Optimization
– Introduction –

Nelly Pustelnik

CNRS, Laboratoire de Physique de l'ENS de Lyon, France

(several slides in this part traced back Tutorial ICASSP 2014 written in collaboration with **Jean-Christophe Pesquet** from Centre de Vision Numérique, CentraleSupelec, University Paris-Saclay, Inria, France.)

Optimization ?

We think that convex optimization is an important enough topic that everyone who uses computational mathematics should know at least a little bit about it. In our opinion, convex optimization is a natural next topic after advanced linear algebra and linear programming.

(Stephen Boyd and Lieven Vandenberghe)

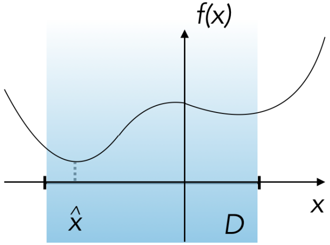
Optimization ?

Whatever people do, at some point they get a craving to organize things in a best possible way. This intention, converted in a mathematical form, turns out to be an optimization problem of certain type.

(Yurii Nesterov)

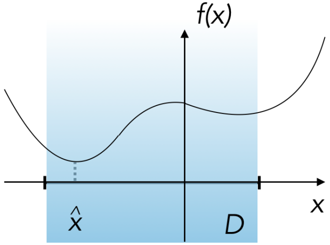
Optimization: Minimization problems

- **Minimization problems** involve :
  - a cost function $f : \mathbb{R}^N \to \mathbb{R}$;
  - a subset D of $\mathbb{R}^N$.



Optimization: Minimization problems

- **Minimization problems** involve :
  - a cost function $f : \mathbb{R}^N \to \mathbb{R}$;
  - a subset D of $\mathbb{R}^N$.

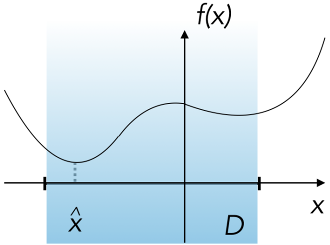


- Goal: We want to

$$
\begin{aligned}
& \text{Find } \widehat{x} \in D \quad \text{such that} \quad (\forall x \in D)\ f(\widehat{x}) \leq f(x) \\
\Leftrightarrow\ & \text{Find } \widehat{x} \in D \quad \text{such that} \quad f(\widehat{x}) = \inf_{x \in D} f(x) \\
\Leftrightarrow\ & \text{Find } \widehat{x} \in \operatorname*{Argmin}_{x \in D} \ f(x).
\end{aligned}
$$

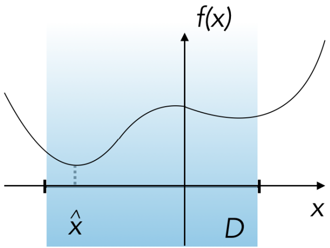
Optimization: Maximization problems

- **Maximization problems** involving :
  - a reward function $f : \mathbb{R}^N \to \mathbb{R}$;
  - a subset D of $\mathbb{R}^N$.



Optimization: Maximization problems

- **Maximization problems** involving :
  - a reward function $f\colon \mathbb{R}^N \to \mathbb{R}$;
  - a subset D of $\mathbb{R}^N$.



- Goal: We want to

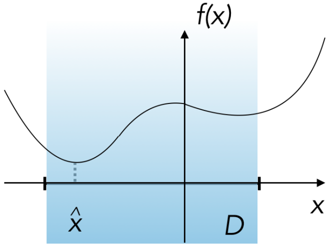
$$
\begin{aligned}
& \text{Find } \widehat{x} \in D \quad \text{such that} \quad (\forall x \in D)\ f(\widehat{x}) \geq f(x) \\
\Leftrightarrow \quad & \text{Find } \widehat{x} \in D \quad \text{such that} \quad (\forall x \in D)\ -f(\widehat{x}) \leq -f(x) \\
\Leftrightarrow \quad & \text{Find } \widehat{x} \in \underset{x \in D}{\operatorname{Argmin}} \ (-f(x)).
\end{aligned}
$$

Optimization: Maximization problems

- **Maximization problems** involving :
  - a reward function $f : \mathbb{R}^N \to \mathbb{R}$;
  - a subset D of $\mathbb{R}^N$.

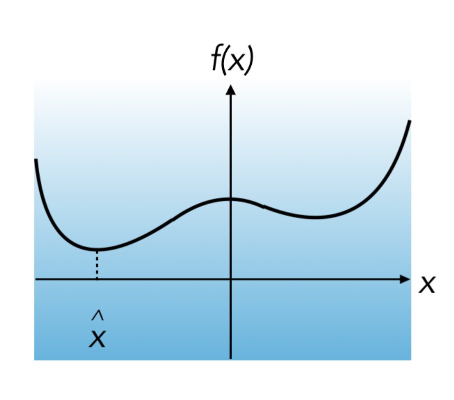
Without loss of generality, we can focus on minimization problems

Find $\widehat{x} \in \underset{x \in D}{\operatorname{Argmin}} \; f(x)$.



