Capacity based MCDA and XAI

Jian-Zhang Wu

Research Fellow, Deakin University
jianzhang.wu@deakin.edu.au

March 7, 2023

Outline

Capacity and Nonlinear Integral

Integral + Measure

Measure → Additivity: $\mu(A \cup B) = \mu(A) + \mu(B)$ for disjoint subsets A, B. → Independent → Measure only describes the **importance** of criteria.

Nonlinear/Fuzzy Integral + Capacity (Fuzzy/Nonadditive Measure)

Capacity → Nonadditivity: $\mu(A \cup B) =$ (**or** $\leq, \geq$) $\mu(A) + \mu(B)$ for disjoint subsets A, B. → Dependent/Interactive → Capacity can describe the **importance as well as interaction** of criteria.

Interaction among Criteria

Additivity (=)→ Independent/**Zero** interaction;
Subadditivity ($\leq$) → Substitutive/**Negative** interaction; (Pro, Sta)
Superadditivity ($\geq$)→ Complementary/**Positive** interaction; (Eng, Eco)

Complexity Simplification & Our works

Measure involves n coefficients for n criteria: $\mu(i), i \in N$.
Capacity involves $2^n - 2$ coefficients for n criteria: $\mu(A), A \subset N, A \neq \emptyset$.

Simplify the exponential complexity

- **Particular families** with less coefficients, e.g., λ-capacity, p-symmetric capacity, k-additive capacity, and k-maxitive capacity.

  Michel Grabisch. "K-order additive discrete fuzzy measures and their representation." Fuzzy sets and systems 92.2 (1997): 167-189.

- **Identification methods** by using optimization models, like Least Square model, TOMASO method, NAROR method, etc.

  Michel Grabisch, Ivan Kojadinovic, and Patrick Meyer. "A review of methods for capacity identification in Choquet integral based multi-attribute utility theory: Applications of the Kappalab R package." European journal of operational research 186.2 (2008): 766-785.

Complexity Simplification & Our works

Our works

- **Particular families:** *k*-interactive capacity, *k*-minitive capacity, Lower/Upper *k*-order representative capacity, buoyant and antibuoyant capacity

  Gleb Beliakov and Jian-Zhang Wu. "Learning fuzzy measures from data: simplifications and optimisation strategies." Information Sciences 494 (2019): 100-113.

  Jian-Zhang Wu and Gleb Beliakov. "k-order representative capacity." Journal of Intelligent & Fuzzy Systems 38.3 (2020): 3105-3115.

  Gleb Beliakov and Simon James. "Choquet integral-based measures of economic welfare and species diversity." International Journal of Intelligent Systems 37.4 (2022): 2849-2867.

Complexity Simplification & Our works

Our works

- **Particular families:** k-interactive capacity, k-minitive capacity, Lower/Upper k-order representative capacity, buoyant and antibuoyant capacity
  - **k-interactive capacity:** A capacity is called k-interactive if for some chosen $K \in [0,1]$

    $$\mu(A) = K + \frac{a-k-1}{n-k-1}(1-K), \text{for all } A, a > k.$$

    Our rationale is that by **maximising partial entropy** for subsets of cardinality larger than k we obtain nearly additive measure with respect to those subsets.

    $$\mathcal{C}_\mu(\mathbf{x}) = \frac{1-K}{n-k-1}\sum_{i=1}^{n-k-1} x_{(i)} + Kx_{(n-k)} + \sum_{A\subseteq N, a\leq k} \mu(A)g_A(\mathbf{x}).$$

Complexity Simplification & Our works

Our works

- **Particular families:** k-interactive capacity, k-minitive capacity, Lower/Upper k-order representative capacity, buoyant and antibuoyant capacity
  - ***k*-minitive capacity:** A capacity ν is called k-minitive if its necessity Möbius transform satisfies $m_n(A) = 1$ for any A such that $|A| < n - k$ and there exists at least one subset A_0 of exactly $n - k$ elements such that $m_n(A_0) \neq 1$. Dual of k-maxitive capacity ($\mu(A) = \bigvee_{B \subset A} \mu(B), |A| \geq k + 1$)

$$\nu(A) = \bigwedge_{B \supset A} \nu(B), |A| \leq n - k - 1,$$

$$\nu(A) = \bigwedge_{B \supset A, |B| \geq n-k} \nu(B), |A| \leq n - k - 1,$$

$$\nu(A) = \bigwedge_{A \subset B, |B| = |A|+1} \nu, |A| \leq n - k - 1.$$

Complexity Simplification & Our works

Our works

k-maxitive capacity:

$$\mu(A) - \mu(A \backslash \{i\}) \leq y_{A,i}, |A| \geq k+1, i \in A,$$

$$y_{A,i} = \begin{cases} 0 \text{ if } \mu(A) - \mu(A \backslash \{i\}) = 0 \\ 1 \text{ if } \mu(A) - \mu(A \backslash \{i\}) \neq 0 \end{cases}$$

$$\sum_{A} y_{A,i} \leq |A| - 1.$$

k-minitive capacity:

$$\mu(A \cup \{i\}) - \mu(A) \leq y_{A,i}, |A| \leq n-k-1, i \in N \backslash A,$$

$$y_{A,i} = \begin{cases} 0 \text{ if } \mu(A \cup \{i\}) - \mu(A) = 0 \\ 1 \text{ if } \mu(A \cup \{i\}) - \mu(A) \neq 0 \end{cases}$$

$$\sum_{A} y_{A,i} \leq n - |A| - 1.$$

Complexity Simplification & Our works

Our works

- **Particular families:** k-interactive capacity, k-minitive capacity, Lower/Upper k-order representative capacity, buoyant and antibuoyant capacity
  - **Lower/Upper k-order representative capacity:** A capacity μ on N is said to be (lower) k-order representative if there exists a representation of μ, denoted as r_μ,
    $$r_\mu(A) = c, \ \forall A \subseteq N, |A| > k, ,$$
    where c is a given constant value (usually $c = 0$ or 1) or c is a function, $c(A) = f(r_\mu(B) \mid |B| \leq k)$. A capacity μ on N is said to be **upper** k-order representative if for all $A \subseteq N$, $|A| < n - k$, satisfying $r_\mu(A) = c$.

