# Introduction to optimization

G. Ferrari Trecate

Dipartimento di Ingegneria Industriale e dell'Informazione
Università degli Studi di Pavia

Industrial Automation

# Optimization

Optimization is also known as mathematical programming

- *Programming* means planning or building an action plan for solving a problem or tacking a decision
- Optimization falls in the fields of operations research and management science.

# Optimization

## Basic problem

$$\min_{\substack{g_i(x)\leq 0\\ i=1,2,\ldots,m}} \quad f(x)$$

- Variables: $x = [x_1, \ldots, x_n]^{\mathrm{T}}$
- Constraints: $g_i : \mathbb{R}^n \to \mathbb{R}$, $i = 1, 2, \ldots, m$.
- Feasible region

$$X = \{x \in \mathbb{R}^n : g_1(x) \leq 0, \cdots, \ g_m(x) \leq 0\}$$

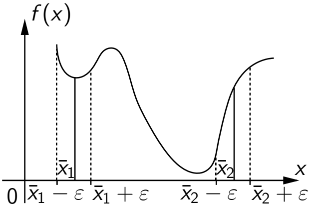
- Feasible solution or feasible point: $x \in X$
- Objective function (or cost): $f : X \to \mathbb{R}$

# Optimization

## Basic problem

$$\min_{\substack{g_i(x)\leq 0\\ i=1,2,\ldots,m}} \quad f(x)$$

- $x^* \in X$ is an *optimal solution* (global minimum point) if $f(x^*) \leq f(x),\ \forall x \in X$
- $\bar{x} \in X$ is a *local optimal solution* (local minimum point) if $\exists \epsilon > 0 : \forall x \in X,\ \|x - \bar{x}\| < \epsilon \Rightarrow f(\bar{x}) \leq f(x)$



# Optimization

## Basic problem

$$\min_{\substack{g_i(x)\leq 0\\ i=1,2,\ldots,m}} \quad f(x)$$

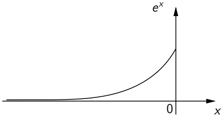
The basic problem can be:

- *infeasible* (if $X = \varnothing$)
- *unbounded* (if $\forall k < 0\, \exists x \in X : f(x) < k$).

even if the basic problem is feasible and bounded, optimal solutions could

- not exist; e.g.

$$\min_{x\leq 0} e^x \quad x \in \mathbb{R}$$




- exist and be not unique (e.g. $f$ costant)

# Optimization

## Basic problem

$$\min_{\substack{g_i(x) \leq 0 \\ i=1,2,\dots,m}} \quad f(x)$$

No easy way to solve the basic problem in its full generality !

- Need of numerical algorithms
- Often, only local optimal solutions can be computed

# Optimization

## Maximum problems

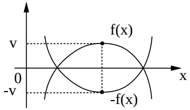
$$\max_{\substack{g_i(x)\leq 0 \\ i=1,2,\ldots,m}} \quad f(x)$$

- the problem is *unbounded* if $\forall k > 0\, \exists x \in X : f(x) > k$.
- $x^* \in X$ is an *optimal solution* (global maximum point) if $f(x^*) \geq f(x),\ \forall x \in X$
- $\bar{x} \in X$is a *local optimal solution* (local maximum point) if $\exists \epsilon > 0 : \forall x \in X,\ \|x - \bar{x}\| < \epsilon \Rightarrow f(\bar{x}) \geq f(x)$

# Conversions in the basic problem form

- Conversions maximum/minumum

$$\max_{x \in X} f(x) = -\min_{x \in X} -f(x)$$

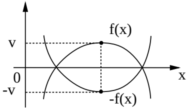



Optimal solutions are the same for both problems

# Conversions in the basic problem form

- Conversions maximum/minumum

$$\max_{x \in X} f(x) = -\min_{x \in X} -f(x)$$




Optimal solutions are the same for both problems

- Conversion form "$\geq$" to "$\leq$" in the constraints

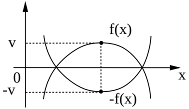
$$\{x \in \mathbb{R}^n : g(x) \geq 0\} = \{x \in \mathbb{R}^n : -g(x) \leq 0\};$$