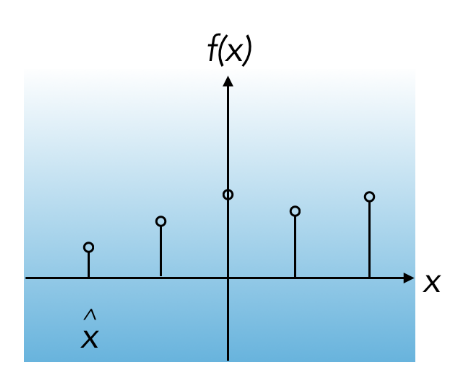
Various types of minimization problems

- Find $\boxed{\widehat{x} \in \underset{x \in D}{\mathrm{Argmin}} \; f(x)}$ with $D = \mathbb{R}^N$: **unconstrained problem**



Various types of minimization problems

- Find $\boxed{\widehat{x} \in \underset{x \in D}{\mathrm{Argmin}} \; f(x)}$ with *D* **countable**: **discrete optimization**



Various types of minimization problems

- Find $\boxed{\widehat{x} \in \underset{x \in D}{\mathrm{Argmin}} \; f(x)}$ with D **being equality or inequality constraints**

  - Example: Optimization problem with P equality constraints

$$D = \left\{ x \in \mathbb{R}^N \mid (\forall i \in \{1, \ldots, P\}) \quad g_i(x) = 0 \right\}$$

  where $g_i \colon \mathbb{R}^N \to \mathbb{R}$.

Various types of minimization problems

- Find $\boxed{\widehat{x} \in \underset{x \in D}{\mathrm{Argmin}} \; f(x)}$ with D **being equality or inequality constraints**

  - Example: Optimization problem with P equality constraints

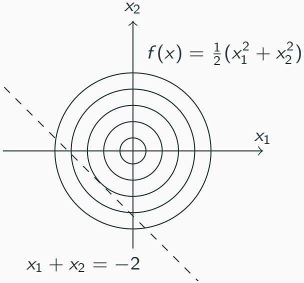
$$D = \left\{ x \in \mathbb{R}^N \mid (\forall i \in \{1, \dots, P\}) \quad g_i(x) = 0 \right\}$$

where $g_i \colon \mathbb{R}^N \to \mathbb{R}$.

  - Particular case: linear (or affine) constraints

$$\begin{aligned} g_i(x) &= \langle a_i \mid x \rangle + b_i \\ &= \sum_{n=1}^{N} a_{i,n} x_n + b_i \end{aligned}$$

where $a_i \in \mathbb{R}^N$ and $\mathrm{b}_i \in \mathbb{R}$.



Various types of minimization problems

- Find $\boxed{\widehat{x} \in \underset{x \in D}{\mathrm{Argmin}} \; f(x)}$ with D **being equality or inequality constraints**

  - Example: Optimization problem with P inequality constraints

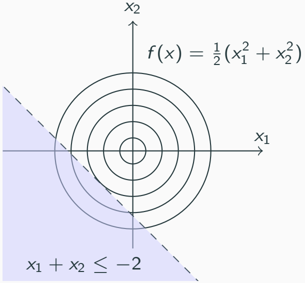
$$D = \big\{x \in \mathbb{R}^N \mid (\forall i \in \{1, \dots, P\}) \quad g_i(x) \leq 0\big\}$$

where $g_i \colon \mathbb{R}^N \to \mathbb{R}$.

  - Particular case: linear (or affine) constraints

$$\begin{aligned} g_i(x) &= \langle a_i \mid x \rangle + b_i \\ &= \sum_{n=1}^{N} a_{i,n} x_n + b_i \end{aligned}$$

where $a_i \in \mathbb{R}^N$ and $\mathrm{b}_i \in \mathbb{R}$.

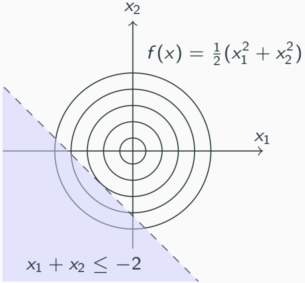


Various types of minimization problems

- Find $\boxed{\widehat{x} \in \underset{x \in D}{\operatorname{Argmin}} \; f(x)}$ with D **being equality or inequality constraints**
  - Example: Optimization problem with P equality constraints

$$D = \left\{ x \in \mathbb{R}^N \mid (\forall i \in \{1, \ldots, P\}) \quad g_i(x) \leq 0 \right\}$$

where $g_i : \mathbb{R}^N \to \mathbb{R}$.