Complexity Simplification & Our works

Our works

- **Particular families:** k-interactive capacity, k-minitive capacity, Lower/Upper k-order representative capacity, buoyant and antibuoyant capacity
  - **buoyant (antibuoyant):**
    A capacity μ is buoyant if

    $$\Delta_i(A \cup \{i,j\} \leq \Delta_j(A \cup \{j\} \forall i,j \notin A.$$

    A capacity μ is antibuoyant if

    $$\Delta_i(A \cup \{i,j\} \geq \Delta_j(A \cup \{j\} \forall i,j \notin A.$$

Complexity Simplification & Our works

Our works

- **Identification methods:**
  - LP and MIP based methods,
  - Maximum entropy principle with AHP,
  - Compromise principle methods,
  - MCCPI based methods,
  - Explicit preference oriented method.

Complexity Simplification & Our works

Our works

- **Identification methods:**
  - Maximum entropy principle with AHP

    The overall importance of each criterion is obtained by the maximum eigenvector method of AHP.

    The Choquet integral-based equivalent alternative curve can help the decision maker estimate the interaction degrees between criteria.

    According to the maximum fuzzy measure entropy principal, a nonlinear programming is constructed to identify the interaction indices among criteria.

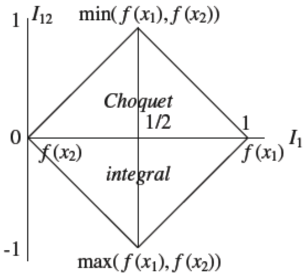
    J.-Z. Wu, and Q. Zhang, 2-order additive fuzzy measure identification method based on diamond pairwise comparison and maximum entropy principle", Fuzzy Optimization and Decision Making 9(4), 435-453 (2010).

Complexity Simplification & Our works

Our works

- **Identification methods:**
  - Maximum entropy principle with AHP

Fig. 1 Interpretation of the Choquet integral



Complexity Simplification & Our works

Our works

- **Identification methods:**
  - Maximum entropy principle with AHP

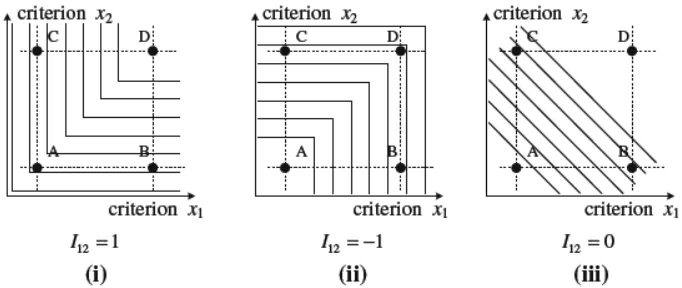


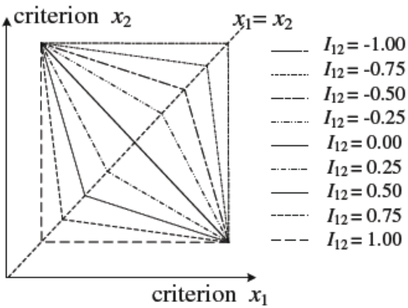
Fig. 4 Different cases of the equivalent alternative curve

Complexity Simplification & Our works

Our works

- **Identification methods:**
  - Maximum entropy principle with AHP

Fig. 5 The equivalent alternative curve for some interaction



Complexity Simplification & Our works

Our works

- **Identification methods:**
  - Maximum entropy principle with AHP

$$\max\ z=\sum_{i=1}^{n}\sum_{A\subset X\setminus x_i}\frac{(|X|-|A|-1)!|A|!}{|X|!}h\left[I_i'-\frac{1}{2}\sum_{x_j\in X\setminus\{A\cup\{x_i\}\}}I_{ij}'+\frac{1}{2}\sum_{x_j\in A}I_{ij}'\right]$$

$$\begin{cases}\max\left(\frac{I_i'I_{ij}}{(n-1)I_i^{ij}},\frac{I_j'I_{ij}}{(n-1)I_j^{ij}}\right)\le I_{ij}'\le 0 & \text{if } I_{ij}\le 0,\\ 0\le I_{ij}'\le\min\left(\frac{I_i'I_{ij}}{(n-1)I_i^{ij}},\frac{I_j'I_{ij}}{(n-1)I_j^{ij}}\right) & \text{if } I_{ij}\ge 0,\\ I_i'=w_i, & \\ i,j=1,2,\ldots,n. & \text{and } i\ne j.\end{cases}\quad(30)$$

Complexity Simplification & Our works

Our works

- **Identification methods:**
  - Compromise principle methods

    The basic aim of the compromise principle employed in this paper is to seek the capacity by which each alternative has a relatively equal chance to reach as close as possible to its highest reachable overall evaluation.

    Three types of capacity identification methods – the simple arithmetic average based compromise method, the least squares based compromise method and the linear programming based compromise method – are proposed.

    Jian-Zhang Wu et al. "Compromise principle based methods of identifying capacities in the framework of multicriteria decision analysis." Fuzzy Sets and Systems 246 (2014): 91-106.

Complexity Simplification & Our works

Our works

- **Identification methods:**
  - Compromise principle methods

$$(\text{LPCM1}):\ \min \sum_{i=1}^{m} \varepsilon_i$$

$$\begin{cases} \mu(\emptyset) = 0, \quad \mu(N) = 1, \\ \mu(A) \leqslant \mu(B) \quad \text{for } \forall A, B \subseteq N,\ A \subseteq B, \\ I_\mu(\{i\}) - I_\mu(\{j\}) \geqslant \delta_{Sh}, \\ \vdots \\ I_\mu(\{i,j\}) - I_\mu(\{k,l\}) \geqslant \delta_I, \\ \vdots \\ C_\mu(d_1) + \varepsilon_1 = H_1, \\ C_{m_\mu}(d_2) + \varepsilon_2 = H_2, \\ \vdots \\ C_\mu(d_m) + \varepsilon_m = H_m \end{cases}$$

Complexity Simplification & Our works

Our works

- **Identification methods:**
  - MCCPI based methods

    The multicriteria correlation preference information (MCCPI) is a group of 2-D preference information which can be described and obtained by the refined diamond diagram.

