Multi-objective optimisation

Lorenzo Castelli, Università degli Studi di Trieste.

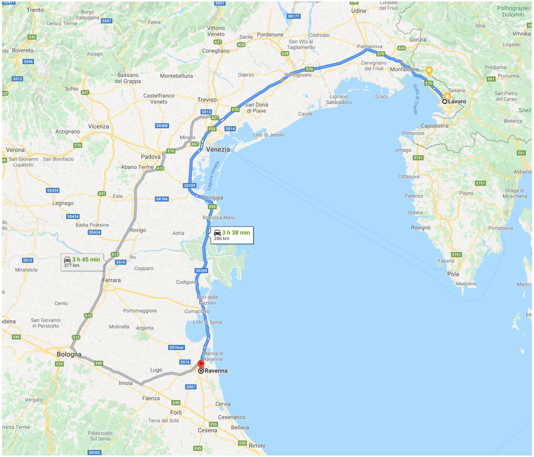
UNIVERSITÀ
DEGLI STUDI DI TRIESTE

Preferences

To make explicit the preferences of a decision maker it is better to make a direct comparison between all the possible alternatives.
The decision maker is asked to compare all pairs of alternatives.
For each pair the result of the comparison can be one of four cases:

- I prefer the alternative $\boldsymbol{A}$ to $\boldsymbol{B}$
- I prefer the alternative $\boldsymbol{B}$ to $\boldsymbol{A}$
- I am indifferent between $\boldsymbol{A}$ and $\boldsymbol{B}$
- I am not able to compare $\boldsymbol{A}$ and $\boldsymbol{B}$

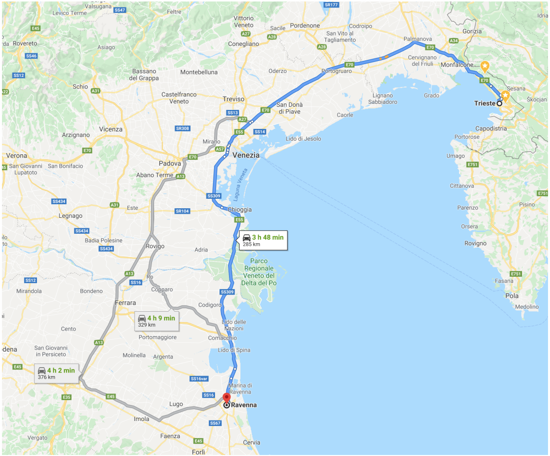
Dominance



3 h 38 min
286 km
3 h 45 min
377 km
Lavoro
Ravenna
Venezia
Bologna
Ferrara
Padova
Treviso
Vicenza
Verona
Chioggia
Udine
Trento
Monfalcone
Palmanova
Portogruaro
Capodistria
Pola
Rovigo
Imola
Lugo
Faenza
Forlì
Cesena
Cervia
Rimini

Multi-objective optimisation

Incomparability



Indifference

The term “indifference” indicates the total interchangeability of one decision with another.

- Indifference between $\boldsymbol{A}$ and $\boldsymbol{B}$ means that $\boldsymbol{A}$ can be replaced with $\boldsymbol{B}$ and vice versa and the decision maker finds the two options equivalent.
- Likewise, if it were also to be indifferent between $\boldsymbol{B}$ and $\boldsymbol{C}$, it could replace $\boldsymbol{B}$ with $\boldsymbol{C}$ and vice versa.
- This interchangeability means that $\boldsymbol{C}$ could also replace $\boldsymbol{A}$ and vice versa.
- So there is indifference also between $\boldsymbol{A}$ and $\boldsymbol{C}$, as to say that **indifference is a transitive relationship**.

Incomparability

Incomparability is not a transitive relationship.

- You can always imagine a situation where two alternatives A and B are not comparable because A is much better than B for one criterion but B is much better than A for another criterion.
- For example A could be a decision that involves high costs but also excellent quality and B instead could be a very cheap but also low quality choice.
- Similarly it could happen between A and C, but this does not imply that B and C are also not comparable.
- It could happen that B is preferred to C because it is a little cheaper than C and of a slightly better quality than C.

