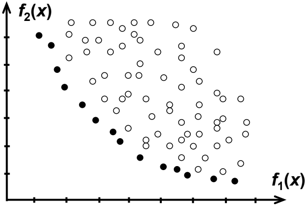
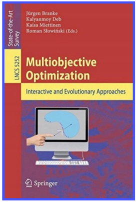




Introduction to Multiple Objective Optimization
Classical Optimization Methods

Miłosz Kadziński

Institute of Computing Science
Poznan University of Technology, Poland

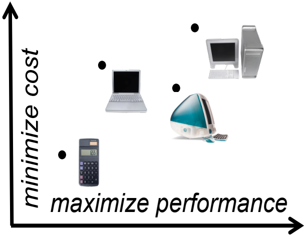
Single vs. Multiple Objective Optimization

Single Objective Optimization

The superiority of a solution over other solutions can be easily determined by comparing their objective function values

Multiple Objective Optimization

- Consider **several objective simultaneously**
- Objectives are conflicting (good quality is not cheap)
- All objectives cannot be optimized simultaneously
- Need for **considering the trade-offs** (compromises) **between objectives**



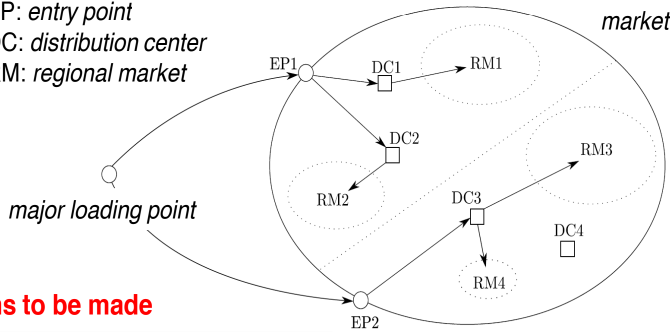
Examples

- Finance: *minimize* **risk** and *maximize* **return**
- Business: *minimize* **cost** and *minimize* **environmental impact**
- Health care: *maximize* **X-ray dose to tumor** and *minimize* **X-ray dose to healthy tissues**

Practical Example of Multiple Objective Optimization (1)

Multi-national company aims to serve a market in the South East Europe

EP: *entry point*
DC: *distribution center*
RM: *regional market*



Decisions to be made

- Selection of entry points
- Choice of transport means
- Location of distribution centers
- Determination of associated flows

Objectives

- *min* total costs (transport, operations)
- *min* amount of emissions of CO_2
- *min* amount of PM emmissions

Practical Example of Multiple Objective Optimization (2)

Health-care facility location in highly developed city such as Hong Kong

- Health-care facilities are essential to all communities
- Important in urban planning
- Steady growth of Hong Kong population
- Complex socio-ecological system

Objectives

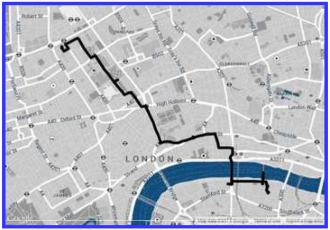
- *min* cost of building new facilities
- *min* the population that falls outside the coverage range
- *max* the equity of accessability

Yahoo! Researchers Have Developed A GPS Algorithm To Find 'Emotionally Pleasant' Routes

Amit Chowdhry Contributor

Tech enthusiast, born in Ann Arbor and educated at Michigan State

shortest path



beautiful/scenic path

A general formulation of multiple objective optimization problems:

Minimize $f(x) = (f_1(x), f_2(x), \ldots, f_M(x))$ ⇨ *multiple (M) objectives*

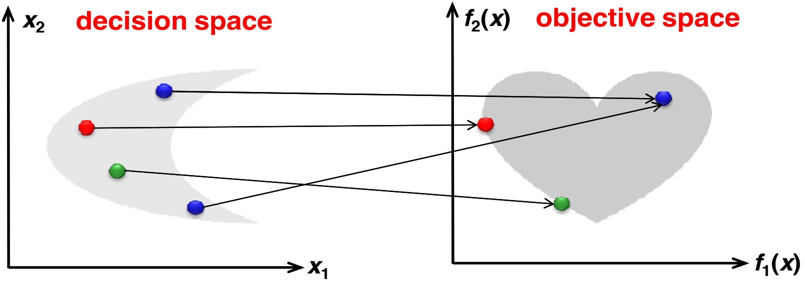
subject to $x \in S$, ⇨ *constraints*

$g_j(x) \geq 0, \quad j = 1, \ldots, J$

$h_k(x) = 0 \quad k = 1, \ldots, K$

$x_i^L \leq x_i \leq x_i^U, \quad i = 1, \ldots, N$

where $x=(x_1, x_2, \ldots, x_N)$ is a solution (represented by decision variables) and $S \subset R^N$ is called the feasible set

