Identification of a 2-additive bi-capacity by using mathematical programming

Julien Ah-Pine*, Brice Mayag♯ and Antoine Rolland*

University of Lyon* and University of Paris Dauphine♯

ADT 2013 - Bruxelles - 14th of november 2013

Motivations

- Multi-criteria Decision Making (MCDM) using aggregation operators

Motivations

- Multi-criteria Decision Making (MCDM) using aggregation operators
- The Choquet integral (CI) wrt a capacity is a versatile approach :

Motivations

- Multi-criteria Decision Making (MCDM) using aggregation operators
- The Choquet integral (CI) wrt a capacity is a versatile approach :
  - A capacity allows modelling the interactions between criteria

Motivations

- Multi-criteria Decision Making (MCDM) using aggregation operators
- The Choquet integral (CI) wrt a capacity is a versatile approach :
  - A capacity allows modelling the interactions between criteria
  - A capacity assigns non-negative weights to all subsets of criteria : $2^n - 1$ values to set

Motivations

- Multi-criteria Decision Making (MCDM) using aggregation operators
- The Choquet integral (CI) wrt a capacity is a versatile approach :
  - A capacity allows modelling the interactions between criteria
  - A capacity assigns non-negative weights to all subsets of criteria : $2^n - 1$ values to set
  - The scores should be in a commensurable unipolar (non-negative) scale

Motivations

- Multi-criteria Decision Making (MCDM) using aggregation operators
- The Choquet integral (CI) wrt a capacity is a versatile approach :
  - A capacity allows modelling the interactions between criteria
  - A capacity assigns non-negative weights to all subsets of criteria : $2^n - 1$ values to set
  - The scores should be in a commensurable unipolar (non-negative) scale
- In many situations, human decision makings are easier using bipolar scales (negative, neutral and positive parts) :

Motivations

- Multi-criteria Decision Making (MCDM) using aggregation operators
- The Choquet integral (CI) wrt a capacity is a versatile approach :
  - A capacity allows modelling the interactions between criteria
  - A capacity assigns non-negative weights to all subsets of criteria : $2^n - 1$ values to set
  - The scores should be in a commensurable unipolar (non-negative) scale
- In many situations, human decision makings are easier using bipolar scales (negative, neutral and positive parts) :
  - Bi-capacities (BC) allows extending CI to bipolar scales

Motivations

- Multi-criteria Decision Making (MCDM) using aggregation operators
- The Choquet integral (CI) wrt a capacity is a versatile approach :
  - A capacity allows modelling the interactions between criteria
  - A capacity assigns non-negative weights to all subsets of criteria : $2^n - 1$ values to set
  - The scores should be in a commensurable unipolar (non-negative) scale
- In many situations, human decision makings are easier using bipolar scales (negative, neutral and positive parts) :
  - Bi-capacities (BC) allows extending CI to bipolar scales
  - But the nb of weights to set is bigger : $3^n - 1$

Contributions

- We use the bipolar Choquet integral (BCI) as an aggregation operator in MCDM

Contributions

- We use the bipolar Choquet integral (BCI) as an aggregation operator in MCDM
- Due to the combinatorial complexity, we restrict ourselves to 2-additive BC (2A-BC) which require $2n^2 + 1$ weights to set

Contributions

- We use the bipolar Choquet integral (BCI) as an aggregation operator in MCDM
- Due to the combinatorial complexity, we restrict ourselves to 2-additive BC (2A-BC) which require $2n^2 + 1$ weights to set
- Then, we propose to elicit the 2A-BC modelling the preference model of a decision maker. The elicitation process is based on already judged alternatives (either fictitious or real)

Contributions

- We use the bipolar Choquet integral (BCI) as an aggregation operator in MCDM
- Due to the combinatorial complexity, we restrict ourselves to 2-additive BC (2A-BC) which require $2n^2 + 1$ weights to set
- Then, we propose to elicit the 2A-BC modelling the preference model of a decision maker. The elicitation process is based on already judged alternatives (either fictitious or real)
- We propose optimization problems to elicit the values of a 2A-BC

Contributions

- We use the bipolar Choquet integral (BCI) as an aggregation operator in MCDM
- Due to the combinatorial complexity, we restrict ourselves to 2-additive BC (2A-BC) which require $2n^2 + 1$ weights to set
- Then, we propose to elicit the 2A-BC modelling the preference model of a decision maker. The elicitation process is based on already judged alternatives (either fictitious or real)
- We propose optimization problems to elicit the values of a 2A-BC
- Our framework presents several common points with supervised machine learning tasks and we underline these points

Outline

Outline

Notations

- $N = \{1, \dots, n\}$ is a finite set of n criteria

Notations

- $N = \{1, \ldots, n\}$ is a finite set of n criteria
- x is an alternative defined by its score distribution $x = (x_1, \ldots, x_n)$ where $x_i \in \mathbb{R}$ is the value of x for the criterion i

Notations

- $N = \{1, \ldots, n\}$ is a finite set of n criteria
- x is an alternative defined by its score distribution $x = (x_1, \ldots, x_n)$ where $x_i \in \mathbb{R}$ is the value of x for the criterion i
- $2^N := \{S \subseteq N\}$ is the set of subsets of N

Notations

- $N = \{1, \dots, n\}$ is a finite set of n criteria
- x is an alternative defined by its score distribution $x = (x_1, \dots, x_n)$ where $x_i \in \mathbb{R}$ is the value of x for the criterion i
- $2^N := \{S \subseteq N\}$ is the set of subsets of N
- $3^N := \{(A, B) \in 2^N \times 2^N : A \cap B = \emptyset\}$ is the set of couples of subsets of N with an empty intersection

Notations

- $N = \{1, \ldots, n\}$ is a finite set of n criteria

- x is an alternative defined by its score distribution $x = (x_1, \ldots, x_n)$ where $x_i \in \mathbb{R}$ is the value of x for the criterion i

- $2^N := \{S \subseteq N\}$ is the set of subsets of N
- $3^N := \{(A, B) \in 2^N \times 2^N : A \cap B = \emptyset\}$ is the set of couples of subsets of N with an empty intersection
- On 3^N, the relation $\sqsubseteq$ is such that $(A_1, A_2) \sqsubseteq (B_1, B_2) \Leftrightarrow A_1 \subseteq B_1 \wedge B_2 \subseteq A_2$

Examples of 2^N, 3^N and instances of $\sqsubseteq$

- $N = \{1, 2, 3\}$

Examples of 2^N, 3^N and instances of $\sqsubseteq$

- $N = \{1, 2, 3\}$
- $2^N = \{\emptyset, \{1\}, \{2\}, \{3\}, \{1, 2\}, \{1, 3\}, \{2, 3\}, \{1, 2, 3\}\}$

Examples of 2^N, 3^N and instances of $\sqsubseteq$

- $N = \{1, 2, 3\}$
- $2^N = \{\emptyset, \{1\}, \{2\}, \{3\}, \{1, 2\}, \{1, 3\}, \{2, 3\}, \{1, 2, 3\}\}$
- 3^N is given by pairs indicated by $\surd$ in the following indicator matrix :

(A, B)	$\emptyset$	$\{1\}$	$\{2\}$	$\{3\}$	$\{1,2\}$	$\{1,3\}$	$\{2,3\}$	$\{1,2,3\}$
$\emptyset$	$\surd$	$\surd$	$\surd$	$\surd$	$\surd$	$\surd$	$\surd$	$\surd$
$\{1\}$	$\surd$	.	$\surd$	$\surd$	.	.	$\surd$	.
$\{2\}$	$\surd$	$\surd$	.	$\surd$	.	$\surd$	.	.
$\{3\}$	$\surd$	$\surd$	$\surd$	.	$\surd$	.	.	.
$\{1,2\}$	$\surd$	.	.	$\surd$	.	.	.	.
$\{1,3\}$	$\surd$	.	$\surd$	.	.	.	.	.
$\{2,3\}$	$\surd$	$\surd$	.	.	.	.	.	.
$\{1,2,3\}$	$\surd$	.	.	.	.	.	.	.

