(N, v) \tag{9}$$

where $\phi(N, v) = (\phi_1(N, v), \phi_2(N, v), \cdots, \phi_n(N, v))$ is the Shapley value of (N, v).

- To prove the theorem, it is sufficient to show that the following function Q is also a potential function.

$$Q(N, v) = \sum_{T \subseteq N, T \neq \emptyset} \frac{\alpha_T}{|T|}, Q(\emptyset, v) = 0 \tag{10}$$

where α_T are the coefficients in equation (4) with respect to v. That is,

$$\sum_{i \in N} D^i Q(N, v) = v(N). \tag{11}$$

- Note that

$$\begin{aligned} D^i Q(N, v) &= Q(N, v) - Q(N \setminus \{i\}, v) \\ &= \sum_{T \subseteq N, T \neq \emptyset} \frac{\alpha_T}{|T|} - \sum_{T \subseteq N \setminus \{i\}, T \neq \emptyset} \frac{\alpha_T}{|T|} \\ &= \sum_{T \subseteq N, i \in T} \frac{\alpha_T}{|T|} \end{aligned}$$

  where the right-hand side of the last equation is precisely $\phi_i(v)$.

- Now, the theorem follows from the uniquness of the potential function (which implies $Q = P$).