Methods and Applications of Multiobjective Optimization

A review of
The Normal Boundary Intersection Approach
of Indraneel Das and John E. Dennis
and some 'variations' or extensions
by

Trond Steihaug
Department of Informatics
University of Bergen, Norway
and
Humboldt Universität zu Berlin

May 4, 2005

Introduction

Multiobjective or *Multicriteria* optimization

$$\min_{x \in \mathcal{V}} F(x) \equiv \begin{bmatrix} f_1(x) \\ \vdots \\ f_m(x) \end{bmatrix}$$

where $m \geq 2$ and

$$\mathcal{V} = \{x \in \mathbb{R}^n | c_i(x) = 0 \; i \in \mathcal{E}, c_i(x) \geq 0 \; i \in \mathcal{I}\}.$$

The constraints should not be more "difficult" than the available algorithm can handle the problem of solving the (single objective) problem

$$\min_{x \in \mathcal{V}} f_i(x), \quad i = 1, \ldots, m.$$

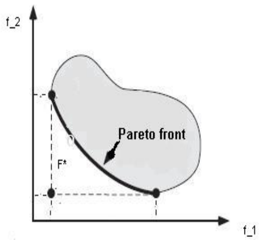
Terminology (1)

A point $x^* \in \mathcal{V}$ is said to be *locally* Pareto optimal if and only if

$$f_i(x) \leq f_i(x^*) \text{ for all } 1 \leq i \leq m \text{ and } x \in \mathcal{V} \cap \mathcal{N}(x^*) \Rightarrow x = x^*.$$

A point $x^* \in \mathcal{C}$ is said to be **globally** Pareto optimal if and only if

$$f_i(x) \leq f_i(x^*) \text{ for all } 1 \leq i \leq m \text{ and } x \in \mathcal{V} \Rightarrow x = x^*.$$

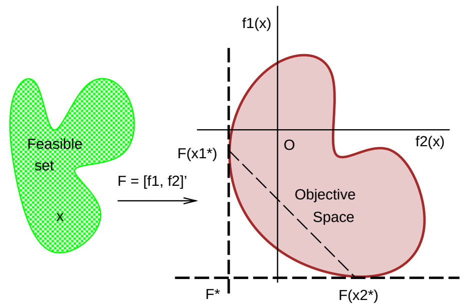


Typically there is an entire curve or surface of Pareto points

Terminology(2)

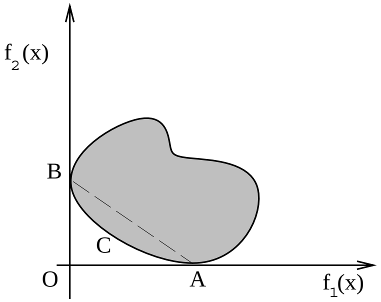
The shadow minimum or utopia point F^* is defined as the vector of the individual global (single objective) $f_i^* \equiv f_i(x_i^*)$, $F^* = (f_1^*, \ldots, f_m^*)^T$ where

$$x_i^* = \text{argmin}\{f_i(x) | x \in \mathcal{V}\}$$



Terminology (3)

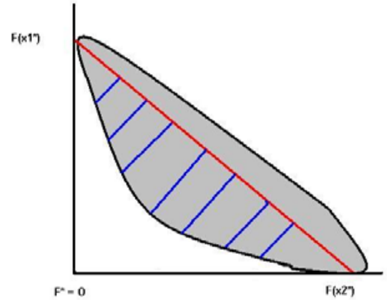
The set of obtainable vectors $\mathcal{F} = \{F(x) | x \in \mathcal{V}\} \subset \mathbb{R}^m$. $\mathcal{F}$ is the objective space. $\partial\mathcal{F}$ is the boundary of $\mathcal{F}$. The set of all Pareto optimal points $\mathcal{P} \subset \partial\mathcal{F}$.



A is $F(x_2^*)$, B is $F(x_1^*)$, C is a Pareto (global) optimal point, O is F^* (from now assumed to be 0). The points on the line A to B is the convex hull of individual minima (CHIM) also called the Utopia line.

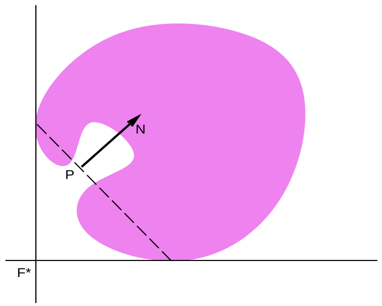
Basic Idea of the NBI method

The intersection point between the boundary $\partial\mathcal{F}$ and the normal pointing toward F^* emanating from any point in the CHIM is a Pareto optimal point



unless it happens . . . .

