Equivalent Representations of a Set Function with Applications to Game Theory and Multicriteria Decision Making

by

Jean-Luc MARICHAL
University of Liège, Belgium

Michel GRABISCH
Thomson-CSF, Orsay, France

Marc ROUBENS
University of Liège, Belgium

Set functions in game theory and multicriteria decision making

Consider a set function $v : 2^N \longrightarrow \mathbb{R}$, where $N = \{1, \ldots, n\}$.

- Game theory: $v(\emptyset) = 0$

  N is a set of players

  v is the characteristic function of a game

  $v(S)$ = worth or power of $S \subseteq N$.

- Multicriteria decision making:

  $v(\emptyset) = 0$, $v(N) = 1$, $S \subseteq T \Rightarrow v(S) \leq v(T)$

  N is a set of criteria

  $v(S)$ = weight of importance of $S \subseteq N$.

Any real valued set function v can be assimilated unambiguously with a pseudo-Boolean function $f : \{0,1\}^n \longrightarrow \mathbb{R}$

$$v \quad \longleftrightarrow \quad f$$

If e_S represents the characteristic vector of S in $\{0,1\}^n$, we have

$$v(S) = f(e_S), \quad S \subseteq N.$$

Theorem (Hammer and Rudeanu, 1968)

Any pseudo-Boolean function $f : \{0,1\}^n \to \mathbb{R}$ has a unique expression as a multilinear polynomial in n variables:

$$f(x) = \sum_{T \subseteq N} a(T) \prod_{i \in T} x_i, \quad x \in \{0,1\}^n,$$

where $a(T) \in \mathbb{R}$.

$$v(S) = \sum_{T \subseteq S} a(T), \quad S \subseteq N.$$

In combinatorics, a viewed as a set function on N is called the Möbius transform of v.

$$a(S) = \sum_{T \subseteq S} (-1)^{s-t} v(T), \quad S \subseteq N$$

where $s = |S|$ and $t = |T|$.

a is a representation of v, and conversely : defining one of the two allows to compute the other without ambiguity.

Definition

A set function $w : 2^N \to \mathbb{R}$ is a representation of v if there exists an invertible transform $\mathcal{T}$ such that

$$w = \mathcal{T}(v) \quad \text{and} \quad v = \mathcal{T}^{-1}(w).$$

Banzhaf and Shapley power indices

Given $i \in N$, it may happen that

- $v(i) = 0$,
- $v(T \cup i) \gg v(T)$ for many $T \subseteq N \setminus i$

The overall importance of $i \in N$ should not be solely determined by $v(i)$, but also by all $v(T \cup i)$ such that $T \subseteq N \setminus i$.

The marginal contribution of i in coalition $T \subseteq N \setminus i$ is defined by

$$v(T \cup i) - v(T)$$

A *power index* for i is given by an average value of the marginal contributions of i alone in all coalitions.

- The Banzhaf power index (1965) :

$$\begin{aligned} \phi_{\mathrm{B}}^v(i) &:= \frac{1}{2^{n-1}} \sum_{T \subseteq N \setminus i} [v(T \cup i) - v(T)] \\ &= \sum_{T \ni i} \frac{1}{2^{t-1}}\, a(T). \end{aligned}$$

- The Shapley power index (1953) :

$$\begin{aligned} \phi_{\mathrm{Sh}}^v(i) &:= \frac{1}{n} \sum_{t=0}^{n-1} \frac{1}{\binom{n-1}{t}} \sum_{\substack{T \subseteq N \setminus i \\ |T| = t}} [v(T \cup i) - v(T)] \\ &= \sum_{T \subseteq N \setminus i} \frac{(n-t-1)!\, t!}{n!} [v(T \cup i) - v(T)] \\ &= \sum_{T \ni i} \frac{1}{t}\, a(T). \end{aligned}$$

Banzhaf and Shapley interaction indices

The difference

$$a(ij) = v(ij) - v(i) - v(j)$$

seems to reflect a degree of interaction between i and j :

- positive interaction $a(ij) > 0$: i and j should cooperate
- negative interaction $a(ij) < 0$: i and j should not cooperate
- no interaction $a(ij) = 0$: i and j can act independently

As for power indices, an interaction index should consider not only $v(i), v(j), v(ij)$ but also the worths of all subsets not containing i and j. We may say that i and j have interest to cooperate if

$$v(T \cup ij) - v(T \cup i) > v(T \cup j) - v(T), \quad T \subseteq N \setminus ij.$$

i and j can act independently in case of $=$ and have no interest to cooperate in case of $<$.