Examples of 2^N, 3^N and instances of $\sqsubseteq$

- $N = \{1, 2, 3\}$
- $2^N = \{\emptyset, \{1\}, \{2\}, \{3\}, \{1, 2\}, \{1, 3\}, \{2, 3\}, \{1, 2, 3\}\}$
- 3^N is given by pairs indicated by $\surd$ in the following indicator matrix :

(A, B)	$\emptyset$	$\{1\}$	$\{2\}$	$\{3\}$	$\{1,2\}$	$\{1,3\}$	$\{2,3\}$	$\{1,2,3\}$
$\emptyset$	$\surd$	$\surd$	$\surd$	$\surd$	$\surd$	$\surd$	$\surd$	$\surd$
$\{1\}$	$\surd$	.	$\surd$	$\surd$	.	.	$\surd$	.
$\{2\}$	$\surd$	$\surd$	.	$\surd$	.	$\surd$	.	.
$\{3\}$	$\surd$	$\surd$	$\surd$	.	$\surd$	.	.	.
$\{1,2\}$	$\surd$	.	.	$\surd$	.	.	.	.
$\{1,3\}$	$\surd$	.	$\surd$	.	.	.	.	.
$\{2,3\}$	$\surd$	$\surd$	.	.	.	.	.	.
$\{1,2,3\}$	$\surd$	.	.	.	.	.	.	.

- $(\{1, 2\}, \{3\})$

Examples of 2^N, 3^N and instances of $\sqsubseteq$

- $N = \{1, 2, 3\}$
- $2^N = \{\emptyset, \{1\}, \{2\}, \{3\}, \{1, 2\}, \{1, 3\}, \{2, 3\}, \{1, 2, 3\}\}$
- 3^N is given by pairs indicated by $\surd$ in the following indicator matrix :

(A, B)	$\emptyset$	$\{1\}$	$\{2\}$	$\{3\}$	$\{1,2\}$	$\{1,3\}$	$\{2,3\}$	$\{1,2,3\}$
$\emptyset$	$\surd$	$\surd$	$\surd$	$\surd$	$\surd$	$\surd$	$\surd$	$\surd$
$\{1\}$	$\surd$	.	$\surd$	$\surd$	.	.	$\surd$	.
$\{2\}$	$\surd$	$\surd$	.	$\surd$	.	$\surd$	.	.
$\{3\}$	$\surd$	$\surd$	$\surd$	.	$\surd$	.	.	.
$\{1,2\}$	$\surd$	.	.	$\surd$	.	.	.	.
$\{1,3\}$	$\surd$	.	$\surd$	.	.	.	.	.
$\{2,3\}$	$\surd$	$\surd$	.	.	.	.	.	.
$\{1,2,3\}$	$\surd$	.	.	.	.	.	.	.

- $(\{1, 2\}, \{3\}) \sqsubseteq (\{1, 2\}, \{\emptyset\})$

Examples of 2^N, 3^N and instances of $\sqsubseteq$

- $N = \{1, 2, 3\}$
- $2^N = \{\emptyset, \{1\}, \{2\}, \{3\}, \{1,2\}, \{1,3\}, \{2,3\}, \{1,2,3\}\}$
- 3^N is given by pairs indicated by $\surd$ in the following indicator matrix :

(A, B)	$\emptyset$	$\{1\}$	$\{2\}$	$\{3\}$	$\{1,2\}$	$\{1,3\}$	$\{2,3\}$	$\{1,2,3\}$
$\emptyset$	$\surd$	$\surd$	$\surd$	$\surd$	$\surd$	$\surd$	$\surd$	$\surd$
$\{1\}$	$\surd$	.	$\surd$	$\surd$	.	.	$\surd$	.
$\{2\}$	$\surd$	$\surd$	.	$\surd$	.	$\surd$	.	.
$\{3\}$	$\surd$	$\surd$	$\surd$	.	$\surd$	.	.	.
$\{1,2\}$	$\surd$	.	.	$\surd$	.	.	.	.
$\{1,3\}$	$\surd$	.	$\surd$	.	.	.	.	.
$\{2,3\}$	$\surd$	$\surd$	.	.	.	.	.	.
$\{1,2,3\}$	$\surd$	.	.	.	.	.	.	.

- $(\{1,2\}, \{3\}) \sqsubseteq (\{1,2\}, \{\emptyset\}) \sqsubseteq (\{1,2,3\}, \{\emptyset\})$

Bi-capacity (BC)

- A set function $\nu : 3^N \rightarrow \mathbb{R}$ is a **bi-capacity (BC)** on 3^N if it satisfies the following two conditions [Grabisch and Labreuche, 2005b], [Grabisch and Labreuche, 2008] :

Bi-capacity (BC)

- A set function $\nu : 3^N \to \mathbb{R}$ is a **bi-capacity (BC)** on 3^N if it satisfies the following two conditions [Grabisch and Labreuche, 2005b], [Grabisch and Labreuche, 2008] :
  - Null weight for empty sets :

$$\nu(\emptyset, \emptyset) = 0$$

Bi-capacity (BC)

- A set function $\nu : 3^N \to \mathbb{R}$ is a **bi-capacity (BC)** on 3^N if it satisfies the following two conditions [Grabisch and Labreuche, 2005b], [Grabisch and Labreuche, 2008] :
  - Null weight for empty sets :

$$\nu(\emptyset, \emptyset) = 0$$

  - Monotonicity : $\forall (A_1, A_2), (B_1, B_2) \in 3^N$:

$$(A_1, A_2) \sqsubseteq (B_1, B_2) \Rightarrow \nu(A_1, A_2) \leq \nu(B_1, B_2)$$

Bi-capacity (BC)

- A set function $\nu : 3^N \to \mathbb{R}$ is a **bi-capacity (BC)** on 3^N if it satisfies the following two conditions [Grabisch and Labreuche, 2005b], [Grabisch and Labreuche, 2008] :
  - Null weight for empty sets :

$$\nu(\emptyset, \emptyset) = 0$$

  - Monotonicity : $\forall (A_1, A_2), (B_1, B_2) \in 3^N$:

$$(A_1, A_2) \sqsubseteq (B_1, B_2) \Rightarrow \nu(A_1, A_2) \leq \nu(B_1, B_2)$$