. . . unless it happens . . .

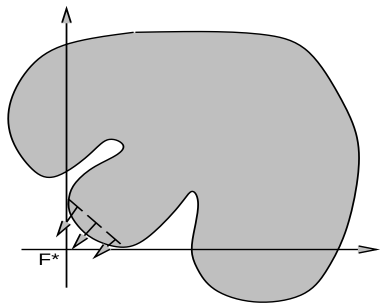


. . . to lay on a sufficiently concave part of the boundary. **Claim** by D&D: The Pareto optimal surface (in the objective space) is convex in almost every application found in the literature.

...but why use NBI when it is convex and a convex combination of objectives will work....

. . . unless it happens . . .

that the algorithm returns a local solution of $\min\{f_i(x), x \in \mathcal{V}\}$ and not the global.



The computed convex hull of the individual local minima is the not the CHIM.

Time to do some mathematics

Formulation of the Problem

Let Φ be an $m \times m$ matrix where column i is $F(x_i^*) - F^*$. Note that $\Phi_{ij} \geq 0$ and $\Phi_{ii} = 0$.

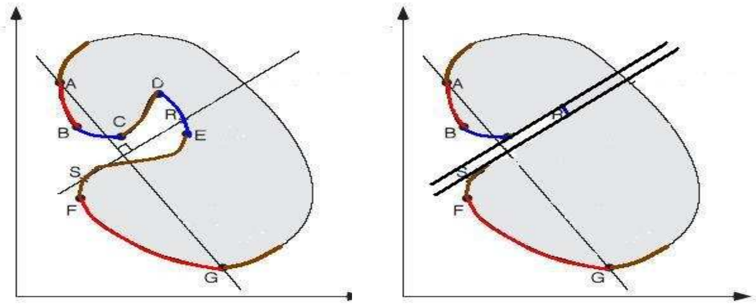
CHIM is now $\{y = \Phi\beta \in \mathbb{R}^m | \sum_{i=1}^m \beta_i = 1, \beta_i \geq 0\}$.

Let $\hat{n}$ be the unit normal to the CHIM simplex pointing toward the origin.

For a given point y in the CHIM the (half) line is $y + t\hat{n}, (t \geq 0)$. The point on this line and ∂F closest to the origin is the follow subproblem

$$\max_{x,t} t$$
$$\text{s.t. } y + t\hat{n} = F(x)$$
$$x \in \mathcal{V}$$

NBI points



A and G are called anchor points.

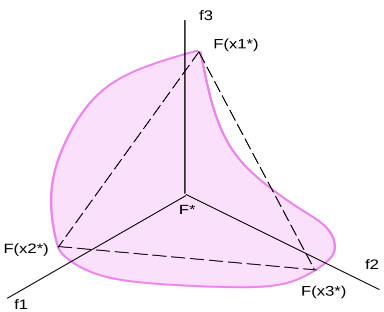
Arcs AB and FG: Global Pareto points. Arcs BC and DE: Local Pareto points

Arcs CD and EF are neither R is a local Pareto point and a NBI point.

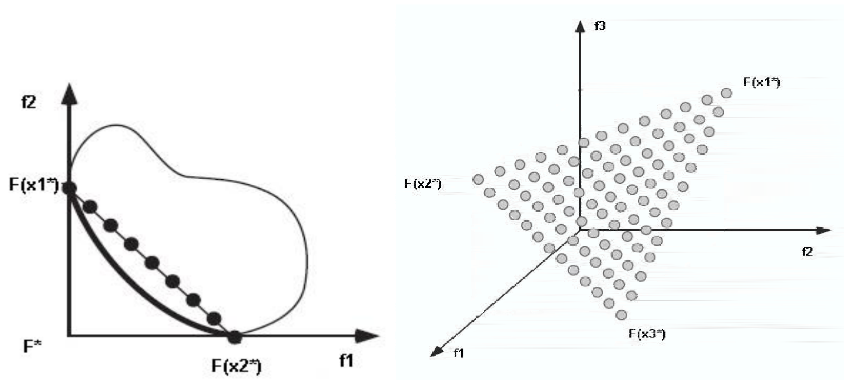
For every Pareto optimal point there exists a NBI point unless . . . .