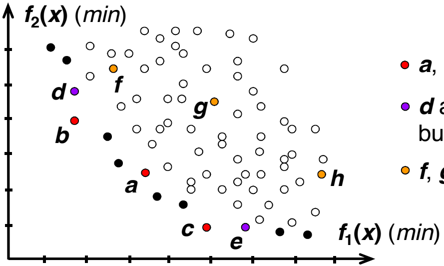


Important Definitions in MOO (2) – Pareto Optimality

In multiple objective optimization, the superiority of a solution over another solutions is determined by the dominance relation (transitive and incomplete)

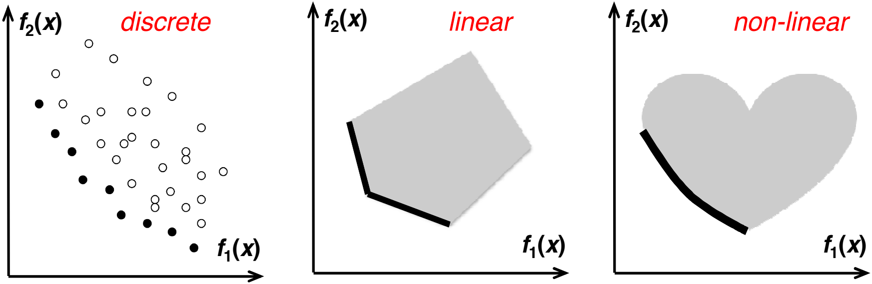
Vilfredo Pareto

- A feasible solution $\boldsymbol{x} \in \boldsymbol{S}$ is called **Pareto optimal (PO)** if there does not exist another solution $\boldsymbol{y} \in \boldsymbol{S}$ such that:
  $f_i(\boldsymbol{y}) \leq f_i(\boldsymbol{x})$ for all $i = 1, \ldots, M$, *and*
  $f_i(\boldsymbol{y}) < f_i(\boldsymbol{x})$ for some $i \in \{1, \ldots, M\}$
- A feasible solution $\boldsymbol{x} \in \boldsymbol{S}$ is called **weakly Pareto optimal (WPO)** if there does not exist another solution $\boldsymbol{y} \in \boldsymbol{S}$ such that $f_i(\boldsymbol{y}) < f_i(\boldsymbol{x})$ for all $i = 1, \ldots, M$

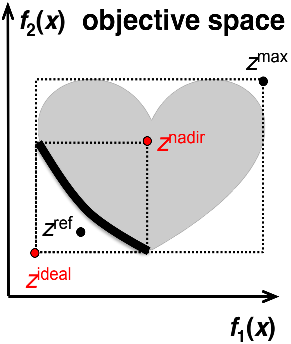


- ***a***, ***b***, and ***c*** are Pareto optimal
- ***d*** and ***e*** are weakly Pareto optimal, but are not Pareto optimal
- ***f***, ***g***, and ***h*** are dominated

- **Pareto-optimal set**: all Pareto-optimal solutions in the decision space (the non-dominated set of the entire feasible decision space *S*)
- **Pareto-front**: all Pareto-optimal solutions in the objective space (the "boundary' in the objective space defined by the set of all points (objective function vectors) mapped from the Pareto-optimal set)
- **There can exist infinitely many PO solutions**, all of them mathematically incomparable (cannot be compared without additional information)
- Other terms used: *efficient/compromise/non-dominated solution*