Ex. 1 : $\nu(\{1,2\},\{3\}) \leq \nu(\{1,2\},\{\emptyset\})$ ("less criteria in the - part")

Bi-capacity (BC)

- A set function $\nu\ : 3^N \to \mathbb{R}$ is a **bi-capacity (BC)** on 3^N if it satisfies the following two conditions [Grabisch and Labreuche, 2005b], [Grabisch and Labreuche, 2008] :
  - Null weight for empty sets :

$$\nu(\emptyset, \emptyset) = 0$$

  - Monotonicity : $\forall (A_1, A_2), (B_1, B_2) \in 3^N$:

$$(A_1, A_2) \sqsubseteq (B_1, B_2) \Rightarrow \nu(A_1, A_2) \leq \nu(B_1, B_2)$$

Ex. 1 : $\nu(\{1,2\},\{3\}) \leq \nu(\{1,2\},\{\emptyset\})$ ("less criteria in the - part")
Ex. 2 : $\nu(\{1,2\},\emptyset) \leq \nu(\{1,2,3\},\{\emptyset\})$ ("more criteria in the + part")

Bi-capacity (BC)

- A set function $\nu\ : 3^N \rightarrow \mathbb{R}$ is a **bi-capacity (BC)** on 3^N if it satisfies the following two conditions [Grabisch and Labreuche, 2005b], [Grabisch and Labreuche, 2008] :
  - Null weight for empty sets :

$$\nu(\emptyset, \emptyset) = 0$$

  - Monotonicity : $\forall (A_1, A_2), (B_1, B_2) \in 3^N$:

$$(A_1, A_2) \sqsubseteq (B_1, B_2) \Rightarrow \nu(A_1, A_2) \leq \nu(B_1, B_2)$$

    Ex. 1 : $\nu(\{1,2\},\{3\}) \leq \nu(\{1,2\},\{\emptyset\})$ ("less criteria in the - part")
    Ex. 2 : $\nu(\{1,2\},\emptyset) \leq \nu(\{1,2,3\},\{\emptyset\})$ ("more criteria in the + part")
- ν is said to be normalized if in addition, it holds :

$$\nu(N, \emptyset) = 1 \wedge \nu(\emptyset, N) = -1$$

Bipolar Möbius transform (BMT) of a BC

- A BC ν can be associated to its **bipolar Möbius transform (BMT)** denoted b and defined by [Fujimoto, 2004, Fujimoto, 2007] :

$$
\begin{aligned}
b(A_1, A_2) &:= \sum_{\substack{B_1 \subseteq A_1 \\ B_2 \subseteq A_2}} (-1)^{|A_1 \setminus B_1| + |A_2 \setminus B_2|} \nu(B_1, B_2) \\
&= \sum_{(\emptyset, A_2) \sqsubseteq (B_1, B_2) \sqsubseteq (A_1, \emptyset)} (-1)^{|A_1 \setminus B_1| + |A_2 \setminus B_2|} \nu(B_1, B_2)
\end{aligned}
$$

Bipolar Möbius transform (BMT) of a BC

- A BC ν can be associated to its **bipolar Möbius transform (BMT)** denoted b and defined by [Fujimoto, 2004, Fujimoto, 2007] :

$$
\begin{aligned}
b(A_1, A_2) \quad &:= \sum_{\substack{B_1 \subseteq A_1 \\ B_2 \subseteq A_2}} (-1)^{|A_1 \setminus B_1| + |A_2 \setminus B_2|} \nu(B_1, B_2) \\
&= \sum_{(\emptyset, A_2) \sqsubseteq (B_1, B_2) \sqsubseteq (A_1, \emptyset)} (-1)^{|A_1 \setminus B_1| + |A_2 \setminus B_2|} \nu(B_1, B_2)
\end{aligned}
$$

- There is a one-to-one relation between ν and b and the converse relation is given by :

$$
\nu(A_1, A_2) := \sum_{\substack{B_1 \subseteq A_1 \\ B_2 \subseteq A_2}} b(B_1, B_2)
$$

Bipolar Möbius transform (BMT) of a BC

- A BC ν can be associated to its **bipolar Möbius transform (BMT)** denoted b and defined by [Fujimoto, 2004, Fujimoto, 2007] :

$$\begin{aligned} b(A_1, A_2) &:= \sum_{\substack{B_1 \subseteq A_1 \\ B_2 \subseteq A_2}} (-1)^{|A_1 \setminus B_1| + |A_2 \setminus B_2|} \nu(B_1, B_2) \\ &= \sum_{(\emptyset, A_2) \sqsubseteq (B_1, B_2) \sqsubseteq (A_1, \emptyset)} (-1)^{|A_1 \setminus B_1| + |A_2 \setminus B_2|} \nu(B_1, B_2) \end{aligned}$$

- There is a one-to-one relation between ν and b and the converse relation is given by :

$$\nu(A_1, A_2) := \sum_{\substack{B_1 \subseteq A_1 \\ B_2 \subseteq A_2}} b(B_1, B_2)$$

B1 Note that the property $\nu(\emptyset, \emptyset) = 0$ translates into :

$$b(\emptyset, \emptyset) = 0$$

k-additivity

- A BC requires $3^n - 1$ values to be set

k-additivity

- A BC requires $3^n - 1$ values to be set
- To reduce this complexity, we can use the concept of *k*-additivity [Grabisch and Labreuche, 2005a, Grabisch and Labreuche, 2005b] and [Grabisch and Labreuche, 2008]

k-additivity

- A BC requires $3^n - 1$ values to be set
- To reduce this complexity, we can use the concept of k-additivity [Grabisch and Labreuche, 2005a, Grabisch and Labreuche, 2005b] and [Grabisch and Labreuche, 2008]
- Given a positive integer $k < n$, a BC ν is k-**additive** iff the two following conditions are satisfied [Fujimoto, 2007] :

k-additivity

- A BC requires $3^n - 1$ values to be set
- To reduce this complexity, we can use the concept of k-additivity [Grabisch and Labreuche, 2005a, Grabisch and Labreuche, 2005b] and [Grabisch and Labreuche, 2008]
- Given a positive integer $k < n$, a BC ν is k-**additive** iff the two following conditions are satisfied [Fujimoto, 2007] :

  B2 $b = 0$ if the nb of criteria involved in (A_1, A_2) is greater than k :

$$\forall (A_1, A_2) \in 3^N : |A_1 \cup A_2| > k \Rightarrow b(A_1, A_2) = 0$$

k-additivity

- A BC requires $3^n - 1$ values to be set
- To reduce this complexity, we can use the concept of k-additivity [Grabisch and Labreuche, 2005a, Grabisch and Labreuche, 2005b] and [Grabisch and Labreuche, 2008]
- Given a positive integer $k < n$, a BC ν is k-**additive** iff the two following conditions are satisfied [Fujimoto, 2007] :

  B2 $b = 0$ if the nb of criteria involved in (A_1, A_2) is greater than k :

$$\forall (A_1, A_2) \in 3^N : |A_1 \cup A_2| > k \Rightarrow b(A_1, A_2) = 0$$

  B3 $b \neq 0$ for at least one pair such that the nb of criteria involved is k :

$$\exists (A_1, A_2) \in 3^N : |A_1 \cup A_2| = k \wedge b(A_1, A_2) \neq 0$$

2-additive BC

- 1-additive BC assume the criteria are independent and boil down to linear aggregation operators

