Decision Methods and Models

Master's Degree in Computer Science

Roberto Cordone

DI - Università degli Studi di Milano

Schedule:	Thursday 16.30 - 18.30 in Aula Magna (CS department)
	Friday 12.30 - 14.30 in classroom 301
Office hours:	on appointment
E-mail:	roberto.cordone@unimi.it
Web page:	https://homes.di.unimi.it/cordone/courses/2024-mmd/2024-mmd.html
Ariel site:	https://myariel.unimi.it/course/view.php?id=4467

Weighted sum method

The idea is to aggregate all indicators into a convex combination

The result is a sufficient condition for a point to be globally Paretian

Theorem

If $w_l > 0$ for all $l \in P$, $\sum_{l \in P} w_l = 1$ and x° is a globally optimal point of

$$\min_{x \in X} z_w(x) = \sum_{l \in P} w_l f_l(x)$$

then x° is a globally Paretian point for f in X

The proof is by contradiction: suppose that x° is not globally Paretian:

$$\exists x' \in X : x' \prec x^\circ \Rightarrow \begin{cases} f_l(x') & \leq f_l(x^\circ) \quad \forall l \in P \\ \exists \bar{l} \in P : f_{\bar{l}}(x') & < f_{\bar{l}}(x^\circ) \end{cases}$$

This implies $\sum_{l \in P} w_l f_l(x') < \sum_{l \in P} w_l f_l(x^\circ)$, against the optimality of x° □

The auxiliary problem is parametric, to be solved for all vectors w such that $w_l > 0$ for all $l \in P$ and $\sum_{l \in P} w_l = 1$ (∞^{p-1} *values*)

Consequently, the solutions provided form a hypersurface of ∞^{p-1} points

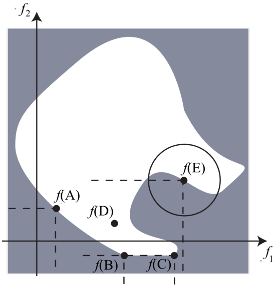
Properties

The combination is convex, but the weights w_l are strictly positive

Why only positive weights?

If weights equal to zero are allowed, the proof does not hold, because weakly Paretian solutions satisfy the condition, even if dominated

$$\begin{cases} f_{\bar{l}}(x') < f_{\bar{l}}(x^\circ) \\ f_l(x') = f_l(x^\circ) \quad \forall l \in P \setminus \{\bar{l}\} \end{cases} \quad \text{and } w_{\bar{l}} = 0 \Rightarrow \sum_{l \in P} w_l f_l(x') = \sum_{l \in P} w_l f_l(x^\circ)$$



C is a global optimum point for $z_{[0,1]}(x) = 0 \cdot f_1(x) + 1 \cdot f_2(x) = f_2$

Support of a Paretian solution

Let $W = \left\{ w \in \mathbb{R}^p : w_l > 0 \; \forall l \in P, \sum_{l \in P} w_l = 1 \right\}$ be the weight space

Given a Paretian solution $x^\circ \in X^\circ$, the support $\mathrm{Supp}(x^\circ) \subseteq W$ is the set of weight vectors w such that $x^\circ \in \arg\min_{x \in X} z_w(x)$

- continuous problems usually have ∞^{p-1} Paretian solutions and ∞^{p-1} weight vectors: the support of a Paretian solution often includes a single vector
- finite or combinatorial problems have a finite number of solution: the support of a Paretian solution often is a region in the weight space
- in general, however, unsupported solutions exist: they are Paretian solutions with empty support $\mathrm{Supp}(x^\circ) = \emptyset$

The weighted sum method finds only supported solutions

Weighted sum method and KKT conditions

The weighted sum method and the KKT conditions are clearly related

- they both build a convex combination of the indicators

but the weighted sum method requires

1. to compute the globally optimal points,
   instead of candidate locally optimal points
2. to impose strictly positive weights,
   instead of nonnegative ones