    The common principle of the proposed models is to minimize the different kinds of deviations between theMCCPI and themost desired 2-additive capacity(ies).

    Jian-Zhang Wu et al. "2-additive capacity identification methods from multicriteria correlation preference information." IEEE Transactions on Fuzzy Systems 23.6 (2015): 2094-2106.

Complexity Simplification & Our works

Our works

- **Identification methods:**
  - MCCPI based methods

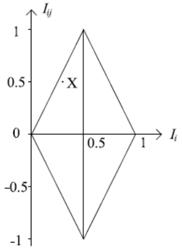


Fig. 1. Diamond diagram.

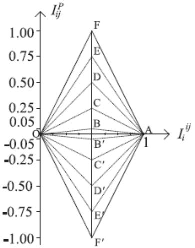


Fig. 2. Refined diamond diagram.

Complexity Simplification & Our works

Our works

- **Identification methods:**
  - MCCPI based methods

$$\mathbf{P} = [p_{ij}]_{n\times n} = \begin{bmatrix} - & I_{21}^P & \cdots & I_{n1}^P \\ I_{21}^P & - & \cdots & I_{n2}^P \\ \vdots & \vdots & \vdots & \vdots \\ I_{n1}^P & I_{n2}^P & \cdots & - \end{bmatrix}$$

$$\mathbf{R} = [r_{ij}]_{n\times n} = \begin{bmatrix} \frac{1}{2} & I_1^{12} & \cdots & I_1^{1n} \\ I_2^{21} & \frac{1}{2} & \cdots & I_2^{2n} \\ \vdots & \vdots & \vdots & \vdots \\ I_n^{n1} & I_2^{n2} & \cdots & \frac{1}{2} \end{bmatrix} \rightarrow \begin{bmatrix} I_1 \\ I_2 \\ \vdots \\ I_n \end{bmatrix}$$

$$\rightarrow \begin{bmatrix} - & I_{21} & \cdots & I_{n1} \\ I_{21} & - & \cdots & I_{n2} \\ \vdots & \vdots & \vdots & \vdots \\ I_{n1} & I_{n2} & \cdots & - \end{bmatrix}.$$

Complexity Simplification & Our works

Our works

- **Identification methods:**
  - MCCPI based methods

$$
\begin{aligned}
(\mathrm{LS-II'}): \min Z_2 = 2\sum_{i=1}^{n-1}\sum_{j>i}^{n} \Bigg[&\left(r_{ij} - \frac{I_i}{I_i + I_j} \right)^2 \\
&+ \left(p_{ij} - \frac{I_{ij}}{I_i + I_j} \right)^2 \Bigg]
\end{aligned}
$$

$$
\begin{cases}
\sum_{i=1}^{n} I_i = 1, i = 1, 2, \ldots, n \\
\mathbf{A}(I_1, I_{12}, I_{13}, \ldots, I_{1n})^{\mathrm{T}} \geq 0 \\
\mathbf{A}(I_i, I_{1i}, \ldots, I_{[i-1][i]}, I_{[i][i+1]}, \ldots, I_{in})^{\mathrm{T}} \geq 0 \\
\quad \text{for } i = 2, 3, \ldots, n-1 \\
\mathbf{A}(I_n, I_{1n}, I_{2n}, \ldots, I_{[n-1][n]})^{\mathrm{T}} \geq 0
\end{cases}
$$

where $\mathbf{A} = [a_{ij}]_{2^{n-1} \times n}$, $a_{ij} = 1$ if $j = 1$ and

$(-1)^{\mathrm{bitget}(\mathrm{dec2bin}(i-1), j-1)} \left(1 \ldots\right)$ if $i = 2, 3, \ldots$

Complexity Simplification & Our works

Our works

- **Identification methods:**
  - MCCPI based methods

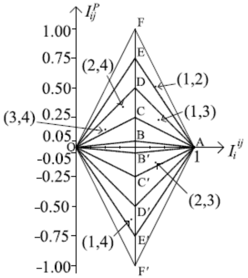


Fig. 3. Two-dimensional pairwise comparisons of the four criteria.

Complexity Simplification & Our works

Our works

- **Identification methods:**
  - MCCPI based methods

SCALE OF 2-D MCCPI

Relative importance of criterion i to j	
Nine categories	Scale value of I_i^{ij}
extremely less important	0.0625
very strongly less important	0.1875
strongly less important	0.3125
slightly less important	0.4375
equally important	0.5000
slightly more important	0.5625
strongly more important	0.6875
very strongly more important	0.8125
extremely more important	0.9375
Partial interaction between i and j	
Nine categories	Scale value of I_{ij}^P
extremely positive	$1.75\min(I_i^{ij}, 1 - I_i^{ij})$
very strongly positive	$1.25\min(I_i^{ij}, 1 - I_i^{ij})$
strongly positive	$0.75\min(I_i^{ij}, 1 - I_i^{ij})$
slightly positive	$0.3\min(I_i^{ij}, 1 - I_i^{ij})$
almost zero	0.000
slightly negative	$-0.3\min(I_i^{ij}, 1 - I_i^{ij})$
strongly negative	$-0.75\min(I_i^{ij}, 1 - I_i^{ij})$
very strongly negative	$-1.25\min(I_i^{ij}, 1 - I_i^{ij})$
extremely negative	$-1.75\min(I_i^{ij}, 1 - I_i^{ij})$

Complexity Simplification & Our works

Our works

- **Identification methods:**
  - Explicit preference oriented method

    The maximum and minimum empty set interaction principles based capacity identification methods,which can be considered as the comprehensive interaction trend preference information oriented capacity identification methods.

    Use MGLP to take decision maker's explicit preference information on interaction and the importance as objective function, and construct the nonempty set interaction indices based capacity identification method.

    Jian-Zhang Wu, Endre Pap and Aniko Szakal. "Two kinds of explicit preference information oriented capacity identification methods in the context of multicriteria decision analysis." International Transactions in Operational Research 25.3 (2018): 807-830.