2-additive BC

- 1-additive BC assume the criteria are independent and boil down to linear aggregation operators
- Thus we deal with 2-additive BC (2A-BC) to have a richer aggregation function. Henceforth, b is the BMT of a 2A-BC ν

2-additive BC

- 1-additive BC assume the criteria are independent and boil down to linear aggregation operators
- Thus we deal with 2-additive BC (2A-BC) to have a richer aggregation function. Henceforth, b is the BMT of a 2A-BC ν
- In this case, only subsets with at most two criteria matter

2-additive BC

- 1-additive BC assume the criteria are independent and boil down to linear aggregation operators
- Thus we deal with 2-additive BC (2A-BC) to have a richer aggregation function. Henceforth, b is the BMT of a 2A-BC ν
- In this case, only subsets with at most two criteria matter
- To lighten the notations we will take :

$$b_{i|} = b(\{i\}, \emptyset) \quad ; \quad b_{ij|} = b(\{i, j\}, \emptyset) \quad ; \quad b_{i|j} = b(\{i\}, \{j\})$$

Example of the BMT of a 2A-BC

- $N = \{1, 2, 3\}$
- We assume b is a BMT of a 2A-BC ν
- $\surd$ are (possibly) non null elements and x are (necessarily) null elements :

$$
\begin{array}{cc}
 & \begin{array}{cccccccc} \emptyset & 1 & 2 & 3 & 12 & 13 & 23 & 123 \end{array} \\
b = \begin{array}{c} (A,B) \\ \emptyset \\ 1 \\ 2 \\ 3 \\ 12 \\ 13 \\ 23 \\ 123 \end{array} &
\left(\begin{array}{cccccccc}
\surd & \surd & \surd & \surd & \surd & \surd & \surd & x \\
\surd & . & \surd & \surd & . & . & x & . \\
\surd & \surd & . & \surd & . & x & . & . \\
\surd & \surd & \surd & . & x & . & . & . \\
\surd & . & . & x & . & . & . & . \\
\surd & . & x & . & . & . & . & . \\
\surd & x & . & . & . & . & . & . \\
x & . & . & . & . & . & . & .
\end{array} \right)
\end{array}
$$

Example of the BMT of a 2A-BC

- $N = \{1, 2, 3\}$
- We assume b is a BMT of a 2A-BC ν
- $\surd$ are (possibly) non null elements and x are (necessarily) null elements :

$b =$

(A, B)	$\emptyset$	1	2	3	12	13	23	123
$\emptyset$	$\surd$	$\surd$	$\surd$	$\surd$	$\surd$	$\surd$	$\surd$	x
1	$\surd$	.	$\surd$	$\surd$	.	.	x	.
2	$\surd$	$\surd$	.	$\surd$	.	x	.	.
3	$\surd$	$\surd$	$\surd$	.	x	.	.	.
12	$\surd$	.	.	x	.	.	.	.
13	$\surd$	.	x	.	.	.	.	.
23	$\surd$	x	.	.	.	.	.	.
123	x	.	.	.	.	.	.	.

- The nb of elements to be set for b reduces from 3^n to $2n^2 + 1$ (27 vs 19 in this example)

2-additive BC (cont'd)

P1 The **monotonicity** property reduces to [Mayag et al, 2012] :

$$\forall (A,B) \in 3^N, \forall k \in A : b_{k|} + \sum_{j \in B} b_{k|j} + \sum_{i \in A \setminus k} b_{ik|} \geq 0$$

$$\forall (A,B) \in 3^N, \forall k \in A : b_{|k} + \sum_{j \in B} b_{j|k} + \sum_{i \in A \setminus k} b_{|ik} \leq 0$$

2-additive BC (cont'd)

P1 The **monotonicity** property reduces to [Mayag et al, 2012] :

$$\forall (A,B) \in 3^N, \forall k \in A : b_{k|} + \sum_{j \in B} b_{k|j} + \sum_{i \in A \setminus k} b_{ik|} \geq 0$$

$$\forall (A,B) \in 3^N, \forall k \in A : b_{|k} + \sum_{j \in B} b_{j|k} + \sum_{i \in A \setminus k} b_{|ik} \leq 0$$

P2 The **normalization** property reduces to [Mayag et al, 2012] :

$$\sum_{i \in N} b_{i|} + \sum_{\{i,j\} \subseteq N} b_{ij|} = 1$$

$$\sum_{i \in N} b_{|i} + \sum_{\{i,j\} \subseteq N} b_{|ij} = -1$$

Bipolar Choquet integral (BCI) wrt a 2A-BC

P3 The **bipolar Choquet integral (BCI)** wrt b denoted $\mathcal{C}_b$ is given by :

$$\begin{aligned}\mathcal{C}_b(x) &= \sum_{i=1}^{n} b_{i|}\ x_i^+ + \sum_{i=1}^{n} b_{|i}\ x_i^- + \sum_{i,j=1}^{n} b_{i|j}\ (x_i^+ \wedge x_j^-) \\ &\quad + \sum_{\{i,j\}\subseteq N} b_{ij|}\ (x_i^+ \wedge x_j^+) + \sum_{\{i,j\}\subseteq N} b_{|ij}\ (x_i^- \wedge x_j^-)\end{aligned}$$

where $x_i^+ = \begin{cases} x_i & \text{if } x_i > 0 \\ 0 & \text{if } x_i \leq 0 \end{cases}$ and $x_i^- = \begin{cases} -x_i & \text{if } x_i < 0 \\ 0 & \text{if } x_i \geq 0 \end{cases}$

and $a \wedge b = \min(a, b)$

Outline

Elicitation process

- The preferences model of the DM is modeled by a 2A-BC ν represented by its BMT b

Elicitation process

- The preferences model of the DM is modeled by a 2A-BC ν represented by its BMT b
- The DM provides a subset of alternatives $X' \subseteq X$ and $\forall x \in X'$ we are given :

Elicitation process

- The preferences model of the DM is modeled by a 2A-BC ν represented by its BMT b
- The DM provides a subset of alternatives $X' \subseteq X$ and $\forall x \in X'$ we are given :
  - the scores distribution $x = (x_1, \ldots, x_n)$
  - the overall score $S(x)$ that is the aggregated value for x

Elicitation process

- The preferences model of the DM is modeled by a 2A-BC ν represented by its BMT b
- The DM provides a subset of alternatives $X' \subseteq X$ and $\forall x \in X'$ we are given :
  - the scores distribution $x = (x_1, \ldots, x_n)$
  - the overall score $S(x)$ that is the aggregated value for x
- Let us denote by T the set of pairs $\{((x_1, \ldots, x_n), S(x))\}_{x \in X'}$

Elicitation process

- The preferences model of the DM is modeled by a 2A-BC ν represented by its BMT b
- The DM provides a subset of alternatives $X' \subseteq X$ and $\forall x \in X'$ we are given :
  - the scores distribution $x = (x_1, \ldots, x_n)$
  - the overall score $S(x)$ that is the aggregated value for x
- Let us denote by T the set of pairs $\{((x_1, \ldots, x_n), S(x))\}_{x \in X'}$
- No extra information is provided

