협력 게임 이론

tonextpage
https://github.com/tonextpage

Contents

1 협력 게임과 코어

1.1 협력 게임의 할당

n을 자연수라 하고, 전체 참가자의 집합을 $N = \{1, 2, \cdots, n\}$으로 두자. N의 부분집합, 즉 참가자 중 일부를 모은 집합을 **연합(coalition)**이라 하고, N을 **대연합(grand coalition)**이라 하자. N의 모든 연합의 집합은 N의 멱집합 (power set)이므로 이를 2^N으로 표기하자.

정의 1.1

$v(\varnothing) = 0$을 만족하는 함수 $v : 2^N \to \mathbb{R}$을 **특성함수(characteristic function)**라 한다. 각 연합 $S \subseteq N$에 대하여 $v(S)$를 연합 S의 **가치(worth)**라 한다.

특성함수는 각 연합이 생산하는 성과를 측정한다. 연합과 특성함수가 주어진 경우의 게임을 생각하자.

정의 1.2

효용 양도 가능 협력 게임(cooperative game with transferable utility)은 참가자의 집합 N과 특성함수 v의 쌍 (N, v)이다.

효용 양도 가능 협력 게임을 간단히 **협력 게임**이라 하자. 참가자의 집합이 맥락상 분명할 때, 협력 게임 (N, v)를 간단히 v로 표기하자. 연합 $\{i, j, \cdots, k\}$는 중괄호를 생략하여 $i, j, \cdots, k$로 표기하자.

정의 1.3

연합 S의 **보수 분배(payoff distribution)**는 실벡터 $(x_i)_{i \in S}$이다.

보수 분배 $x = (x_1, x_2, \cdots, x_n) \in \mathbb{R}^N$과 연합 $S \subseteq N$에 대하여 $x(S) = \sum_{i \in S} x_i$라 하자. $x(\varnothing) = 0$으로 정하자.

협력 게임에서, 각 참가자의 보수는 보수 분배의 요소로 표현되므로 특성함수과 관계없이 임의로 정의할 수 있다. 그러나 (게임의 설계자인) 우리와 협력 게임의 참가자는 연합과 개인의 성과를 비교하여 보수를 구성하고자 하는 욕망을 갖고 있다. 협력 게임 (N, v)에서 다음 두 조건을 만족하는 보수 분배를 생각하자.

(a) x는 **개인에게 합리적(individually rational)**이다: 모든 $i \in N$에 대하여 $x_i \geq v(i)$.

(b) x는 **효율적(efficient)**이다: $x(N) = v(N)$.

위의 두 조건을 만족하는 보수 분배를 할당(imputation)이라 하자. 할당은 대연합의 성과 $v(N)$을 단독 행동의 성과보다 크도록 분배하는 보수 분배이다. 협력 게임 v에서 모든 할당의 집합을 $I(v)$라 하자.

예시 1.4

서로소(disjoint)인 두 연합 S와 T에 대하여 $v(S \cup T) = v(S) + v(T)$를 만족하는 협력 게임 v를 **가법적(additive)**이라 한다. 가법적 협력 게임에서는 모든 연합 S에 대하여

$$v(S) = \sum_{i \in S} v(i)$$

이므로 연합의 성과는 1인 연합의 성과의 합으로 특정된다. 어떤 할당 x에 대하여 $x_j > v(j)$인 $j \in N$가 존재한다면

$$v(N) = x(N) = \sum_{i \in N} x_i = x_j + \sum_{i \in N \setminus \{j\}} x_i > v(j) + \sum_{i \in N \setminus \{j\}} v(i) = \sum_{i \in N} v(i) = v(N)$$

이므로 모순이다. 따라서 모든 $i \in N$에 대하여 $x_i = v(i)$이므로 $I(v) = \{(v(1), v(2), \cdots, v(n))\}$이다.

협력 게임 v에서 $I(v) \neq \varnothing$일 필요충분조건은

$$v(N) \geq \sum_{i \in N} v(i) \tag{1.1}$$

이다. 부등식 1.1을 만족하는 협력 게임을 **본질적(essential)**이라 한다. 부등식 1.1의 부정이 성립한다고 가정하고 $x \in I(v)$라 하면

$$x(N) = v(N) < \sum_{i \in N} v(i) \leq \sum_{i \in N} x_i = x(N)$$

이므로 $I(v) = \varnothing$이다. 역으로 부등식 1.1이 성립한다고 가정하자. 각 $i \in N$에 대하여 보수 분배 f^i를

$$f_i = \begin{cases} v(j) & (j \neq i\text{인 경우}) \\ v(N) - \sum_{k \in N \setminus \{i\}} v(k) & (j = i\text{인 경우}) \end{cases}$$

로 정의하자. f^i가 효율적임은 자명하다. f^i가 개인에게 합리적임은 $f_i^i \geq v(i)$임을 보이면 충분하다:

$$f_i^i = v(N) - \sum_{k \in N \setminus \{i\}} v(k) \geq \sum_{k \in N} v(k) - \sum_{k \in N \setminus \{i\}} v(k) = v(i)$$

다음 정리에서 본질적 협력 게임의 할당 집합 $I(v)$가 f^i에 의해 특정됨을 알 수 있다.

정리 1.5

본질적 협력 게임 v에서 $I(v)$는 n개의 점 $f^1, f^2, \cdots, f^n$의 볼록 껍질(convex hull)이다.

증명. 먼저 $I(v)$가 볼록집합(convex set)임을 보이자. $x, y \in I(v)$, $0 \leq \lambda \leq 1$이라 하자. $z = \lambda x + (1-\lambda)y$라 하면

(a) $z_i = \lambda x_i + (1-\lambda)y_i \geq \lambda v(i) + (1-\lambda)v(i) = v(i)$

(b) $z(N) = \lambda x(N) + (1-\lambda)y(N) = \lambda v(N) + (1-\lambda)v(N) = v(N)$

이므로 $I(v)$는 볼록집합이다. v가 본질적 협력 게임이므로 $I(v)$는 집합 $\{f^i \mid i \in N\}$의 볼록 껍질을 포함한다. 이제 $x \in I(v)$라 하자. 각 $i \in N$에 대하여 $\alpha_i = x_i - v(i) \geq 0$라 하면

$$x = (v(1), v(2), \cdots, v(n)) + (\alpha_1, \alpha_2, \cdots, \alpha_n)$$

이고

$$\alpha := \sum_{j \in N} \alpha_j = x(N) - \sum_{j \in N} v(j) = v(N) - \sum j \in N v(j) \geq 0$$

이다. 만약 $\alpha = 0$이면 임의의 $i \in N$에 대하여 $x = f^i$이다. 따라서 $\alpha \neq 0$이라 가정하고 $\lambda_j = \alpha_j / \alpha$라 하자. 그러면 $0 \leq \lambda_j \leq 1$, $\sum_{j \in N} \lambda_j = 1$이고

$$\begin{aligned} \sum_{j \in N} \lambda_j f_i^j &= \sum_{j \in N \setminus \{i\}} \lambda_j v(i) + \lambda_i \left[v(N) - \sum_{j \in N \setminus \{i\}} v(j) \right] \\ &= v(i) \sum_{j \in N \setminus \{i\}} \lambda_j + \lambda_i [\alpha + v(i)] \\ &= v(i) \sum_{j \in N} \lambda_j + \lambda_i \alpha = v(i) + \alpha_i = x_i \end{aligned}$$

이므로 $x = \sum_{j \in N} \lambda_j f^j$이다. □

예시 1.6

$(\{1,2,3\}, v)$를 다음과 같은 3인 협력 게임이라 하자.

$$v(1) = v(3) = 0, \quad v(2) = 3, \quad v(1,2,3) = 5$$

그러면 $f^1 = (2,3,0)$, $f^2 = (0,5,0)$, $f^3 = (0,3,2)$이고 $I(v)$는 $\mathbb{R}^3$에서 세 점 f^1, f^2, f^3을 잇는 삼각형이다.

협력 게임에서 한 연합이 구성되었을 때, 그 연합의 일원은 자신의 성과 내에서 최대한 많은 보상을 받고 싶어한다. 이에 연합 내에서 보수를 비교하는 기준을 다음과 같이 정의하자.

정의 1.7

(N, v)를 협력 게임이라 하자. $y, z \in I(v)$이고 $S \in 2^N \setminus \{\varnothing\}$이라 하자. 다음 두 조건을 만족하면 연합 S에서 **y가 z를 압도한다(y dominates z via coalition S)**고 하고, 이를 $y \succ_S z$로 표기한다.

(i) 모든 $i \in S$에 대하여 $y_i > z_i$이다.

(ii) $y(S) \leq v(S)$

연합 S에서 할당 y가 z를 압도함은 S의 각 참가자가 받는 보수가 더 크면서, S 내에서의 협력만으로 보수를 충당할 수 있음을 의미한다. 연합 S에 대하여 집합 $D(S)$를

$$D(S) = \{z \in I(v) \mid y \succ_S z \text{인 } y \in I(v) \text{가 존재한다.}\}$$

로 정의하자. 이 집합에 속하는 할당은 항상 어떤 다른 할당에 비해 우월하지 않으므로 연합 S에 속한 참가자는 $D(S)$의 할당을 거부한다. 집합 $I(V) \setminus \bigcup_{S \in 2^N \setminus \{\varnothing\}} D(S)$에 속한 할당을 **압도당하지 않는다(undominated)**고 한다.