The feasible region does not change

# Conversions in the basic problem form

- Conversions maximum/minumum

$$\max_{x \in X} f(x) = -\min_{x \in X} -f(x)$$




Optimal solutions are the same for both problems

- Conversion form "$\geq$" to "$\leq$" in the constraints

$$\{x \in \mathbb{R}^n : g(x) \geq 0\} = \{x \in \mathbb{R}^n : -g(x) \leq 0\};$$

The feasible region does not change

- Conversion from "$=$" to inequalities in the constraints

$$\{x \in \mathbb{R}^n : g(x) = 0\} = \{x \in \mathbb{R}^n : g(x) \leq 0,\ g(x) \geq 0\};$$

An equality constraint is replaced by two inequality constraints

# Classes of optimization problems

## Basic problem

$$\min_{\substack{g_i(x)\le 0\\ i=1,2,\dots,m}} \quad f(x)$$

- $f$ is quadratic if $f(x) = x^{\mathrm{T}}Qx + c^{\mathrm{T}}x$ ($Q$ matrix, $c$ vector)
- $f$ is linear if $f(x) = c^{\mathrm{T}}x$
- $f$ is affine if $f(x) = c^{\mathrm{T}}x + b$ ($b$ constant)

# Classes of optimization problems

## Basic problem

$$\min_{\substack{g_i(x)\leq 0\\ i=1,2,\ldots,m}} \quad f(x)$$

- $f$ is quadratic if $f(x) = x^{\mathrm{T}}Qx + c^{\mathrm{T}}x$ ($Q$ matrix, $c$ vector)
- $f$ is linear if $f(x) = c^{\mathrm{T}}x$
- $f$ is affine if $f(x) = c^{\mathrm{T}}x + b$ ($b$ constant)

## Notable problems for which efficient algorithms exist

- $f$ is convex and $g_i$ are convex $\Rightarrow$ *convex* programming
- if $f$ is quadratic and $g_i$ are affine $\Rightarrow$ *quadratic* programming
- if $f$ is linear and $g_i$ are affine $\Rightarrow$ *linear* programming

If the variables must also verify $x \in \mathbb{Z}^n$ we have an *integer* programming problem (mixed-integer programming problem if only a subset of variables is constrained to integer values)

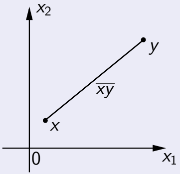
# Convex programming

## Definition

Given two points $x, y \in \mathbb{R}^n$, the set

$$\overline{xy} = \{\lambda x + (1-\lambda)y : \lambda \in [0,1]\}$$

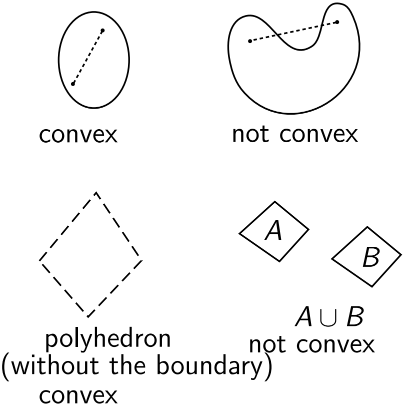
is a *segment* joining $x$ and $y$.




## Definition

The set $X \subseteq \mathbb{R}^n$ is *convex* if $\forall x, y \in X$ one has $\overline{xy} \in X$.