Elicitation process

- The preferences model of the DM is modeled by a 2A-BC ν represented by its BMT b
- The DM provides a subset of alternatives $X' \subseteq X$ and $\forall x \in X'$ we are given :
  - the scores distribution $x = (x_1, \dots, x_n)$
  - the overall score $S(x)$ that is the aggregated value for x
- Let us denote by T the set of pairs $\{((x_1, \dots, x_n), S(x))\}_{x \in X'}$
- No extra information is provided
- No interaction loop with the DM is assumed

Taking into account preference relations

- Our goal is to identify a 2A-BC that is consistent with T

Taking into account preference relations

- Our goal is to identify a 2A-BC that is consistent with T
- Being consistent with T implies that we need to respect the preference relations provided by T as much as possible

Taking into account preference relations

- Our goal is to identify a 2A-BC that is consistent with T
- Being consistent with T implies that we need to respect the preference relations provided by T as much as possible

C1 The BCI $\mathcal{C}_b$ wrt b should satisfy the given preference relations :

$$\forall x, x' \in X', x \neq x' : S(x) - S(x') \geq 0 \Rightarrow \mathcal{C}_b(x) - \mathcal{C}_b(x') \geq 0$$

Taking into account preference relations

- Our goal is to identify a 2A-BC that is consistent with T
- Being consistent with T implies that we need to respect the preference relations provided by T as much as possible

C1 The BCI $\mathcal{C}_b$ wrt b should satisfy the given preference relations :

$$\forall x, x' \in X', x \neq x' : S(x) - S(x') \geq 0 \Rightarrow \mathcal{C}_b(x) - \mathcal{C}_b(x') \geq 0$$

- However it might happen that it is impossible to satisfy these constraints for all pairs :

Taking into account preference relations

- Our goal is to identify a 2A-BC that is consistent with T
- Being consistent with T implies that we need to respect the preference relations provided by T as much as possible

C1 The BCI $\mathcal{C}_b$ wrt b should satisfy the given preference relations :

$$\forall x, x' \in X', x \neq x' : S(x) - S(x') \geq 0 \Rightarrow \mathcal{C}_b(x) - \mathcal{C}_b(x') \geq 0$$

- However it might happen that it is impossible to satisfy these constraints for all pairs :
  - Our restriction to 2A-BC might be too strong and more general BC could better fit the problem

Taking into account preference relations

- Our goal is to identify a 2A-BC that is consistent with T
- Being consistent with T implies that we need to respect the preference relations provided by T as much as possible

C1 The BCI $\mathcal{C}_b$ wrt b should satisfy the given preference relations :

$$\forall x, x' \in X', x \neq x' : S(x) - S(x') \geq 0 \Rightarrow \mathcal{C}_b(x) - \mathcal{C}_b(x') \geq 0$$

- However it might happen that it is impossible to satisfy these constraints for all pairs :
  - Our restriction to 2A-BC might be too strong and more general BC could better fit the problem
  - The DM could provide scores and preferences that present incoherences and which are not fully representable by any BC

Taking into account preference relations (cont'd)

C2 If C1 cannot be satisfied, we replace it with more flexible constraints :

$$\forall x, x' \in X', x \neq x' : S(x) - S(x') \geq 0 \Rightarrow \mathcal{C}_b(x) - \mathcal{C}_b(x') \geq -\xi_{xx'}$$

where $\xi_{xx'} \geq 0$ are **slack variables** of the model

Taking into account preference relations (cont'd)

C2 If C1 cannot be satisfied, we replace it with more flexible constraints :

$$\forall x, x' \in X', x \neq x' : S(x) - S(x') \geq 0 \Rightarrow \mathcal{C}_b(x) - \mathcal{C}_b(x') \geq -\xi_{xx'}$$

where $\xi_{xx'} \geq 0$ are **slack variables** of the model

- $\xi_{xx'}$ allow inconsistencies :

Taking into account preference relations (cont'd)

C2 If C1 cannot be satisfied, we replace it with more flexible constraints :

$$\forall x, x' \in X', x \neq x' : S(x) - S(x') \geq 0 \Rightarrow \mathcal{C}_b(x) - \mathcal{C}_b(x') \geq -\xi_{xx'}$$

where $\xi_{xx'} \geq 0$ are **slack variables** of the model

- $\xi_{xx'}$ allow inconsistencies :
  - if $\xi_{xx'} = 0$ then b provides no incoherence with (x, x')

Taking into account preference relations (cont'd)

C2 If C1 cannot be satisfied, we replace it with more flexible constraints :

$$\forall x, x' \in X', x \neq x' : S(x) - S(x') \geq 0 \Rightarrow \mathcal{C}_b(x) - \mathcal{C}_b(x') \geq -\xi_{xx'}$$

where $\xi_{xx'} \geq 0$ are **slack variables** of the model

- $\xi_{xx'}$ allow inconsistencies :
  - if $\xi_{xx'} = 0$ then b provides no incoherence with (x, x')
  - if $0 < \xi_{xx'}$ then b cannot reproduce the preference relation for (x, x')

Taking into account preference relations (cont'd)

C2 If C1 cannot be satisfied, we replace it with more flexible constraints :

$$\forall x, x' \in X', x \neq x' : S(x) - S(x') \geq 0 \Rightarrow \mathcal{C}_b(x) - \mathcal{C}_b(x') \geq -\xi_{xx'}$$

where $\xi_{xx'} \geq 0$ are **slack variables** of the model

- $\xi_{xx'}$ allow inconsistencies :
  - if $\xi_{xx'} = 0$ then b provides no incoherence with (x, x')
  - if $0 < \xi_{xx'}$ then b cannot reproduce the preference relation for (x, x')

L1 We could also impose $\mathcal{C}_b$ to be bounded :

$$\forall x \in X' \ : lb \leq \mathcal{C}_b(x) \leq ub$$

The split approach

C3 We add a variable $\varepsilon \geq 0$ that reflects the difference between $\mathcal{C}_b(x)$ and $\mathcal{C}_b(x')$:

$$\forall x, x' \in X', x \neq x' : S(x) - S(x') \geq 0 \Rightarrow \mathcal{C}_b(x) - \mathcal{C}_b(x') \geq \varepsilon$$

The split approach

C3 We add a variable $\varepsilon \geq 0$ that reflects the difference between $\mathcal{C}_b(x)$ and $\mathcal{C}_b(x')$:

$$\forall x, x' \in X', x \neq x' : S(x) - S(x') \geq 0 \Rightarrow \mathcal{C}_b(x) - \mathcal{C}_b(x') \geq \varepsilon$$

- We want ε to be as big as possible :

$$\max \varepsilon$$

The split approach

C3 We add a variable $\varepsilon \geq 0$ that reflects the difference between $\mathcal{C}_b(x)$ and $\mathcal{C}_b(x')$:

$$\forall x, x' \in X', x \neq x' : S(x) - S(x') \geq 0 \Rightarrow \mathcal{C}_b(x) - \mathcal{C}_b(x') \geq \varepsilon$$