예시 1.8

협력 게임 $(N = \{1,2,3\}, v)$를

$$v(S) = \begin{cases} 2 & (S = \{1,2\}) \\ 1 & (S = N) \\ 0 & (S \neq \{1,2\}, N) \end{cases}$$

으로 정의하자. 그러면 $I(v)$는 세 점 $(1,0,0)$, $(0,1,0)$, $(0,0,1)$을 꼭짓점으로 하는 삼각형이고, $D(\{1,2\}) = \{x \in I(v) \mid x_3 > 0\}$이다. $\{1,2\}$가 아닌 모든 연합 S에 대해서는 $D(S) = \varnothing$이다.

연습문제 1.9

협력 게임 (N, v)에서 $|S| = 1$이거나 $S = N$이면 $D(S) = \varnothing$임을 보여라.

풀이

일반성을 잃지 않고 $S = \{1\}$이라 하자. 할당 y, z에 대하여 $z \in D(S)$이고 $y \succ_S z$라 가정하자. y는 S에서 z를 압도하는 할당이므로 $y_1 \geq v(1)$이고 $y_1 \leq v(1)$이다. 따라서

$$v(1) = y_1 > z_1 \geq z_1 \geq v(1)$$

이 성립하므로 모순이다. 이제 $S = N$이라 하자. 할당 y와 z가 앞선 상황과 같은 조건을 만족하면 모든 $i \in N$에

대하여 $y_i \geq v(i)$, $z_i \geq v(i)$, $y_i > z_i$이므로

$$v(N) = y(N) = \sum_{i \in N} y_i > \sum_{i \in N} z_i = z(N) = v(N)$$

에서 모순이다. 그러므로 $|S| = 1$이거나 $S = N$이면 $D(S) = \varnothing$이다. □

1.2 코어와 D-코어

협력 게임 이론에서는 대연합을 형성하도록 하는 조건을 모색하는데 중점을 둔다. (이 조건을 협력 게임의 **해(solution)**라 한다.) 특정 참가자들이 대연합보다 작은 연합, 즉 파벌을 형성하지 않도록 하는 보수 분배의 집합을 **코어**라 한다.

정의 1.10

협력 게임 (N, v)의 **D-코어(unDominated core)**는 집합

$$DC(v) := I(v) \setminus \bigcup_{S \in 2^N \setminus \{\varnothing\}} D(S),$$

즉 압도당하지 않는 할당의 집합이다. 협력 게임 (N, v)의 **코어(core)**는 다음과 같이 정의한다.

$$C(v) := \{x \in I(v) \mid \text{모든 } S \subseteq N\text{에 대하여 } x(S) \geq v(S).\}$$

$C(v) \neq \varnothing$이면 대연합 N이 아닌 연합(파벌) S가 형성될 이유는 없다. $x \in C(v)$에 대하여 연합 S에 배정된 보수의 합 $x(S)$가 연합 S를 형성하여 생산하는 가치 $v(S)$보다 부족하지 않기 때문이다.

예시 1.11

예시 1.8에서 정의한 협력 게임의 D-코어는 공집합이 아니지만 코어는 공집합이다.

예시 1.12

실수 $\alpha \in [0, 1]$에 대하여 3인 협력 게임 $(N = \{1, 2, 3\}, v)$를 다음과 같이 정의하자.

$$v(S) = \begin{cases} 1 & (|S| = 3) \\ \alpha & (|S| = 2) \\ 0 & (|S| \leq 1) \end{cases}$$

협력 게임 (N, v)의 코어 $C(v)$는

$$x_i \geq 0, \quad x_1 + x_2 + x_3 = 1, \quad x_1 + x_2 \geq \alpha, \quad x_2 + x_3 \geq \alpha, \quad x_3 + x_1 \geq \alpha$$

를 만족하는 벡터 (x_1, x_2, x_3)의 집합이다. 만약 $\alpha > 2/3$이면

$$x_1 + x_2 + x_3 = \frac{1}{2}[(x_1 + x_2) + (x_2 + x_3) + (x_3 + x_1)] \geq 3\alpha > 2$$

이다. 따라서 $C(v) \neq \varnothing$일 필요충분조건은 $\alpha \leq 2/3$이다.

코어와 D-코어 사이에는 다음 관계가 성립한다.

정리 1.13

협력 게임 v에서 $C(v) \subseteq DC(v)$이다.

증명. 협력 게임 v에 대하여 $x \in I(v) \setminus DC(v)$라 하자. 그러면 연합 $S \neq \varnothing$과 할당 y가 존재하여 $y \succ_S x$이다. 따라서 $v(S) \geq y(S) > x(S)$가 성립하므로 $x \notin C(v)$이다. □

협력 게임의 코어는 선형 연립부등식으로부터 구할 수 있는 다포체이다. 특히 D-코어는 볼록 집합이다. (연습문제 1.16) 코어와 D-코어가 같아지도록 하는 충분조건은 다음과 같다.

정리 1.14

협력 게임 (N, v)가

$$\text{모든 연합 } S \subseteq N\text{에 대하여 } v(N) \geq v(S) + \sum_{i \in N \setminus S} v(i) \tag{1.2}$$

를 만족하면 $DC(v) = C(v)$이다.

증명. 정리 1.13으로부터 $DC(v) \subseteq C(v)$임을 보이면 충분하다. 먼저 다음을 보이자.

보조정리

할당 x가 어떤 연합 S에 대하여 $x(S) < v(S)$를 만족하면 $y \succ_S x$인 할당 y가 존재한다.

y를 다음과 같이 정의하자.

$$y_i = \begin{cases} x_i + \dfrac{v(S) - x(S)}{|S|} & (i \in S\text{인 경우}) \\ v(i) + \dfrac{v(N) - v(S) - \sum_{i \in N \setminus S} v(i)}{|N \setminus S|} & (i \notin S\text{인 경우}) \end{cases}$$

그러면 조건 1.2로부터 모든 $i \in N \setminus S$에 대하여 $y_i \geq v(i)$이므로 y는 $y \succ_S x$를 만족하는 할당이다. 이제 $DC(v) \subseteq C(v)$임을 보이기 위해 $x \in DC(v)$라 하자. x를 압도하는 할당은 존재하지 않으므로 모든 $S \in 2^N \setminus \{\varnothing\}$에 대하여 $x(S) \geq v(S)$이 되어 $x \in C(v)$이다. □

참고. 협력 게임 v의 코어가 공집합이 아니면 조건 1.2를 만족한다. $x \in C(v)$이면 모든 $S \in 2^N \setminus \{\varnothing\}$에 대하여 $x(S) \geq v(S)$이므로

$$v(N) = x(N) = x(S) + \sum_{i \in N \setminus S} x(i) \geq v(S) + \sum_{i \in N \setminus S} v(i)$$

가 성립한다. 그러므로 이 경우에 D-코어와 코어는 서로 같다.

현실적인 상황에서 유도되는 협력 게임 v는 다음 조건을 만족한다.

$$\text{임의의 서로소인 두 연합 } S\text{와 } T\text{에 대하여 } v(S \cup T) \geq v(S) + v(T). \tag{1.3}$$

조건 1.3을 만족하는 협력 게임을 **초가법적(superadditive)**라 한다. 조건 1.3은 조건 1.2를 함의하므로 정리 1.14는 초가법적 협력 게임에서도 성립한다. □

연습문제 1.15

정리 1.14의 역은 성립하는가?

풀이

$DC(v) \neq \varnothing$이면 역도 성립한다. ([3] Proposition 2.1) $C(v) = DC(v)$라 가정하자. $DC(v) \neq \varnothing$이므로 $x \in C(v)$라 하자. 그러면 임의의 연합 $S \subseteq N$에 대하여

$$v(N) = x(N) = x(S) + \sum_{i \in N \setminus S} x(i) \geq v(S) + \sum_{i \in N \setminus S} v(i)$$

이 성립한다. □

연습문제 1.16

협력 게임 (N, v)의 D-코어가 볼록 집합임을 보여라.

풀이

([3] 참고) 협력 게임 (N, v)에 대하여 특성함수 v^*를 $v^*(\varnothing) = 0$이고, $S \in 2^N \setminus \{\varnothing\}$에 대하여

$$v^*(S) = \min\left\{v(S), v(N) - \sum_{i \notin S} v(i)\right\}$$

로 정의하자. 다음을 먼저 보이자.

보조정리

$DC(v) \neq \varnothing$이면 $DC(v) = C(v^*)$이다.