Decisions

For each ordered pair of decisions $(\boldsymbol{A}, \boldsymbol{B})$ the relationship that is established between the two decisions assumes a value among the following four

$$\prec \quad \succ \quad \sim \quad ?$$

where

$$\boldsymbol{A} \prec \boldsymbol{B} \Rightarrow \boldsymbol{A} \text{ is preferred over } \boldsymbol{B}$$
$$\boldsymbol{A} \succ \boldsymbol{B} \Rightarrow \boldsymbol{B} \text{ is preferred over } \boldsymbol{A}$$
$$\boldsymbol{A} \sim \boldsymbol{B} \Rightarrow \boldsymbol{A} \text{ and } \boldsymbol{B} \text{ are indifferent}$$
$$\boldsymbol{A} \; ? \; \boldsymbol{B} \Rightarrow \boldsymbol{A} \text{ and } \boldsymbol{B} \text{ are not comparable}$$

Decision set

- We can define for each set $\boldsymbol{X}$ of feasible decisions a structure of preferences such that for each ordered pair of decisions $\boldsymbol{x} \in \boldsymbol{X}$ and $\boldsymbol{y} \in \boldsymbol{X}$ we have $\boldsymbol{xRy}$ where $\boldsymbol{R} \in \{\prec, \succ, \sim, ?\}$.
- Assuming that preferences are consistent, if $\boldsymbol{x} \prec \boldsymbol{y}$ occurs, then solution $\boldsymbol{y}$ must be discarded.

These considerations lead to the following definition.

Definition

Decisions $\boldsymbol{x}$ are said to be non-dominated, or efficient, or Pareto-optimal, if there is no decision $\boldsymbol{y}$ to be preferred to $\boldsymbol{x}$.

Non-dominated decisions

- Consistency implies the existence of at least one non-dominated decision (if $\boldsymbol{X}$ is a finite set).
- The problem in reality is that there are normally many non-dominated decisions and more refined criteria must be established to prefer one decision over all (in the end you have to choose one and only one decision).
- Note also that the presence of pairs of incomparable decisions increases the number of non-dominated decisions.
- If there were no incomparable pairs, the non-dominated decisions would all be indifferent to each other and therefore the problem of the final choice would not exist.

Multi-objective problems

Let $\boldsymbol{X}$ be the set of feasible decisions. Assume that $\boldsymbol{m}$ different objectives are defined on $\boldsymbol{X}$, which take the form of $\boldsymbol{m}$ objective functions $\boldsymbol{f_i}(\boldsymbol{x}) : \boldsymbol{X} \to \mathbb{R}$ to be minimised. These functions define the following preference structure:

$$\boldsymbol{x} \prec \boldsymbol{y} \Leftrightarrow \boldsymbol{f_i}(\boldsymbol{x}) \leq \boldsymbol{f_i}(\boldsymbol{y}), \boldsymbol{i} \in \{\mathbf{1}, \ldots, \boldsymbol{m}\}, \text{ and } \boldsymbol{f}(\boldsymbol{x}) \neq \boldsymbol{f}(\boldsymbol{y})$$
$$\boldsymbol{x} \succ \boldsymbol{y} \Leftrightarrow \boldsymbol{f_i}(\boldsymbol{x}) \geq \boldsymbol{f_i}(\boldsymbol{y}), \boldsymbol{i} \in \{\mathbf{1}, \ldots, \boldsymbol{m}\}, \text{ and } \boldsymbol{f}(\boldsymbol{x}) \neq \boldsymbol{f}(\boldsymbol{y})$$
$$\boldsymbol{x} \sim \boldsymbol{y} \Leftrightarrow \boldsymbol{f}(\boldsymbol{x}) = \boldsymbol{f}(\boldsymbol{y})$$
$$\boldsymbol{x} \ ? \ \boldsymbol{y} \text{ otherwise}$$

