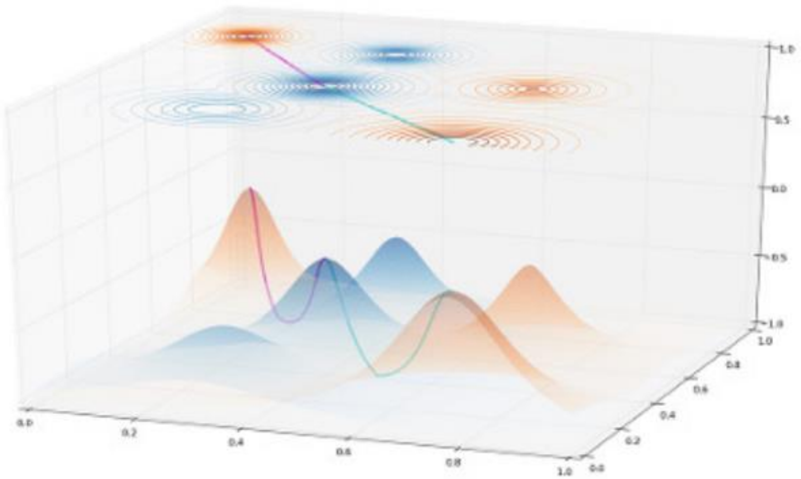
Unit: Efficiency and level sets



Orange Mountains $f_1(x_1, x_2) \rightarrow \max$

Blue Mountains $f_1(x_1, x_2) \rightarrow \max$

Pareto optimization: All Definitions

Decision space $\mathbb{S}$, Feasible decision space $\mathcal{X}$

Objective functions $f_1 : \mathbb{S} \to \mathbb{R}, f_2 : \mathbb{S} \to \mathbb{R}, \ldots, f_m : \mathbb{S} \to \mathbb{R}$.

Or as a vector valued function: $\mathbf{f}(\mathcal{X}) \to \mathbb{R}^m$

Image of $\mathcal{X}$ under $\mathbf{f}$:
$\mathcal{Y} = \mathbf{f}(\mathcal{X}) = \{\mathbf{y} \in \mathbb{R}^m \mid \text{exists } x \in \mathcal{X} : \mathbf{f}(x) = \mathbf{y}\}$

Pareto dominance:
$\forall \mathbf{y}^1, \mathbf{y}^2 \in \mathbb{R}^m : \mathbf{y}^1 \prec \mathbf{y}^2 \Leftrightarrow \mathbf{y}^1 \leq \mathbf{y}^2 \wedge \mathbf{y}^1 \neq \mathbf{y}^2$.

We define a preorder in the feasible decision space $\mathcal{X}$:
$\forall \mathbf{x}^1, \mathbf{x}^2 \in \mathcal{X} : \mathbf{x}^1 \preceq \mathbf{x}^2 :\Leftrightarrow \mathbf{f}(\mathbf{x}^1) \leq \mathbf{f}(\mathbf{x}^2)$

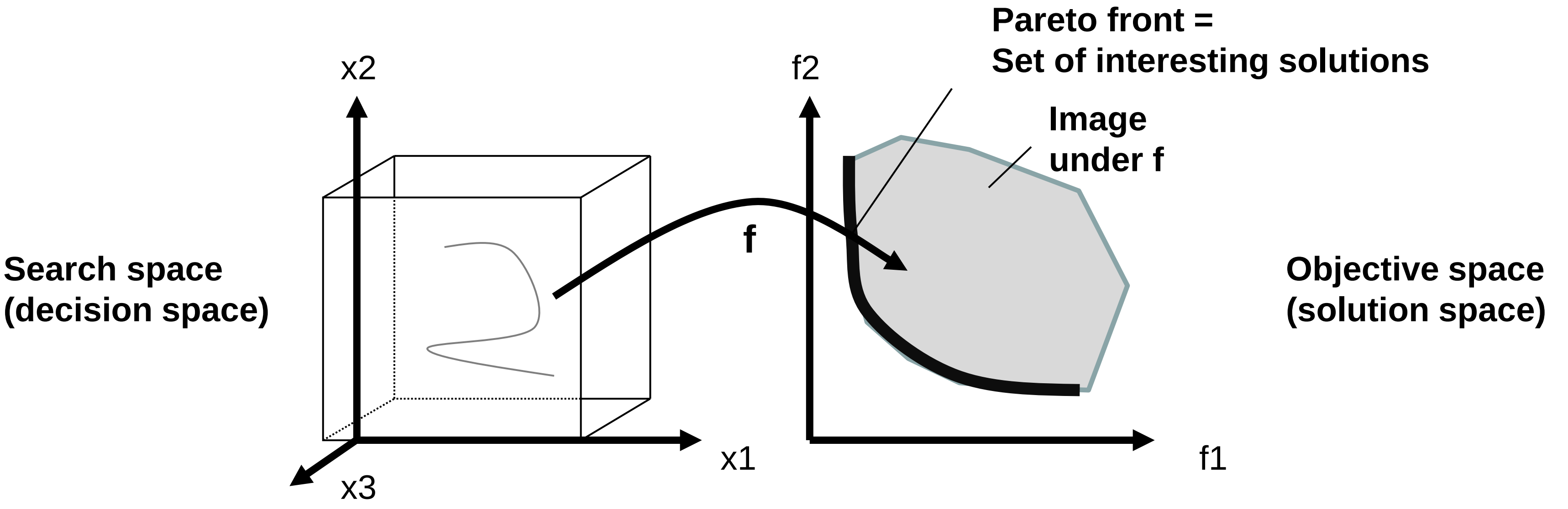
$$\mathbf{x}^1 \prec \mathbf{x}^2 :\Leftrightarrow \mathbf{f}(\mathbf{x}^1) \prec \mathbf{f}(\mathbf{x}^2) \quad \prec$$

$\leq$: *weak componentwise order. In every component smaller or equal*

Matthias Ehrgott: Multicriteria Optimization: Springer 2005
Open Access:
Emmerich, M. T., & Deutz, A. H. (2018). A tutorial on multiobjective optimization: fundamentals and evolutionary methods. *Natural computing*, *17*(3), 585-609.
https://link.springer.com/article/10.1007/s11047-018-9685-y

Learning Goals

1. Correct definition related to multiobjective *optimization*: Efficient set, Pareto front, weak efficient set, strict efficient set, strictly non-dominated set, weakly non-dominated set.
2. Shapes of Pareto fronts: Classification convex/concave and invariances
3. Identification of efficient sets based on contour plots and level sets



Pareto optimization: All Definitions

Efficient point: A point $x \in \mathcal{X}$ is called efficient, iff not exists $x' \in \mathcal{X}$ with $x' \prec x$

Efficient set $\mathcal{X}_E$: Set of all efficient points in $\mathcal{X}$

Nondominated point: A point $\mathbf{y} \in \mathcal{Y}$ is called nondominated (or Pareto optimum), iff not exists $\mathbf{y}' \in \mathcal{Y}$ with $\mathbf{y}' \prec \mathbf{y}$

Nondominated set or Pareto front $\mathcal{Y}_N$: The set of all nondominated points in $\mathcal{Y}$ is called the Pareto front or nondominated set.

Weakly efficient and nondominated set

A point x is weakly efficient, if it there is no other point x' in $\mathcal{X}$ with $f_1(x') < f_1(x) \wedge \ldots \wedge f_m(x') < f_m(x)$.

A point x is strictly efficient, if it there is no other point x' in $\mathcal{X}$ with $x' \preceq x$.

The weakly (strictly) efficient set $\mathcal{X}_{wE}$ ($\mathcal{X}_{sE}$) is the set of all weakly (strictly) efficient points.

A point in $\mathbf{y} \in \mathcal{Y}$ is called weakly non-dominated, iff there is no point in $\mathbf{y}' \in \mathcal{Y}$ such that $y_1{}' < y_1 \wedge \ldots \wedge y_m{}' < y_m$.

The weakly non-dominated set $\mathcal{Y}_{wN}$ is the set of all weakly nondominated solutions in $\mathcal{Y}$.