$v^*(N) = v(N)$이고, 정의로부터 $v^*(i)$는 $v(i)$와 $v(N) - \sum_{j \neq i} v(j)$의 임의의 볼록 결합보다 작거나 같으므로

$$\begin{aligned}\sum_{i \in N} v^*(i) &\leq \frac{|N|-1}{|N|} \sum_{i \in N} v(i) + \frac{1}{|N|} \sum_{i \in N} \left[v(N) - \sum_{j \neq i} v(j)\right] \\ &= \frac{|N|-1}{|N|} \sum_{i \in N} v(i) + v(N) - \frac{|N|-1}{|N|} \sum_{i \in N} v(i) = v(N) = v^*(N)\end{aligned}$$

이 성립한다. 협력 게임 v^*는 본질적이므로 $I(v^*) \neq \varnothing$이다. 또한 $v^*(i) = v(i)$일 필요충분조건은 $I(v) \neq \varnothing$이다:

$$v^*(i) = v(i) \iff v(i) \leq v(N) - \sum_{j \neq i} v(j) \iff v(N) \geq \sum_{j \in N} v(j) \iff I(v) \neq \varnothing.$$

이로부터 $I(v) \neq \varnothing$일 필요충분조건은 $I(v^*) = I(v)$임을 알 수 있다. 이제 $DC(v) \neq \varnothing$이라 가정하면 $I(v) \neq \varnothing$이므로 $I(v) = I(v^*)$이다. $x \notin DC(v)$라 하자. 그러면 연합 S와 할당 y가 존재하여 $y \succ_S x$이다. y가 할당이므로

$$y(S) = y(N) - \sum_{i \notin S} y_i \leq v(N) - \sum_{i \notin S} v(i)$$

이 성립하고, $y(S) \leq v(S)$이므로 $y(S) \leq v^*(S)$이다. 따라서 $y \in I(v^*)$이고 협력 게임 (N, v^*)에서 $y \succ_S x$이다. 따라서 $x \notin DC(v^*)$이므로 $DC(v^*) \subseteq DC(v)$가 성립한다. $DC(v) \subseteq DC(v^*)$임은 자명하므로 정리 1.14에 의해

$$DC(v) = DC(v^*) = C(v^*)$$

이다. 협력 게임의 코어가 볼록 집합임은 쉽게 확인할 수 있으므로 $DC(v)$ 역시 볼록 집합이다. □

1.3 단순 협력 게임

정의 1.17

단순 협력 게임(simple game)은 다음 두 조건을 만족하는 협력 게임 (N, v)이다.

(i) 임의의 연합 S에 대하여 $v(S) \in \{0, 1\}$이다.

(ii) $v(N) = 1$

$v(S) = 1$을 만족하는 연합 S를 **승리 연합(winning coalition)**, 그렇지 않은 연합을 **패배 연합(losing coalition)**이라 한다.

정의 1.18

단순 협력 게임 (N, v)에 대하여

- 모든 진부분집합이 패배 연합인 승리 연합을 **최소 승리 연합(minimal winning coalition)**이라 한다.
- $v(S) = 1 \iff i \in S$를 만족하는 참가자 i를 **독재자(dictator)**라 한다.
- 모든 승리 연합에 속해있는 참가자 i를 **거부권자(veto player)**라 하고, 모든 거부권자의 집합을 $\mathrm{veto}(v)$로 나타낸다. 즉 $\mathrm{veto}(v) = \bigcap\{S \in 2^N \mid v(S) = 1\}$이다.

각 $i \in N$에 대하여 i번째 성분만 1이고 나머지 성분은 모두 0인 $\mathbb{R}^N$의 벡터를 e^i라 하자.

예시 1.19

$i \in N$이라 하자. **독재 게임(dictator game)** δ_i를

$$\delta_i(S) = 1 \iff i \in S$$

인 단순 협력 게임이라 하자. 그러면 $\mathrm{veto}(\delta_i) = \{i\}$이다. $x \in I(v)$이면 모든 $j \neq i$에 대하여 $x_j \geq v(j) = 0$이고

$$1 = v(i) \leq x_i \leq x(N) = v(N) = 1$$

이므로 $x = e^i$이다. 따라서 $I(\delta_i) = \{e^i\}$이고, $C(\delta_i) = DC(\delta_i) = \{e^i\}$이다.

예시 1.20

$|N|$이 홀수일 때, **다수결 게임(majority game)** (N, v)를

$$v(S) = 1 \iff |S| \geq \frac{|N|}{2}$$

인 단순 협력 게임이라 하자. $x \in C(v)$라 하자. $|S| = |N| - 1$이면 $v(S) = 1$이므로 $x(S) = 1$이다. 크기가 $|N| - 1$인 연합은 $|N|$개 있으므로 $\sum_{S:|S|=|N|-1} x(S) = |N|$이다. 한편,

$$\sum_{S:|S|=|N|-1} x(S) = \sum_{S:|S|=|N|-1} \sum_{i \in S} x_i = (|N| - 1)x(N) = |N| - 1$$

이므로 이는 모순이다. 따라서 다수결 게임의 코어는 공집합이다.

단순 협력 게임의 코어는 거부권자가 존재할 때에만 공집합이 아니다. 게다가 코어에 속하는 보수 분배는 대연합의 가치 1을 거부권자에게만 분배한다.

정리 1.21

(N, v)를 단순 협력 게임이라 하자.

(1) $C(v)$는 집합 $\{e^i \mid i \in \text{veto}(v)\}$의 볼록 껍질이다.

(2) $\text{veto}(v) = \varnothing$이고 $\{i \in N \mid v(i) = 1\} = \{k\}$이면 $C(v) = \varnothing$이고 $DC(v) = \{e^k\}$이다. 그렇지 않으면, $DC(v) = C(v)$이다.

증명.

(1) $i \in \text{veto}(v)$이고 $S \in 2^N \setminus \{\varnothing\}$이라 하자. $i \in S$이면 $e^i(S) = 1 \geq v(S)$이고, $i \notin S$이면 $e^i(S) = 0 = v(S)$이다. $e^i(N) = 1 = v(N)$이므로 $e^i \in C(v)$이다. $C(v)$는 볼록집합이므로 집합 $\{e^i \mid i \in \text{veto}(v)\}$의 볼록 껍질을 포함한다. 역으로 $x \in C(v)$라 하자. $i \notin \text{veto}(v)$이면 $x_i = 0$임을 보이면 충분하다. 어떤 $i \notin \text{veto}(v)$에 대하여 $x_i > 0$이라 가정하자. i는 거부권자가 아니므로 $i \notin S$이고 $v(S) = 1$인 연합 S가 존재한다. 그러면

$$x(S) = x(N) - x(N \setminus S) \leq 1 - x_i < 1 = v(S)$$

이므로 $x \in C(v)$임에 모순이다. 따라서 $C(v)$는 집합 $\{e^i \mid i \in \text{veto}(v)\}$의 볼록 껍질이다.

(2) $\text{veto}(v) = \varnothing$이면 (1)에 의해 $C(v) = \varnothing$이다. 여기에 k가 집합 $\{i \in N \mid v(i) = 1\}$의 유일한 원소이면 $I(v) = \{e^k\}$이므로 $DC(v) = \{e^k\}$이다. $\text{veto}(v) = \varnothing$이고 $\{i \in N \mid v(i) = 1\} = \varnothing$이면 조건 1.2를 만족하므로 정리 1.14에 의해 코어와 D-코어는 같다. $\text{veto}(v) = \varnothing$이고 $|\{i \in N \mid v(i) = 1\}| \geq 2$이면 $I(v) = \varnothing$이므로 $C(v) = DC(v) = \varnothing$이다. 이제 $\text{veto}(v) \neq \varnothing$이면 (1)에 의해 $C(v) \neq \varnothing$이므로 $C(v) = DC(v)$이다. □

예시 1.22

T를 공집합이 아닌 연합이라 하자. T**–만장일치 게임(**T**–unanimity game)** u_T를

$$u_T(S) = 1 \iff T \subseteq S$$

인 단순 협력 게임이라 하자. $\text{veto}(u_T) = T$이므로 정리 1.21로부터 $C(u_T)$와 $DC(u_T)$ 모두 집합 $\{e^i \mid i \in T\}$의 볼록 껍질과 같다.

예시 1.23

3인 협력 게임 $(\{1, 2, 3\}, v)$를

$$v(S) = \begin{cases} 1 & (S = \{1\}, \{2, 3\}, \{1, 2, 3\}) \\ 0 & (\text{그 외의 경우}) \end{cases}$$

로 정의하자. 그러면 $\text{veto}(v) = \varnothing$, $C(v) = \varnothing$, $DC(v) = \{(1, 0, 0)\}$이다. (N, v)는 초가법적이 아니고 조건 1.2를 만족하지 않는다.

1.4 안정 집합

존 폰 노이만과 모겐스턴은 협력 게임에서 다음과 같은 해를 고려했다.

정의 1.24

협력 게임 v에 대하여 집합 $A \subseteq I(v)$가 다음 두 조건을 만족하면 A를 **안정 집합(stable set)**이라 한다.

(i) **내적 안정성(internal stability)**: $x, y \in A$이면 x는 y를 압도하지 않는다.

(ii) **외적 안정성(external stability)**: $x \in I(v) \setminus A$이면 x를 압도하는 $y \in A$가 존재한다.

안정 집합은 할당의 비교 관점에서 '안정적'이다. 한 협력 게임에 여러 안정 집합이 존재할 수 있으며, 안정성은 집합이 갖는 성질이므로 안정 집합의 적절한 선택이 필요하다. 코어가 존재하는 협력 게임은 1.5에서 완벽하게 특정되지만, 안정 집합의 존재성은 부분적으로만 해결되었다.