Complexity Simplification & Our works

Our works

- **Identification methods:**
  - Explicit preference oriented method

$$(\text{Min-ESI}) : \min I_{\mathrm{p}}^{m_\mu}(\emptyset)$$

or

$$(\text{Max-ESI}) : \max I_{\mathrm{p}}^{m_\mu}(\emptyset)$$

$$\begin{cases} m_\mu(\emptyset) = 0, \sum_{A \subseteq N} m_\mu(A) = 1, \\ \sum_{B \subseteq A,\, i \in B} m_\mu(B) \geq 0,\ \forall A \subseteq N,\ \forall i \in A, \\ I_{\mathrm{Sh}}^{m_\mu}(\{i\}) - I_{\mathrm{Sh}}^{m_\mu}(\{j\}) \geq \delta_{\mathrm{Sh}}, \\ \qquad \vdots \\ I_{\mathrm{Sh}}^{m_\mu}(\{i, j\}) - I_{\mathrm{Sh}}^{m_\mu}(\{k, l\}) \geq \delta_{\mathrm{I}}, \\ \qquad \vdots \end{cases}$$

where $I_{\mathrm{p}}^{\mu}(\emptyset)$ is a probabilistic interaction index of the empty set such that $p_B^{\emptyset}(N) \neq 0, \forall B \subseteq N$, δ_{Sh} and δ_{I} are positive indifference thresholds.

Complexity Simplification & Our works

Our works

- **Identification methods:**
  - Explicit preference oriented method

$$(\text{NSIIO}) : \max z = \sum_{A \in \mathcal{K}^+} w_A d_A^+ + \sum_{A \in \mathcal{K}^-} w_A d_A^- + \sum_{A \in \mathcal{K}^0} (d_A^- + d_A^+)$$

$$\begin{cases} \mu(\emptyset) = 0, \mu(N) = 1, \\ \mu(A) \leq \mu(B) \text{ for } \forall A, B \subseteq N, A \subseteq B, \\ I_{\text{Sh}}^{\mu}(\{i\}) - I_{\text{Sh}}^{\mu}(\{j\}) \geq \delta_{\text{Sh}}, \\ \qquad \vdots \\ I_{\text{Sh}}^{\mu}(\{i, j\}) - I_{\text{Sh}}^{\mu}(\{k, l\}) \geq \delta_{\text{I}}, \\ \qquad \vdots \\ I_{\text{Sh}}^{\mu}(A) - d_A^+ + d_A^- = 0 \quad \mathit{for}\ \forall A \in K, \\ d_A^+, d_A^- \geq 0, d_A^+ \times d_A^- = 0 \quad \mathit{for}\ \forall A \in K, \end{cases}$$

Preference Representation & Our works

Preference Representation

- How to measure the **interaction kind and density**?
  - Simultaneous interaction index.
- How to represent the decision **preference on interaction** of criteria?
  - Interval range or comparison form.

Michel Grabisch. "Alternative representations of discrete fuzzy measures for decision making." International Journal of Uncertainty, Fuzziness and Knowledge-Based Systems 5.05 (1997): 587-607.

Michel Grabisch and Marc Roubens. "Application of the Choquet integral in multicriteria decision making." Fuzzy Measures and Integrals-Theory and Applications (2000): 348-374.

Katsushige Fujimoto, Ivan Kojadinovic and Jean-Luc Marichal. "Axiomatic characterizations of probabilistic and cardinal-probabilistic interaction indices." Games and Economic Behavior 55.1 (2006): 72-99.

Preference Representation & Our works

Our works

- **Interaction index:**
  - Nonadditivity index → Bipartition interaction index.
  - Nonmodularity index → Comprehensive nonmodularity index.

Jian-Zhang Wu and Gleb Beliakov. "Nonadditivity index and capacity identification method in the context of multicriteria decision making." Information Sciences 467 (2018): 398-406.

Jian-Zhang Wu and Gleb Beliakov. "Nonmodularity index for capacity identifying with multiple criteria preference information." Information Sciences 492 (2019): 164-180.

Jian-Zhang Wu and Gleb Beliakov. "Probabilistic bipartition interaction index of multiple decision criteria associated with the nonadditivity of fuzzy measures." International Journal of Intelligent Systems 34.2 (2019): 247-270.

Gleb Beliakov and Jian-Zhang Wu. "The axiomatic characterisations of non-modularity index." International Journal of General Systems 49.7 (2020): 675-688.

Gleb Beliakov, Simon James and Jian-Zhang Wu. Discrete fuzzy measures. Springer International Publishing, 2020.

Preference Representation & Our works

Our works

- **Interaction index:**
  - Nonadditivity index
    A capacity μ on N is additive, if $\mu(B \cup C)$ = $\mu(B) + \mu(C)$ for arbitrary pair of nonempty disjoint subsets A, B of N. Let $A = B \cup C$, $\mu(A)$ = $\mu(A \backslash C) + \mu(C)$, by collecting all the bipartitions equally, we have the nonadditivity index as

$$n_\mu(A) = \mu(A) - \frac{1}{2^{|A|-1} - 1} \sum_{C \subset A} \mu(C).$$

    Nonadditivity has some many good properties:
    - $\star$-additivity: If a capacity μ on N is $\star$-additive, then $n_\mu(A) \overset{\star}{=} 0, \forall A \subseteq N, |A| \geq 2$. If a capacity μ on N is $\star$-additive, then $n_\mu(A), \overset{\star}{=} 0, \forall A \subseteq N, |A| \geq 2$.
    - Uniform Range: $-1 \leq n_\mu(A) \leq 1, \forall A \subseteq N$.
    - Maximality and Minimality: $n_\mu(A) = 1 \;\Leftrightarrow\; \mu(S) = \varepsilon_A^+(S)$, $S \subseteq A$, and $n_\mu(A) = -1 \;\Leftrightarrow\; \mu(S) = \varepsilon_A^-(S),\; S \cap A \neq \emptyset$.