→ **Continuous optimization**

Constrained and unconstrained minimization problems

- Reformulation using **indicator function**

$$\text{Find } \widehat{x} \in \underset{x \in D}{\mathrm{Argmin}}\ f(x) \quad \Leftrightarrow \quad \text{Find } \widehat{x} \in \underset{x \in \mathbb{R}^N}{\mathrm{Argmin}}\ f(x) + \iota_D(x)$$

where

$$(\forall x \in \mathbb{R}^N) \quad \iota_D(x) = \begin{cases} 0 & \text{if } x \in D \\ +\infty & \text{otherwise.} \end{cases}$$

Constrained and unconstrained minimization problems

- Reformulation using **indicator function**

$$\text{Find } \widehat{x} \in \underset{x \in D}{\operatorname{Argmin}} \; f(x) \quad \Leftrightarrow \quad \text{Find } \widehat{x} \in \underset{x \in \mathbb{R}^N}{\operatorname{Argmin}} \; f(x) + \iota_D(x)$$

where

$$(\forall x \in \mathbb{R}^N) \quad \iota_D(x) = \begin{cases} 0 & \text{if } x \in D \\ +\infty & \text{otherwise.} \end{cases}$$

- or equivalently

$$\text{Find } \widehat{x} \in \underset{x \in \mathbb{R}^N}{\operatorname{Argmin}} \; \widetilde{f}(x)$$

where

$$(\forall x \in \mathbb{R}^N) \quad \widetilde{f}(x) = \begin{cases} f(x) & \text{if } x \in D \\ +\infty & \text{otherwise.} \end{cases}$$

Constrained and unconstrained minimization problems

- Reformulation using **indicator function**

$$\text{Find } \widehat{x} \in \underset{x \in D}{\operatorname{Argmin}} \; f(x) \quad \Leftrightarrow \quad \text{Find } \widehat{x} \in \underset{x \in \mathbb{R}^N}{\operatorname{Argmin}} \; f(x) + \iota_D(x)$$

where

$$(\forall x \in \mathbb{R}^N) \quad \iota_D(x) = \begin{cases} 0 & \text{if } x \in D \\ +\infty & \text{otherwise.} \end{cases}$$

Allowing non finite valued functions leads to a unifying view of constrained and unconstrained minimization problems.

Main questions to be addressed

1. **Existence/uniqueness** of a solution $\widehat{x}$?

Main questions to be addressed

1. **Existence/uniqueness** of a solution $\widehat{x}$?
2. Characterization of solutions: **necessary/sufficient conditions** for $\widehat{x}$ to be a solution.

Main questions to be addressed

1. **Existence/uniqueness** of a solution $\widehat{x}$?
2. Characterization of solutions: **necessary/sufficient conditions** for $\widehat{x}$ to be a solution.
3. Designing an **algorithm** to approximate a solution in the frequent case when no closed form solution is available, i.e. building a sequence $(x_n)_{n\in\mathbb{N}}$ of $\mathbb{R}^N$ such that

$$\lim_{n\to+\infty} x_n = \widehat{x}.$$

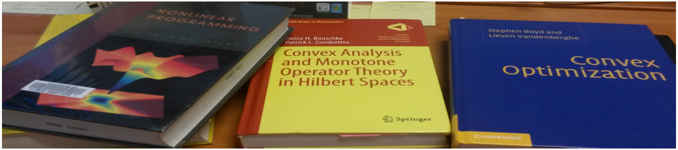
Main questions to be addressed

1. **Existence/uniqueness** of a solution $\widehat{x}$?
2. Characterization of solutions: **necessary/sufficient conditions** for $\widehat{x}$ to be a solution.
3. Designing an **algorithm** to approximate a solution in the frequent case when no closed form solution is available, i.e. building a sequence $(x_n)_{n\in\mathbb{N}}$ of $\mathbb{R}^N$ such that

$$\lim_{n\to+\infty} x_n = \widehat{x}.$$

4. Evaluation of the performance of the optimization algorithm:
   - **Convergence rate** or at least convergence of the iterates
     Example: If there exists $\rho \in]0,1[$ and $n^* \in \mathbb{N}$ such that $(\forall n \geq n^*)$ $\|x_{n+1} - \widehat{x}\| \leq \rho\|x_n - \widehat{x}\|$, then ($Q$-)*linear* convergence rate.
   - **Robustness** to numerical errors
   - Amenability to **parallel/distributed implementations**.

Reference books



- **D. Bertsekas**, Nonlinear programming, Athena Scientic, Belmont, Massachussets, 1995.
- **Y. Nesterov**, Introductory Lectures on Convex Optimization: A Basic Course, Springer, 2004.
- **S. Boyd and L. Vandenberghe**, Convex optimization, Cambridge University Press, 2004.
- **J. Nocedal and S. Wright**, Numerical Optimization, Springer, 2006.
- **H. H. Bauschke and P. L. Combettes**, Convex Analysis and Monotone Operator Theory in Hilbert Spaces, Springer, New York, 2011.
- **A. Beck**, First-Order Methods In Optimization, MOS-SIAM, 2019.