- The **ideal point** (vector) $z^{ideal} \in R^N$ has:
  $\boldsymbol{z_i^{ideal} = min_{x \in S}\ f_i(x)}$ for all $i=1,\dots,M$
  - *lower bound of the Pareto front (non-existent)*
- The **nadir point** (vector) $z^{nadir} \in R^N$ has:
  $\boldsymbol{z_i^{nadir} = max_{x \in S \text{ is PO}}\ f_i(x)}$ for all $i=1,\dots,M$
  - *upper bound of the Pareto front*
- The **max point** (vector) $z^{max} \in R^N$ has:
  $\boldsymbol{z_i^{max} = max_{x \in S}\ f_i(x)}$ for all $i=1,\dots,M$
  - *maximum objective function values of the entire objective space*
  - *often used as an estimate of the nadir point*

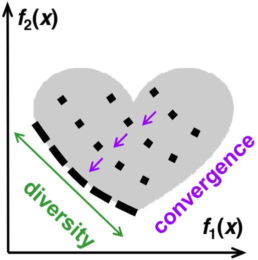


- **Decision Maker** (DM) is an expert in the application area able to express preferences related to the objectives (no need to expertize in optim.)
- The **reference point** $\boldsymbol{z^{ref} \in R^N}$ (one way for the DM to express preferences):
  $\boldsymbol{z_i^{ref}}$ indicates an aspiration level (desired value for the objective)

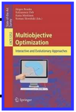
What Means Solving a MOO Problem?

Means different things for different people:

- **Find all PO solutions**
  - Theoretical approach, not feasible in practice
- **Approximate PO set**
  - As close as possible to Pareto-optimal front
  - Good diversity (representatives in all parts of PO set)
  - Can also be an approximation of some part of PO set
- **Find the most preferred PO solution**
  - Requires preferences from DM

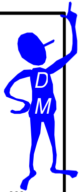


J. Branke, K. Deb, K. Miettinen, R. Słowiński, Multiobjective Optimization: Interactive and Evolutionary Approaches. *Springer*, Berlin, 2008



Different Types of MOO Methods

Multiple objective optimization methods are *classified w.r.t. to the role of DM*:



No preference methods where the participation of DM is not needed

A priori methods where the DM articulates preferences before optimization and the best solution according to the given preferences is found

- DM can tell what kind of solutions he wants
- Must have understanding about his preferences, method, objectives, feasibility

Interactive methods allow the DM to guide the search by alternating optimization and preference articulation iteratively

- DM guides the search, only solutions that are interesting to him are computed
- Need active participation from the DM (who may be busy)

A posteriori methods aim to generate a representative set of PO solutions and the DM chooses the best one among them

our focus

A Posteriori MOO Methods

A posteriori methods = representative set of all PO solutions + decision making

- Approximating the complete Pareto optimal set may take time
- It may be hard to know beforehand how big a representation is dense enough
- It may be hard to choose from a large representation of Pareto-optimal set
- Visualization is easy only in problems with 2 or 3 objectives

Most a posteriori methods fall into two classes:

- **Evolutionary methods** evolve a population of solutions simultaneously into a representative set of PO solutions *next lecture*
- **Classical methods** solve multiple single-objective optimization problems that each produce one PO solution at a time ***still today***
  - Usually, multiple runs needed to obtain a set of Pareto optimal solutions
  - Usually, problem knowledge is necessary
  - **Scalarizing methods** convert a MOO problem into a single-objective one

Weighted Sum Method (WSM) – A Priori Perspective

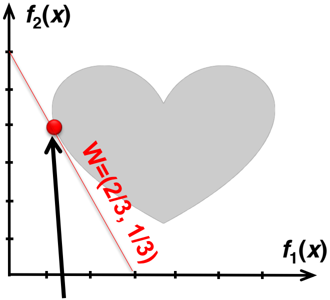
Scalarize a set of objectives into a single objective by adding objectives values multiplied by respective weights w_i, $i=1,\dots,M$:

Minimize $f(x) = (f_1(x),\dots, f_M(x))$ **subject to** $x \in S$ ⇨ **Minimize** $F(x) = \sum_{i=1,\dots,M} w_i \cdot f_i(x)$ **subject to** $x \in S$

Used as a priori method:

- Objective weight chosen in proportion to the relative importance
- $w_i > w_j$ means that the DM appreciates more the improvement on f_i than on f_j
- *Advantage*: simple
- *Disadvantage*: it is difficult to set the weight vector to obtain a solution in a desired region in the objective space