Preference Representation & Our works

Our works

- **Interaction index:**
  - Probabilistic bipartition index
    The marginal bipartition interaction is defined as

$$\hat{\Delta}_A \mu(B) = \mu(B \cup A) + \mu(B) - \frac{1}{2^{|A|-1} - 1} \sum_{\emptyset \neq C \subset A} \mu(B \cup C),$$

The Shapley bipartition interaction index is defined as

$$\hat{I}^{\mu}_{\mathrm{Sh}}(A) = \sum_{B \subseteq N \setminus A} \frac{1}{|N| - |A| + 1} \binom{|N| - |A|}{|B|}^{-1} \hat{\Delta}_A \mu(B).$$

The Banzhaf bipartition interaction index is defined as

$$\hat{I}^{\mu}_{\mathrm{Ba}}(A) = \sum_{B \subseteq N \setminus A} \frac{1}{2^{(|N| - |A|)}} \hat{\Delta}_A \mu(B).$$

Preference Representation & Our works

Our works

- **Interaction index:**
  - Nonmodularity index
    By the Corollary 2.23 in Michel Grabisch, Set Functions, Games and Capacities in Decision Making. Springer, Berlin, New York, 2016:
    A capacity is supermodular on N iff for any $A \subseteq N, i, j \notin A, i \neq j$:

$$\Delta_{ij}\mu(A) = \mu(A \cup \{i,j\}) - \mu(A \cup \{i\}) - \mu(A \cup \{j\}) + \mu(A) \geq 0.$$

    By collecting all the second order derivatives equally, we have the nonmodularity index as

$$d_\mu(A) = \mu(A) - \frac{1}{|A|} \sum_{\{i\} \subset A} [\mu(\{i\}) + \mu(A \backslash \{i\})].$$

    Nonmodularity index also has some many good properties such as $\star$-modularity, uniform range, maximality and minimality.

Preference Representation & Our works

Our works

- **Interaction index:**
  - Comprehensive nonmodularity index

The marginal nonmodularity index is defined as

$$\begin{aligned}\bar{\Delta}_B^{\mu}(A) =& [\mu(A \cup B) - \mu(B)] \\ & - \frac{1}{|A|} \sum_{\{i\} \subset A} [(\mu(\{i\} \cup B) - \mu(B)) + (\mu((A \backslash \{i\}) \cup B) - \mu(B))].\end{aligned}$$

The Shapley and Banzhaf nonmodularity interaction index are defined as

$$d_{\text{Sh}}^{\mu}(A) = \sum_{B \subseteq N \backslash A} \frac{1}{|N| - |A| + 1} \binom{|N| - |A|}{|B|}^{-1} \bar{\Delta}_B^{\mu}(A),$$

$$d_{\mathrm{Ba}}^{\mu}(A) = \sum_{B \subseteq N \backslash A} \frac{1}{2^{(|N| - |A|)}} \bar{\Delta}_B^{\mu}(A).$$

Preference Representation & Our works

Our works—axioms for nonadditivity index

For a linear function of μ, $I_\mu(A) = \sum_{B\subseteq A} \alpha_B \mu(B)$.

- Equality for All subsets (EAS): $\alpha_B = \alpha_C \neq 0, \forall B, C \subset A$.
- Additivity (Add):
  $\mu(A) = \mu(B) + \mu(C)\ \forall B, C \subset A, B \cap C = \emptyset, B \cup C = A \Rightarrow I_\mu(A) = 0$.
- Maximality (Max): $\mu(A) = 1$ and $\mu(B) = 0, \forall B \subsetneq A \Rightarrow I_\mu(A) = 1$.
- Minimality (Min): $\mu(A) = 1$ and $\mu(B) = 1, \forall \emptyset \neq B \subset A \Rightarrow I_\mu(A) = -1$.

Theorem

The linear function $I_\mu(A)$ is the nonadditivity index $n_\mu(A)$ if and only if one of following conditions holds:

(1) *$I_\mu(A)$ has properties (EAS), (Max) and (Add)*

(2) *$I_\mu(A)$ has properties (EAS), (Max) and (Min)*

Preference Representation & Our works

Our works

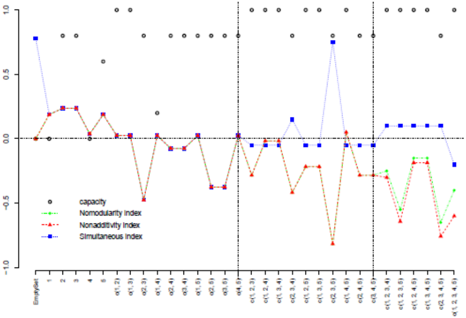


Figure 4: Three types of Banzhaf interaction indices of maximum orness index.

Preference Representation & Our works

Our works—some comparisons

- **Representation tools and methods:**
  - Aiding figures and scales of pairwise comparison MCCPI.
  - Inconsistency recognition: MGLP model (deviation variables).
  - Matrix (from capacity to representation), Marginal contribution (first derivative) and Graphic representations and Computing programmings.

Jian-Zhang Wu et al. "2-additive capacity identification methods from multicriteria correlation preference information." IEEE Transactions on Fuzzy Systems 23.6 (2015): 2094-2106.

Jian-Zhang Wu and Gleb Beliakov. "Nonadditive robust ordinal regression with nonadditivity index and multiple goal linear programming." International Journal of Intelligent Systems 34.7 (2019): 1732-1752.

Jian-Zhang Wu and Gleb Beliakov. "Marginal contribution representation of capacity-based multicriteria decision making." International Journal of Intelligent Systems 35.3 (2020): 373-400.

Preference Representation & Our works

Our works

- **Representation tools and methods:**
  - Aiding figures and scales of pairwise comparison MCCPI.

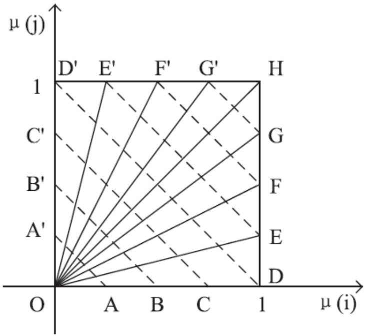


Fig. 2. The diagram of nonadditivity index type MCCPI.