The weakly non-dominated set $\mathcal{Y}_{wN}$ is the image of $\mathcal{X}_{wE}$ under $\mathbf{f}$,
that is $\mathcal{Y}_{wN} = \mathbf{f}(\mathcal{X}_{wE})$

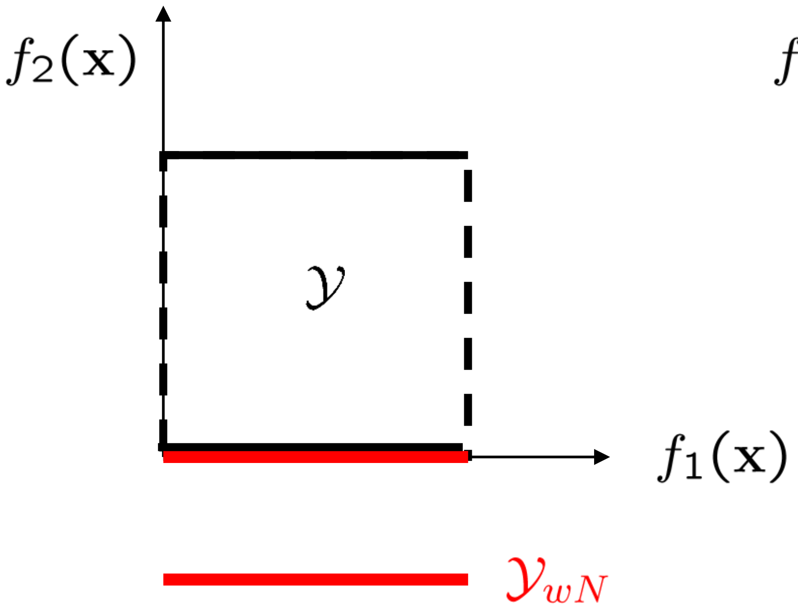
Weak non-domination vs. non-domination

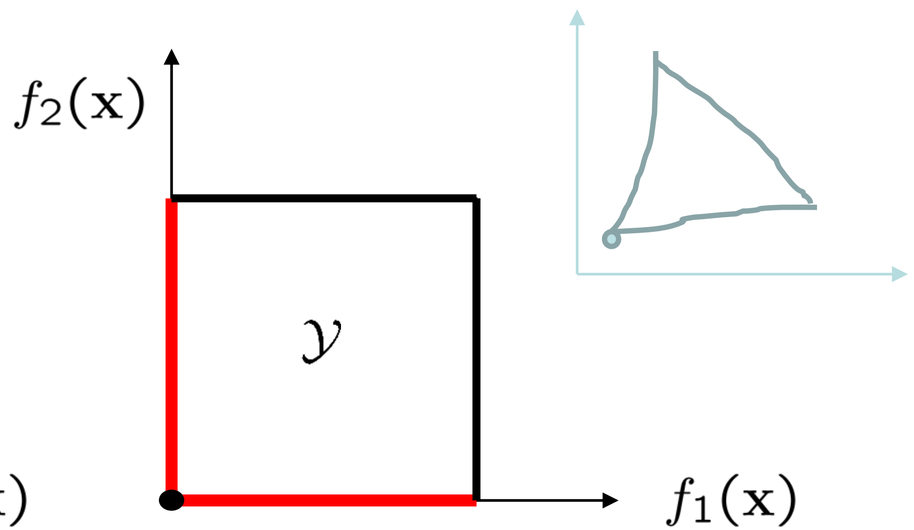
Consider the set $\mathcal{Y} = \{\mathbf{y} \in \mathbb{R}^2 | 0 < y_1 < 1, 0 \leq y_2 \leq 1\}$:

The non-dominated set $\mathcal{Y}_N$ is empty, while $\mathcal{Y}_{wN}$ is not.

Consider the closed square $\mathcal{Y} = \{\mathbf{y} \in \mathbb{R}^2 | 0 \leq y_1 \leq 1, 0 \leq y_2 \leq 1\}$

We have $\mathcal{Y}_N = \{0\}$ and $\mathcal{Y}_{wN} = \{\mathbf{y} \in \mathcal{Y} | y_1 = 0 \vee y_2 = 0\}$



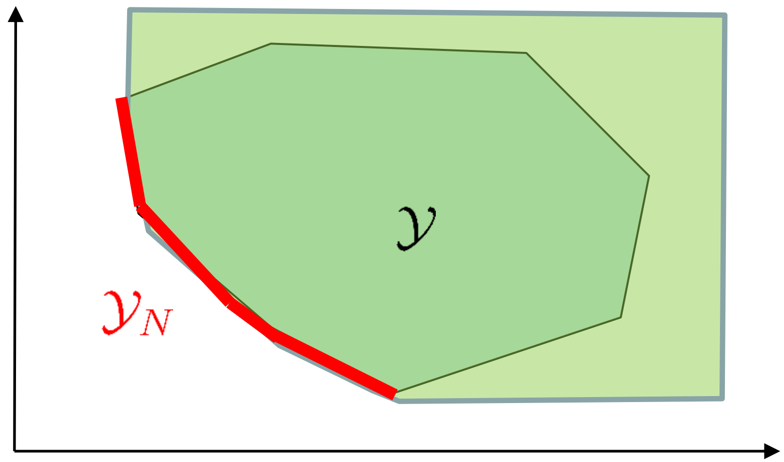


Convex and concave PF: precise definition

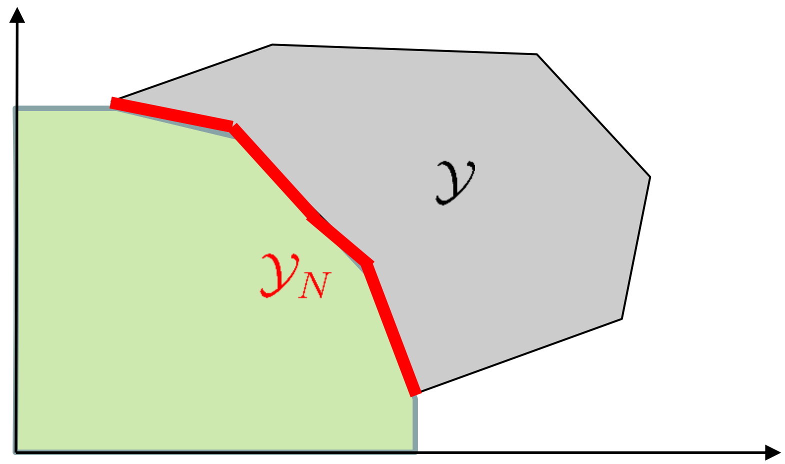
A Pareto front $\mathcal{Y}$ is said to be convex, if $\mathcal{Y} \oplus \mathbb{R}^m_{\geq}$ is a convex set.

A Pareto front $\mathcal{Y}$ is said to be concave if $\mathcal{Y} \oplus \mathbb{R}^m_{\leq}$ is a convex set.