- We want ε to be as big as possible :

$$\max \varepsilon$$

- subject to :
  - B1 : $b(\emptyset, \emptyset) = 0$
  - B2 : 2-additivity
  - P1 : monotonicity
  - P2 : normalization
  - P3 : computation of $\mathcal{C}_b$
  - L1 : $\mathcal{C}_b$ bounded
  - C3 : preference relations "with ε"

The split approach

C3 We add a variable $\varepsilon \geq 0$ that reflects the difference between $\mathcal{C}_b(x)$ and $\mathcal{C}_b(x')$:

$$\forall x, x' \in X', x \neq x' : S(x) - S(x') \geq 0 \Rightarrow \mathcal{C}_b(x) - \mathcal{C}_b(x') \geq \varepsilon$$

- We want ε to be as big as possible :

$$\max \varepsilon$$

- subject to :
  - B1 : $b(\emptyset, \emptyset) = 0$
  - B2 : 2-additivity
  - P1 : monotonicity
  - P2 : normalization
  - P3 : computation of $\mathcal{C}_b$
  - L1 : $\mathcal{C}_b$ bounded
  - C3 : preference relations "with ε"
- This is a linear program

A flexible version of the split approach

- In case of inconsistencies, we use the following objective function :

$$\max \left(\varepsilon - \sum_{x,x' : S(x) \geq S(x')} \xi_{xx'} \right)$$

- subject to :
  - B1 : $b(\emptyset, \emptyset) = 0$
  - B2 : 2-additivity
  - P1 : monotonicity
  - P2 : normalization
  - P3 : computation of $\mathcal{C}_b$
  - L1 : $\mathcal{C}_b$ bounded
  - C4 : preference relations "with $\varepsilon - \xi_{xx'}$"

$$\forall x, x' \in X', x \neq x' : S(x) - S(x') \geq 0 \Rightarrow \mathcal{C}_b(x) - \mathcal{C}_b(x') \geq \varepsilon - \xi_{xx'}$$

A flexible version of the split approach

- In case of inconsistencies, we use the following objective function :

$$\max\left(\varepsilon - \sum_{x,x':S(x)\geq S(x')} \xi_{xx'}\right)$$

- subject to :
  - B1 : $b(\emptyset,\emptyset)=0$
  - B2 : 2-additivity
  - P1 : monotonicity
  - P2 : normalization
  - P3 : computation of $\mathcal{C}_b$
  - L1 : $\mathcal{C}_b$ bounded
  - C4 : preference relations "with $\varepsilon - \xi_{xx'}$"

$$\forall x,x' \in X', x \neq x' : S(x) - S(x') \geq 0 \Rightarrow \mathcal{C}_b(x) - \mathcal{C}_b(x') \geq \varepsilon - \xi_{xx'}$$

- This is a linear program

Regression like approach

- We could also minimize the residual sum of squares
- We define the following objective function :

$$\min \sum_{x \in X'} (S(x) - \mathcal{C}_b(x))^2$$

Regression like approach

- We could also minimize the residual sum of squares
- We define the following objective function :

$$\min \sum_{x \in X'} (S(x) - \mathcal{C}_b(x))^2$$

- subject to :
  - B1 : $b(\emptyset, \emptyset) = 0$
  - B2 : 2-additivity
  - P1 : monotonicity
  - P2 : normalization
  - P3 : computation of $\mathcal{C}_b$
  - L1 : $\mathcal{C}_b$ bounded
  - C1 : preference relations

Regression like approach

- We could also minimize the residual sum of squares
- We define the following objective function :

$$\min \sum_{x \in X'} (S(x) - \mathcal{C}_b(x))^2$$

- subject to :
  - B1 : $b(\emptyset, \emptyset) = 0$
  - B2 : 2-additivity
  - P1 : monotonicity
  - P2 : normalization
  - P3 : computation of $\mathcal{C}_b$
  - L1 : $\mathcal{C}_b$ bounded
  - C1 : preference relations
- This is a quadratic program

A flexible version of the regression like approach

- In case of incoherences, we optimize the following objective function :

$$\min \left(\sum_{x \in X'} (S(x) - \mathcal{C}_b(x))^2 + \sum_{x,x' : S(x) \geq S(x')} \xi_{xx'} \right)$$

- subject to :
  - B1 : $b(\emptyset, \emptyset) = 0$
  - B2 : 2-additivity
  - P1 : monotonicity
  - P2 : normalization
  - P3 : computation of $\mathcal{C}_b$
  - L1 : $\mathcal{C}_b$ bounded
  - C2 : preference relations "with $-\xi_{xx'}$"

A flexible version of the regression like approach

- In case of incoherences, we optimize the following objective function :

$$\min \left(\sum_{x \in X'} (S(x) - \mathcal{C}_b(x))^2 + \sum_{x,x' : S(x) \geq S(x')} \xi_{xx'} \right)$$

- subject to :
  - B1 : $b(\emptyset, \emptyset) = 0$
  - B2 : 2-additivity
  - P1 : monotonicity
  - P2 : normalization
  - P3 : computation of $\mathcal{C}_b$
  - L1 : $\mathcal{C}_b$ bounded
  - C2 : preference relations “with $-\xi_{xx'}$”
- This is a quadratic program

Outline

An example without incoherence

- It concerns the grades (scores) obtained by 7 students (alternatives) for $n = 5$ subjects (criteria) : statistics (S), probability (P), economics (E), management (M), and English (En).
- The scores are given in a bipolar scale $[-4, 4]$

An example without incoherence

- It concerns the grades (scores) obtained by 7 students (alternatives) for $n = 5$ subjects (criteria) : statistics (S), probability (P), economics (E), management (M), and English (En).
- The scores are given in a bipolar scale $[-4, 4]$

Stu.	S	P	E	M	En
a	4	-3	-3	-3	4
b	4	-3	4	-3	-3
c	-3	-3	4	-3	4
d	4	4	-3	-3	-3
e	-3	-3	4	4	-3
f	-3	-3	4	-3	-3
g	-3	-3	-3	-3	4

An example without incoherence

- It concerns the grades (scores) obtained by 7 students (alternatives) for $n = 5$ subjects (criteria) : statistics (S), probability (P), economics (E), management (M), and English (En).
- The scores are given in a bipolar scale $[-4, 4]$

Stu.	S	P	E	M	En	S
a	4	-3	-3	-3	4	1
b	4	-3	4	-3	-3	0.5
c	-3	-3	4	-3	4	0
d	4	4	-3	-3	-3	-0.5
e	-3	-3	4	4	-3	-1
f	-3	-3	4	-3	-3	-1.5
g	-3	-3	-3	-3	4	-2

An example without incoherence

- It concerns the grades (scores) obtained by 7 students (alternatives) for $n = 5$ subjects (criteria) : statistics (S), probability (P), economics (E), management (M), and English (En).
- The scores are given in a bipolar scale $[-4, 4]$

Stu.	S	P	E	M	En	*S*	*split*
a	4	-3	-3	-3	4	1	1.68
b	4	-3	4	-3	-3	0.5	1.04
c	-3	-3	4	-3	4	0	0.41
d	4	4	-3	-3	-3	-0.5	-0.23
e	-3	-3	4	4	-3	-1	-0.86
f	-3	-3	4	-3	-3	-1.5	-1.5
g	-3	-3	-3	-3	4	-2	-2.14