Therefore, the former returns a (often much) smaller region

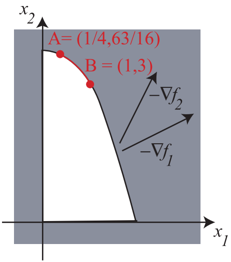
Example

The combined antigradient $\nabla z_w = w\nabla f_1 + (1-w)\nabla f_2$
is in the open cone identified by the original antigradients

Solve the problem graphically

(*or apply KKT keeping only the globally optimal solutions*)

$$\begin{aligned} \min z_w(x) &= w(-2x_1 - x_2) + (1-w)(-x_1 - 2x_2) \\ g_1(x) &= -x_1 \geq 0 \\ g_2(x) &= -x_2 \geq 0 \\ g_3(x) &= x_1^2 + x_2 - 4 \geq 0 \end{aligned}$$



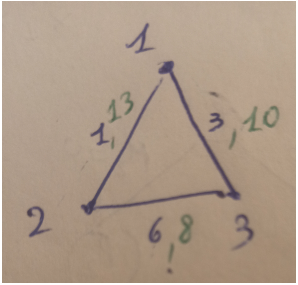
We obtain the parabola arc from $A = \left(\frac{1}{4}, \frac{63}{16}\right)$ to $B = (1,3)$,
without the two extreme points

Not the full Paretian region: $X^{WS} \subset X^{\circ}$

A combinatorial example

Given a complete graph of three vertices and two cost functions, find the minimum spanning tree

$f(x)$	(1,2)	(1,3)	(2,3)
f_1	1	3	6
f_2	13	10	8



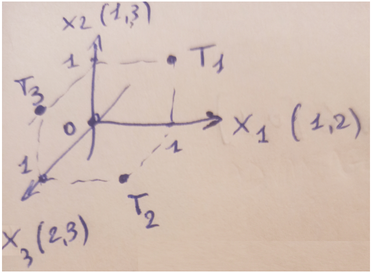
There are three feasible solutions

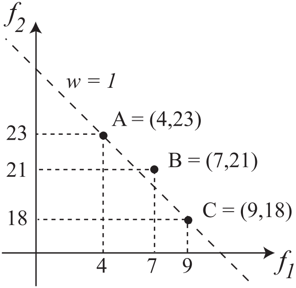
X	f_1	f_2
$T_1 = \{(1,2),(1,3)\}$	4	23
$T_2 = \{(1,2),(2,3)\}$	7	21
$T_3 = \{(1,3),(2,3)\}$	9	18

Applying the definition, all solutions are Paretian: $X^\circ = \{T_1, T_2, T_3\}$

A combinatorial example

All solutions correspond to impacts with an empty lower left quadrant





The inverse transformation method yields the whole Paretian region:

$$X^{IT} = X^{\circ} = \{T_1, T_2, T_3\}$$

The KKT conditions return all feasible solutions (locally Paretian)

A combinatorial example

The weighted sum method solves the auxiliary parametric problem

$$\min z_w(x) = w\ f_1(x) + (1-w)\ f_2(x)$$

that is

$$\min z_w(x) = (13-12w)\,x_{12} + (1-7w)\,x_{13} + (8-2w)\,x_{23}$$

where x is a spanning tree

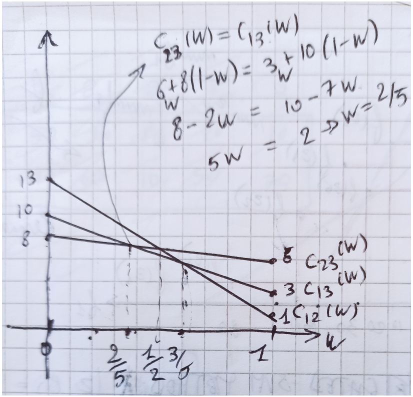
It is a minimum spanning tree with parametric costs on the edges $c_{ij}(w)$