Convex pareto front



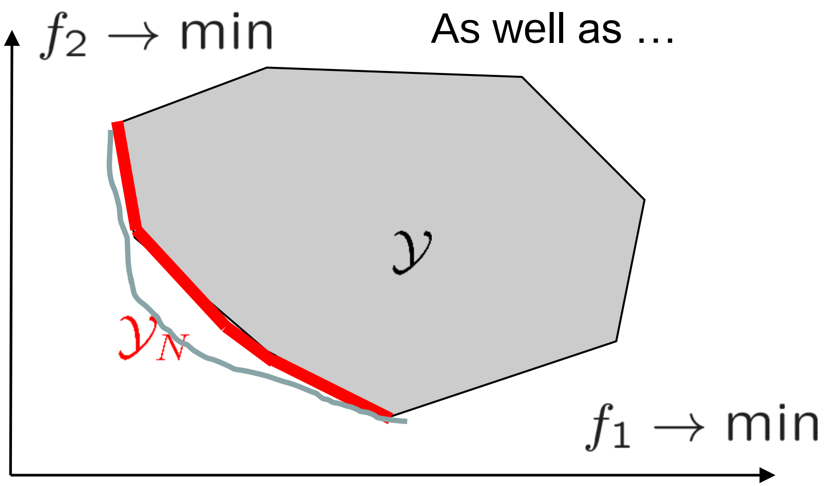
Concave pareto front



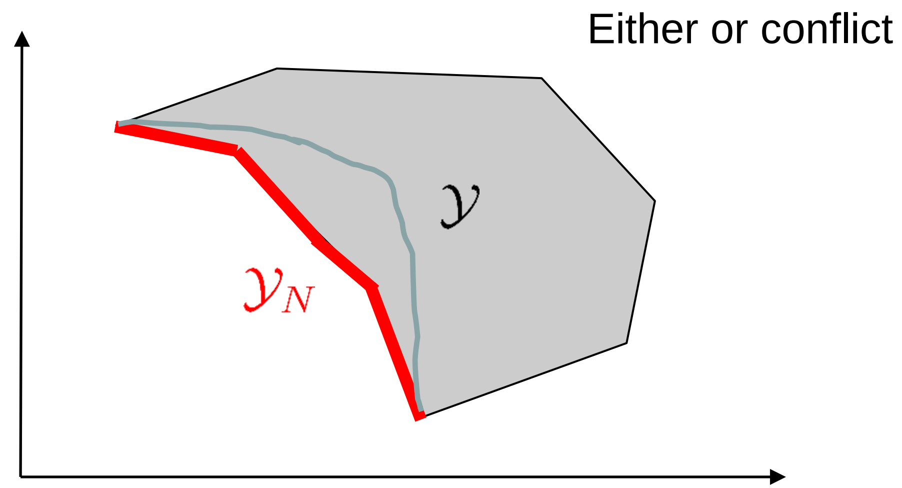
… convex: monotonically decreasing, slope is increasing
… concave: “ “ is decreasing (gets more negative)

Different shapes of Pareto fronts

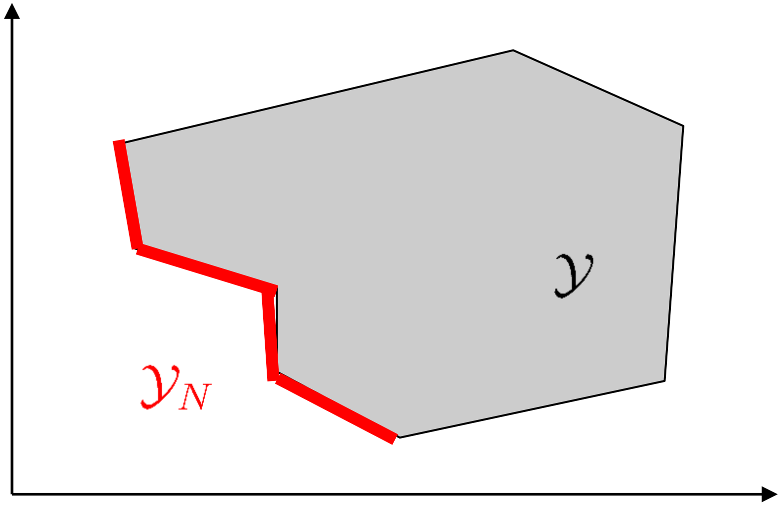
Convex pareto front



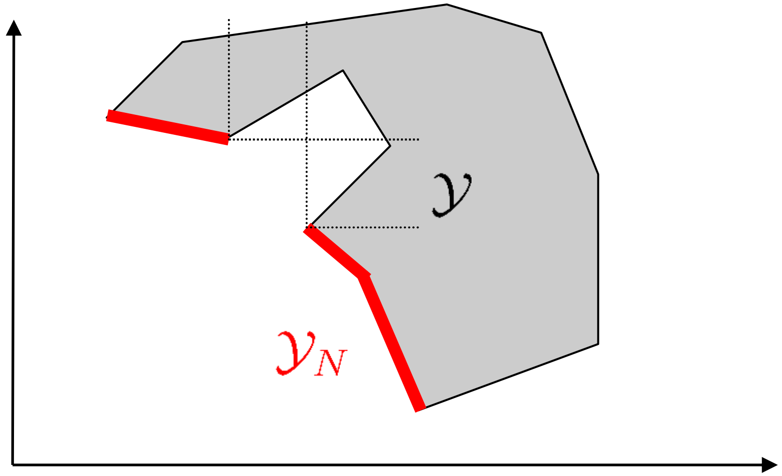
Concave pareto front



PF that is neither convex nor concave.