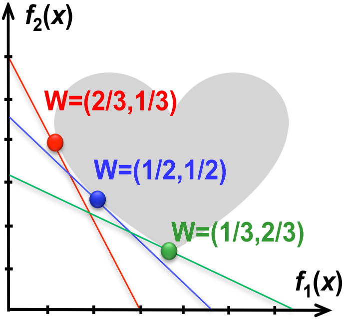
M. Ehrgott, Multicriteria Optimization, 2nd Edition. Springer, 2005



Pareto-optimal solution found with **$W=(W_1,W_2)=(2/3,1/3)$**

Weighted Sum Method (WSM) – A Posteriori Perspective

- Any a priori method can be turned into an a posteriori method by producing **an evenly spaced set in the set of preferences**
- WSM involves parameters meaningful for DM
- For each weight, potentially different solution can be obtained
- Weights can be generated from a uniform distribution using, e.g., Hit-And-Run



Properties of the Weighted Sum Method:

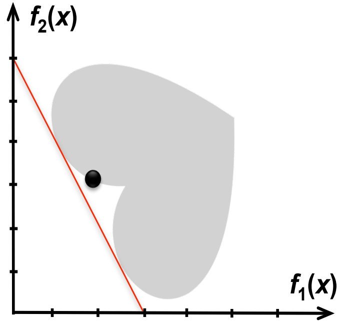
- If $w_i \geq 0$ for all $i=1,\ldots,M$ and $w_j > 0$ for some $j=1,\ldots,M$, then an identified solution is weakly Pareto-optimal (WPO)
- If $w_i > 0$ for all $i=1,\ldots,M$, then an identified solution is Pareto-optimal (PO)

Weighted Sum Method (WSM) – Major Weakness

- There may be Pareto-optimal (PO) solutions that do not optimize any weight vector
- WSM cannot find certain PO solutions in case of a non-convex objective space
- **Non-supported PO (efficient) solutions**

Assume choosing a husband

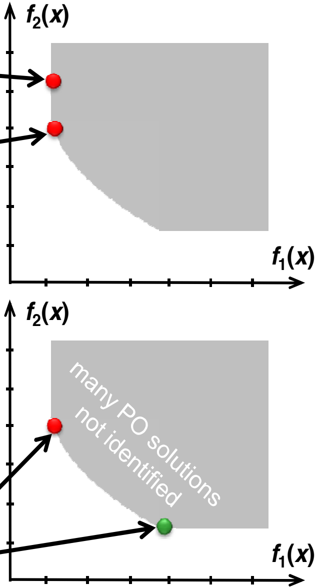
Objective (*max*)	**Marcin**	**Kuba**	**Michał**
Appearance	1	5	10
Cooking	10	5	1
House keeping	10	5	1
Cleanliness	10	5	1



Intuitively appealing compromise **Kuba** is not chosen with any given weights

- **Individual minima** for some i=1,…,M:
  $f_i(y) \leq f_i(x)$ for all $x \in S$
- Notation: $f(x) = (f_1(x), f_2(x), \ldots, f_M(x))$
- **Lexicographic optimality** (1)
  $f(y) \leq_{lex} f(x)$ for all $x \in S$
  $f_1(y) \leq f_1(x)$ for all $x \in S$,
  $f_2(y) \leq f_2(x)$ for all $x \in S$ such that $f_1(y) = f_1(x)$,
  etc.

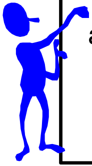
- **Lexicographic optimality** (2)
  $f^{\pi}(y) \leq_{lex} f^{\pi}(x)$ for all $x \in S$ and *some permutation* f^{π} of $(f_1, f_2, \ldots, f_M)$
- Optimize objectives in different lexicographic orders, while keeping the already optimized objective values at their minima



Optimized for lexicographic order (f_1, f_2)

Optimized for lexicographic order (f_2, f_1)