An example without incoherence

- It concerns the grades (scores) obtained by 7 students (alternatives) for $n = 5$ subjects (criteria) : statistics (S), probability (P), economics (E), management (M), and English (En).
- The scores are given in a bipolar scale $[-4, 4]$

Stu.	S	P	E	M	En	*S*	*split*	*split flex*
a	4	-3	-3	-3	4	1	1.68	1.02
b	4	-3	4	-3	-3	0.5	1.04	0.38
c	-3	-3	4	-3	4	0	0.41	-0.25
d	4	4	-3	-3	-3	-0.5	-0.23	-0.89
e	-3	-3	4	4	-3	-1	-0.86	-1.53
f	-3	-3	4	-3	-3	-1.5	-1.5	-2.16
g	-3	-3	-3	-3	4	-2	-2.14	-2.8

An example without incoherence

- It concerns the grades (scores) obtained by 7 students (alternatives) for $n = 5$ subjects (criteria) : statistics (S), probability (P), economics (E), management (M), and English (En).
- The scores are given in a bipolar scale $[-4, 4]$

Stu.	S	P	E	M	En	*S*	*split*	*split flex*	*rss*
a	4	-3	-3	-3	4	1	1.68	1.02	1
b	4	-3	4	-3	-3	0.5	1.04	0.38	0.5
c	-3	-3	4	-3	4	0	0.41	-0.25	0
d	4	4	-3	-3	-3	-0.5	-0.23	-0.89	-0.5
e	-3	-3	4	4	-3	-1	-0.86	-1.53	-1
f	-3	-3	4	-3	-3	-1.5	-1.5	-2.16	-1.5
g	-3	-3	-3	-3	4	-2	-2.14	-2.8	-2

An example without incoherence

- It concerns the grades (scores) obtained by 7 students (alternatives) for $n = 5$ subjects (criteria) : statistics (S), probability (P), economics (E), management (M), and English (En).
- The scores are given in a bipolar scale $[-4, 4]$

Stu.	S	P	E	M	En	*S*	*split*	*split flex*	*rss*	*rss flex*
a	4	-3	-3	-3	4	1	1.68	1.02	1	1
b	4	-3	4	-3	-3	0.5	1.04	0.38	0.5	0.5
c	-3	-3	4	-3	4	0	0.41	-0.25	0	0
d	4	4	-3	-3	-3	-0.5	-0.23	-0.89	-0.5	-0.5
e	-3	-3	4	4	-3	-1	-0.86	-1.53	-1	-1
f	-3	-3	4	-3	-3	-1.5	-1.5	-2.16	-1.5	-1.5
g	-3	-3	-3	-3	4	-2	-2.14	-2.8	-2	-2

An example with an incoherence

- We change S into S' and the only difference is $S'(g) = 0.5$ while $S(g) = -2$. There is an incoherence between c and g : $S'(g) > S'(c)$ while g is Pareto dominated by c

Stu.	S	P	E	M	En	S'
a	4	-3	-3	-3	4	1
b	4	-3	4	-3	-3	0.5
c	-3	-3	4	-3	4	0
d	4	4	-3	-3	-3	-0.5
e	-3	-3	4	4	-3	-1
f	-3	-3	4	-3	-3	-1.5
g	-3	-3	-3	-3	4	**0.5**

An example with an incoherence

- We change S into S' and the only difference is $S'(g) = 0.5$ while $S(g) = -2$. There is an incoherence between c and g : $S'(g) > S'(c)$ while g is Pareto dominated by c

Stu.	S	P	E	M	En	S'	*split*
a	4	-3	-3	-3	4	1	.
b	4	-3	4	-3	-3	0.5	.
c	-3	-3	4	-3	4	0	.
d	4	4	-3	-3	-3	-0.5	.
e	-3	-3	4	4	-3	-1	.
f	-3	-3	4	-3	-3	-1.5	.
g	-3	-3	-3	-3	4	**0.5**	.

An example with an incoherence

- We change S into S' and the only difference is $S'(g) = 0.5$ while $S(g) = -2$. There is an incoherence between c and g : $S'(g) > S'(c)$ while g is Pareto dominated by c

Stu.	S	P	E	M	En	S'	*split*	*split flex*
a	4	-3	-3	-3	4	1	.	0.22
b	4	-3	4	-3	-3	0.5	.	-0.28
c	-3	-3	4	-3	4	0	.	-0.78
d	4	4	-3	-3	-3	-0.5	.	-1.28
e	-3	-3	4	4	-3	-1	.	-1.78
f	-3	-3	4	-3	-3	-1.5	.	-2.28
g	-3	-3	-3	-3	4	**0.5**	.	-0.78

An example with an incoherence

- We change S into S' and the only difference is $S'(g) = 0.5$ while $S(g) = -2$. There is an incoherence between c and g : $S'(g) > S'(c)$ while g is Pareto dominated by c

Stu.	S	P	E	M	En	S'	*split*	*split flex*	*rss*
a	4	-3	-3	-3	4	1	.	0.22	.
b	4	-3	4	-3	-3	0.5	.	-0.28	.
c	-3	-3	4	-3	4	0	.	-0.78	.
d	4	4	-3	-3	-3	-0.5	.	-1.28	.
e	-3	-3	4	4	-3	-1	.	-1.78	.
f	-3	-3	4	-3	-3	-1.5	.	-2.28	.
g	-3	-3	-3	-3	4	**0.5**	.	-0.78	.

An example with an incoherence

- We change S into S' and the only difference is $S'(g) = 0.5$ while $S(g) = -2$. There is an incoherence between c and g : $S'(g) > S'(c)$ while g is Pareto dominated by c

Stu.	S	P	E	M	En	S'	*split*	*split flex*	*rss*	*rss flex*
a	4	-3	-3	-3	4	1	.	0.22	.	1.12
b	4	-3	4	-3	-3	0.5	.	-0.28	.	0.62
c	-3	-3	4	-3	4	0	.	-0.78	.	0.12
d	4	4	-3	-3	-3	-0.5	.	-1.28	.	-0.5
e	-3	-3	4	4	-3	-1	.	-1.78	.	-1
f	-3	-3	4	-3	-3	-1.5	.	-2.28	.	-1.5
g	-3	-3	-3	-3	4	**0.5**	.	-0.78	.	0.12

An example with an incoherence

- We change S into S' and the only difference is $S'(g) = 0.5$ while $S(g) = -2$. There is an incoherence between c and g : $S'(g) > S'(c)$ while g is Pareto dominated by c

Stu.	S	P	E	M	En	S'	*split*	*split flex*	*rss*	*rss flex*
a	4	-3	-3	-3	4	1	.	0.22	.	1.12
b	4	-3	4	-3	-3	0.5	.	-0.28	.	0.62
c	-3	-3	4	-3	4	0	.	-0.78	.	0.12
d	4	4	-3	-3	-3	-0.5	.	-1.28	.	-0.5
e	-3	-3	4	4	-3	-1	.	-1.78	.	-1
f	-3	-3	4	-3	-3	-1.5	.	-2.28	.	-1.5
g	-3	-3	-3	-3	4	**0.5**	.	-0.78	.	0.12