Disconnected Pareto front

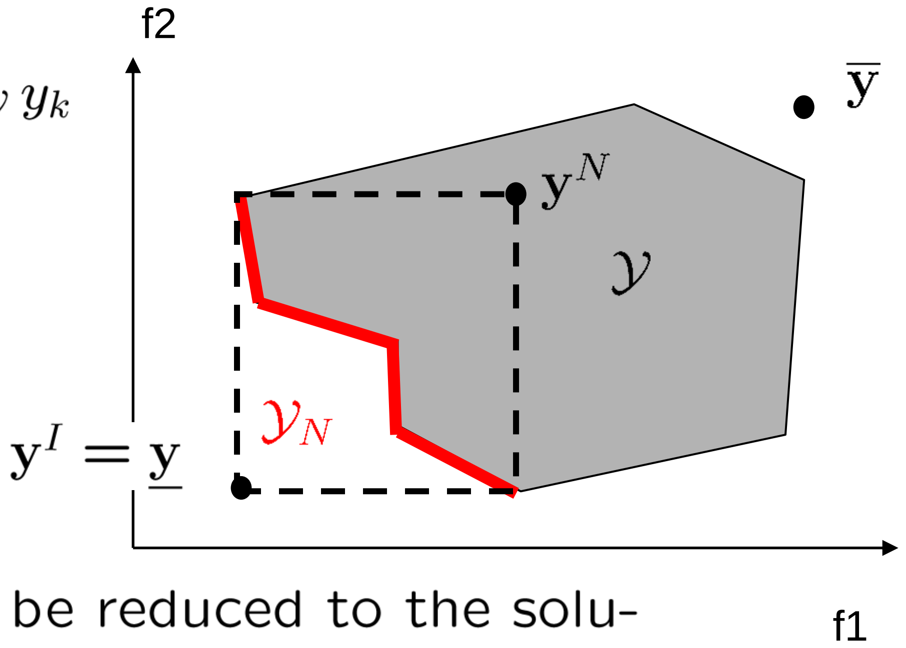


Special points

Ideal vector: $y_k^I := \underline{y}_k := \min_{\mathbf{y}\in\mathcal{Y}} y_k$

Maximal point: $\overline{y}_k = \max_{\mathbf{y}\in\mathcal{Y}} y_k$

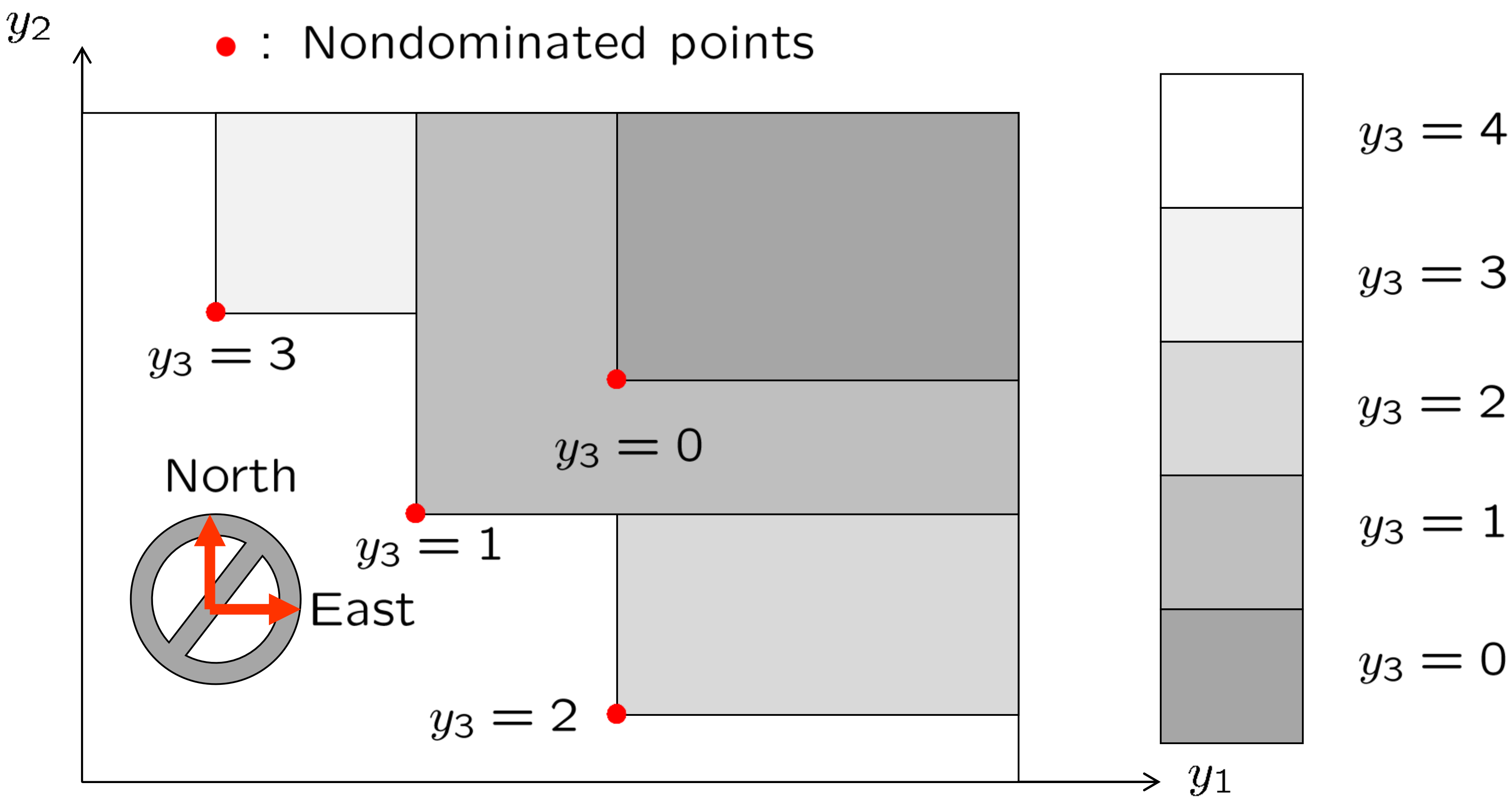
Nadir point: $y_k^N = \max_{\mathbf{y}\in\mathcal{Y}_N} y_k$



Computation of ideal point can be reduced to the solution of m single-objective optimization problems

The computation of the Nadir point is a very difficult problem and no efficient method for computing $\mathbf{y}^N$ is known for $m > 2$, yet.

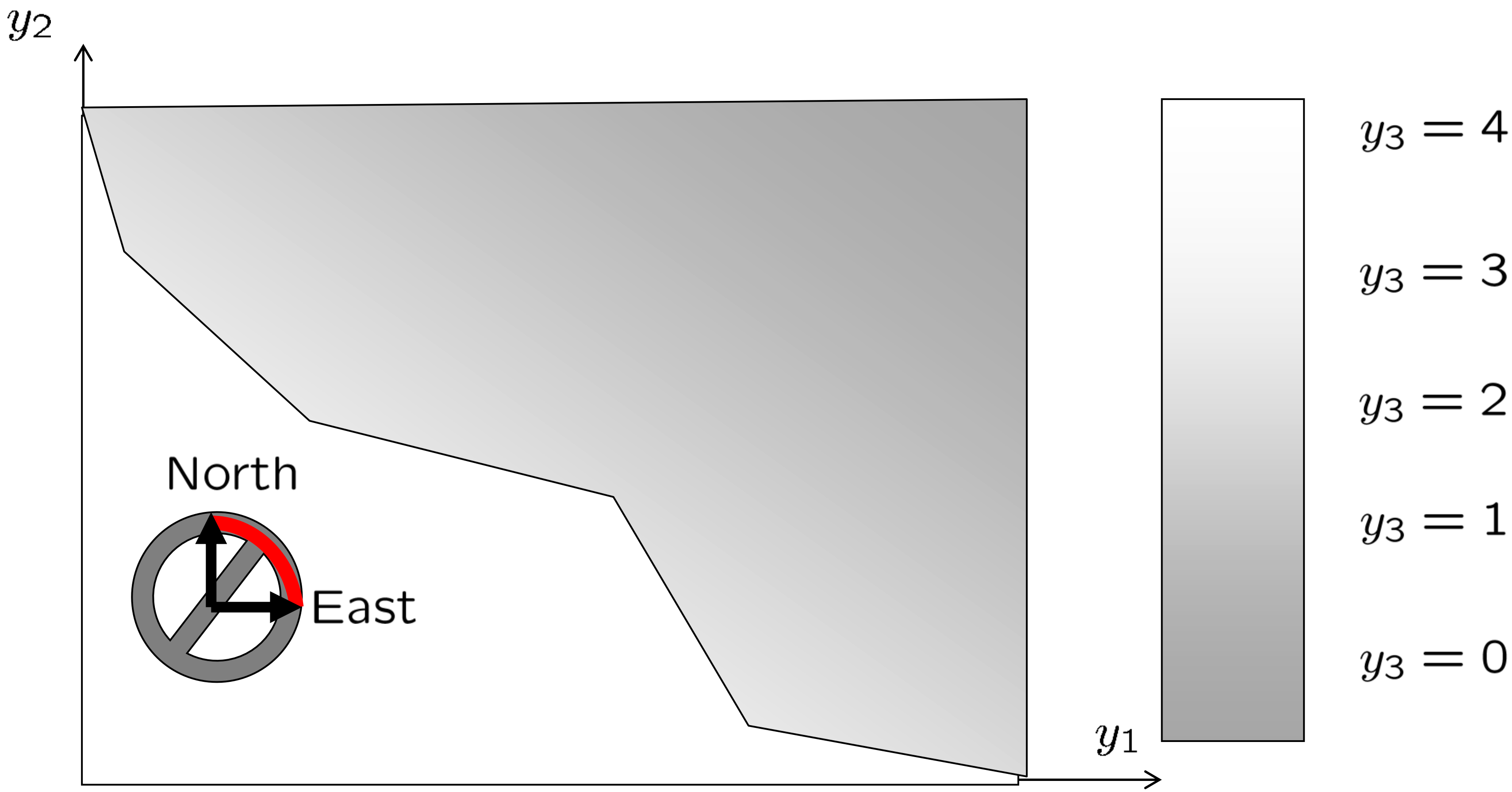
3-D Attainment surface, dominated space



3D Attainment surface: Useful for visualizing finite non-dominated sets in 3-D

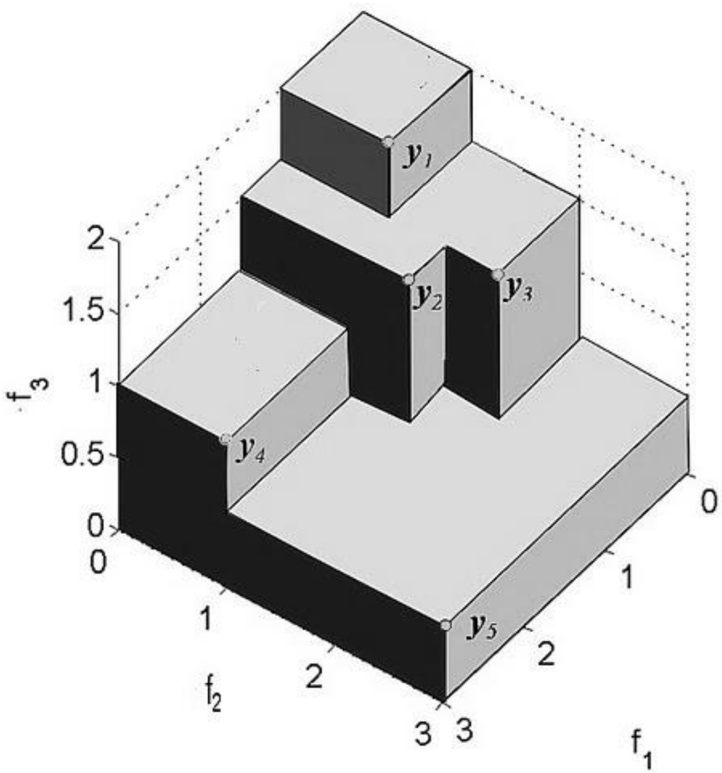
'Steps' into direction north to east.

