Optimization

Màster de Fonaments de Ciència de Dades

Lecture 1. Optimization

Introduction

- **What is Optimization?** Given a system or process, find the best solution to this process within (or not) constraints.
- **Objective Function:** Indicator of "goodness" of the solution of the optimization problem, e.g., cost, profit, time, etc.
- **Decision Variables:** Variables that influence process behavior and can be adjusted for optimization.
- We are interested in a **systematic approach** to the optimization process, and to make it as efficient as possible.
- **Optimization** is also called: **Mathematical Programming,** or **Operations Research.**

Current applications

- In modern times, (linear and nonlinear) optimization is used in optimal engineering design, finance, statistics and many other fields.
- Think of:
  - designing a car with minimal air resistance,
  - designing a bridge of minimal weight that still meets essential specifications,
  - defining a stock portfolio where the risk is minimal and the expected return high,...
- **Rule of thumb:** If you can make a mathematical model of your decision problem, then you can *try to optimize* it!

First introductory examples

Problem. *Find the solution of*

$$\text{minimize} \quad f(x, y) = (e^x - 1)^2 + (y - 1)^2.$$

This is an example of an unconstrained optimization problem.

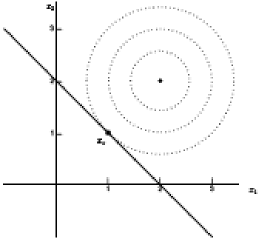
The set where we must look for the solution, the feasible set, is the entire two-dimensional space $\mathbb{R}^2$.

The solution is $(x, y)^T = (0, 1)^T$, since the function value is zero only at this point and is positive elsewhere.

In this problem, the objective function is f and the decision variables are x, y.

First introductory examples

Problem. *Find the point on the line $x + y = 2$ that is closest to the point $(2, 2)^T$.*



The mathematical model can be written as

$$
\begin{array}{ll}
\text{minimize} & f(x, y) = (x-2)^2 + (y-2)^2, \\
\text{subject to} & x + y = 2.
\end{array}
$$

The solution is $(x, y)^T = (1, 1)^T$.

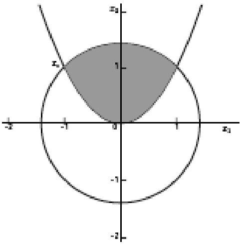
In this example, the objective function is f, the decision variables are x, y, and the (feasible) set is defined by an equality.

First introductory examples

Problem. *Find the point such that:*

$$\begin{array}{ll} \text{minimize} & f(x,y) = x, \\ \text{subject to} & x^2 \leq y, \\ & x^2 + y^2 \leq 2. \end{array}$$

In this example, the (feasible) set where we must look for the solution is given by two constraints defined by inequalities.



The solution (optimal point) is $(x,y)^T = (-1,1)^T$.

Optimization viewpoints

- **Mathematician** - characterization of theoretical properties of optimization, convergence, existence, local convergence rates.
- **Numerical Analyst** - implementation of optimization method for efficient and "practical" use. Concerned with fast computations, numerical stability, performance.
- **User** - applies optimization method to real problems. Concerned with reliability, robustness, efficiency, diagnosis, and recovery from failure.
- Optimization is a **fast moving research field**. Currently, there are over 30 journals devoted to optimization with roughly 200 published papers/month.
- In **this course**, we will see only the most **basic concepts, results, and procedures.**

Some classical optimization problems - I

1. Dido's (or isoperimetric) problem. Among all closed plain curves of a given length, find the one that encloses the largest area.

2. Heron's problem. Given two points A and B on the same side of a line L, find a point D on L such that the sum of the distances form A to D and from D to B is a minimum.

3. Snell's law of refraction. Given two points A and B on either side of a horizontal line L separating two (homogeneous) different media, find a point D on L such that the time it takes for a light ray to traverse the path ADB is a minimum.
   Note: In an inhomogeneous medium, light travels from one point to another along the path requiring the shortest time ($v_i = c/n_i$).

4. Euclid (Elements, 4th cent. B.C.). In a given triangle ABC inscribe a parallelogram $ADEF$ ($EF||AB, DE||AC$) of maximal area.

5. Steiner. In the plane of a triangle, find a point (Fermat point) such that the sum of its distances to the vertices of the triangle is minimal

Some classical optimization problems - II

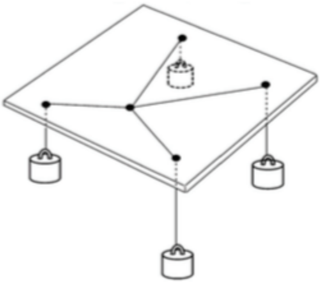
6. Find the maximum of the product of two numbers whose sum is given.
7. Find the maximal area of a right triangle whose small sides have constant sum.
8. In a given circle find a rectangle of maximal area.
9. In a given sphere find a cylinder of maximal volume.
10. Of all rectangular parallelepipeds inscribed in a sphere find the one of maximal volume.
11. Of all rectangular parallelepipeds with square base inscribed in a sphere find the one of maximal volume.
12. The Brachistochrone. Let two points A and B be given in a vertical plane. Find the curve that a point M, moving on a path AMB must follow such that, starting from A with zero velocity, it reaches B in the shortest time under its own gravity.