Multi-objective problems

We say that a decision $\boldsymbol{x}$ dominates a decision $\boldsymbol{y}$ if $\boldsymbol{f_i}(\boldsymbol{x}) \leq \boldsymbol{f_i}(\boldsymbol{y})$ for each $\boldsymbol{i} \in \{1, \ldots, \boldsymbol{m}\}$ and thar there exists an objective $\boldsymbol{k}$ such that $\boldsymbol{f_k}(\boldsymbol{x}) < \boldsymbol{f_k}(\boldsymbol{y})$. If there is only one non-dominated decision $\boldsymbol{x}^*$, then we would have

$$\boldsymbol{f_i}(\boldsymbol{x}^*) \leq \boldsymbol{f_i}(\boldsymbol{x}), \boldsymbol{i} \in \{1, \ldots, \boldsymbol{m}\}, \boldsymbol{x} \in \boldsymbol{X}$$

and such a decision would obviously be the best possible in every respect and there would be nothing more to say. We therefore suppose that there is more than one non-dominated decision.

Pareto solutions

- In accordance with the criterion of rationality dominated decisions must certainly be discarded.
- If objectives $\boldsymbol{f}_1, \dots, \boldsymbol{f}_m$ are actually all the objectives and reflect the preferences of a single decision maker, then it is natural to assume that '$\boldsymbol{x}$ is rationally preferred to $\boldsymbol{y}$' if $\boldsymbol{x}$ dominates $\boldsymbol{y}$.

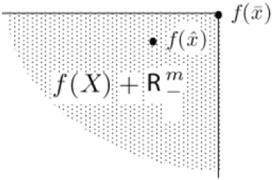
  In case of multiple decision makers, it may not be universally acceptable to reject decision $\boldsymbol{y}$ because of $\boldsymbol{x}$'s dominance. A decision maker whose objective function was indifferent between $\boldsymbol{x}$ and $\boldsymbol{y}$ might not like an improvement from all other decision makers and therefore would not automatically prefer $\boldsymbol{x}$ to y.
- In the following, however, we will not deal with this second aspect and we will always think that the decision maker is unique and the objective functions express his preferences.

Pareto solutions

- In this perspective, the only decisions that can be taken into consideration are the Pareto optima.
- These are usually represented geometrically in the image space of the objective functions.
- For each decision $\boldsymbol{x} \in \boldsymbol{X}$ consider points $\boldsymbol{f_1}(\boldsymbol{x}), \ldots, \boldsymbol{f_m}(\boldsymbol{x})$ and define $\boldsymbol{f}(\boldsymbol{X})$ as the union of all such points.
- All possible objective function values therefore correspond to elements of the set $\boldsymbol{f}(\boldsymbol{X})$

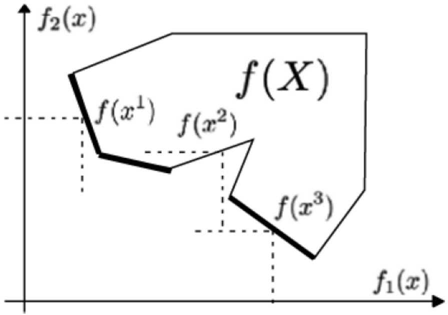
Pareto solutions

- Given a decision $\overline{\boldsymbol{x}}$, consider $\boldsymbol{f}(\overline{\boldsymbol{x}})$, attach to point $\boldsymbol{f}(\overline{\boldsymbol{x}})$ the negative orthant $\mathbb{R}^{\boldsymbol{m}}_{-}$ to define the set $\boldsymbol{f}(\overline{\boldsymbol{x}}) + \mathbb{R}^{\boldsymbol{m}}_{-}$.
- If there are decisions $\hat{\boldsymbol{x}}$ such that $\boldsymbol{f}(\hat{\boldsymbol{x}}) \in \boldsymbol{f}(\overline{\boldsymbol{x}}) + \mathbb{R}^{\boldsymbol{m}}_{-}$ and $\boldsymbol{f}(\hat{\boldsymbol{x}}) \neq \boldsymbol{f}(\overline{\boldsymbol{x}})$, $\hat{\boldsymbol{x}}$ dominates $\overline{\boldsymbol{x}}$ by definition.