The problem can be solved with Kruskal's algorithm

- sort the edges by increasing costs
- include the edges that do not close loops

A parametric version of Kruskal's algorithm

- let the weight vector w vary in the weight space W, here $(0, 1)$
- describe the costs $c_{ij}(w)$ of the three edges as a function of w

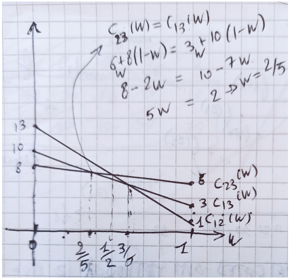


A parametric version of Kruskal's algorithm

- find the regions in W where each arc is the cheapest, second, etc.

$$c_{23}(w) = c_{13}(w) \Leftrightarrow w = \frac{2}{5}$$
$$c_{13}(w) = c_{12}(w) \Leftrightarrow w = \frac{3}{5}$$
$$c_{23}(w) = c_{12}(w) \Leftrightarrow w = \frac{1}{2}$$

- apply Kruskal's algorithm to each region
  - if $w \in (0, 2/5]$, select $(2,3)$ and $(1,3)$
  - if $w \in [2/5, 3/5]$, select $(1,3)$; then:
    - if $w \in [2/5, 1/2]$, select $(2,3)$
    - if $w \in [1/2, 3/5]$, select $(1,2)$
  - if $w \in [3/5, 1)$, select $(1,2)$ and $(1,3)$



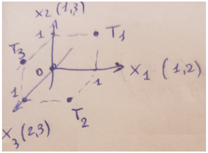
A parametric version of Kruskal's algorithm

In summary

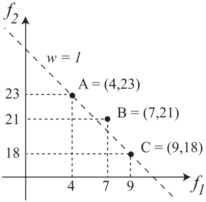
- if $w \in (0, 2/5]$, $x = \{(2,3)(1,3)\} = T_3$
- if $w \in [2/5, 1/2]$, $x = \{(1,3),(2,3)\} = T_3$
- if $w \in [1/2, 3/5]$, $x = \{(1,3)(1,2)\} = T_1$
- if $w \in [3/5, 1)$, $x = \{(1,2)(1,3)\} = T_1$

Solution T_2 is not found: it is unsupported $(\mathrm{Supp}\,(T_2) = \emptyset)$

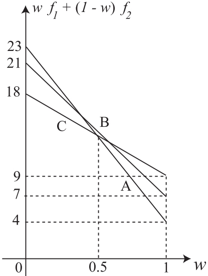
Indeed, T_2 is nonoptimal for any convex combination of the indicators, even if it can be a good compromise



X



$f(X)$



$z : W \to \mathbb{R}$

Note: the profiles z_w on the right refer to solutions, not edges

Weighted sum method and Multi-Attribute Utility Theory

Under suitable conditions, an additive utility function exist

$$\max_{x \in X} u(f(x)) = \sum_{l \in P} w_l \tilde{u}(f_l(x))$$

that is very similar to the auxiliary problem

$$\min_{x \in X} z_w(x) = \sum_{l \in P} w_l f_l(x)$$

Is there a relation with the weighted sum method?

Not really, since in *MAUT*

- $u(f(x))$ assumes a weak order on impacts, not a partial order
- the weights w_l have a fixed value in W
- the normalised utilities $\tilde{u}_l$ are nonlinear and yield indifference curves, that can reach unsupported solutions, unlike straight lines

But the basic concepts keep returning under different forms

Advantages and disadvantages

The weighted sum method has several advantages

- it is absolutely general
- it is intuitive
- it usually allows to simply extend single-objective algorithms

but also disadvantages

- it requires a parametric version of the algorithm
- it requires to find all globally optimal solutions
- it finds only supported solutions: $X^{\mathrm{WS}} \subseteq X^{\circ}$; moreover, as p increases, the fraction of supported solutions decreases:

$$\lim_{p \to +\infty} \frac{|X^{\mathrm{WS}}|}{|X^{\circ}|}$$

What about sampling W?