3-D Attainment Surface, Continuous

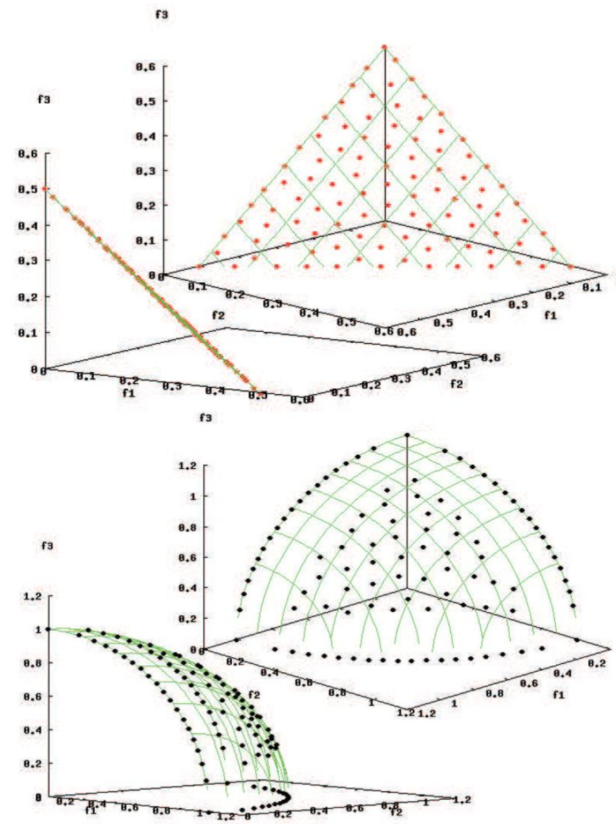


The slope of the attainment surface is always in the direction north-northeast-east

Pareto front in three dimensons



Visualization of finite PF with 5 points.



3-D continuous Pareto fronts and approximations to them with 70 points.

Here maximization is considered: Dominance cones are the negative orthants

Optima seeking using contour plots

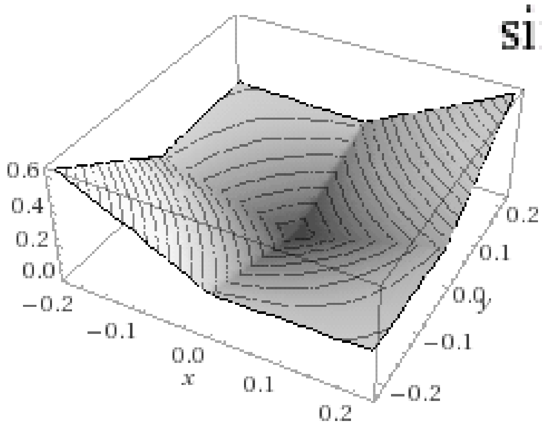
Contour plots help to localize optimizers of single-objective problems.

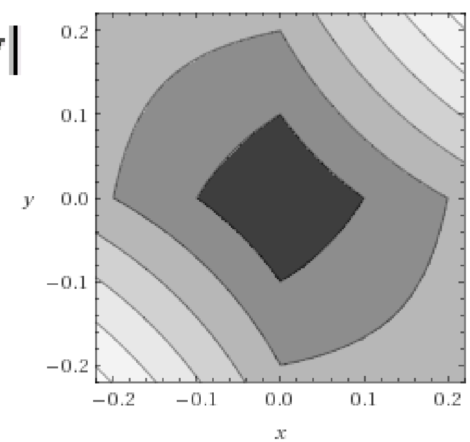
Often, they provide an intuition for reasoning about optima for higher dimensional functions.

A level set is informally defined as a set of arguments (variable settings) for which the function obtains the same value.

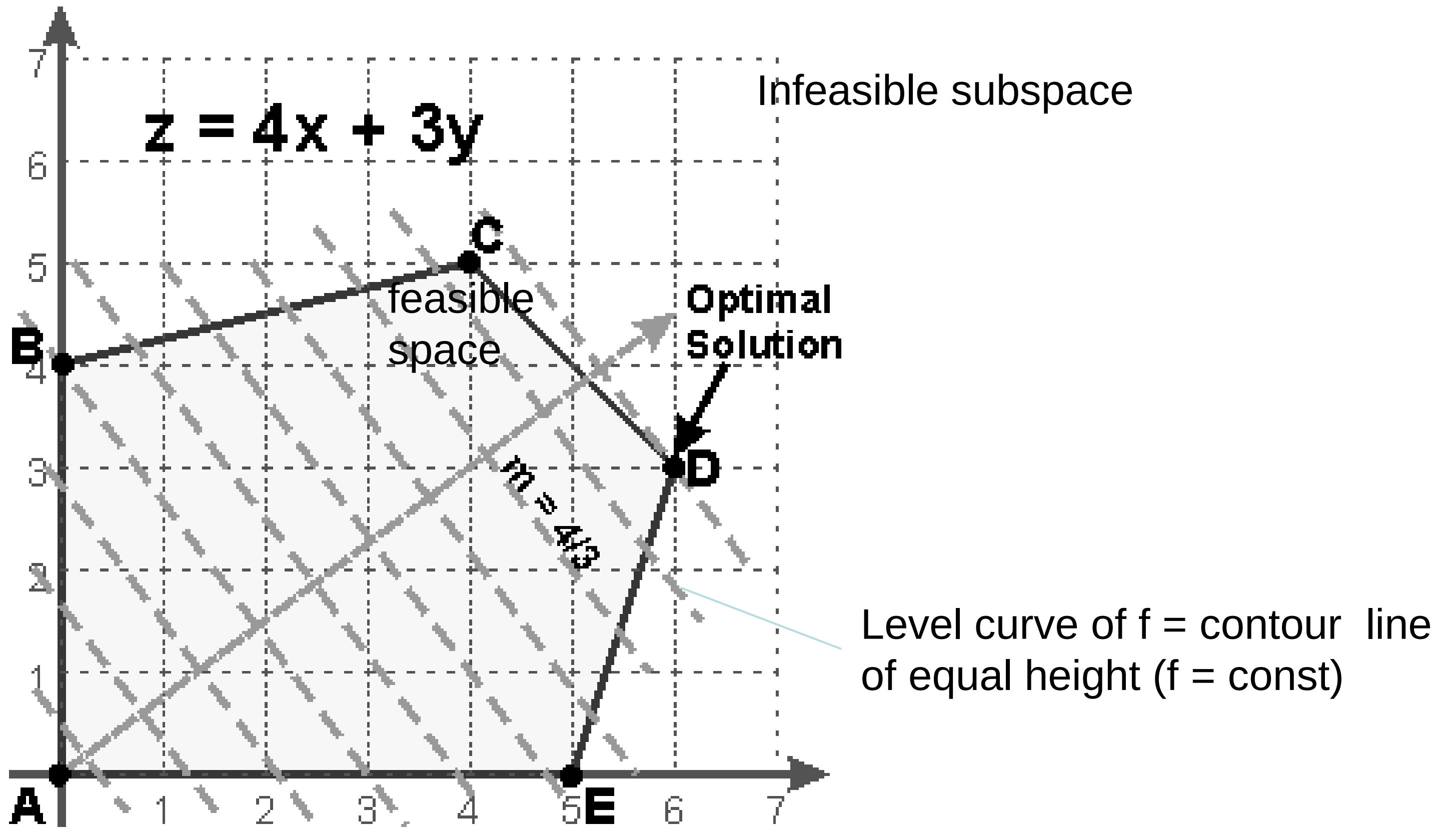
A contour is a connected part of a level set of a 2-dimensional function.