- For *split flex* and *rss flex*, $\xi_{gc} > 0$ while for other pairs the slack variables are null

Outline

Relationships with supervised Learning (SL)

The addressed task has many similarities with supervised learning (SL) tasks in machine learning (ML)

Relationships with supervised Learning (SL)

The addressed task has many similarities with supervised learning (SL) tasks in machine learning (ML)

- The set T is similar as a training set in SL

Relationships with supervised Learning (SL)

The addressed task has many similarities with supervised learning (SL) tasks in machine learning (ML)

- The set T is similar as a training set in SL
- Eliciting the parameters of the preference model (b) and the aggregation operator ($\mathcal{C}_b$) is similar as inferring the parameters of a SL model (like the coefficients of the linear regression)

Relationships with supervised Learning (SL)

The addressed task has many similarities with supervised learning (SL) tasks in machine learning (ML)

- The set T is similar as a training set in SL
- Eliciting the parameters of the preference model (b) and the aggregation operator ($\mathcal{C}_b$) is similar as inferring the parameters of a SL model (like the coefficients of the linear regression)
- Allowing incoherences in our elicitation framework is similar as permitting errors in SL models. Thus the flexible versions of the split and the regression like methods are similar to SL approaches

Relationships with supervised Learning (SL)

The addressed task has many similarities with supervised learning (SL) tasks in machine learning (ML)

- The set T is similar as a training set in SL
- Eliciting the parameters of the preference model (b) and the aggregation operator ($\mathcal{C}_b$) is similar as inferring the parameters of a SL model (like the coefficients of the linear regression)
- Allowing incoherences in our elicitation framework is similar as permitting errors in SL models. Thus the flexible versions of the split and the regression like methods are similar to SL approaches
- 2A-BC is a less vast family of preference models than unconstrained BC. Thus constraining the BC to be 2-additive is like choosing a family of hypothesis with potentially greater bias but lower variance (we expect better generalization for unseen alternatives)

Relationships with supervised Learning (SL) (cont'd)

- As for 2-additivity, we integrate the constraint B2 but not B3 :

Relationships with supervised Learning (SL) (cont'd)

- As for 2-additivity, we integrate the constraint B2 but not B3 :

  B2 $b = 0$ if the nb of criteria involved in (A_1, A_2) is greater than k :

$$\forall (A_1, A_2) \in 3^N : |A_1 \cup A_2| > k \Rightarrow b(A_1, A_2) = 0$$

  B3 $b \neq 0$ for at least one pair such that the nb of criteria involved is k :

$$\exists (A_1, A_2) \in 3^N : |A_1 \cup A_2| = k \wedge b(A_1, A_2) \neq 0$$

Relationships with supervised Learning (SL) (cont'd)

- As for 2-additivity, we integrate the constraint B2 but not B3 :

  B2 $b = 0$ if the nb of criteria involved in (A_1, A_2) is greater than k :

$$\forall (A_1, A_2) \in 3^N : |A_1 \cup A_2| > k \Rightarrow b(A_1, A_2) = 0$$

  B3 $b \neq 0$ for at least one pair such that the nb of criteria involved is k :

$$\exists (A_1, A_2) \in 3^N : |A_1 \cup A_2| = k \wedge b(A_1, A_2) \neq 0$$

- Thus, the model can be a 1-additive BC and not necessarily a 2A-BC

Relationships with supervised Learning (SL) (cont'd)

- As for 2-additivity, we integrate the constraint B2 but not B3 :

  B2 $b = 0$ if the nb of criteria involved in (A_1, A_2) is greater than k :

$$\forall (A_1, A_2) \in 3^N : |A_1 \cup A_2| > k \Rightarrow b(A_1, A_2) = 0$$

  B3 $b \neq 0$ for at least one pair such that the nb of criteria involved is k :

$$\exists (A_1, A_2) \in 3^N : |A_1 \cup A_2| = k \wedge b(A_1, A_2) \neq 0$$

- Thus, the model can be a 1-additive BC and not necessarily a 2A-BC
- But, this is not guaranteed : we need to add a penalty term to favor sparse solutions

Relationships with supervised Learning (SL) (cont'd)

- As for 2-additivity, we integrate the constraint B2 but not B3 :
  - B2 $b = 0$ if the nb of criteria involved in (A_1, A_2) is greater than k :

$$\forall (A_1, A_2) \in 3^N : |A_1 \cup A_2| > k \Rightarrow b(A_1, A_2) = 0$$

  - B3 $b \neq 0$ for at least one pair such that the nb of criteria involved is k :

$$\exists (A_1, A_2) \in 3^N : |A_1 \cup A_2| = k \wedge b(A_1, A_2) \neq 0$$

- Thus, the model can be a 1-additive BC and not necessarily a 2A-BC
- But, this is not guaranteed : we need to add a penalty term to favor sparse solutions
- The slack variables $\xi_{xx'}$ and the constraints C1, C2, C3, C4 have relationships with SVM (margin, structured SVM)

Outline

Conclusion and future work

- We proposed to use the BCI wrt a 2A-BC as a preference model in MCDM
- We introduced two kinds of optimization problems to elicit a 2A-BC
- Our models allows dealing with inconsistencies
- Our setting and elicitation model has several common points with supervised learning

Conclusion and future work

- We proposed to use the BCI wrt a 2A-BC as a preference model in MCDM
- We introduced two kinds of optimization problems to elicit a 2A-BC
- Our models allows dealing with inconsistencies
- Our setting and elicitation model has several common points with supervised learning

- As for ongoing and future work :
  - Further exploit ML concepts in MCDM like adding a penalty term to have sparse b
  - Have a better understanding of the behavior of b and ν provided by the different methods
  - Extend our elicitation framework to integrate other information provided by the DM such as the importance or the interaction between criteria

Some references

Grabisch, M., Labreuche, C. :
Bi-capacities-I : definition, Möbius transform and interaction.
Fuzzy Sets and Systems **151**(2) (2005) 211 – 236

Grabisch, M., Labreuche, C. :
Bi-capacities-II : the Choquet integral.
Fuzzy Sets and Systems **151**(2) (2005) 237 – 259

Grabisch, M., Labreuche, C. :
A decade of application of the Choquet and Sugeno integrals in multi-criteria decision aid.
4OR **6**(1) (2008) 1–44

Fujimoto, K. :
New characterizations of k-additivity and k-monotonicity of bi-capacities.
In : Joint 2nd Int. Conf. on Soft Computing and Intelligent Systems and 5th International Symposium on Advanced Intelligent Systems. (2004)

Fujimoto, K., Murofushi, T., Sugeno, M. :
k-additivity and c-decomposability of bi-capacities and its integral.
Fuzzy Sets Syst. **158**(15) (August 2007) 1698–1712

Mayag, B., Rolland, A., Ah-Pine, J. :
Elicitation of a 2-additive bi-capacity through cardinal information on trinary actions.
In : IPMU (4). (2012) 238–247

Thank you for your attention ! Questions ?