# Examples




$\mathbb{R}^n$ is convex

# Convexity and intersection

**Proposition (try to prove it at home !)**

The intersection of two convex sets is a convex set

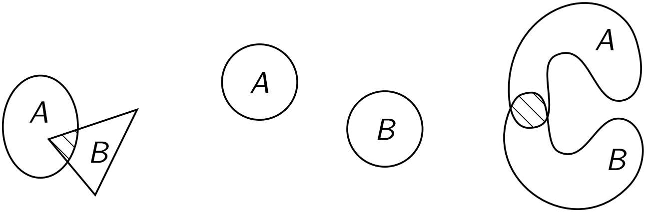
It follows that the empty set is convex

# Convexity and intersection

**Proposition (try to prove it at home !)**

The intersection of two convex sets is a convex set

It follows that the empty set is convex



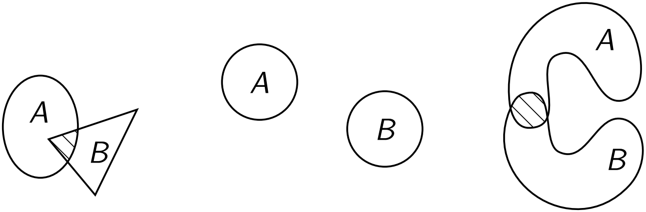

$A \cap B$ convex　　$A \cap B = \emptyset$ convex　　$A \cup B$ not convex, $A \cap B$ convex

# Convexity and intersection

**Proposition (try to prove it at home !)**

The intersection of two convex sets is a convex set

It follows that the empty set is convex




$A \cap B$ convex

$A \cap B = \emptyset$ convex

$A \cup B$ not convex
$A \cap B$ convex

The union of two convex sets is not convex, in general

# Convex functions

## Definition

A function $f : X \to \mathbb{R}$ on a convex set $X \subseteq \mathbb{R}^n$ is convex if $\forall x, y \in X$ and $\forall \lambda \in [0, 1]$ one has

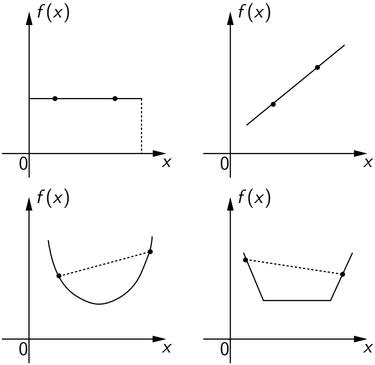
$$f(\lambda x + (1 - \lambda)y) \leq \lambda f(x) + (1 - \lambda)f(y)$$

# Convex functions

## Definition

A function $f : X \to \mathbb{R}$ on a convex set $X \subseteq \mathbb{R}^n$ is convex if $\forall x, y \in X$ and $\forall \lambda \in [0, 1]$ one has

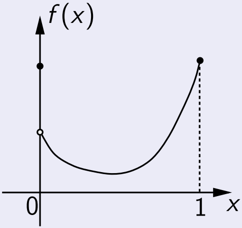
$$f(\lambda x + (1 - \lambda)y) \leq \lambda f(x) + (1 - \lambda)f(y)$$



# Convexity and smoothness

## Convexity and continuity

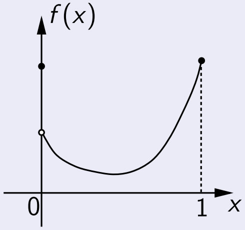
A convex function $f : X \to \mathbb{R}, X \subseteq \mathbb{R}^n$ is continuous in the interior of $X$.



# Convexity and smoothness

## Convexity and continuity

A convex function $f : X \to \mathbb{R}, X \subseteq \mathbb{R}^n$ is continuous in the interior of $X$.



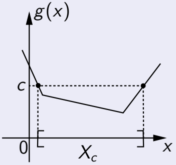

## Theorem - convexity test for smooth functions

Let $X \subseteq \mathbb{R}^n$ be open and convex and let $f : X \to \mathbb{R}$ be a $\mathcal{C}^2$ function. Then, $f$ is convex only if the Hessian matrix $H(x)$ is positive semidefinite $\forall x \in X$. In particular, if $X \subseteq \mathbb{R}$ and $f \in \mathcal{C}^2$, then $f$ is convex only if $\frac{d^2 f(x)}{dx^2} \geq 0, \forall x \in X$.

# Convex functions and sets

## Theorem

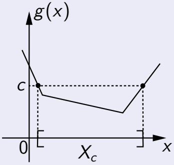
Let $g : \mathbb{R}^n \rightarrow \mathbb{R}$ be a convex function and take $c \in \mathbb{R}$. Then, the level set $X_c = \{x \in \mathbb{R}^n : g(x) \leq c\}$ is convex.