Pareto solutions

In this figure, points $\boldsymbol{f}(\boldsymbol{x_1})$ and $\boldsymbol{f}(\boldsymbol{x_3})$ are Pareto optima, while $\boldsymbol{f}(\boldsymbol{x_2})$ is not. The Pareto optima are all the points on the boundary of f (X) highlighted with a solid thick black line.



Pareto solutions - Example

- We have to build power plants to meet the energy need of a region.
- 5 sites have been identified and the possible plants for each site.
- The powers obtainable from the plants and the construction costs were respectively estimated as $\boldsymbol{P} = (\mathbf{50, 35, 30, 25, 60})$ (MW) and $\boldsymbol{C} = (\mathbf{20, 16, 13, 6, 12})$ (M€).
- We need to decide in which site to build which plant.
- The overall power is the sum of the powers of the plants built and the same is true for the cost.

We initially want to evaluate which decisions are efficient. Then the choice will be made between the efficient solutions based on the estimated power demand and the funds available.

Pareto solutions - Example

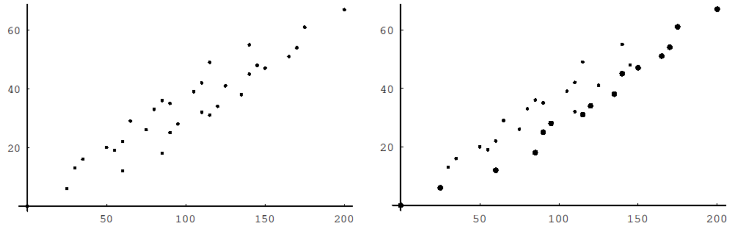
We can represent the problem as

$$X = \{0,1\}^5 \quad f_1(x) = \sum_{i=1}^{5} P_i x_i \quad f_2(x) = \sum_{i=1}^{5} C_i x_i$$

where f_1 has to be maximised, while f_2 minimised.

- There are **32** alternatives as shown in Figure (a) (in abscissa the power and in ordinate the cost)
- In Figure (b) the **15** Pareto optimal solutions are highlighted.
- The choice of the decision maker will have to fall between these **15** solutions depending on the cost-benefit evaluation you want to do.

Pareto solutions - Example



(a) (b)

Pareto solutions - Example

The Pareto solutions are

$$\{0,0,0,0,0\},\{0,0,0,1,0\},\{0,0,0,0,1\},\{0,0,0,1,1\},\{0,0,1,0,1\},$$
$$\{0,1,0,0,1\},\{0,0,1,1,1\},\{0,1,0,1,1\},\{1,0,0,1,1\},\{1,0,1,0,1\},$$
$$\{0,1,1,1,1\},\{1,0,1,1,1\},\{1,1,0,1,1\},\{1,1,1,0,1\},\{1,1,1,1,1\},$$

where the associated objective function values are

$$\{0,0\},\{25,6\},\{60,12\},\{85,18\},\{90,25\},$$
$$\{95,28\},\{115,31\},\{120,34\},\{135,38\},\{140,45\},$$
$$\{150,47\},\{165,51\},\{170,54\},\{175,61\},\{200,67\}.$$

If, for instance, we need a power between $\mathbf{120}$ and $\mathbf{150}$ MW, we could evaluate the following four options $\{120,34\},\{135,38\},\{140,45\},\{150,47\}$.

Pareto solutions

- Pareto solutions can be very numerous, even infinite in continuous problems, and therefore the problem arises of how to choose a particular decision within the Pareto optima.
- While restricting the choice between non-dominated solutions is a target that can be left to the analyst, the choice of the final decision pertains to the decision maker.
- For this reason various methods have been suggested to generate Pareto optima by interacting with the decision maker.