Preference Representation & Our works

Our works

- **Representation tools and methods:**
  - Inconsistency recognition: MGLP model (deviation variables).

$$
\min \sum_{r=1}^{30} d_r^+ + d_r^-
$$

$$
\text{s.t.} \begin{cases}
C_\mu(A_6) - C_\mu(A_2) - d_1^+ + d_1^- = 0.1, \\
\dots \\
C_\mu(A_7) - C_\mu(A_1) - d_4^+ + d_4^- = 0.1, \\
C_\mu(A_8) - C_\mu(A_{10}) - d_5^+ + d_5^- = 0.1, \\
I_\mu(\{1\}) - I_\mu(\{2\}) - d_6^+ + d_6^- = 0.1, \\
\dots \\
I_\mu(\{4\}) \geq I_\mu(\{5\}) - d_8^+ + d_8^- = 0.1, \\
\dots \\
I_\mu(\{1,2,4,5\}) - d_{17}^+ + d_{17}^- = 0.05, \\
I_\mu(\{1,3,4,5\}) - d_{18}^+ + d_{18}^- = 0.05, \\
I_\mu(\{2,3\}) - d_{19}^+ + d_{19}^- = -0.05, \\
\dots \\
I_\mu(\{1,4,5\}) - I_\mu(\{1,2,4,5\}) - d_{28}^+ + d_{28}^- = 0.1, \\
I_\mu(\{1,2,3,5\}) - I_\mu(\{1,3,4,5\}) - d_{29}^+ + d_{29}^- = 0.1, \\
\dots \\
\mu(\emptyset) = 0, \mu(\{N\}) = 1, \\
\mu(\{i\} \cup B) - \mu(B) \geq 0, \forall i \in N, B \subseteq N \setminus \{i\}.
\end{cases} \quad (2)
$$

Preference Representation & Our works

Our works

- **Representation tools and methods:**
  - Inconsistency recognition: MGLP model (deviation variables).

$d_1^{-*} = 0.0546,\ d_4^{-*} = 0.0689,\ d_5^{-*} = 0.3970,\ d_8^{-*} = 0.0182,\ d_{17}^{-*} = 0.05,\ d_{18}^{-*} = 0.05,\ d_{28}^{-*} = 0.05,\ d_{28}^{-*} = 0.05,$
which means that some inconsistency exits in the preference constraints.

According to the third step in Section 4, the corresponding constraints are adjusted as:

$$C_\mu(A_6) \geq C_\mu(A_2) - 0.0546,$$

$$C_\mu(A_7) \geq C_\mu(A_1) - 0.0689,$$

$$C_\mu(A_8) \geq C_\mu(A_{10}) - 0.3970.$$

$$I_\mu(\{4\}) \geq I_\mu(\{5\}) + 0.1 - 0.0182,$$

$$I_\mu(\{1,2,4,5\}) - d_{17}^{+} + d_{17}^{-} = 0.05 - 0.05,$$

$$I_\mu(\{1,3,4,5\}) - d_{18}^{+} + d_{18}^{-} = 0.05 - 0.05,$$

$$I_\mu(\{1,4,5\}) - I_\mu(\{1,2,4,5\}) - d_{28}^{+} + d_{28}^{-} = 0.05,$$

$$I_\mu(\{1,2,3,5\}) - I_\mu(\{1,3,4,5\}) - d_{29}^{+} + d_{29}^{-} = 0.05.$$

Then all the preference constraints are consistent.

Preference Representation & Our works

Our works

- **Representation tools and methods:**
  - Matrix (from capacity to representation) representation and Computing programmings

- The transform matrix of Möbius representation is $\mathbf{M} = [m_{ij}]_{2^n \times 2^n}$, $i, j = 0, \ldots, n$,

$$m_{ij} = \begin{cases} (-1)^{(|arg\ pos(i)| - |arg\ pos(j)|)} & \text{if } arg\ pos(j) \subseteq arg\ pos(i) \\ 0 & \text{otherwise} \end{cases} \quad (3)$$

- The transform matrix of the co-Möbius representation is $\mathbf{C} = [c_{ij}]_{2^n \times 2^n}$, $i, j = 0, \ldots, n$,

$$c_{ij} = \begin{cases} (-1)^{|arg\ pos(i) \setminus arg\ pos(j)|} & \text{if } arg\ pos(j) \cup arg\ pos(i) = N \\ 0 & \text{otherwise.} \end{cases} \quad (4)$$

- The transform matrix of Shapley simultaneous interaction index is $\mathbf{S} = [s_{ij}]_{2^n \times 2^n}$, $i, j = 0, \ldots, n$,

$$s_{ij} = \frac{(-1)^{|arg\ pos(i) \setminus arg\ pos(j)|}}{n - |arg\ pos(i)| + 1} \binom{n - |arg\ pos(i)|}{|arg\ pos(j) \setminus arg\ pos(i)|}^{-1} \quad (5)$$

- The transform matrix of Banzhaf simultaneous interaction index is $\mathbf{B} = [b_{ij}]_{2^n \times 2^n}$, $i, j = 0, \ldots, n$,

$$b_{ij} = \left(\frac{1}{2}\right)^{n - arg\ pos(i)} (-1)^{|arg\ pos(i) \setminus arg\ pos(j)|} \quad (6)$$

Preference Representation & Our works

Our works

- **Representation tools and methods:**
  - Marginal contribution (first derivative) representations and Computing programmings.