- Sampling further reduces the subset found, and it can be inefficient (*finding the same solution for several different weight vectors*)

The ϵ-constraint method

The idea is to keep one indicator and quality constraints on the others

The result is a necessary condition for a point to be globally Paretian
(*sufficient and necessary to be weakly Paretian*)

Theorem
If x° is globally Paretian for f in X, $\epsilon_l = f_l(x^\circ)$ and $\ell \in P$,
then x° is globally optimal for

$$\begin{array}{rcll} \min z_\epsilon(x) & = & f_\ell(x) & \\ x & \in & X & \\ f_l & \leq & \epsilon_l & \quad l \in P \setminus \{\ell\} \end{array}$$

The proof is by contradiction: suppose that x° is not globally optimal:

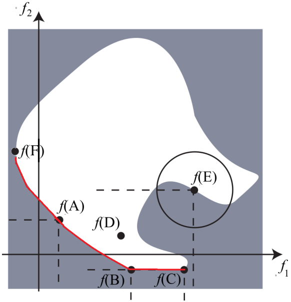
$$\exists x' \in X : \begin{cases} f_\ell(x') < f_\ell(x^\circ) \\ f_l(x') \leq \epsilon_l = f_l(x^\circ) \quad l \in P \setminus \{\ell\} \end{cases} \quad \Rightarrow x' \prec x^\circ$$

against the paretianity of x° □

The auxiliary problem is parametric, to be solved for all vectors $\epsilon \in \mathbb{R}^{p-1}$
(∞^{p-1} *values*)

Consequently, the solutions provided form a hypersurface of ∞^{p-1} points

Example



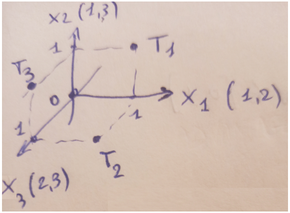
Replacing $\min f_1(x)$ with $f_1(x) \leq \epsilon_1$ and solving

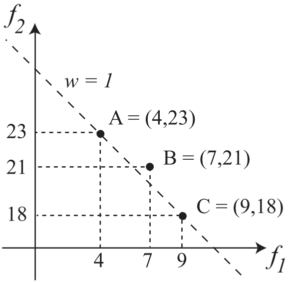
$$\min f_2(x)$$
$$\begin{array}{rcl} x & \in & X \\ f_1(x) & \leq & \epsilon_1 \end{array}$$

yields

- for small ϵ_1, no solution
- for larger ϵ_1, solutions mapping onto the arc from $f(F)$ to $f(B)$
- for large ϵ_1, solutions mapping onto part or all of segment $\overline{f(B)f(C)}$

Example





In this case, the ϵ-constraint method returns all Paretian solutions

For example, min f_2 with $x \in X$ and $f_1 \leq \epsilon$ yields

- for $\epsilon < 4$, no solution
- for $4 \leq \epsilon < 7$, solution T_1
- for $7 \leq \epsilon < 9$, solution T_2
- for $9 \leq \epsilon$, solution T_3

The same holds solving with respect to $f_1(x)$

ϵ-constraint method and lexicographic preference

The lexicographic method also focuses on a single indicator

Is there a relation with the ϵ-constraint method?

Not really, since lexicographic preference

- assumes a total order on impacts, not a partial order
- discriminates optimal impacts based on the secondary indicators
- does not impose aspiration levels ϵ_l on the secondary indicators

But the basic concepts keep returning under different forms

Advantages and disadvantages

The weighted sum method has several advantages

- it is absolutely general
- it is rather intuitive

but also disadvantages

- it requires a parametric version of the algorithm
- the additional constraints often make the basic algorithm unviable
- it requires to find all globally optimal solutions
- it finds also weakly Paretian solutions: $X^{\epsilon c} \supseteq X^{\circ}$;
  this can be refined changing the reference indicator ℓ
  and intersecting the regions obtained

What about sampling W?

- Sampling can be inefficient
  (*finding the same solution for several different weight vectors*)
- It yields an underestimate of an overestimate of X°