Linear combination

The problem is reduced to a single objective function through an aggregation as a linear combination with positive coefficients

$$\boldsymbol{F}(\boldsymbol{x}) := \sum_{\boldsymbol{i}} \alpha_{\boldsymbol{i}} \boldsymbol{f}_{\boldsymbol{i}}(\boldsymbol{x}), \quad \text{where } \alpha_{\boldsymbol{i}} > \mathbf{0}, \tag{1}$$

and we solve

$$\min\{\boldsymbol{F}(\boldsymbol{x}) : \boldsymbol{x} \in \boldsymbol{X}\} \tag{2}$$

Aggregating different criteria into a single objective function as in (1) poses the following questions:

1. given some values $\alpha_{\boldsymbol{i}} > \mathbf{0}$, do you always get a Pareto solution by solving (2)?
2. given a Pareto optimum $\hat{\boldsymbol{x}}$, do there always exist some values $\alpha_{\boldsymbol{i}} > \mathbf{0}$ to get this $\hat{\boldsymbol{x}}$?

Linear combination – Question 1

The positive aspect of the aggregation is that $\min\{\boldsymbol{F}(\boldsymbol{x}) : \boldsymbol{x} \in \boldsymbol{X}\}$ gives rise to a Pareto optimum. In fact, if $\boldsymbol{y}$ is a solution dominated by $\boldsymbol{x}$, we have by the definition of dominance

$$\alpha_i f_i(x) \leq \alpha_i f_i(y) \quad i \in \{1, \ldots, m\} \quad \alpha_k f_i(x) < \alpha_i f_k(y).$$

By summing up, we get

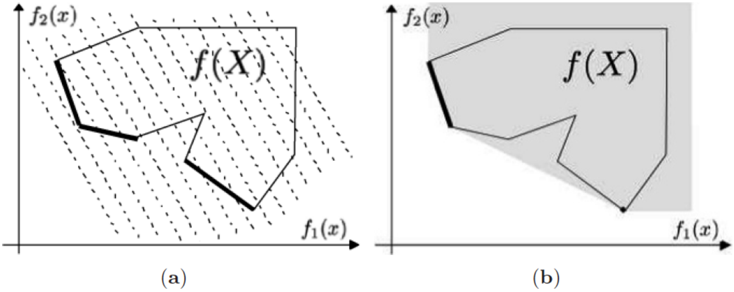
$$\sum_i \alpha_i f_i(x) < \sum_i \alpha_i f_i(y)$$

and therefore no dominated solution can be optimal for (2).

Linear combination - Question 2

- On the question of whether any Pareto optimum can be obtained from an appropriate choice of $\boldsymbol{\alpha_i}$ coefficients, unfortunately the answer is not positive in general.
- We note that solving (2) is equivalent to minimising the linear function $\sum_i \boldsymbol{\alpha_i y_i}$ for $\boldsymbol{y} \in \boldsymbol{f}(\boldsymbol{X})$ (see Figure (a)).
- It is also known that the minima of a linear function on a set $\boldsymbol{Y}$ belong (in addition to $\boldsymbol{Y}$) to the frontier of the convex envelope of $\boldsymbol{Y}$.
- So those Pareto optima that are not on the frontier of the convex envelope cannot be generated by (2).
- More exactly, since only non-negative $\boldsymbol{\alpha_i}$ coefficients are allowed, solving (2) only generates optima that lie on the frontier of the convex envelope of $\boldsymbol{f}(\boldsymbol{X}) + \mathbb{R}_+^{\boldsymbol{m}}$. In Figure (b) the convex envelope of $\boldsymbol{f}(\boldsymbol{X}) + \mathbb{R}_+^{\boldsymbol{m}}$ and the Pareto optima (solid thick black line) that can be generated with this method.

Linear combination – Question 1&2



Linear combination - Example

The feasible set is $X = \{0, 1\}^5$ and the objective functions are

$$f_1(x) = 50x_1 + 35x_2 + 30x_3 + 25x_4 + 60x_5$$
$$f_2(x) = 20x_1 + 16x_2 + 13x_3 + 6x_4 + 12x_5$$

Since f_1 must be maximised, in (1) we must consider $-f_1$. Moreover, since there are only two objectives, the coefficients α_1 and α_2 can be replaced by α and $1 - \alpha$. Then problem (2) becomes:

