Preferences Aggregation: the MAUT approach

University Paris Dauphine
LAMSADE
FRANCE

Chapter 3

Outline

1. Introduction to MCDA
2. Some simple models
3. Multi Attribute Utility Theory

Outline

MultiCriteria Decision Analysis (MCDA)

- A Decision Maker (DM) is facing a decision problem, i.e., the DM has to deal with multiple alternatives and has to compare them.
- Alternatives are described on several attributes.
- A criterion is an attribute with a preference relation (monotonic attribute).
- Criteria cannot be reduced to one criterion as they are potentially in conflict.

MCDA formal model: Inputs

- A set of alternatives $X = X_1 \times X_2 \times \cdots \times X_n$ evaluated on a finite set $N = \{1, \ldots, n\}$ of criteria.
- There exists preferences on the values of each criterion i (utility function, qualitative preference relation $\succsim_i$, . . .)
- A representation of the importance of each criterion or set of criteria (weights, importance relation, . . .)

MCDA formal model: a treatment

- Using the input information, elaborate a decision rule allowing to compare two different alternatives, i.e.,

$$\left.\begin{array}{l} x = (x_1, \ldots, x_n) \\ y = (y_1, \ldots, y_n) \end{array}\right\} \Longrightarrow x \succsim y \text{ or } y \succsim x$$

Example (A classic example of Grabisch et al. (2010))

	1 : Mathematics (M)	2 : Statistics (S)	3 : Language (L)
a	16	13	7
b	16	11	9
c	6	13	7
d	6	11	9

How to rank these four students ?

Example (Compare two bikes on three attributes)

	Speed	Robustness	Price
Mountain bike	20 km/h	Good	500 €
Race bike	35 km/h	Middle	1000 €

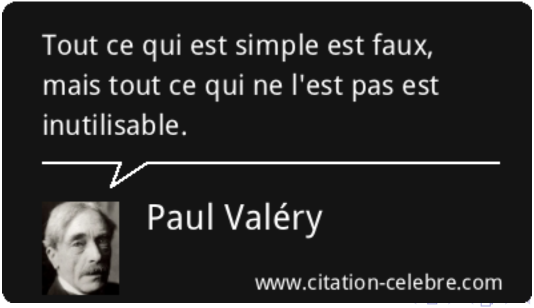
MultiCriteria Decision Aiding (MCDA): Difficulties

- MultiCriteria Decision Aiding is not so easy: it is not an easy task
- Every method has advantages and inconveniences : there is no “best method”
- All methods have structural bias.

Paul Valery (Artist, Writer, Poet, Philosopher (1871-1945))

- Tout ce qui est simple est faux, mais tout ce qui ne l'est pas est inutilisable
- Everything simple is false. Everything complex is unusable.

  It summarizes the difficulty of a task: if we make things as simple as possible, we probably forget many particular cases; if we try to take into account all the cases, the result becomes so complex that nobody can understand how it works anymore.



Three types of problems in MCDA

- Choice Problem: choose the “best” alternative(s).
- Ranking Problem: rank the alternatives from the “best” to the “worst”.
- Sorting Problem: sort the alternatives into pre-defined categories (in general ordered categories)

Outline

PARETO DOMINANCE

An alternative is preferred to another one if it is considered to be better on all the criteria.

$$x \succsim y \Longleftrightarrow [\forall i \in N, x_i \succsim_i y_i]$$

Example

	Speed	Robustness	Price
Bike A	10 km/h	Good	600 €
Bike B	20 km/h	Good	550 €
Bike C	19m/h	Very Good	800 €

$$B \succsim A \text{ and } not(A \succsim B) \Longrightarrow B \succ A$$

$$not(B \succsim C) \text{ and } not(C \succsim B)$$

Pareto dominance **is not so interesting**

Dominance

- An alternative $x = (x_1, \ldots, x_n)$ dominates an alternative $y = (y_1, \ldots, y_n)$ if $\forall i \in N, x_i \succsim_i y_i$.
- An alternative $x = (x_1, \ldots, x_n)$ strictly dominates an alternative $y = (y_1, \ldots, y_n)$ if $\forall i \in N, x_i \succsim_i y_i$ and $\exists i_0 \in N, x_{i_0} \succ_{i_0} y_{i_0}$.

Definition

The Pareto front is the set of all non-dominated alternatives.

Remark

- The optimal solution is necessary in the Pareto front
- In general, the Pareto front may be poor, i.e., it is not really different to the whole set of alternatives.