An *interaction* index for the pair $i, j \in N$ is given by the average of the marginal interaction between i and j, conditioned to the presence of elements of the coalition $T \subseteq N \setminus ij$: (Murofushi and Soneda, 1993)

$$I_{\mathrm{B}}^v(ij) := \frac{1}{2^{n-2}} \sum_{T \subseteq N \setminus ij} [v(T \cup ij) - v(T \cup i) - v(T \cup j) + v(T)]$$

$$I_{\mathrm{Sh}}^v(ij) := \sum_{T \subseteq N \setminus ij} \frac{(n-t-2)!\, t!}{(n-1)!} [v(T \cup ij) - v(T \cup i) - v(T \cup j) + v(T)].$$

Interaction indices among a combination S of players or criteria has been introduced by Grabisch and Roubens (1998) as natural extensions of the case $|S| = 2$:

- The *Banzhaf interaction index* :

$$I_{\rm B}^v(S) := \frac{1}{2^{n-s}} \sum_{T\subseteq N\setminus S} \sum_{L\subseteq S} (-1)^{s-l} v(L\cup T), \quad S\subseteq N$$

- The *Shapley interaction index* :

$$I_{\rm Sh}^v(S) := \sum_{T\subseteq N\setminus S} \frac{(n-t-s)!\, t!}{(n-s+1)!} \sum_{L\subseteq S} (-1)^{s-l} v(L\cup T), \quad S\subseteq N$$

(characterized by Grabisch and Roubens, 1998)

We have

$$\begin{aligned} I_{\rm B}^v(S) &= \sum_{T\supseteq S} (\frac{1}{2})^{t-s}\, a(T) \\ I_{\rm Sh}^v(S) &= \sum_{T\supseteq S} \frac{1}{t-s+1}\, a(T) \end{aligned}$$

and

$$\phi_{\rm B}^v(i) = I_{\rm B}^v(i) \quad \text{and} \quad \phi_{\rm Sh}^v(i) = I_{\rm Sh}^v(i)$$

for all $i \in N$.

Theorem

The interaction indices $I_{\rm B}$ and $I_{\rm Sh}$, viewed as set functions from 2^N to $\mathbb{R}$ are equivalent representations of v.

Multilinear extension of pseudo-Boolean functions

Let $S \subseteq N$. The S-derivative of the pseudo-Boolean function f at $x \in \{0,1\}^n$, denoted $\Delta_S\, f(x)$, is defined inductively as

$$\begin{aligned}
\Delta_i\, f(x) &:= f(x\,|\,x_i=1) - f(x\,|\,x_i=0), \\
\Delta_{ij}\, f(x) &:= \Delta_i(\Delta_j\, f)(x) = \Delta_j(\Delta_i\, f)(x), \\
&\;\;\vdots \\
\Delta_S\, f(x) &:= \Delta_i(\Delta_{S\setminus i}\, f)(x)
\end{aligned}$$

Definition The multilinear extension (MLE) of a pseudo-Boolean function f (or a game v) is defined by

$$g(x) := \sum_{T\subseteq N} a(T) \prod_{i\in T} x_i, \quad x \in [0,1]^n$$

see Hammer and Rudeanu (1968) and Owen (1972).

The S-derivative of g is defined inductively in the same way as for f :

$$\Delta_S\, g(x) = \sum_{T\supseteq S} a(T) \prod_{i\in T\setminus S} x_i, \quad x \in [0,1]^n$$

In particular, we have

$$I_{\mathrm{B}}(S) = \sum_{T\supseteq S} \frac{1}{2^{t-s}}\, a(T) = \int_{[0,1]^n} (\Delta_S\, g)(x)\, dx.$$

Note that it has been proved that

$$I_{\mathrm{B}}(S) = \frac{1}{2^n} \sum_{x\in\{0,1\}^n} (\Delta_S\, f)(x).$$

Recall that

$$\Delta_S\, g(x) = \sum_{T \supseteq S} a(T) \prod_{i \in T \setminus S} x_i, \quad x \in [0,1]^n$$

Setting $\underline{x} := (x, \ldots, x)$, we have

$$\Delta_S\, g(\underline{x}) = \sum_{T \supseteq S} a(T)\, x^{t-s}, \quad x \in [0,1]^n$$

Consequently, we have, for all $S \subseteq N$,

$$\begin{aligned} a(S) &= (\Delta_S\, g)(\underline{0}) \\ I_{\mathrm{B}}(S) &= \sum_{T \supseteq S} \frac{1}{2^{t-s}}\, a(T) = (\Delta_S\, g)(\underline{1/2}) \\ I_{\mathrm{Sh}}(S) &= \sum_{T \supseteq S} \frac{1}{t-s+1}\, a(T) = \int_0^1 (\Delta_S\, g)(\underline{x})\, dx. \end{aligned}$$