Theorem 10. *Let μ be a capacity on $N, A \subseteq N, |A| \geq 2$, then*

$$\begin{aligned} n_\mu(A) &= \frac{1}{2^{|A|}-2} \sum_{C \subset A} \sum_{k=1}^{|C|} \left(\Delta_{a_{\pi(k)}} \mu(A_{\pi(|A|-|C|+k-1)}) - \Delta_{a_{\pi(k)}} \mu(C_{\pi(k-1)})\right) \\ &= \frac{1}{2^{|A|}-2} \sum_{C \subset A} \sum_{k=1}^{|C|} \left(\Delta_{a_{\pi(k)}} \mu(\overline{A}_{\pi(k)}) - \Delta_{a_{\pi(k)}} \mu(\overline{C}_{\pi(k)})\right). \end{aligned} \tag{16}$$

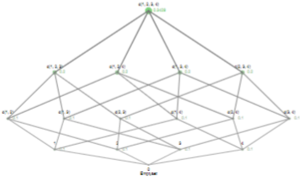
$$\begin{aligned} orness(C_\mu) &= \sum_{A \subset N} \frac{(n-|A|)!|A|!}{n!(n-1)} \sum_{i=1}^{|A|} \Delta_{a_{\pi(i)}} \mu(A_{\pi(i-1)}) \\ &= \sum_{A \subset N} \frac{(n-|A|)!|A|!}{n!(n-1)} \sum_{i=1}^{|A|} \Delta_{a_{\pi(i)}} \mu(\overline{A}_{\pi(i)}), \end{aligned} \tag{31}$$

$$orness(C_m) = \sum_{A \subset N} \frac{(n-|A|)!|A|!}{n!(n-1)} \sum_{C \subseteq A \setminus \{a_1\}} (-1)^{|A \setminus C|-1} \Delta_{\{a_1\}} \mu(C).$$

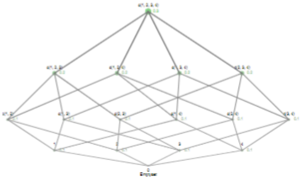
Preference Representation & Our works

Our works

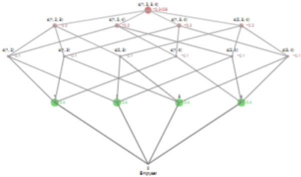
- **Representation tools and methods:**
  - Graphic representations and Computing programmings.



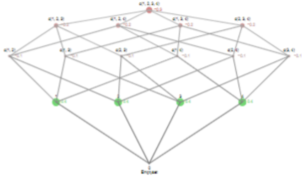
(a) μ_1 with nonadditivity index.



(b) μ_1 with nonmodularity index.



(c) $\overline{\mu}_1$ with nonadditivity index.



(d) $\overline{\mu}_1$ with nonmodularity index.

Capacity and eXplainable Artificial Intelligence

Treat AI resutls as capacity ML targets

- Complexity Reduce (sparse capacity)
- Preference Representation
- Inconsistence Recognition (MGLP)
- Random Generation

Random Generation & Our works

Random Generation

Elías F Combarro and Pedro Miranda. "Identification of fuzzy measures from sample data with genetic algorithms." Computers & Operations Research 33.10 (2006): 3046-3066.

Elías F Combarro, Irene Díaz and P. Miranda. "On random generation of fuzzy measures." Fuzzy Sets and Systems 228 (2013): 64-77.

Our works

Gleb Beliakov, "On random generation of supermodular capacities." IEEE Transactions on Fuzzy Systems 30.1 (2020): 293-296.

Gleb Beliakov and Jian-Zhang Wu. "Random generation of capacities and its application in comprehensive decision aiding." Information Sciences 577 (2021): 424-435.

Gleb Beliakov, FJ Cabrerizo, E Herrera-Viedma and JZ Wu . "Random generation of k-interactive capacities." Fuzzy Sets and Systems 430 (2022): 48-55.

(MinimalsPlus method, Directed acyclic graph DAG, Markov chain random walk, acceptance/rejection approach)

Random Generation & Our works

Our works

LP and convex combination with preference constraints

Algorithm 2: Linear programming based capacity random generation.

Input: Decision criteria set N; boundary and monotonicity conditions, denoted $\mathcal{B}$; linear constraints on decision preference information, denoted $\mathcal{D}$.

Output: $\overline{k}$ randomly generated capacities, denoted by matrix $\mathbf{U}_{\overline{k}\times 2^N}$.

```
for k in 1 : k̄ do
    Set all nonempty proper subsets in N as P,
    for i in 1 : (2^n − 2) do
        Randomly choose a subset A ∈ P,
        Calculate L_A = min μ(A) s.t. B and D,   /* can use m_μ(A) or I_μ(A) */
        Calculate U_A = max μ(A) s.t. B and D,   /* can use m_μ(A) or I_μ(A) */
        Set μ(A) to a random value r in [L_A, U_A],   /* can use m_μ(A) or I_μ(A) */
        Set P = P \ A,
        Add constraint: μ(A) = r into D.   /* can use m_μ(A) or I_μ(A) */
    end
    Store the generated capacity, denoted μ_k, as the k-th row of U.
end
Return U.
```

Random Generation & Our works

Our works

LP and convex combination with preference constraints

Algorithm 4: Random convex combination of pseudo extreme capacities.

Input: Decision criteria set N; boundary and monotonicity conditions, denoted by BM; linear constraints modeling decision preference information, denoted DP.
Output: Extreme capacities matrix $\mathbf{V}$; $\overline{k}$ randomly generated capacities, denoted by matrix $\mathbf{U}_{\overline{k}\times 2^N}$.

```
/* generate pseudo extreme capacities */
```

for l *in* $1:\overline{l}$ **do**
  For random s, $1 \leqslant s \leqslant 2^n$, select s random subsets of N, denoted $A_1, ..., A_s$.
  Set μ_l= arg max$(\sum_{B\in\mathcal{A}} \mu(B) - \sum_{B\notin\mathcal{A}} \mu(B))$ s.t. BM and DP, where arg max z returns the capacity reaching the maximization of z, $\mathcal{A} = \{A | A_1 \subseteq A, ..., \text{ or } A_s \subseteq A\}$.
  Store μ_l as the l-th row of $\mathbf{V}$.
end
for k *in* $1:\overline{k}$ **do**
  For random t, $2 \leqslant t \leqslant |\mathbf{V}|$, randomly select t extreme capacities in $\mathbf{V}$, denoted $v_1, ..., v_t$.
  Randomly generate t positive nubmers $\lambda_1, ..., \lambda_t$ with their sum equals to 1.
  Set $\mu_k = \lambda_1 v_1 + \cdots + \lambda_t v_t$, `/* convex combination of extreme capacities */`
  Store μ_k as the k-th row of $\mathbf{U}$.
end
Return $\mathbf{U}$ and $\mathbf{V}$.

Random Generation & Our works

Our works

LP and convex combination with preference constraints

Algorithm 5: Random convex combination of vertex capacities.