Epsilon Constraint Method (ECM) – A Priori Perspective

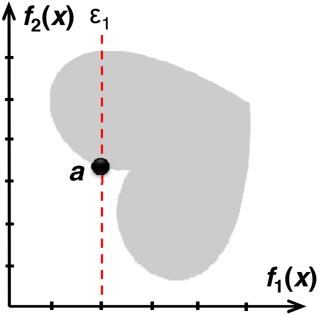


Optimize one of the objectives f_i for some $i=1,\ldots,M$
and restrict the remaining objectives within specified bounds:

$$\begin{aligned} &\textbf{\textit{Minimize}} \quad f_i(x) \\ &\textbf{\textit{subject to}} \; x \in S \\ &\qquad\qquad f_j(x) \le \varepsilon_j \quad \text{for } j=1,\ldots,M \text{ and } j \ne i \end{aligned}$$

Keep f_2 as the objective: **Minimize** $f_2(x)$
s.t. $x \in S$
Treat f_1 as the constraint: $f_1(x) \le \varepsilon_1$

- *Advantage*: applicable to convex or non-convex problems
- *Disadvantage*: the vector ε has to be chosen carefully so that it is within the extreme values of the individual objective function



objective bound ε_1 means that we want a solution x with $f_1(x) \le \varepsilon_1$

Epsilon Constraint Method (ECM) – A Posteriori

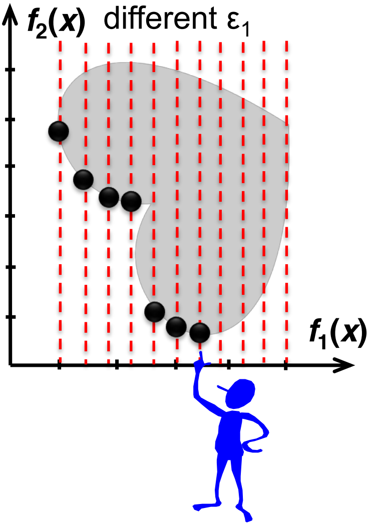
Solve iteratively for different values of ε_j for $j=1,\dots,M$ and $j \neq i$

Minimize $f_i(x)$

_s.t._ $x \in S$

$f_j(x) \leq \varepsilon_j$ for $j=1,\dots,M$ and $j \neq i$

- ε_j can be chosen **equally spaced** within an interval from by the extreme values of objective f_j
- ε_j can be set based on the solution identified in the previous iteration y (ε_j slightly less than $f_j(y)$) – ***backward loop***



Properties of the Epsilon Constraint Method:

- For any $\varepsilon = (\varepsilon_1, \ldots, \varepsilon_{i-1}, \varepsilon_{i+1}, \ldots, \varepsilon_M) \in R^{M-1}$, an identified solution is weakly Pareto-optimal (WPO)
- PO solution can be found by means of lexicographic optimization (or… see next slide)
- Let $x \in S$ be a PO solution; then by setting $f_j(x) = \varepsilon_j$ solution x can be identified (each PO solution can be found with ECM)

Y. Haimes, L. Lasdon, D. Wismer, On a bicriterion formulation of the problems of integrated system identification and system optimization. *IEEE Transac. on Systems, Man and Cybernetics*, 1, 296-297, 1971

Augmented Scalarizing Function

- Many methods (e.g., ECM) guarantee only that possibly not unique solutions obtained for a given set of parameters are merely weakly Pareto-optimal (WPO)
- To guarantee Pareto-optimal (PO) solutions, so-called **augmentation term** is often added to the objective function

$$\rho \cdot \sum_{i=1,\ldots,M} f_i(x) \quad \text{or} \quad \rho \cdot \sum_{i=1,\ldots,M} (f_i(x) - z_i^{ref})$$

  where $\rho > 0$ is a small positive constant
- More often than not, augmented versions are used in practice

Augmented Scalarizing Function:

Minimize $f_i(x) + \rho \cdot \sum_{i=1,\ldots,M} f_i(x)$

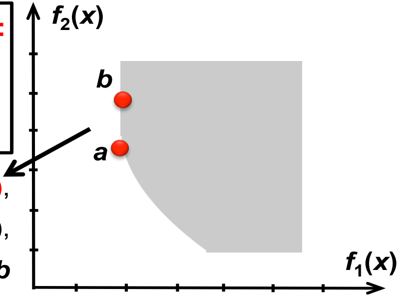
s.t. $x \in S$

$f_j(x) \leq \varepsilon_j$ for $j=1,\ldots,M$ and $j \neq i$

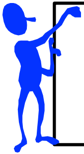
for ***a*** and ***b***: $\rho \cdot \sum_{i=1,\ldots,M} f_i(a) < \rho \cdot \sum_{i=1,\ldots,M} f_i(b)$,

so in case they both minimize $f_i(x)$,

It pays off to return ***a*** rather than ***b***



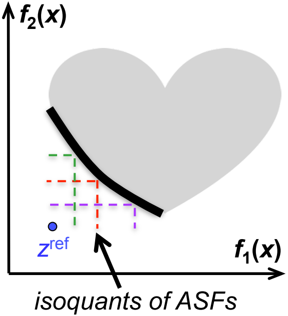
Achievement Scalarizing Function (ASF)



Minimize the weighted Chebyshev distance (i.e., maximal weighted distance on any objective) from the reference point $\mathbf{z}^{\mathbf{ref}}$:

$$\textbf{Minimize} \ \max_{i=1,\dots,M} w_i \cdot (f_i(x) - z_i^{\mathrm{ref}})$$
$$\textbf{subject to} \ x \in S$$

- $\mathbf{z}^{\mathbf{ref}}$ indicates the aspiration levels (desired values that the DM would like to have) for all objectives
- $\mathbf{z}^{\mathbf{ref}}$ can be set to the ideal point $\mathbf{z}^{\mathbf{ideal}}$ (no DM's preferences)
- *Property: for any* $z^{\mathrm{ref}} \in R^N$ and any weight vector w an optimal solution is weakly Pareto-optimal (WPO)
- *Advantage*: when using the ideal point $\mathbf{z}^{\mathbf{ideal}}$ all PO solution can be found
- *Disadvantage*: requires $\mathbf{z}^{\mathbf{ref}}$ (or $\mathbf{z}^{\mathbf{ideal}}$)



A. Wierzbicki. On the completeness and constructiveness of parametric characterizations to vector optimiz. problems. *OR Spektrum*, 8, 73-87, 1986

Augmented Achievement Scalarizing Function (ASF)

- Many methods (e.g., ASF or ECM) guarantee only that possibly not unique solutions obtained for a given set of parameters are merely weakly Pareto-optimal (WPO)

- To guarantee Pareto-optimal (PO) solutions, so-called **augmentation term** is often added to the objective function

$$\rho \cdot \sum_{i=1,\dots,M} f_i(x) \quad \text{or} \quad \rho \cdot \sum_{i=1,\dots,M} (f_i(x) - z_i^{ref})$$

  where $\rho > 0$ is a small positive constant
- More often than not, augmented versions are used in practice

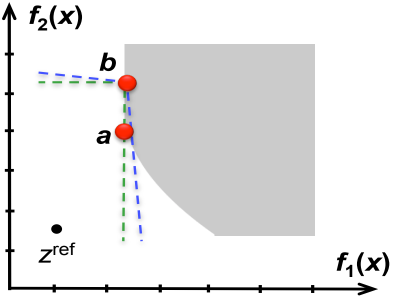
Agumented ASF:

Minimize $\max_{i=1,\dots,M} w_i \cdot (f_i(x) - z_i^{ref}) + \rho \cdot \sum_{i=1,\dots,M} (f_i(x) - z_i^{ref})$

subject to $x \in S$

without augmentation
(***a*** and ***b*** - the same distance)

with augmentation
(***a*** is closer to z^{ref} than ***b***)



Multiple Objective Optimization – Summary

Multiple Objective Optimization is area of Multiple Criteria Decision Making:

- Mathematical optimization problems involving **more than one objective function to be optimized simultaneously**
- **Thousands of applications** in engineering (optimal control, optimal design, process optimization), economics (monetary policy, portfolio selection, production possibilities frontier), logistics, urban planning, …
- **Basic concepts** (objectives, constraints, spaces, solutions, points, …)
- **A priori**, interactive, and **a posteriori** methods
- **Classical** a posteriori **methods** (WSM, ECM) for MOO involve multiple runs with various parameters to approximate Pareto frontier
- **Evolutionary methods** evolve a population of solutions simultaneously