Weighted sum

Let be $x = (x_1, \dots, x_n)$ and $y = (y_1, \dots, y_n)$ two alternatives such that $x_i, y_i \in \mathbb{R}$, $\forall i \in N$. Let be w_i the weight associated to the criterion i.

$$x \succsim y \Longleftrightarrow \sum_{i=1}^{n} w_i \; x_i \geq \sum_{i=1}^{n} w_i \; y_i$$

Example

	Speed	Robustness	Price
Bike A	8/20	18/20	12/20
Bike B	18/20	8/20	12/20
Bike C	12/20	12/20	12/20

$$w_S > w_R \Longrightarrow B \succsim A$$

$$w_R > w_S \Longrightarrow A \succsim B$$

$$\forall w_R, w_S, \text{ we have } A \succsim C \text{ or } B \succsim C$$

The majority rule

Let be $x = (x_1, \ldots, x_n)$ and $y = (y_1, \ldots, y_n)$ two alternatives. x is preferred to y if it is considered "good" on a majority of criteria.

$$x \succsim y \Longleftrightarrow \left|\{i \in N : x_i \succsim_i y_i\}\right| \geq \left|\{i \in N : y_i \succsim_i x_i\}\right|$$

Example

	Speed	Robustness	Price
Bike A	10 km/h	Good	600 €
Bike B	20 km/h	Good	550 €
Bike C	19m/h	Very Good	800 €

$$B \succsim C$$

Example (Majority rule)

	Speed	Robustness	Price
Bike A	20 km/h	Very Good	600 €
Bike B	15 km/h	Good	500 €
Bike C	25m/h	Bad	550 €

Which bike do you choose?

Example (Majority rule)

	Speed	Robustness	Price
Bike A	20 km/h	Very Good	600 €
Bike B	15 km/h	Good	500 €
Bike C	25m/h	Bad	550 €

$$A \succsim B$$

$$B \succsim C$$

$$C \succsim A$$

$\Longrightarrow$ Condorcet Paradox

Two main approaches in MCDA

- Multi Attribute Utility Theory: A quantitative approach "aggregate then compare" (scoring)

$$x \succsim y \iff U(x_1, \ldots, x_n) \geq U(y_1, \ldots, y_n)$$

- Outranking: qualitative approach "compare then aggregate"

$$x \succsim y \iff \left|\{i \in N : x_i \succsim_i y_i\}\right| \rhd \left|\{i \in N : y_i \succsim_i x_i\}\right|$$

Outline

Principle

Le X be a set of alternatives evaluated on a finite set of n criteria $N = \{1, \ldots, n\}$. In general, we set $X = X_1 \times X_2 \times \ldots \times X_n$.

Le be $\succsim_X$ a complete preorder on X (preferences of a DM).

- $\succsim_X$ are supposed to be representable by an overall utility function:

$$\forall x, y \in X, \quad x \succsim_X y \Leftrightarrow F(U(x)) \geq F(U(y))$$

  where

  - $x = (x_1, \ldots, x_n)$ and $y = (y_1, \ldots, y_n)$
  - $U(x) = (u_1(x_1), \ldots, u_n(x_n))$
  - $u_i : X_i \to \mathbb{R}$ is a marginal utility function or simply called utility function or a scale on $\mathbb{R}$
  - $F : \mathbb{R}^n \to \mathbb{R}$ an aggregation function

- F is generally characterized by a parameter vector θ (weight vector,...).

Problems

1. How to choose the aggregation function F?
2. How to construct the marginal utility functions $u_i : X_i \to \mathbb{R}$?
3. The marginal utility functions $u_i : X_i \to \mathbb{R}$ should have a signification for the decision maker (see measurement theory):
    - Ordinal scales: Differences between values have no importance (e.g. a rank). They can represent orders and pre-orders.
    - Cardinal scales: Differences between values may be meaningful.
        - *Interval scales* : absolute differences between values are important.

The additive model

- $\succsim_X$ are supposed to be representable by an overall utility function:

$$\forall x \in X, \quad F(u_1(x_1), \ldots, u_n(x_n)) = \sum_{i=1}^{n} u_i(x_i)$$

- This model is equivalent to the existence of weights w_i, $i = 1, \ldots, n$, such that

$$\forall x \in X, \quad F(u_1(x_1), \ldots, u_n(x_n)) = \sum_{i=1}^{n} w_i \; u_i(x_i)$$

The additive model

- A simple method
- Additive value function involves compensation between criteria, i.e., a bad performance on a criterion *i* could be compensated by a good performance on another criterion.

  See e.g. students evaluation based on the weighted sum.
- In the weighted sum, weights represent, in reality, the substitution rate between criteria.

The additive model

- In the weighted sum, weights represent, in reality, the substitution rate between criteria.
  - E.g. for $n = 2$, $w_1 = b\ w_2$ means the DM is indifferent between these two alternatives $(0, b)$ and $(1, 0)$, i.e., $(0, b) \sim (1, 0)$.
  - There is a total compensation between "bad" performances and "good" performances.

  If we have $(a, b) \sim (a - \delta, b + \gamma)$ then gain of γ compensates the loss of δ.

  Indeed we have

  $$a\ w_1 + b\ w_2 = (a - \delta)\ w_1 + (b + \gamma)\ w_2$$
  $$\iff \delta\ w_1 = \gamma\ w_2$$
  $$\iff \frac{w_1}{w_2} = \frac{\gamma}{\delta}$$

  Implicitly, this implies that all the criteria could be express indirectly in the same unit (€, seconds, . . .).

The additive model

- Requires to normalize the criteria. In general, we set $\forall i \in N, u_i : X_i \mapsto [0, 1]$.