$$\sin(4\,x\,y) + |x| + |y|$$



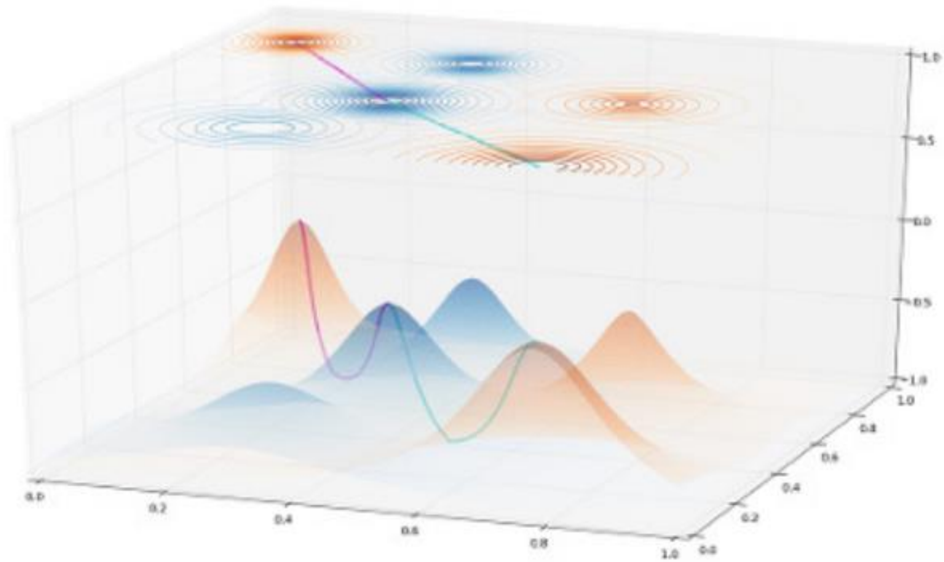


Finding efficient set using level sets (contours): Single objective optimization, linear case



Draw constraint boundaries $g_i(\mathbf{x}) = 0$ and contours for $f(\mathbf{c}) \equiv C$ for different constants C.

Hillclimbing in Multiobjective Landscapes

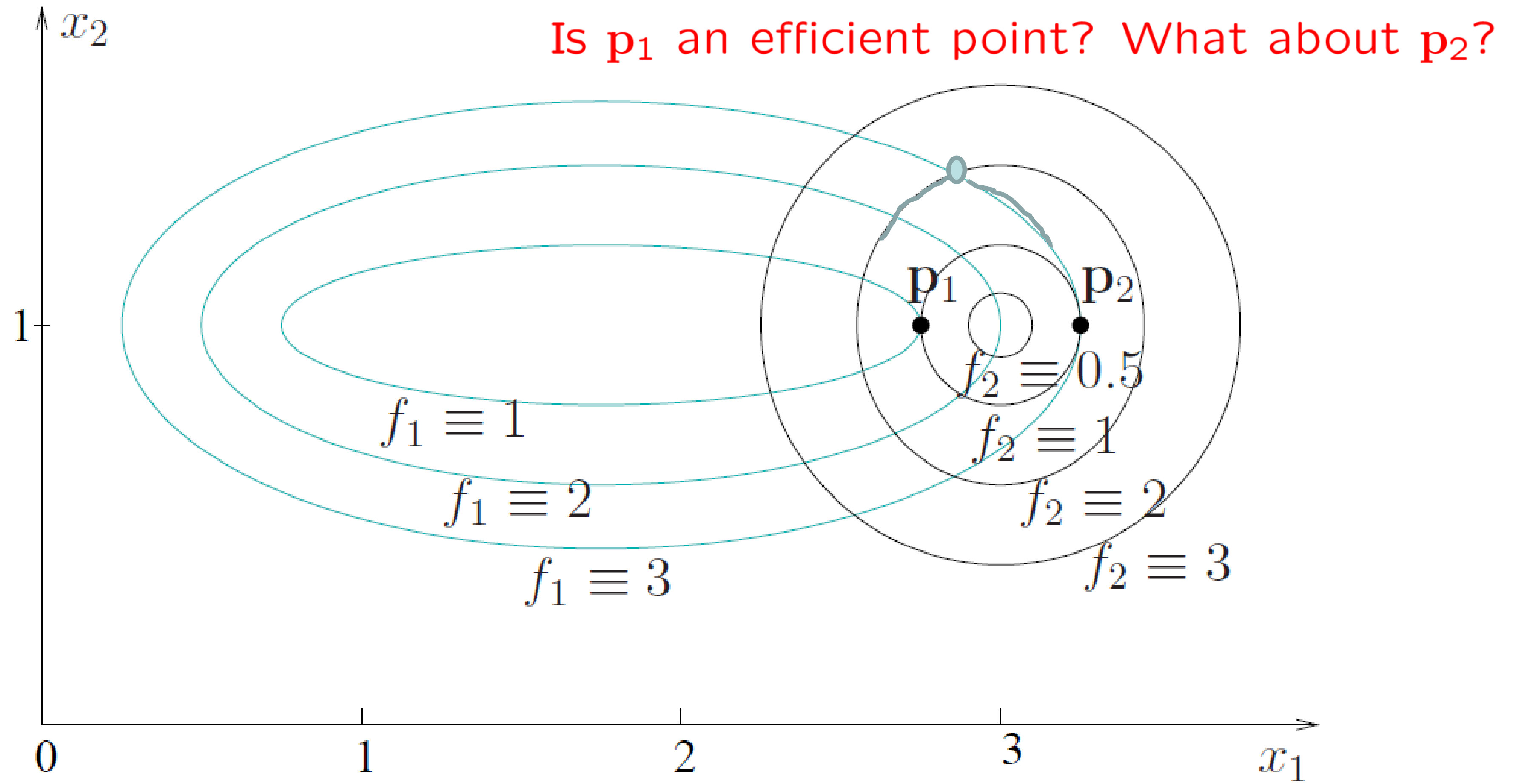


Orange Mountains $f_1(x_1, x_2) \rightarrow \max$

Blue Mountains $f_1(x_1, x_2) \rightarrow \max$

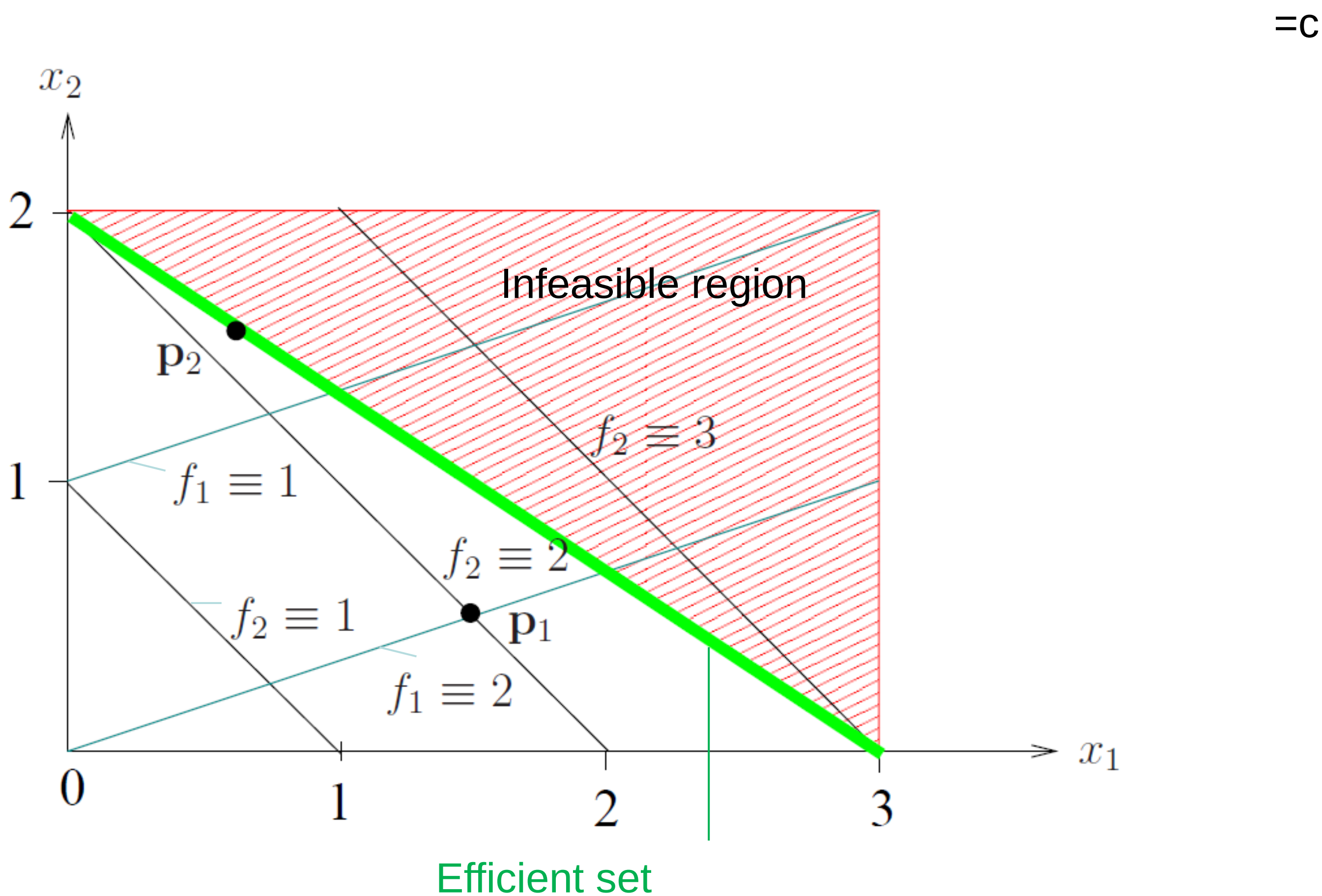
Finding efficient points using contour plots