정리 1.25

v를 단순 협력 게임이라 하고, S를 최소 승리 연합이라 하자. 집합 Δ^S를

$$\Delta_S := \{x \in I(v) \mid \text{모든 } i \notin S\text{에 대하여 } x_i = 0\}$$

이라 하자. $\Delta^S \neq \varnothing$이면 Δ^S는 안정 집합이다.

증명. $\Delta^S \neq \varnothing$이면 모든 $i \notin S$에 대하여 $v(i) = 0$임은 쉽게 확인할 수 있다.

(i) $x, y \in \Delta^S$와 연합 T에 대하여 $x \succ_T y$라 하자. $i \notin S$에 대하여 $x_i = y_i = 0$이므로 $T \subseteq S$여야 한다. 만약 $T \subsetneq S$이면 S의 최소성에 의해 $v(T) = 0$이므로 $x(T) = 0$이다. 따라서 모든 $i \in T$에 대하여 $x_i = 0$이므로 $x \succ_T y$임에 모순이다. 이제 $T = S$라 하자. $x(N \setminus S) = \sum_{i \notin S} x_i = 0$이므로 $x(S) = x(S) + x(N \setminus S) = x(N) = 1$이고, 같은 논리로 $y(S) = 1$이다. 이는 $x \succ_T y$임에 모순이다.

(ii) $x \in I(v) \setminus \Delta^S$라 하자. 그러면 어떤 $i \notin S$에 대하여 $x_i > 0$이므로 $\alpha := x(N \setminus S) > 0$이다. 보수 분배 y를

$$y_i = \begin{cases} x_i + \dfrac{\alpha}{|S|} & (i \in S) \\ 0 & (i \notin S) \end{cases}$$

로 정의하자. 그러면 $y \in \Delta^S$이고 $y(S) = \sum_{i \in S} x_i + \alpha = x(N) = 1$이므로 $y \succ_S x$이다.

(i), (ii)로부터 Δ^S가 안정 집합임을 알 수 있다. □

D-코어와 안정 집합 사이의 관계는 다음과 같다.

정리 1.26

협력 게임 (N, v)에 대하여

(1) D-코어는 모든 안정 집합의 부분집합이다.

(2) A와 B가 서로 다른 안정 집합이면 $A \not\subseteq B$이다.

(3) D-코어가 안정 집합이면 이는 유일한 안정 집합이다.

증명. (1)과 (2)는 외적 안정성에 의해 성립한다. (3)은 (1)과 (2)에 의해 성립한다. □

안정 집합은 참가자의 행동 기준(standard of behavior)를 규정한다. 다음 예시를 보자.

예시 1.27

$(\{1,2,3\}, v)$를 3인 다수결 게임이라 하자. (예시 1.20) 집합 X를

$$X := \left\{\left(\frac{1}{2}, \frac{1}{2}, 0\right), \left(\frac{1}{2}, 0, \frac{1}{2}\right), \left(0, \frac{1}{2}, \frac{1}{2}\right)\right\}$$

라 하자. X의 내적 안정성은 쉽게 확인할 수 있다. X의 외적 안정성을 보이기 위해 $z \in I(v) \setminus X$라 하자. 그러면 어느 두 참가자 i와 j에 대하여 $z_i < 1/2$이므로 연합 $\{i,j\}$에서 z를 압도하는 $x \in X$를 택할 수 있다. 따라서 X는 안정 집합이다. X은 두 참가자가 연합을 이루어 보상을 얻은 후 이를 서로 반씩 가지는 행동 양식으로 해석 가능하다. 이제 $c \in [0, 1/2)$이라 하고 정해진 차가자 i에 대하여

$$Y_{i,c} := \{y \in I(v) \mid y_i = c\}$$

라 하자. $Y_{i,c}$의 내적 안정성 역시 쉽게 확인할 수 있다. 일반성을 잃지 않고 $i = 3$이라 하자. $Y_{i,c}$의 외적 안정성을 보이기 위해 $z \in I(v) \setminus Y_{i,c}$라 하자. $z_3 > c$이면 $z_1 + z_2 < 1 - c$이므로 할당 $y = (y_1, y_2, y_3)$을

$$y_1 = z_1 + \frac{(1-c) - (z_1 + z_2)}{2}, \quad y_2 = z_2 + \frac{(1-c) - (z_1 + z_2)}{2}, \quad y_3 = c$$

라 하면 $y \succ_{\{1,2\}} z$이다. $z_3 < c$이고 $z_1 \leq z_2$이면 $(1-c, 0, c) \succ_{\{1,3\}} z$이다. 따라서 $Y_{i,c}$ 역시 안정 집합이다. $Y_{i,c}$는 참가자 i를 정해 무조건적으로 c만큼의 보상을 받게끔 하는 행동 양식으로 해석 가능하다.

연습문제 1.28 (글러브 게임 각색)

어느 보물 상자는 두 열쇠 α와 β가 적어도 하나씩 있어야 열리는 구조이다. 세 사람 A, B, C는 각각 열쇠 α, β, α를 갖고 있다. 그러므로 보물 상자를 열기 위해서는 B가 A 또는 C와 같은 편이 되어야 한다. 이를 다음과 같이 정의된 3인 협력 게임 $(\{1,2,3\}, v)$로 기술하자.

$$v(S) = \begin{cases} 1 & (S = \{1,2\}, \{2,3\}, \{1,2,3\}) \\ 0 & (\text{그 외의 경우}) \end{cases}$$

(a) $e^2 = (0,1,0)$이 아닌 할당은 어떤 다른 할당에 압도됨을 보여라.

(b) 이 협력 게임의 코어와 D-코어를 구하여라.

(c) D-코어가 안정 집합이 아님을 보여라.

(d) 집합 $B = \{(x, 1-2x, x) \mid 0 \leq x \leq 1/2\}$이 안정 집합임을 보여라.

풀이

(a) 할당 $x = (x_1, x_2, x_3)$가 e^2이 아니면 $x_2 < 1$이다.

  (i) $x_3 > 0$인 경우: $(x_1 + x_3/2, x_2 + x_3/2, 0) \succ_{\{1,2\}} x$이다.

  (ii) $x_3 = 0$인 경우: $(0, x_2 + x_1/2, x_1/2) \succ_{\{2,3\}} x$이다.

(b) $\mathrm{veto}(v) = \{2\}$이므로 정리 1.14와 정리 1.21로부터 $C(v) = DC(v) = \{e^2\}$이다.

(c) v의 정의에 의해 어떤 할당이 연합 S에서 다른 할당을 압도한다면 S는 2를 반드시 포함해야 한다. 그러므로 D-코어는 외적 안정성을 갖지 않는다.

(d) 편의를 위해 실수 $x \in [0, 1/2]$에 대하여 $b(x) = (x, 1-2x, x) \in B$라 하자.

(i) 내적 안정성: (c)와 같은 논리로 어떤 연합 S에 대하여 $b(x) \succ_S b(y)$이면

$$b(x)_2 > b(y)_2 \implies 1 - 2x > 1 - 2y \implies x < y \implies S = \{2\}$$

이다. $v(2) = 0$이므로 이는 불가능하다.

(ii) 외적 안정성: 할당 $x = (x_1, x_2, x_3)$가 B에 속하지 않을 필요충분조건은 $x_1 \neq x_3$이다. 그러한 할당 x에 대하여 $y = (x_1 + x_3)/2$라 하면 $b(y) \in B$이고 $b(y) \succ_{\{1,2\}} x$ ($x_1 < x_3$인 경우) 또는 $b(y) \succ_{\{2,3\}} x$ ($x_1 > x_3$인 경우)이다. □

1.5 코어의 존재성

여기서는 협력 게임의 코어가 공집합이 아닐 필요충분조건을 논한다. 코어는 단순히 선형 연립부등식의 해로 정의되지만, 협력 게임의 사고 방식으로부터 구체적인 필요충분조건을 유도하고자 한다. 협력 게임 (N, v)의 연합 S에 대하여 **특성벡터(characteristic vector)** $e^S \in \mathbb{R}^N$을 다음과 같이 정의하자.

$$e_i^S = \begin{cases} 1 & (i \in S) \\ 0 & (i \notin S) \end{cases}$$

함수 $\lambda : 2^N \setminus \{\varnothing\} \to [0, \infty)$가

$$\sum_{S \subseteq N} \lambda(S) e^S = e^N$$

을 만족하면 λ를 **균형 사상(balanced map)**이라 한다. 공집합이 아닌 연합의 모임 B가 어떤 균형 사상 λ에 대하여

$$B = \{S \in 2^N \setminus \{\varnothing\} \mid \lambda(S) > 0\}$$

을 만족하면 B를 **균형 연합 모임(balanced collection)**이라 한다.

균형 사상은 참가자의 시간 분배의 관점에서 해석할 수 있다. 각 참가자마다 정해진 단위 시간이 주어진다고 생각하자. 참가자 i는 자신이 속할 수 있는 연합(파벌)마다 시간을 투자할 수 있다. 이러한 '시간 배분'이 어떤 균형 사상 λ로 기술되면 적절한 배분으로 보는 것이다. $\lambda(S)$는 연합 S가 형성되어 존재하는 시간으로 생각할 수 있고, 균형 사상은 각 참가자가 자신이 가진 한 단위의 시간을 제각기 다른 연합에 정확하게 배분함을 의미한다.