Input: Decision criteria set N; the number of capacities to be included in each convex combination, denoted t, $(t > 1)$; boundary and monotonicity conditions, denoted BM; linear constraints of decision preference information, denoted DP.

Output: $\bar{l}$ extreme capacities, denoted by matrix $\mathbf{V}_{\bar{l}\times 2^N}$; $\bar{k}$ randomly generated capacities, denoted by matrix $\mathbf{U}_{\bar{k}\times 2^N}$.

```
/* generate vertex capacities */
```

for l *in* $1:\bar{l}$ **do**

- Randomly select a subset of the constraints index set, denoted Q',
- Construct the linear program according to (10) and denote the optimal capacity obtained as μ_l,
- Store μ_l as the l-th row of $\mathbf{V}$.

end

for k *in* $1:\bar{k}$ **do**

- Randomly select t extreme capacity in $\mathbf{V}$, denoted $v_1, ..., v_t$
- Randomly generate t positive nubmers $\lambda_1, ..., \lambda_t$ with their sum equal to 1.
- Set $\mu_k = \lambda_1 v_1 + \cdots + v_t\lambda_t$, `/* convex combination of extreme capacities */`
- Store μ_k as the k-th row of $\mathbf{U}$

end

Return $\mathbf{U}$ and $\mathbf{V}$.

Random Generation & Our works

Our works

LP and convex combination with preference constraints

Table 4: Results of Algorithm 5 with different coefficients.

	Max Orness	Min Orness	Max Entropy	Min Entropy	Max Value	Min Value
$\lvert Q'\rvert = 10, \bar{l} = 100, t = 5$	0.8969550	0.0331210	1.3172188	0.1667198	1.0000000	0.2500000
$\lvert Q'\rvert = 10, \bar{l} = 100, t = 8$	0.8543053	0.1165865	1.3444648	0.4791961	0.9963015	0.2714749
$\lvert Q'\rvert = 10, \bar{l} = 150, t = 5$	0.9146329	0.0663548	1.3080448	0.1603575	1.0000000	0.2500000
$\lvert Q'\rvert = 10, \bar{l} = 150, t = 8$	0.8834968	0.1417478	1.3455874	0.5802006	0.9863421	0.2703662
$\lvert Q'\rvert = 16, \bar{l} = 100, t = 5$	0.9703060	0.0000000	1.2701887	0.0000000	1.0000000	0.2500000
$\lvert Q'\rvert = 16, \bar{l} = 100, t = 8$	0.9319932	0.0467202	1.3480901	0.3359108	1.0000000	0.2770809
$\lvert Q'\rvert = 16, \bar{l} = 150, t = 5$	0.9732916	0.0220063	1.2758855	0.1132478	1.0000000	0.2500000
$\lvert Q'\rvert = 16, \bar{l} = 150, t = 8$	0.9395918	0.0593133	1.3420079	0.3576877	0.9985646	0.2500000

Table 7: Result about the maximum split method

	a	b	c	d	e	f	g	Max Split	PD
Model 6	14.86000	14.25000	13.54000	12.91000	12.17000	11.56000	11.00000	0.56000	0%
Algorithm 2	14.53643	13.92836	13.35919	12.81803	12.20145	11.60074	11.04004	0.54116	3%
Algorithm 4	14.60879	14.04283	13.46777	12.78880	12.15260	11.59637	11.00000	0.55623	1%
Algorithm 5	14.41518	13.91518	13.41518	12.91518	12.37170	11.87170	11.37170	0.50000	11%

Capacity and eXplainable Artificial Intelligence

Update AI Algorithms with Capacity

- Principal Component Analysis
- Decision Tree and Random Forest
- Nonlinear Regression and Logic Regression
- Support Vector Machine and Kernel Transformation
- Artificial Neural Network and Deep Learning

Principal Component Analysis & Our works

Based on the result of PCA, the importance of items i, denoted as RI_i, can be **partially** estimated by:

$$PI_i = \sum_{l=1}^{n_o} \lambda_l |v_{li}| \quad (9)$$

and the **comprehensive/overall** importance of item i can be estimated by:

$$CI_i = \frac{\sum_{l=1}^{n_o} \lambda_l |v_{li}|}{\sum_{i=1}^{n} \sum_{l=1}^{n_o} \lambda_l |v_{li}|}. \quad (10)$$

For the interaction between two original items i, j, it can be **partially** estimated from the correlation coefficients of the scaled data $\overline{S}$ as:

$$PC_{ij} = \frac{cov(\overline{s}_i, \overline{s}_j)}{sd(\overline{s}_i) sd(\overline{s}_j)}, \quad (11)$$

and their **comprehensive** correlation coefficient can be given as:

$$CC_{ij} = \frac{cov((v_{i1}, ..., v_{in_o}), (v_{j1}, ..., v_{jn_o}))}{sd(v_{i1}, ..., v_{in_o}) sd(v_{j1}, ..., v_{jn_o})}, \quad (12)$$

where $cov()$ is the covariance function and $sd()$ is the standard deviation function, hence the correlation coefficients are basically the Pearson correlation coefficients.

Decision Tree and Random Forest & Our works

Algorithm 1 Capacity random forest (CRF) algorithm for ranking decisions.

Input: Decision criteria set N, learning set L, the tree number t of the forest, the largest number k of criteria in a tree.
Output: The capacity random forest and the prediction of new instance.

```
/* Training of the capacity random forest. */
```

for S *in* $\mathcal{P}(N)$ **do**
  if $|S| \geq 2$ *and* $|S| \leq k$ **then**
    Use capacity identification method to get the optimal capacity on S, denoted as μ_S.
    Calculate the performance of μ_S, denoted as $f(\mu_S)$.
  end
  Based on performances $f(\mu_S)$, identify the appearance frequency of each tree $q(\mu_S)$ and randomly generate t trees to get the capacity random forest.
end

```
/* Predicting by the capacity random forest. */
```

for *every new unknow instance* **do**
  Calculate the outcomes of all the decision trees of given new instance.
  Aggregate these outcome into the collective prediction evaluation.
end
Get the final ranking orders of all new instances according to their collective predictions.

Thanks & Questions

Thanks for listening.
Any questions?