Linear combination - Example

$$\min_{x\in\{0,1\}^5} -\alpha(50x_1+35x_2+30x_3+25x_4+60x_5)+$$
$$(1-\alpha)(20x_1+16x_2+13x_3+6x_4+12x_5)=$$

$$\min_{x\in\{0,1\}^5}(20-70\alpha)x_1+(16-51\alpha)x_2+(13-43\alpha)x_3+$$
$$(6-31\alpha)x_4+(12-72\alpha)x_5=$$

$$\min_{x_1\in\{0,1\}}(20-70\alpha)x_1+\min_{x_2\in\{0,1\}}(16-51\alpha)x_2+$$
$$\min_{x_3\in\{0,1\}}(13-43\alpha)x_3+\min_{x_4\in\{0,1\}}(6-31\alpha)x_4+\min_{x_5\in\{0,1\}}(12-72\alpha)x_5$$

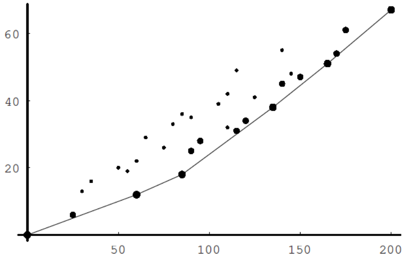
Linear combination - Example

The minima are 0 or 1 depending on the sign of the coefficient of the variable. Therefore we have, for $\mathbf{0} \leq \alpha \leq \mathbf{1}$ (the values indicated correspond to the values of α for which the coefficients become zero):

	x	$\{f_1(x), f_2(x)\}$
$0 \leq \alpha < 1/6$	$\{0, 0, 0, 0, 0\}$	$\{0, 0\}$
$1/6 < \alpha \leq 6/31$	$\{0, 0, 0, 0, 1\}$	$\{60, 12\}$
$6/31 < \alpha \leq 2/7$	$\{0, 0, 0, 1, 1\}$	$\{85, 18\}$
$2/7 < \alpha \leq 13/43$	$\{1, 0, 0, 1, 1\}$	$\{135, 38\}$
$13/43 < \alpha \leq 16/51$	$\{1, 0, 1, 1, 1\}$	$\{165, 51\}$
$16/51 < \alpha \leq 1$	$\{1, 1, 1, 1, 1\}$	$\{200, 67\}$

Linear combination - Example

6 Pareto optima are obtained. However, we know there are 15 Pareto solutions, therefore 9 Pareto optima cannot be generated by the convex combination. In particular, only one of the four solutions previously highlighted is included among the six solutions generated!

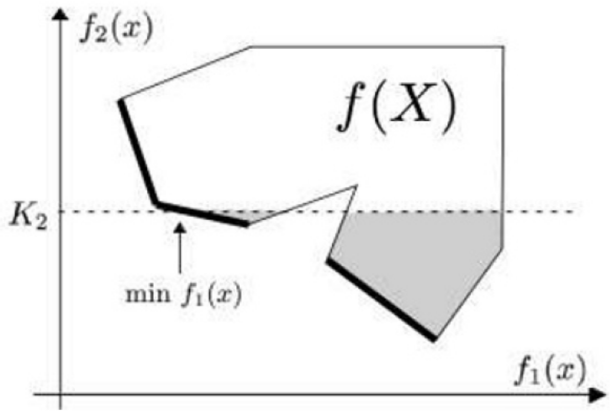


Constraints on objectives

In this method, out of the m objective functions, only one is maintained as an objective and the other $m - 1$ are transformed into constraints, setting threshold values $K_2, \ldots, K_m$ above which (if the functions are to be minimised) solutions are not allowed:

$$
\begin{aligned}
\min & f_1(x) \\
& f_2(x) \leq K_2 \\
& f_3(x) \leq K_3 \\
& \vdots \\
& f_m(x) \leq K_m \\
& x \in X
\end{aligned}
$$

Constraints on objectives



Constraints on objectives - Question 1

Q: Do we always get a Pareto optimal solution?