. . . unless it happens . . .

that $m \geq 3$



Choosing the y in $\max\{t | y + t\hat{n} = F(x), x \in \mathcal{V}\}$



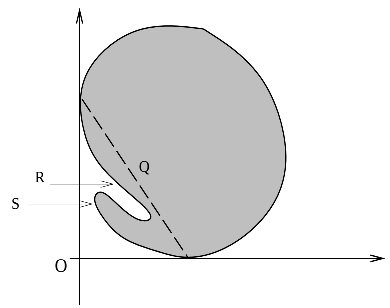
The Optimizer

Most (all) optimization routines for general nonlinear problems will in general find only local solutions. The type of algorithm will determine what kind of problems that can be solved.

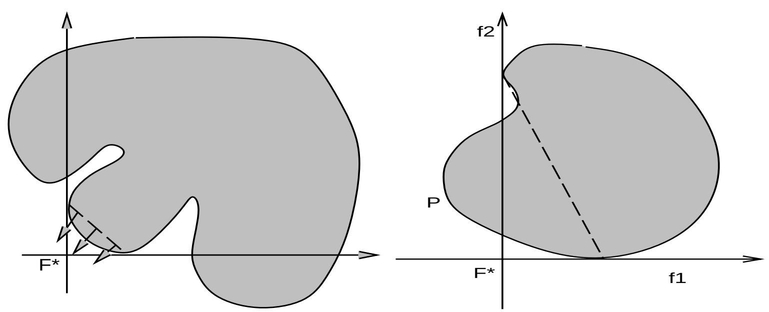
Matlab `fmincon` will handle linear and box constraints separately:

$$
\begin{aligned}
& \min_x f(x) \\
\text{s.t.} \quad & c_i(x) = 0, i \in \mathcal{E} \\
& c_i(x) \geq 0, i \in \mathcal{I} \\
& Ax = b, \quad \widehat{A}x \geq \hat{b} \\
& \ell \leq x \leq u.
\end{aligned}
$$

Local v.s. global solution of $\max\{t | y + t\hat{n} = F(x), x \in \mathcal{V}\}$



Local v.s. global solution of $\min\{f_i(x), x \in \mathcal{V}\}$



The computed convex hull of the individual local minima is the not the CHIM.

Example (N.Kroll)

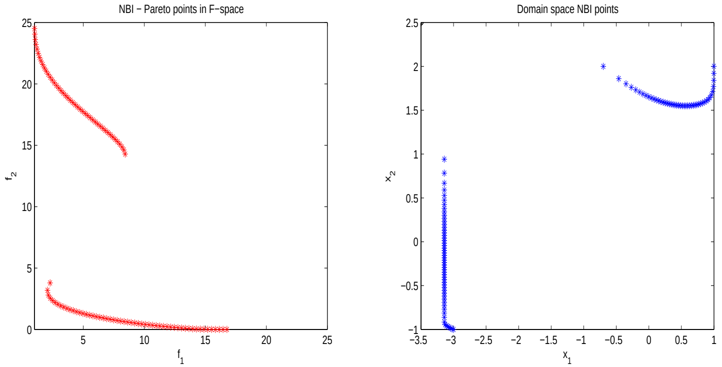
Only box constraints $m = 2$, $n = 2$, $\ell = -\pi \leq x_i \leq \pi = u$

$$
F(x_1, x_2) = \begin{bmatrix} 1 + (\phi_1(1,2) - \phi_1(x_1,x_2))^2 + (\phi_2(1,2) - \phi_2(x_1,x_2))^2 \\ (x_1 + 3)^2 + (x_2 + 1)^2 \end{bmatrix}
$$

Here $\phi_1(x_1,x_2) = 1/2\sin(x_1) - 2\cos(x_1) + \sin(x_2) - 3/2\cos(x_2)$ and $\phi_1(x_1,x_2) = 3/2\sin(x_1) - \cos(x_1) + 2\sin(x_2) - 1/2\cos(x_2)$. Solved with NBI using a very fine mesh of the CHIM.