Contour plots can sometimes be used to find efficient points in bi-objective optimization graphically.



Tangential points of contours are often efficient point

Finding the efficient set in IR^2: Example!!!!!



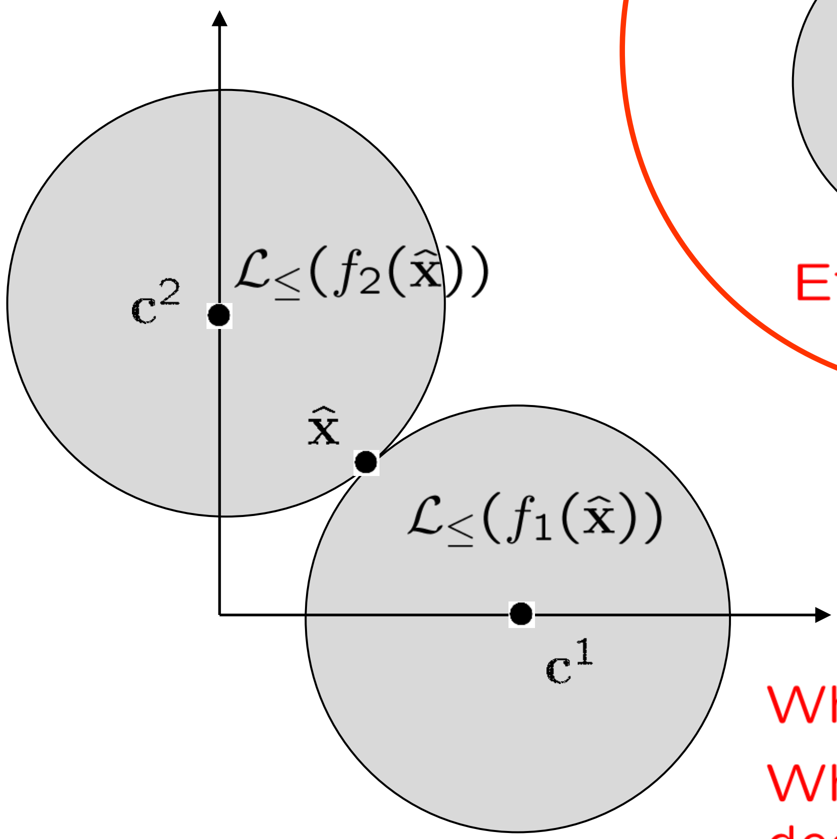
Indicate region that is dominated by $\hat{\mathbf{p}}_1$.

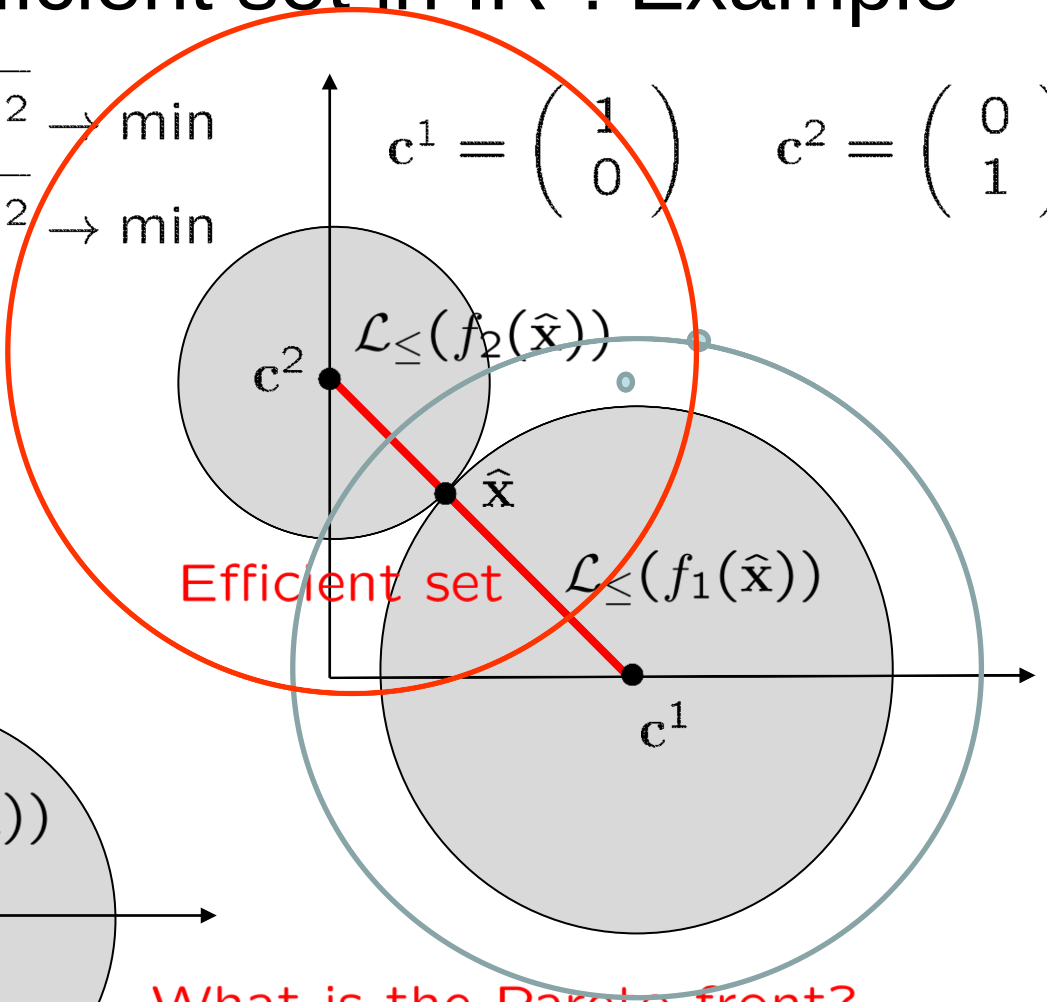
Finding the efficient set in IR^2: Example

$$f_1(\mathbf{x}) = \sqrt{\sum_{i=1}^{2}(x_i - \mathbf{c}_i^1)^2} \rightarrow \min$$

$$f_2(\mathbf{x}) = \sqrt{\sum_{i=1}^{2}(x_i - \mathbf{c}_i^2)^2} \rightarrow \min$$

$$\mathbf{c}^1 = \begin{pmatrix} 1 \\ 0 \end{pmatrix} \qquad \mathbf{c}^2 = \begin{pmatrix} 0 \\ 1 \end{pmatrix}$$





What is the Pareto front?

What about three dimensional decision spaces?

Facility location problems

Take home messages

Important definitions in Pareto optimization are the (weakly, strictly) efficient set, Pareto front, ideal/nadir point, (feasible) decision/objective space

Pareto fronts can be convex or concave, connected or disconnected

Theorems on level sets can be used to identify (globally) efficient points analytically; they are useful for reasoning about the location of the efficient set;

Often optima occur at the constraint boundary; In particular, for linear problems this is the case. In 2-D contour plots can be used to identify efficient solutions at the boundary.