A: Not always. The solution of problem (3) is certainly not dominated if it is the only optimum. Unfortunately, the uniqueness of an optimum is not a property that can always be easily verified. It could happen that by solving (3) we get an optimum $\hat{\boldsymbol{x}}$, while, without our knowledge, there is another optimum $\overline{\boldsymbol{x}}$, such that

$$\boldsymbol{f}_1(\hat{\boldsymbol{x}}) = \boldsymbol{f}_1(\overline{\boldsymbol{x}}), \quad \boldsymbol{f}_2(\overline{\boldsymbol{x}}) < \boldsymbol{f}_2(\hat{\boldsymbol{x}}) \leq \boldsymbol{K}_2$$

The optimum $\hat{\boldsymbol{x}}$ is dominated.

Constraints on objectives - Question 2

Q: given a Pareto optimum $\hat{\boldsymbol{x}}$, do they exist values $\boldsymbol{K_i} > \mathbf{0}$ such that $\hat{\boldsymbol{x}}$ is optimal in (3)?

A: Yes. In fact, if $\hat{\boldsymbol{x}}$ is a Pareto optimum solution, just choose

$$\boldsymbol{K_i} := \boldsymbol{f_i}(\hat{\boldsymbol{x}}) \quad \boldsymbol{i} = \mathbf{2}, \ldots, \boldsymbol{m}$$

Then, if there is a solution $\overline{\boldsymbol{x}}$ feasible for problem (3) and such that $\boldsymbol{f_1}(\overline{\boldsymbol{x}}) < \boldsymbol{f_1}(\hat{\boldsymbol{x}})$, we would have that $\overline{\boldsymbol{x}}$ dominates $\hat{\boldsymbol{x}}$, which is contrary to the hypothesis of Pareto optimality of $\hat{\boldsymbol{x}}$. Hence, no Pareto optimum is lost by varying the $\boldsymbol{K_i}$ parameters, regardless of the convexity or non-convexity of $\boldsymbol{f}(\boldsymbol{X})$.

Constraints on objectives - Example

It is necessary to solve, by applying (3) and choosing to turn the cost objective into a constraint,

$$\begin{aligned} \max & \; 50x_1 + 35x_2 + 30x_3 + 25x_4 + 60x_5 \\ & \; 20x_1 + 15x_2 + 13x_3 + 6x_4 + 12x_5 \leq K \\ & \; x \in \{0, 1\}^5 \end{aligned}$$

This is a **KNAPSACK PROBLEM !!!!**

Lexicographic order

- There are cases where objectives have different priorities and you can think of improving an objective only if the ones with higher priority have already been met at best.
- Formally a lexicographic order between elements of $\mathbb{R}^m$ is defined by the following preference relationship $\prec$:

$$\boldsymbol{x} \prec \boldsymbol{y} \Leftrightarrow \exists k \in \{1, \dots, m\} : x_i = y_i, i \in \{1, \dots, k-1\}, x_k < y_k$$

- The lexicographic order is a total order. Given two elements $\boldsymbol{x} \neq \boldsymbol{y}$ in $\mathbb{R}^m$, either $\boldsymbol{x} \prec \boldsymbol{y}$ or $\boldsymbol{y} \prec \boldsymbol{x}$. There are no other alternatives.
- So if the objectives are defined in order of priority $f_1, f_2, \dots, f_m$, we can define the lexicographic optimum as that decision $\boldsymbol{x}^*$ such that the vector $(f_1(\boldsymbol{x}^*), \dots, f_m(\boldsymbol{x}^*))$ is minimal according to the lexicographic order.

Lexicographic order

To get the lexicographic optimum, we need to solve m minimum problems in series. Let $\hat{x}^k$ be the optimum of the k-th minimum problem. Then the $k+1$-th problem is defined by

$$\begin{aligned} \min & f_{k+1}(x) \\ & f_i(x) = f_i(\hat{x}^k) \quad i = 1, \ldots, k \\ & x \in X \end{aligned}$$

A more rapid method consists in solving a single minimum problem obtained as a linear combination of objectives with weights $\alpha_1 \gg \alpha_2 \gg \ldots \gg \alpha_m$. In general, the lexicographic optimum is not obtained in this way but it is only approximated. However, if the problem is discrete and the weights are sufficiently different from each other, the lexicography optimum is obtained.