- First plot shows the computed NBI points (some are not Pareto point)
- Second plot shows the corresponding values in the feasible set (design space)

NBI example (Kroll)



NBI example Das and Dennis

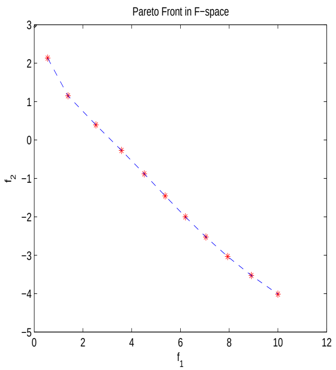
$m = 2$, $n = 5$, linear and nonlinear constraints.

$$F(x) = \begin{bmatrix} \|x\|_2^2 \\ 3x_1 + 2x_2 - \frac{x_3}{3} + \frac{1}{100}(x_4 - x_5)^2 \end{bmatrix}$$

The constraints are

$$x_1 + 2x_2 - x_3 - \frac{x_4}{2} + x_5 = 2$$
$$4x_1 - 2x_2 + \frac{4}{5}x_3 + \frac{3}{5}x_4 + \frac{1}{2}x_5^2 = 0$$
$$\|x\|_2^2 \leq 10$$

NBI example (Das and Dennis)



Coarse mesh on CHIM.

Scaling - Filters - A Reformulation

Several papers from the Multidisciplinary Design and Optimization Laboratory raise concern on scaling and generating only Pareto solutions for the NBI.

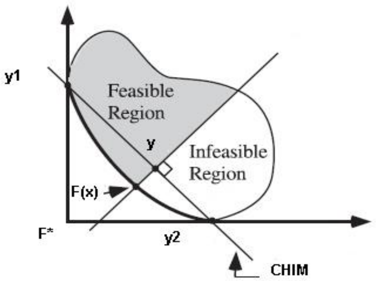
- Rescale F so that $f_i(x_i^*) = 1$ (the anchor points).
- Pareto filter
- Reformulation of the sub problem.

A Pareto filter is a database of the smallest partially ordered computed y in the objective space $y \in \mathcal{F} = \{F(x) | x \in \mathcal{V}\}$.

Normalized Normal Constraint Method

Let $y_i = F(x_i^*)$ and pick reference 'corner' on the CHIM, say y_r. Let y be any (interior)point on the CHIM.

$$\min_{x,t} t$$
$$\text{s.t. } (F(x) - y)^T (y_r - y_k) \leq 0, \quad k \neq r$$
$$t = f_r(x)$$
$$x \in \mathcal{V}$$

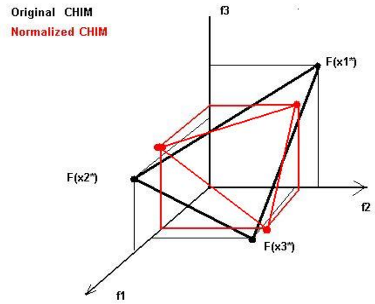


If ... then NBI and NC(Messac 2003) reformulation will generate the same point

Proof: By handwaving.

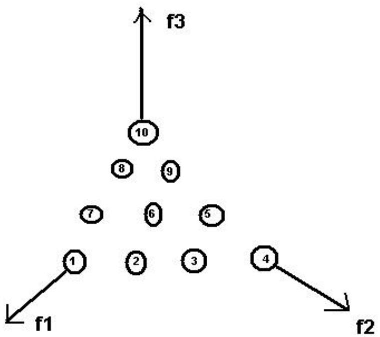
Scaling

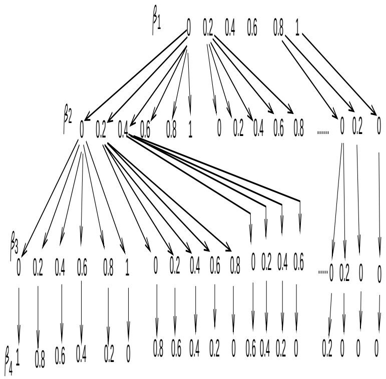
In principle scaling of the anchor points should not effect the computation of the NBI points $\max\{t | y + t\hat{n} = F(x), x \in \mathcal{V}\}$ except for the distribution of the points since $Dy + t\hat{n} = DF(x)$ where D is the scale.



Warmstarts

The number of subproblems to be solved for the NBI and NC depends on the gridding or mesh of the CHIM. It is important to utilize that the subproblems will be 'near' each other (in domain). Normalize and consider $y = \Phi\beta$, $\beta_1, \ldots \beta_m$.





The Road Ahead

Issues not covered:

- No need for the 'exact' normal and D&D use $\hat{n} = -\Phi e$, where e is a column vector of all ones.
- The relationship between NBI and NC and minimization a convex combination of the objectives.

Possible improvements:

- More advanced warm-starts possibilities for the optimizer (more than just the startingpoint)
- Use derivative information (Jacobian of F) to eliminate non-Pareto point.
- Other scalar minimization sub-problems.

The Road Ahead