예시 1.29

(1) $B = \{N_1, N_2, \cdots, N_k\}$를 공집합이 아닌 연합으로 이루어진 N의 분할이라 하자. 사상 λ를

$$\lambda(S) = \begin{cases} 1 & (S \in B) \\ 0 & (\text{그 외의 경우}) \end{cases}$$

라 하면 λ는 균형 사상이고, B는 λ에 대해 균형 연합 모임이 된다.

(2) $N = \{1, 2, 3\}$에 대하여 집합 $B = \{\{1, 2\}, \{1, 3\}, \{2, 3\}\}$은

$$\lambda(S) = \begin{cases} 1/2 & (|S| = 2) \\ 0 & (\text{그 외의 경우}) \end{cases}$$

로 정의된 균형 사상 λ에 대해 균형 연합 모임을 이룬다.

정의 1.30

협력 게임 (N, v)가 모든 균형 사상 $\lambda : 2^N \setminus \{\varnothing\} \to [0, \infty)$에 대하여

$$\sum_{S \subseteq N} \lambda(S) v(S) \leq v(N)$$

을 만족하면 (N, v)를 **균형 게임(balanced game)**이라 한다.

시간 배분의 관점을 고려하여 균형 게임을 해석하자. '균형잡힌' 시간 배분 λ에 대하여 각 참가자가 각 연합에 시간을 투자하여 얻는 가치의 총합이 부등식의 좌변으로 표현된다. 한편 우변은 참가자 전원이 자신의 시간을 대연합에 투자하여 얻는 가치이다. 그러므로 균형 게임은 대연합에 온전히 집중하지 않고 더 작은 연합(파벌)을 만드는 행위가 비효율적임을 의미한다. 여기에 코어의 의미를 덧붙여 생각하면 다음 정리를 자연스럽게 떠올릴 수 있다.

정리 1.31 (Bondareva–Shapley)

협력 게임 (N, v)에 대하여 다음은 서로 동치이다.

(i) $C(v) \neq \varnothing$

(ii) (N, v)는 균형 게임이다.

정리 1.31은 선형계획법의 쌍대성으로부터 유도된다.

보조정리 1.32

$A \in \mathbb{R}^{n \times p}$, $b \in \mathbb{R}^p$, $c \in \mathbb{R}^n$이라 하고, 두 집합 $X = \{x \in \mathbb{R}^n \mid xA \geq b\}$와 $Y = \{y \in \mathbb{R}^p \mid Ay = c, y \geq 0\}$이 모두 공집합이 아니라고 하자. 그러면 다음이 성립한다:

$$\min\{x \cdot c \mid xA \geq b\} = \max\{b \cdot y \mid Ay = c, y \geq 0\}.$$

만약 X와 Y 중 하나가 공집합이면 위의 최댓값와 최솟값은 달성할 수 없다.

정리 1.31의 증명. (N, v)를 협력 게임이라 하자. $n = |N|$에 대하여 $p = 2^n$이라 하고, N의 모든 부분집합을

$$S_1 = \varnothing, S_2, \cdots, S_p = N$$

이라 하자. $k = 1, 2, \cdots, p$에 대하여 e^{S_k}를 k번째 열로 갖는 $n \times p$ 행렬을 A라 하자. $b = (v(S_k))_{k=1}^p \in \mathbb{R}^p$라 하면

$$xA \geq b \iff x \cdot e^{S_k} \geq v(S_k) \iff x(S_k) \geq v(S_k)$$

이므로 $\{x \in \mathbb{R}^n \mid xA \geq b\} \neq \varnothing$일 필요충분조건은 $C(v) \neq \varnothing$이다. 한편

$$Ay = c \iff \sum_{k=1}^{p} y_k e^{S_k} = e^N$$

이므로 집합 $\{y \in \mathbb{R}^p \mid Ay = c, y \geq 0\} \neq \emptyset$이다. (벡터 $(0, 0, \cdots, 0, 1) \in \mathbb{R}^p$이 이 집합에 속한다.) $C(v) \neq \varnothing$일 필요충분조건이

$$v(N) = \min\{x(N) \mid x \in \mathbb{R}^N \text{이고 모든 } S \subseteq N\text{에 대하여 } x(S) \geq v(S).\} = \min\{x \cdot c \mid xA \geq b\}$$

이다. 보조정리 1.32로부터 이는

$$v(N) = \max\left\{\sum_{S\subseteq N}\lambda(S)v(S) \mid \sum_{S\subseteq N}\lambda(S)e^S = e^N, \lambda \geq 0\right\} = \{b\cdot y \mid Ay = c, y \geq 0\}$$

와 동치이고, 이는 (N, v)가 균형 게임임을 의미한다. □

연습문제 1.33

$N = \{1, 2, 3, 4\}$이라 하자. 다음과 같이 정의된 협력 게임 (N, v)의 코어가 공집합임을 보여라.

$$v(S) = \begin{cases} 1 & (S = N) \\ \frac{3}{4} & (S = \{1,2\}, \{1,3\}, \{1,4\}, \{2,3,4\}) \\ 0 & (\text{그 외의 경우}) \end{cases}$$

풀이

사상 $\lambda : 2^N \setminus \{\varnothing\} \to [0, \infty)$를 다음과 같이 정의하자.

$$\lambda(\{1,2\}) = \lambda(\{1,3\}) = \lambda(\{1,4\}) = \frac{1}{3}, \quad \lambda(\{2,3,4\}) = \frac{2}{3}, \quad \text{그 외의 경우에는 } \lambda(S) = 0$$

그러면

$$\sum_S \lambda(S)e^S = \frac{1}{3}(1,1,0,0) + \frac{1}{3}(1,0,1,0) + \frac{1}{3}(1,0,0,1) + \frac{2}{3}(0,1,1,1) = e^N$$

이므로 λ는 균형 사상이고,

$$\sum_S \lambda(S)v(S) = \frac{3}{4}\left(\frac{1}{3} + \frac{1}{3} + \frac{1}{3} + \frac{2}{3}\right) = \frac{5}{4} > 1 = v(N)$$

이므로 (N, v)는 균형 게임이 아니다. 따라서 정리 1.31에 의해 $C(v) = \varnothing$이다.

2 섀플리 값

2.1 섀플리 값의 정의

지금까지는 여러 해 집합의 성질을 확인했다. 이제 각 협력 게임에 하나의 보수 체계를 대응시키는 상황, 즉 단일점 해 (one-point solution)를 생각하자. 참가자의 집합 N에 대한 모든 협력 게임의 집합을 $\mathcal{G}^N$이라 하자.

정의 2.1

$\mathcal{G}^N$ 위의 **값(value)**은 함수 $\psi : \mathcal{G}^N \to \mathbb{R}^N$이다. ψ는 각 협력 게임 v를 보수 분배 $\psi(v)$에 대응시킨다.

협력 게임이 현실적인 협력과 분열의 상황을 반영함을 고려하면, 우리는 협력 게임의 값을 일관적인 규칙으로 정하고 싶다. 그 중 참가자의 기여도에 따라 보수를 분배하는 방법이 가장 직관적이다. (N, v)를 협력 게임이라 하고, $\sigma : N \to N$을 N의 순열(일대일대응)이라 하자. 이제 하나의 연합에 참가자 $\sigma(1), \sigma(2), \cdots, \sigma(n)$이 차례대로 합류한다고 생각하자. 참가자가 합류할 때마다 연합이 확장되고, 추가 참가자에 따른 성과 변화를 특성함수로 측정할 수 있다. 참가자 i보다 먼저 합류한 참가자의 연합을

$$P_\sigma(i) = \{j \in N \mid \sigma^{-1}(j) < \sigma^{-1}(i)\}$$

로 정의하자. 예를 들어 $N = \{1,2,3,4,5\}$이고 순열 $\sigma : N \to N$이

$$\sigma(1) = 2, \quad \sigma(2) = 5, \quad \sigma(3) = 4, \quad \sigma(4) = 1, \quad \sigma(5) = 3$$

이면, 참가자 2가 먼저 연합에 합류하고, 그 다음에 참가자 $5, 4, 1, 3$의 순서로 합류한다. 따라서 $P_\sigma(1) = \{2,5,4\}$이다. 이제 **한계 기여도 벡터(marginal vector)** $m^\sigma = (m_1^\sigma, m_2^\sigma, \cdots, m_n^\sigma)$을

$$m_i^\sigma = v(P_\sigma(i) \cup \{i\}) - v(P_\sigma(i)) \tag{2.1}$$

로 정의하자. 한계 기여도 벡터 m^σ는 합류 순서 σ가 주어질 때, 참가자 i의 합류에 의한 가치 변화(참가자 i의 기여도)를 나타낸다. 이제 가능한 모든 순열 σ에 대한 평균치를 택하면 처음에 목표했던 보수 분배를 구성할 수 있다.