Some conversion formulas derived from the MLE

For the multilinear extension g, the operator Δ_S identifies with the classical S-derivative, that is

$$\Delta_S\, g(x) = \frac{\partial^s\, g(x)}{\partial x_{i_1} \cdots \partial x_{i_s}} \quad \text{where } S = \{i_1, \ldots, i_s\}.$$

The Taylor formula for functions of several variables then can be applied to g. This leads to the equality:

$$g(x) = \sum_{T \subseteq N} \prod_{i \in T} (x_i - y_i)\, \Delta_T\, g(y), \quad x, y \in [0,1]^n. \qquad (*)$$

Replacing x by e_S ($S \subseteq N$) provides:

$$v(S) = \sum_{T \subseteq N} \prod_{i \in T} ((e_S)_i - y)\, (\Delta_T\, g)(\underline{y}), \quad y \in [0,1].$$

$y \to 0, 1/2$: passages from a and I_{B} to v

By successive derivations of $(*)$, we obtain:

$$\Delta_S\, g(x) = \sum_{T \supseteq S} \prod_{i \in T \setminus S} (x_i - y_i)\, \Delta_T g(y), \quad \forall x, y \in [0,1]^n.$$

In particular,

$$(\Delta_S\, g)(\underline{x}) = \sum_{T \supseteq S} (x - y)^{t-s}\, (\Delta_T g)(\underline{y}), \quad \forall x, y \in [0,1].$$

$x, y \to 0, 1/2$: passages between a and I_{B}

Recall that

$$(\Delta_S\, g)(\underline{x}) = \sum_{T\supseteq S} (x-y)^{t-s}\, (\Delta_T g)(\underline{y}), \quad \forall x, y \in [0,1].$$

We then have

$$\begin{aligned} I_{\mathrm{Sh}}(S) &= \int_0^1 (\Delta_S\, g)(\underline{x})\, dx \\ &= \sum_{T\supseteq S} \left[\int_0^1 (x-y)^{t-s}\, dx\right] (\Delta_T g)(\underline{y}) \\ &= \sum_{T\supseteq S} \frac{(1-y)^{t-s+1} - (-y)^{t-s+1}}{t-s+1}\, (\Delta_T g)(\underline{y}) \end{aligned}$$

$y \to 0, 1/2$: passages from a and I_{B} to I_{Sh}

Let $\{B_n\}_{n\in\mathbb{N}}$ be the sequence of Bernoulli numbers:

$$B_0 = 1, B_1 = -\frac{1}{2}, B_2 = \frac{1}{6}, B_3 = 0, B_4 = -\frac{1}{30}, B_5 = 0, B_6 = \frac{1}{42}, \ldots$$

and define the Bernoulli polynomials by

$$B_n(x) = \sum_{k=0}^{n} \binom{n}{k} B_k\, x^{n-k}, \quad \forall n \in \mathbb{N},\ \forall x \in \mathbb{R}.$$

Theorem We have

$$(\Delta_S\, g)(\underline{x}) = \sum_{T\supseteq S} B_{t-s}(x)\, I_{\mathrm{Sh}}(T), \quad \forall x \in [0,1].$$

$x \to 0, 1/2$: passages from I_{Sh} to a and I_{B}

Conversion formulas

$$
\begin{aligned}
v(S) &= \sum_{T\subseteq S} a(T) \\
v(S) &= \sum_{T\subseteq N} (\frac{1}{2})^t (-1)^{|T\setminus S|} I_{\mathrm{B}}(T)
\end{aligned}
$$

$$
\begin{aligned}
a(S) &= \sum_{T\subseteq S} (-1)^{s-t} v(T) \\
a(S) &= \sum_{T\supseteq S} (-\frac{1}{2})^{t-s} I_{\mathrm{B}}(T) \\
a(S) &= \sum_{T\supseteq S} B_{t-s}\, I_{\mathrm{Sh}}(T)
\end{aligned}
$$

$$
\begin{aligned}
I_{\mathrm{B}}(S) &= \sum_{T\supseteq S} (\frac{1}{2})^{t-s} a(T) \\
I_{\mathrm{B}}(S) &= \sum_{T\supseteq S} (\frac{1}{2^{t-s-1}} - 1) B_{t-s}\, I_{\mathrm{Sh}}(T)
\end{aligned}
$$

$$
\begin{aligned}
I_{\mathrm{Sh}}(S) &= \sum_{T\supseteq S} \frac{1}{t-s+1}\, a(T) \\
I_{\mathrm{Sh}}(S) &= \sum_{T\supseteq S} \frac{1+(-1)^{t-s}}{(t-s+1)\, 2^{t-s+1}}\, I_{\mathrm{B}}(T)
\end{aligned}
$$