Additional Material

Level sets and curves

Level sets can be used to visualize $\mathcal{X}_E$, $\mathcal{X}_{wE}$ and $\mathcal{X}_{sE}$ for continuous spaces:

$$\mathcal{L}_{\leq}(f(\hat{\mathbf{x}})) = \{\mathbf{x} \in \mathcal{X} : f(\mathbf{x}) \leq f(\hat{\mathbf{x}})\} : \textit{ Level set}$$

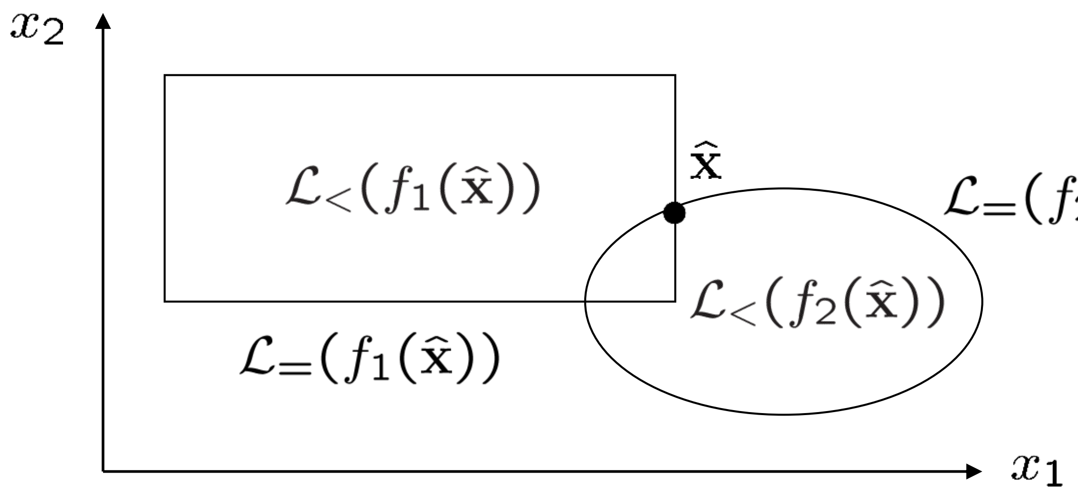
$$\mathcal{L}_{=}(f(\hat{\mathbf{x}})) = \{\mathbf{x} \in \mathcal{X} : f(\mathbf{x}) = f(\hat{\mathbf{x}})\} : \textit{ Level curve}$$

$$\mathcal{L}_{<}(f(\hat{\mathbf{x}})) = \{\mathbf{x} \in \mathcal{X} : f(\mathbf{x}) < f(\hat{\mathbf{x}})\} : \textit{ Strict level set}$$

Draw the level set $\mathcal{L}_{\leq}(f(\mathbf{x}_0))$ for $f(\mathbf{x}) = |\mathbf{1} - \mathbf{x}|^2 = (x_1 - 1)^2 + (x_2 - 1)^2$ and $\mathbf{x}_0 = (1, 0)$ in the x_1, x_2 plane !

Finding Efficient Points by Level Sets: Example 1

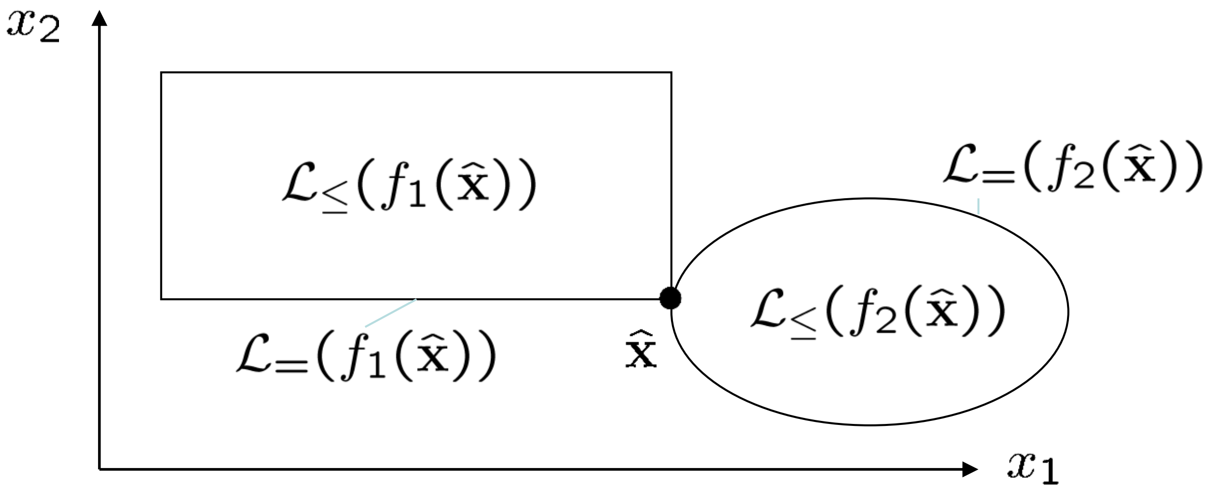
Level sets can be used to determine whether $\hat{\mathbf{x}} \in \mathcal{X}$ is (strictly, weakly) non-dominated or not.



The point $\hat{\mathbf{x}}$ cannot be nondominated! Why ?

Answer: Dominating solutions are in the area where the two strict level sets intersect.

Finding Efficient Points by Level Sets: Example 2



Is $\hat{\mathbf{x}}$ efficient?

Answer: It is not possible to improve f_1 and f_2 at the same time relative to their values in $\hat{\mathbf{x}}$. Therefore, $\hat{\mathbf{x}}$ is efficient.

Level Sets

The point $\hat{\mathbf{x}}$ can only be efficient if its level sets intersect in level curves.

$$\mathbf{x} \text{ is efficient} \Leftrightarrow \bigcap_{k=1}^{m} \mathcal{L}_{\leq}(f_k(\mathbf{x})) = \bigcap_{k=1}^{m} \mathcal{L}_{=}(f_k(\mathbf{x}))$$

The point $\hat{\mathbf{x}}$ can only be weakly efficient if its strict level sets do not intersect.

$$\mathbf{x} \text{ is weakly efficient} \Leftrightarrow \bigcap_{k=1}^{m} \mathcal{L}_{<}(f_k(\mathbf{x})) = \emptyset$$

The point $\hat{\mathbf{x}}$ can only be strictly efficient if its level sets intersect in exactly one point.

$$\mathbf{x} \text{ is strictly efficient} \Leftrightarrow \bigcap_{k=1}^{m} \mathcal{L}_{\leq}(f_k(\mathbf{x})) = \{\mathbf{x}\}$$

Proof: Theorem on efficient points

The point $\hat{\mathbf{x}}$ can only be efficient if its level sets intersect in level curves.

$$\hat{\mathbf{x}} \text{ is efficient} \Leftrightarrow \bigcap_{k=1}^{m} \mathcal{L}_{\leq}(f_k(\widehat{x})) = \bigcap_{k=1}^{m} \mathcal{L}_{=}(f_k(\widehat{x}))$$

Proof:
$\hat{\mathbf{x}}$ is efficient

$\Leftrightarrow$ there is no $\mathbf{x}$ such that both $f_k(\mathbf{x}) \leq f_k(\hat{\mathbf{x}})$ for all $k = 1, \ldots, m$ and $f_k(\mathbf{x}) < f(\hat{\mathbf{x}})$ for at least one $k = 1, \ldots, m$

$\Leftrightarrow$ there is no $\mathbf{x} \in \mathcal{X}$ such that both $\mathbf{x} \in \cap_{k=1}^{m} \mathcal{L}_{\leq}(f(\hat{\mathbf{x}}))$ and $\mathbf{x} \in \mathcal{L}_{<}(f_j(\hat{\mathbf{x}}))$ for some j

$\Leftrightarrow \bigcap_{k=1}^{m} \mathcal{L}_{\leq}(f_k(\widehat{x})) = \bigcap_{k=1}^{m} \mathcal{L}_{=}(f_k(\widehat{x}))$