정의 2.2

$\Pi(N)$을 N 위의 모든 순열의 집합이라 하자. 협력 게임 $(N, v) \in \mathcal{G}^N$의 **섀플리 값(Shapley value)** $\Phi(v)$를 다음과 같이 정의한다:

$$\Phi(v) = \frac{1}{n!} \sum_{\sigma \in \Pi(N)} m^\sigma.$$

예시 2.3

(1) 2인 협력 게임 $(\{1,2\}, v)$의 섀플리 값은 다음과 같다.

$$\Phi(v) = \left(v(1) + \frac{v(N) - v(1) - v(2)}{2}, \quad v(2) + \frac{v(N) - v(1) - v(2)}{2} \right)$$

(2) 가법적 협력 게임 (N, v)의 섀플리 값은 $\Phi(v) = (v(1), v(2), \cdots, v(n))$이다. (예시 1.4 참고)

식 2.1을 대입하여 섀플리 값의 i 성분을 생각하자.

$$\Phi_i(v) = \frac{1}{n!} \sum_{\sigma \in \Pi(N)} [v(P_\sigma(i) \cup \{i\}) - v(P_\sigma(i))] \tag{2.2}$$

σ가 N의 순열이므로 $P_\sigma(i)$는 결국 2.2의 합에서 i를 포함하지 않는 연합 S를 대표하게 되고, 그러한 모든 S에 대해 $v(S \cup \{i\}) - v(S)$를 더하는 것으로 생각할 수 있다. $P_\sigma(i) = S$가 되기 위한 순열 σ의 수는 i 앞의 S를 나열하는 경우의 수 $|S|!$와 i 뒤의 남은 참가자를 나열하는 경우의 수 $(n-1-|S|)!$의 곱이다.

$$\underbrace{\boxed{\quad S \quad}}\ \underbrace{\boxed{i}}\ \underbrace{\boxed{\quad N \setminus (S \cup i) \quad}}$$

S를 먼저 나열하고, i를 배치한 후, 남은 참가자를 배치한다.

그러므로 식 2.2는

$$\begin{aligned} \Phi_i(v) &= \sum_{S \subseteq N : i \notin S} \frac{|S|!(n-1-|S|)!}{n!} [v(S \cup i) - v(S)] \\ &= \sum_{S \subseteq N : i \notin S} \frac{1}{n} \binom{n-1}{|S|}^{-1} [v(S \cup i) - v(S)] \end{aligned} \tag{2.3}$$

로 다시 쓸 수 있다. 식 2.3은 한계 기여도의 기댓값으로 정의된 섀플리 값의 새로운 확률론적 해석을 제시한다. 특정 참가자 i를 포함하지 않는 연합 S를 다음 과정을 따라 구성하자.

1. 연합 S의 크기 $|S|$를 숫자 $0, 1, 2, \cdots, n-1$를 선택하여 정한다. (균등 확률 $1/n$)
2. i를 제외한 참가자의 집합 $N \setminus \{i\}$에서 S에 속할 인원을 정한다. (균등 확률 $\binom{n-1}{|S|}^{-1}$)

이렇게 택한 연합 S에 대하여 i에게 $v(S\cup\{i\})-v(S)$ 만큼의 보수를 제시하자. 이 확률 과정에서 i가 받게 되는 보수의 기댓값은 정확히 섀플리 값과 일치한다.

연습문제 2.4

연습문제 1.28의 3인 협력 게임의 섀플리 값을 구하여라.

풀이

순열 $\sigma \in \Pi(\{1,2,3\}$에 따른 한계 기여도 벡터를 계산하자.

σ	$1\to2\to3$	$1\to3\to2$	$2\to1\to3$	$2\to3\to1$	$3\to1\to2$	$3\to2\to1$
m_1^σ	0	0	1	0	0	0
m_2^σ	1	1	0	0	1	1
m_3^σ	0	0	0	1	0	0

따라서 구하는 섀플리 값은 $(1/6, 2/3, 1/6)$이다.

연습문제 2.5

N의 공집합이 아닌 부분집합 T에 대하여 T**-표준 게임(**T**-standard game)** 1_T를 다음과 같이 정의하자.

$$1_T(S)=\begin{cases}1 & (S=T)\\ 0 & (\text{그 외의 경우})\end{cases}$$

섀플리 값 $\Phi(1_T)$을 구하여라.

풀이

식 2.3을 이용하자. 먼저, $i\in T$이면 $S=T\setminus\{i\}$일 때에만 $1_T(S\cup i)=1$이므로

$$\Phi_i(1_T)=\sum_{S\not\ni i}\frac{|S|!(n-1-|S|)!}{n!}[1_T(S\cup i)-1_T(S)]=\frac{(|T|-1)!(n-|T|)!}{n!}$$

이다. 역으로 $i\notin T$이면 $S=T$일 때에만 $1_T(S)=1$이므로

$$\Phi_i(1_T)=\sum_{S\not\ni i}\frac{|S|!(n-1-|S|)!}{n!}[1_T(S\cup i)-1_T(S)]=-\frac{|T|!(n-1-|T|)!}{n!}$$

이다. 섀플리 값은 T에 속하지 않은 참가자에게 음의 보수를 제시한다. □

2.2 섀플리 값의 성질과 특정 (1)

여기서는 단일점 해가 가졌으면 하는 좋은 성질(공리)을 제시하고, 섀플리 값이 몇몇 공리를 동시에 만족하는 유일한 단일점 해임을 보인다. $\psi:\mathcal{G}^N\to\mathbb{R}^N$을 $\mathcal{G}^N$ 위의 값이라 하자. 효율성은 할당의 정의에서 이미 논하였다.

- **효율성(efficiency, EFF)**: 모든 $v\in\mathcal{G}^N$에 대하여 $\sum_{i=1}^n \psi_i(v)=v(N)$이다.

효율성은 다른 문헌에서 **파레토 최적성(Pareto optimality)** 또는 **파레토 효율성(Pareto efficiency)** 등으로 불린다.

모든 연합 $S\subseteq N$에 대하여 $v(S\cup i)-v(S)=0$을 만족하는 참가자 $i\in N$를 **무임승차자(null-player)**라 하자. 무임승차자는 어떤 연합에 합류시켜도 추가적인 기여를 하지 않는다. 그러므로 우리가 원하는 단일점 해는 무임승차자에게 보수를 지급하지 않아야 한다.

- **무임승차자 배제(null-player property, NP)**: 모든 $v\in\mathcal{G}^N$의 무임승차자 $i\in N$에 대하여 $\psi_i(v)=0$이다.

두 참가자 i와 j가 모든 연합 $S \subseteq N \setminus \{i, j\}$에 대하여 $v(S \cup i) = v(S \cup j)$를 만족하면 i와 j가 **대칭(symmetric)**이라 하자. 즉 i와 j가 대칭이면 한 연합에 각각 합류할 때의 기여도는 서로 같다. 대칭인 두 참가자에게 같은 보수를 지급함이 바람직하다.

- **대칭성(symmetry, SYM)**: 모든 $v \in \mathcal{G}^N$의 대칭인 두 참가자 $i, j \in N$에 대하여 $\psi_i(v) = \psi_j(v)$이다.

마지막 성질은 둘 이상의 협력 게임을 진행할 때를 위한 성질이다. 가령 어느 하루의 오전에 협력 게임 (N, v_1)을 진행하고, 오후에 다른 협력 게임 (N, v_2)를 진행한다고 하자. 각각의 게임을 통해 참가자들이 얻는 보수는 $\psi(v_1) + \psi(w_2)$이다. 이는 협력 게임 $(N, v_1 + v_2)$를 하루에 진행할 때의 보수 $\psi(v_1 + v_2)$와 일치하는 것이 합리적이다.

- **가법성(additivity, ADD)**: 모든 $v_1, v_2 \in \mathcal{G}^N$에 대하여 $\psi(v_1 + v_2) = \psi(v) + \psi(w)$이다.

이제 섀플리 값이 위의 네 성질을 만족함을 보이자.

정리 2.6

섀플리 값 Φ는 EFF, NP, SYM, ADD를 만족한다.

증명. $N = \{1, 2, \cdots, n\}$, $(N, v) \in \mathcal{G}^N$이라 하자. $P_\sigma(\sigma(i)) = \{j \in N \mid \sigma^{-1}(j) < i\}$이므로

$$\begin{aligned} P_\sigma(\sigma(i+1)) &= \{j \in N \mid \sigma^{-1}(j) < i + 1\} \\ &= \{j \in N \mid \sigma^{-1}(j) < i \text{ 또는 } \sigma^{-1}(j) = i\} \\ &= P_\sigma(\sigma(i)) \cup \{\sigma(i)\} \end{aligned}$$

이다. $\sigma : N \to N$이 일대일대응임을 이용하면

$$\begin{aligned} \sum_{i=1}^n \Phi_i(v) &= \sum_{i=1}^n \frac{1}{n!} \sum_{\sigma \in \Pi(N)} [v(P_\sigma(i) \cup i) - v(P_\sigma(i))] \\ &= \frac{1}{n!} \sum_{\sigma \in \Pi(N)} \sum_{i=1}^n [v(P_\sigma(i) \cup i) - v(P_\sigma(i))] \\ &= \frac{1}{n!} \sum_{\sigma \in \Pi(N)} \sum_{i=1}^n [v(P_\sigma(\sigma(i)) \cup \sigma(i)) - v(P_\sigma(\sigma(i)))] \\ &= \frac{1}{n!} \sum_{\sigma \in \Pi(N)} \sum_{i=1}^n [v(P_\sigma(\sigma(i+1))) - v(P_\sigma(\sigma(i)))] \\ &= \frac{1}{n!} \sum_{\sigma \in \Pi(N)} [v(N) - v(\varnothing)] = v(N) \end{aligned}$$