  E.g. For a criterion to be maximized, we could choose the following normalization functions:

  - $u_i(x_i) = \dfrac{x_i}{\max x_i}$
  - $u_i(x_i) = \dfrac{x_i - \min x_i}{\max x_i - \min x_i}$
  - ...

Example (A classic example of Grabisch et al. (2010))

	1 : Mathematics (M)	2 : Statistics (S)	3 : Language (L)
a	16	13	7
b	16	11	9
c	6	13	7
d	6	11	9

How to rank these four students by using an additive model (a weighted sum) by giving your own weights associated to the criteria ?

Example (A classic example of Grabisch et al. (2010))

	1 : Mathematics (M)	2 : Statistics (S)	3 : Language (L)
a	16	13	7
b	16	11	9
c	6	13	7
d	6	11	9

- For a student good in Mathematics, Language is more important than Statistics

$$\Longrightarrow a \prec b,$$

- For a student bad in Mathematics, Statistics is more important than Language

$$\Longrightarrow d \prec c.$$

Are these preferences representable by an additive model ?

The mutual Preferential independence

- The additive model requires to satisfy the mutual Preferential independence axiom, i.e., criteria are independent in the sense of preferences

  $\forall i \in N, \forall z_i, t_i \in X_i, \forall x, y \in X,$

  $$(z_i, x_{N-i}) \succsim (z_i, y_{N-i}) \Leftrightarrow (t_i, x_{N-i}) \succsim (t_i, y_{N-i})$$

  An attribute is preferentially independent from all other attributes when changes in the rank ordering of preferences of other attributes does not change the preference order of the attribute.

MAUT in practice

1. People suppose $\succsim_X$ representable by an overall utility function:

$$x \succsim_X y \Leftrightarrow F(U(x)) \geq F(U(y))$$

2. F is generally characterized by a parameter vector θ (weight vector,...).
3. People ask to the DM some preferential information $\succsim_{X'}$ on a reference subset (learning set) $X' \subseteq X$
4. The parameter vector is constructed so that $\succsim_X$ is an extension of $\succsim_{X'}$.
5. The model obtained in X' will be then automatically extended to X.

The UTA Approach

Principles

- Created by Jacquet Lagreze & Siskos in 1982 (at LAMSADE)
- The UTA (UTilités Additives) method aims at inferring one or more additive value functions from a given ranking on a reference set A_R.
- The method uses special linear programming techniques to assess these functions so that the ranking(s) obtained through these functions on A_R is (are) as consistent as possible with the given one.

The UTA Approach

UTA Principles: Input data

- A set of Criteria N
- A set of alternatives X evaluated on N
- A preorder $\succsim_{X'}$ on $X' \subseteq X$ (not necessary complete)
- For each element $x = (x_1, \ldots, x_n) \in X$, it is assumed that

$$U(x) = \sum_{i=1}^{n} u_i(x_i) \tag{1}$$

where $u_i : X_i \to \mathbb{R}$, $i = 1, \ldots, n$ are marginal utility functions

The UTA Approach

UTA Principles: The model

- $U(x) = \sum_{i=1}^{n} u_i(x_i)$
- For each element $x \in X'$, set

$$V(x) = U(x) + \sigma(x)$$

where $\sigma(x)$ is a nonnegative real value estimating the error of the estimation of the value $U(x)$, i.e., $\sigma(x) = V(x) - U(x)$.

The value $\sigma(x)$ will be minimized by the linear program.

UTA Principles: The linear program to solve

$$\begin{cases} \min \sum_{x \in X'} \sigma(x) \\ \\ V(x) \geq V(y) + \delta \text{ if } x \succ y \\ V(x) = V(y) \text{ if } x \sim y \\ \\ u_i(x_i^{j+1}) - u_i(x_i^j) \geq 0 \text{ if } x_i^{j+1} \succ_i x_i^j \text{ (monotonicity constraints of) } u_i \\ \\ \sigma(x) \geq 0, \forall x \in X' \end{cases}$$

- If the optimal solution is equal to 0 then $\succsim_{X'}$ is representable by (compatible with) an additive model.
- There are many versions of the UTA method

Exercise

We have the following performance matrix of children's diapers.

	1- Performance	2- Composition
A- Joone	17	17
B- Pamp. Prem	16	18
C- Pamp. Baby	10	17
D- Naty	11	19
E- Pamp. Activ.	12	12
F- Carref. Baby	13	11
G- Lupilu	14	8
H- Mots d'enfants	11	9
I- Love & Green	13	8
K- Lotus Baby	14	5
L- Pommette	11	6
M- Lillydoo	12	4

Table: A performance table of children's diapers

Apply the UTA approach when the preferences on the reference set of alternatives is $A \succ B \succ D \succ I \succ L$. You can use a linear programming implemented in Python language.

Some references

- Fishburn, “Utility theory for Decision Making”, 1970, Wiley
- Keeney-Raiffa, “Decisions with multiple objectives preferences and trade-off”, 1976, Wiley
- M. Grabisch, J-C. Marichal, R. Mesiar, and E. Pap. Aggregation functions, volume 127 of Encyclopedia of Mathematics and its Applications. Cambridge University Press, Cambridge, UK, 2009.