# Convex functions and sets

**Theorem**

Let $g : \mathbb{R}^n \to \mathbb{R}$ be a convex function and take $c \in \mathbb{R}$. Then, the level set $X_c = \{x \in \mathbb{R}^n : g(x) \leq c\}$ is convex.




**Proof.** Pick $x, y \in X_c$ and $\lambda \in [0, 1]$ and consider $z = \lambda x + (1 - \lambda)y$: we have to show that $z \in X_c$.
From the convexity of $g$ one has that $g(z) \leq \lambda g(x) + (1 - \lambda)g(y)$. Since $x, y \in X_c$ one has

$$g(z) \leq \lambda g(x) + (1 - \lambda)g(y) \leq \lambda c + (1 - \lambda)c = c$$

that implies $z \in X_c$.

# Convex functions and sets

## Key corollary

Consider the optimization problem

$$\min_{\substack{g_i(x)\leq 0 \\ i=1,2,\dots,m}} \quad f(x)$$

If functions $g_i(x)$, $i = 1, 2, \dots, m$ are convex, then the feasibile region is convex.

# Convex functions and sets

## Key corollary

Consider the optimization problem

$$\min_{\substack{g_i(x)\leq 0 \\ i=1,2,\ldots,m}} \quad f(x)$$

If functions $g_i(x)$, $i = 1, 2, \ldots, m$ are convex, then the feasibile region is convex.

**Proof.** The proof follows from the previous theorem and the fact that convexity is preserved by intersection.

In convex programming, the feasible region is convex

# Fundamental theorem of convex programming

**Theorem**

If $\tilde{x} \in X$ is a *local optimal solution* for the convex programming problem

$$\{\min f(x) : g_i(x) \leq 0, i = 1, 2, \ldots, m\}$$

then $\tilde{x}$ is an *optimal solution*.

**Remarks**

Often one tries to transform a programming problem into a convex programming problem by performing suitable changes of variables

# Fundamental theorem of convex programming

## Remarks

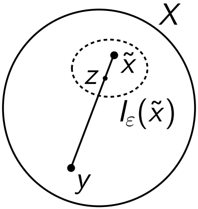
The optimization problem $\{\max f(x) : g_i(x) \leq 0, i = 1, 2, \ldots, m\}$ is not a convex programming problem even if $f$ and $g_i$, $i = 1, 2, \ldots, m$ are convex. Indeed, it is equivalent to $\{-\min -f(x) : g_i(x) \leq 0, i = 1, 2, \ldots, m\}$ where the function $-f(x)$ is concave.

Notable exception: $f(x)$ linear.

# Proof of the theorem

The goal is to show $f(\tilde{x}) \leq f(y)\ \forall y \in X$.
Fix $y \in X$, $y \neq \tilde{x}$ and let $I_\epsilon(\tilde{x})$ be a neighborhood of $\tilde{x}$ such that $z \in I_\epsilon(\tilde{x}) \Rightarrow f(\tilde{x}) \leq f(z)$. Pick $z \in X$ such that $z \in \overline{\tilde{x}y}$, $z \in I_\epsilon(\tilde{x})$ and $z \neq \tilde{x}$.
Such a $z$ exists because

$$z = \lambda\tilde{x} + (1-\lambda)y$$




and

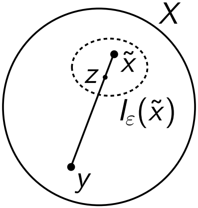
- choosing $\lambda$ sufficiently close to 1 guarantees $z \in I_\epsilon(\tilde{x})$
- choosing $\lambda \neq 1$ guarantees $z \neq \tilde{x}$

# Proof of the theorem

Then,

$$f(\tilde{x}) \underbrace{\leq}_{\text{local optimizer}} f(z) = f(\lambda\tilde{x} + (1-\lambda)y) \leq$$

$$\underbrace{\leq}_{f \text{ convex}} \lambda f(\tilde{x}) + (1-\lambda)f(y)$$




From the last inequality one has

$$(1-\lambda)f(\tilde{x}) \leq (1-\lambda)f(y) \underbrace{\Rightarrow}_{\lambda \neq 1} f(\tilde{x}) \leq f(y)$$