이므로 Φ는 EFF를 만족한다. 무임승차자 $i \in N$에 대해 $v(P_\sigma(i) \cup \{i\}) - v(P_\sigma(i)) = 0$이므로 Φ는 NP를 만족한다. 이제 $i, j \in N$가 대칭이라 하자. 순열 $\sigma \in \Pi(N)$에 대하여 i와 j의 합류 순서를 바꾼 순열을 σ'이라 하자. 즉 i와 j의 자리를 바꾸는 호환(transposition) τ에 대하여 $\sigma' = \tau\sigma\tau^{-1}$이다. $P_\sigma(i) = P_{\sigma'}(j)$이고 두 집합은 i와 j를 포함하지 않으므로 대칭성에 의해

$$\begin{aligned} m_i^\sigma - m_j^{\sigma'} &= [v(P_\sigma(i) \cup i) - v(P_\sigma(i))] - [v(P_{\sigma'}(j) \cup j) - v(P_{\sigma'}(j))] \\ &= v(P_\sigma(i) \cup i) - v(P_{\sigma'}(j) \cup j) = 0 \end{aligned}$$

이고, 그러므로 Φ는 SYM를 만족한다. 마지막으로 $v = v_1 + v_2$라 하면

$$v(S \cup i) - v(S) = [v_1(S \cup i) - v_1(S)] + [v_2(S \cup i) - v_2(S)]$$

이므로 Φ는 ADD를 만족한다. □

사실, 섀플리 값은 EFF, NP, SYM, ADD를 만족하는 유일한 단일점 해이다. 이를 보이기 위해 $\mathcal{G}^N$의 벡터 공간으로서의 성질을 먼저 보이자. 연합 $T \in 2^N \setminus \{\varnothing\}$에 대하여 연습문제 2.5의 표준 게임 1_T를 고려하자. 임의의 $v \in \mathcal{G}^N$에 대하여 $v = \sum_{T \neq \varnothing} v(T) 1_T$이므로 집합 $\mathcal{B}_1 = \{1_T \mid T \in 2^N \setminus \{\varnothing\}\}$은 $\mathcal{G}^N$의 기저이다. 이 기저는 간단하지만 섀플리 값의 성질을 규명하는데 도움이 되지 않는다. 1_T에는 무임승차자가 없기 때문이다. 그 대신 T–만장일치 게임의 모임 $\mathcal{B}_u = \{u_T \mid T \in 2^N \setminus \{\varnothing\}\}$을 고려하자. (예시 1.22 참고) $|\mathcal{B}_1| = |\mathcal{B}_2| = 2^{|N|} - 1$이므로 $\mathcal{B}_u$가 $\mathcal{G}^N$의 기저임을 보이기 위해서는 $\mathcal{B}_u$가 선형 독립(linearly independent)임을 보이면 충분하다. $\sum_{T \neq \varnothing} \beta_T u_T = 0$이라 하고, 어떤 S에 대하여 $\beta_S \neq 0$이라 하자. 그러한 S 중에서 모든 $T \subsetneq S$에 대하여 $\beta_T = 0$인 S를 택하면 $\sum_{T \neq \varnothing} \beta_T u_T(S) = \beta_S \neq 0$이므로 모순이다. 따라서 $\mathcal{B}_u$는 선형 독립이다. 이제 다음을 보이자.

정리 2.7

$\mathcal{G}^N$ 위의 값 ψ가 EFF, NP, SYM, ADD를 모두 만족할 필요충분조건은 $\psi = \Phi$이다.

증명. 정리 2.6에 의해 ψ가 네 성질을 만족한다고 가정하고 $\psi = \Phi$임을 보이면 충분하다. 임의의 협력 게임 $v \in \mathcal{G}^N$는 T–만장일치 게임의 선형 결합으로 (유일하게) 표현할 수 있고, ψ와 Φ 모두 ADD를 만족하므로 모든 $T \in 2^N \setminus \{\varnothing\}$과 $c \in \mathbb{R}$에 대하여 $\psi(cu_T) = \Phi(cu_T)$임을 보이면 충분하다. 임의의 참가자 $i \in N \setminus T$와 $S \subseteq N$에 대하여 $cu_T(S \cup i) - cu_T(S) = 0$이므로 i는 무임승차자이다. ψ와 ϕ는 NP를 만족하므로 모든 $i \in N \setminus T$에 대하여

$$\psi_i(cu_T) = \Phi_i(cu_T) = 0 \tag{2.4}$$

이다. 이제 i와 j를 T에 속하는 서로 다른 참가자라 하자. $\{i, j\}$를 포함하지 않는 연합 S에 대하여

$$cu_T(S \cup i) = cu_T(S \cup j) = 0$$

이므로 i와 j는 cu_T에서 대칭이다. ψ와 Φ는 SYM을 만족하므로 모든 $i, j \in T$에 대하여

$$\psi_i(cu_T) = \psi_j(cu_T), \quad \Phi_i(cu_T) = \Phi_j(cu_T)$$

이다. 따라서 EFF로부터 모든 $i \in T$에 대하여

$$\psi_i(cu_T) = \Phi_i(cu_T) = \frac{c}{|T|} \tag{2.5}$$

이므로 $\psi = \Phi$이다. □

섀플리 값이 갖는 성질을 더 알아보자. 협력 게임 (N, v)의 참가자 i가 모든 $S \subseteq N \setminus \{i\}$에 대하여 $v(S \cup i) - v(S) = v(i)$을 만족하면 i를 **더미(dummy player)**라 하자. 더미 i는 연합에 참가하더라도 추가적인 시너지 효과를 내지 못하므로 기여분 $v(i)$의 보수를 받는 것이 이상적이다.

- **더미 성질(dummy player property, DUM)**: 모든 $v \in \mathcal{G}^N$의 더미 $i \in N$에 대하여 $\psi_i(v) = v(i)$이다.

DUM은 NP를 함의한다. 다음 연습문제를 참고하라.

연습문제 2.8

DUM $\Longrightarrow$ NP가 성립함을 보여라.

풀이

무임승차자 i는 $v(i) = 0$인 더미이므로 DUM은 NP를 함의한다. □

정리 2.9

섀플리 값 Φ는 DUM을 만족한다.

증명. $i \in N$을 협력 게임 (N, v)의 더미라 하면 $P_\sigma(i)$는 i를 포함하지 않으므로

$$\Phi_i(v) = \frac{1}{n!} \sum_{\sigma \in \Pi(N)} [v(P_\sigma(i) \cup i) - v(P_\sigma(i))] = \frac{1}{n!} \sum_{\sigma \in \Pi(N)} v(i) = v(i)$$

이므로 섀플리 값은 DUM을 만족한다. $\square$

참고. 식 2.3을 이용하면

$$\Phi_i(v) = \sum_{S \subseteq N \setminus \{i\}} \frac{1}{n} \binom{n-1}{|S|}^{-1} [v(S \cup i) - v(S)] = \frac{v(i)}{n} \sum_{S \subseteq N \setminus \{i\}} \binom{n-1}{|S|}^{-1} = v(i)$$

이므로 같은 결과를 얻는다. (i를 포함하지 않으면서 크기가 k인 연합의 개수가 $\binom{n-1}{k}$임을 이용하라.)

순열 $\sigma \in \Pi(N)$에 대하여 협력 게임 $v^\sigma \in \mathcal{G}^N$을

$$v^\sigma(S) = v(\sigma^{-1}(S)),$$

함수 $\sigma^* : \mathbb{R}^N \to \mathbb{R}^N$를 $\sigma_i^*(x) = x_{\sigma^{-1}(i)}$로 정의하자.

- **익명성(anonimity, AN)**: 모든 $v \in \mathcal{G}^N$와 $\sigma \in \Pi(N)$에 대하여 $\psi(v^\sigma) = \sigma^*(\psi(v))$이다.

순열 $\sigma : N \to N$이 참가자의 이름(또는 번호)을 바꾼다고 생각하자. 협력 게임 v^σ에서의 연합 S와 협력 게임 v에서의 연합 $\sigma^{-1}(S)$는 구성원의 이름만 다를 뿐 서로 같다. 따라서 v^σ는 연합 S에게 $v(\sigma^{-1}(S))$ 만큼의 가치를 책정한다. $\mathcal{G}^N$ 위의 값 ψ가 익명성을 가지면 각 $i \in N$에 대하여 $\psi_i(v^\sigma) = \sigma_i^*(\psi(v)) = \psi_{\sigma^{-1}(i)}(v)$이므로 참가자 i는 v^σ에서 과거의 이름 $\sigma^{-1}(i)$으로서 v에 대한 보수를 받는다. 즉 익명성은 참가자를 이름, 번호만으로 차별하지 아니함을 의미한다.

AN은 SYM을 함의한다. 다음 연습문제를 참고하라.

연습문제 2.10

AN $\Longrightarrow$ SYM가 성립함을 보여라.

풀이

ψ를 AN을 만족하는 $\mathcal{G}^N$ 위의 값이라 하자. 두 참가자 i와 j가 서로 대칭이라 하고, 순열 σ를 i와 j를 교환하는 호환이라 하자. 연합 S가 $S \subset N \setminus \{i, j\}$이거나 $\{i, j\} \subseteq S$를 만족하면 $\sigma^{-1}(S) = S$이므로 $v^\sigma(S) = v(\sigma^{-1}(S)) = v(S)$이다. 만약 $i \in S$이고 $j \notin S$이면 $S \setminus \{i\} \subseteq N \setminus \{i, j\}$이므로

$$v^\sigma(S) = v(\sigma^{-1}(S)) = v(S \setminus \{i\} \cup \{j\}) = v(S \setminus \{i\} \cup \{i\}) = v(S)$$

이다. 같은 논리로 $i \notin S$이고 $j \in S$이면 $v^\sigma(S) = v(S)$이다. 즉 $v^\sigma = v$이므로

$$\psi_i(v) = \psi_{\sigma^{-1}(j)}(v) = \psi_j(v^\sigma) = \psi_j(v)$$

이 되어 ψ는 SYM을 만족한다. $\square$

정리 2.11

섀플리 값 Φ는 AN을 만족한다.

증명. 먼저 모든 $v \in \mathcal{G}^N$과 $\tau, \sigma \in \Pi(N)$에 대하여

$$\tau^*(m^\sigma(v)) = m^{\tau\sigma}(v^\tau)$$

임은 다음 계산을 통해 확인할 수 있다.

$$\begin{aligned}
\tau^*(m^\sigma(v))_{\tau\sigma(i)} &= m^\sigma(v)_{\sigma(i)} \\
&= v(P_\sigma(\sigma(i)) \cup \sigma(i)) - v(P_\sigma(i)) \\
&= v(\sigma(\{1, 2, \cdots, i\})) - v(\sigma(\{1, 2, \cdots, i-1\})) \\
&= v(\tau^{-1}(\tau\sigma(\{1, 2, \cdots, i\}))) - v(\tau^{-1}(\tau\sigma(\{1, 2, \cdots, i-1\}))) \\
&= v^\tau(\tau\sigma(\{1, 2, \cdots, i\})) - v^\tau(\tau\sigma(\{1, 2, \cdots, i-1\})) \\
&= v^\tau(P_{\tau\sigma}(\tau\sigma(i)) \cup \tau\sigma(i)) - v^\tau(P_{\tau\sigma}(\tau\sigma(i))) \\
&= m^{\tau\sigma}(v^\tau)_{\tau\sigma(i)}
\end{aligned}$$

이제 $v \in \mathcal{G}^N$, $\tau \in \Pi(N)$이라 하자. 사상 $\tau \mapsto \tau\sigma$는 $\Pi(N)$ 위의 전사 사상이므로

$$\Phi(v^\tau) = \frac{1}{n!} \sum_{\sigma \in \Pi(N)} m^\sigma(v^\tau) = \frac{1}{n!} \sum_\sigma m^{\tau\sigma}(v^\tau) = \frac{1}{n!} \sum_\sigma \tau^*(m^\sigma(v)) = \tau^*\left(\frac{1}{n!} \sum_\sigma m^\sigma\right) = \tau^*(\Phi(v))$$

이 성립한다. 따라서 Φ는 AN을 만족한다. □

연습문제 2.12

이 연습문제에서는 조건 EFF, DUM, SYM, ADD 중 하나라도 누락되면 정리 2.7이 성립하지 않음을 보인다. $(N, v) \in \mathcal{G}^N$이라 하자.

(a) 각 $i \in N$에 대하여

$$\psi_i(v) = v(i)$$

라 하면 ψ는 DUM, SYM, ADD를 만족하지만 EFF는 만족하지 않음을 보여라.

(b) 각 $i \in N$에 대하여

$$\psi_i(v) = \sum_{\sigma : \sigma(1) = 1} \frac{1}{(|N| - 1)!} m_i^\sigma$$

라 하면 ψ는 EFF, DUM, ADD를 만족하지만 SYM은 만족하지 않음을 보여라.

(c) 각 $i \in N$에 대하여

$$\psi_i(v) = \frac{v(N)}{|N|}$$

이라 하면 ψ는 EFF, SYM, ADD를 만족하지만 DUM은 만족하지 않음을 보여라.

(d) $D(v)$를 v의 모든 더미의 집합이라 하자. 각 $i \in N$에 대하여

$$\psi_i(v) = \begin{cases} v(i) & (i \in D(v)) \\ \dfrac{1}{|N \setminus D(v)|}\left(v(N) - \displaystyle\sum_{j \in D(v)} v(j)\right) & (i \in N \setminus D(v)) \end{cases}$$

라 하면 ψ는 EFF, DUM, SYM을 만족하지만 ADD는 만족하지 않음을 보여라.

풀이

(a) ψ가 DUM, ADD를 만족함은 자명하다. 두 참가자 $i, j \in N$가 대칭이면 $v(i) = v(j)$이므로 ψ는 SYM을 만족한다. (정의에 $S = \varnothing$을 대입하라.) ψ가 EFF를 만족하지 않음은 표준 게임 1_N으로 보일 수 있다:

$$\sum_{i\in N} \psi_i(v) = \sum_{i\in N} v(i) = 0 \neq 1 = v(N).$$

(b) ψ가 EFF, DUM, ADD를 만족함은 정리 2.6, 2.9와 같은 방법으로 확인할 수 있다. ψ가 SYM을 만족하지 않음을 보이기 위해 $N = \{1,2,3\}$, $T = \{1,2\} \subseteq N$이라 하고 $(N, 1_T)$를 T-표준 게임이라 하자.

$$v(\varnothing \cup i) = v(\varnothing \cup j) = 0, \quad v(1,3) = v(2,3) = 0$$

이므로 두 참가자 1과 2는 대칭이다. 그러나 $\psi_1(1_T) = 0$, $\psi_2(1_T) = 1/2$이므로 ψ는 SYM을 만족하지 않는다.

(c) ψ가 EFF, SYM, ADD를 만족함은 자명하다. 연습문제 2.8로부터 ψ가 NP를 만족하지 않음을 보이면 충분하다. $N = \{1,2,3\}$, $T = \{1,2\} \subseteq N$에 대하여 (N, u_T)를 T-만장일치 게임이라 하자. 모든 $S \subseteq \{1,2\}$에 대하여 $v(S \cup 3) - v(S) = 0$이므로 참가자 3은 무임승차자이다. 그러나 $\psi_3(u_T) = 1/3$이므로 ψ는 DUM을 만족하지 않는다.

(d) ψ가 EFF, DUM을 만족함은 자명하다. ψ가 SYM을 만족함은 다음을 보이면 충분하다.

보조정리

두 참가자가 대칭이면, 둘은 모두 더미이거나 모두 더미가 아니다.

대칭인 두 참가자 $i, j \in N$에 대하여 $i \in D(v)$라 하자. i와 j의 대칭성에 의해 $v(i) = v(j)$이다. 임의의 $S \subseteq N \setminus \{j\}$에 대하여 $S \setminus \{i\} \subseteq N \setminus \{i,j\}$이므로

$$\begin{aligned} v(S \cup j) &= v([S \setminus \{i\} \cup i] \cup j) = v([S \setminus \{i\} \cup j] \cup i) \\ &= v(S \setminus \{i\} \cup j) + v(i) = v(S \setminus \{i\} \cup i) + v(j) \\ &= v(S) + v(j) \end{aligned}$$

이 성립한다. 즉 $j \in D(v)$이다. 이제 ψ가 ADD를 만족하지 않음을 보이자. $N = \{1,2,3\}$에 대하여 $u_{\{1,2\}}$와 $u_{\{2,3\}}$을 각각 만장일치 게임이라 하고, $u = u_{\{1,2\}} + u_{\{2,3\}}$이라 하자.

$$D(u_{\{1,2\}}) = \{3\}, \quad D(u_{\{2,3\}}) = \{1\}, \quad D(u) = \varnothing$$

이므로 $\psi(u_{\{1,2\}})$, $\psi(u_{\{2,3\}})$, $\psi(u)$는 다음과 같다.

$v \in \mathcal{G}^N$	$\psi_1(v)$	$\psi_2(v)$	$\psi_2(v)$
$u_{\{1,2\}}$	$1/2$	$1/2$	0
$u_{\{2,3\}}$	0	$1/2$	$1/2$
u	$2/3$	$2/3$	$2/3$

따라서 ψ는 ADD를 만족하지 않는다.

지금까지 섀플리 값이 갖는 좋은 성질을 확인했고, 아래에서 더 다양한 접근 방식으로 섀플리 값을 탐구할 것이다. 하지만 연습문제 2.4와 2.5에서 알 수 있듯이 섀플리 값은 코어에 속할 필요도 없고, 개인에게 합리적일 필요도 없다는 단점이 있다.

References

[1] Peters, Hans. Game Theory: A Multi-Leveled Approach. 2nd ed., Springer (2015). https://doi.org/10.1007/978-3-662-46950-7

[2] Osborne, Martin J., Ariel Rubinstein. A Course in Game Theory. MIT Press (1994).

[3] Rafels, C., Tijs, S. On the Cores of cooperative games and the stability of the Weber set. Int J Game Theory 26, 491–499 (1997). https://doi.org/10.1007/BF01813887