Some classical optimization problems - III

13. **Exercise 0. (The Fermat point of a set of points)** To be delivered before the end of the course as: `Ex00-YourSurname.pdf`.
Given set of points $y_1,\ldots,y_m$ in the plane, find a point x^* whose sum of weighted distances to the given set of points is minimized.
Mathematically, the problem is

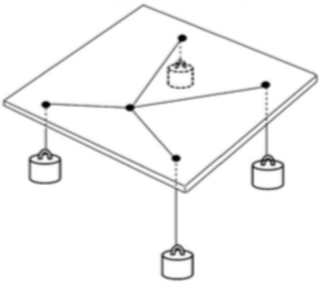
$$\min \sum_{i=1}^{m} w_i \|x^* - y_i\|, \quad \textit{subject to } x^* \in \mathbb{R}^2,$$

where $w_1, \ldots, w_m$ are given positive real numbers.



Exercise 0 cont.

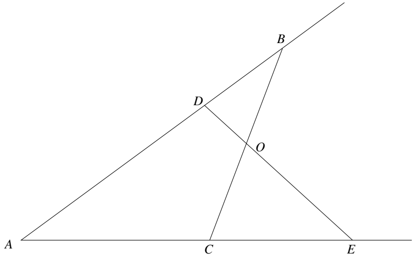
1. *Show that there exists a global minimum for this problem (that it can be realized by means of the mechanical model shown in the figure).*



2. *Is the optimal solution always unique?*
3. *Show that an optimal solution minimizes the potential energy of the mechanical model defined as $\sum_{i=1}^{m} w_i h_i$, where h_i is the height of the ith weight measureed from some reference level.*

Exercise 1. (Smallest area problem) To be delivered before 27-IX-2021 as: `Ex01-YourSurname.pdf`.

14. *Given an angle with vertex A and a point O in its interior. Pass a line BC through the point O that cuts off from the angle a triangle of minimal area*



Hint: proof that for a triangle of minimal area the segments OB and OC should be equal.

The general optimization problem

Definition:
The general nonlinear optimization (NLO) problem can be written as follows:

$$\begin{array}{lll} \min & f(x), & \\ \text{subject to} & g_i(x) = 0, & i \in I = \{1, ..., m\}, \\ & h_j(x) \leq 0, & j \in J = \{1, ..., p\}, \\ & x \in \mathcal{C}, & \end{array}$$

where $x \in \mathbb{R}^n$, $\mathcal{C} \subset \mathbb{R}^n$ is a certain set, and $f, g_1, ..., g_m, h_1, ..., h_p$ are real-valued functions defined on $\mathcal{C}$.

Terminology:

- The function f is called the objective function of the NLO.
- The set $\mathcal{F}$ defined by:

  $$\mathcal{F} = \{x \in \mathcal{C} \;:\; g_i(x) = 0, i = 1, ..., m,\; h_j(x) \leq 0, j = 1, ..., p\},$$

  is called the feasible set (or feasible region).
- If $\mathcal{F} = \emptyset$ then we say that the optimization problem is infeasible.
- If the infimum of f over $\mathcal{F}$ is attained at $x^* \in \mathcal{F}$, then we call x^* an optimal solution of the NLO, and $f(x^*)$ the the optimal (objective) value of the NLO.

Classification of optimization problems

- Unconstrained Optimization: The index sets I and J are empty:

$$g_1 = \ldots = g_m = h_1 = \ldots = h_p = 0,$$

  and $\mathcal{C} = \mathbb{R}^n$.

- Linear Optimization (LO) (Linear programming): The functions $f, g_1, \ldots, g_m, h_1, \ldots, h_p$ are linear (affine: $F(\boldsymbol{x}) = A\boldsymbol{x} + \boldsymbol{b}$) and the set $\mathcal{C}$ either equals to $\mathbb{R}^n$, the positive (negative) orthant $\mathbb{R}^n_+$, or is polyhedral.

- Quadratic Optimization (QO): The objective function f is quadratic:

$$f(\boldsymbol{x}) = \boldsymbol{x}^T Q\boldsymbol{x} + \boldsymbol{c}^T\boldsymbol{x} + d,$$

  all the constraint functions $g_1, \ldots, g_m, h_1, \ldots, h_p$ are linear and the set $\mathcal{C}$ is $\mathbb{R}^n$ or the positive (negative) orthant $\mathbb{R}^n_+$, Q is a $n \times n$ real matrix ($Q \in \mathbb{R}^{n\times n}$), $\boldsymbol{c} \in \mathbb{R}^n$, and $d \in \mathbb{R}$.

- Quadratically Constrained Quadratic Optimization: Same as QO, except that the constraint functions are quadratic.

- Convex Quadratic Optimization (CQO).

- Convex Quadratically Constrained Quadratic Optimization:

- ...

A well known application of Quadratic Optimization: Regression problems

- If a system
$$A\boldsymbol{x} = \boldsymbol{b}, \quad A \in \mathbb{R}^{m\times n}, \quad \boldsymbol{x} \in \mathbb{R}^n, \quad \boldsymbol{b} \in \mathbb{R}^m,$$
has more equations than unknowns ($m > n$), then, in general, it has no solution, but we can compute the least squares solution
$$\boldsymbol{x}^* = \min_{\boldsymbol{x}\in\mathbb{R}^n} \|A\boldsymbol{x} - \boldsymbol{b}\|,$$
for the Euclidean norm $\|\boldsymbol{x}\| = \sqrt{\sum_{i=1}^n x_i^2} = \sqrt{\boldsymbol{x}^T\boldsymbol{x}} \geq 0$.

- Note that
$$\begin{aligned}\|A\boldsymbol{x} - \boldsymbol{b}\|^2 &= (A\boldsymbol{x} - \boldsymbol{b})^T(A\boldsymbol{x} - \boldsymbol{b}) \\ &= \boldsymbol{x}^T A^T A\boldsymbol{x} - 2\boldsymbol{b}^T A\boldsymbol{x} + \|\boldsymbol{b}\|^2.\end{aligned}$$

- Note also that if $A \in \mathbb{R}^{m\times n}$, then $A^TA \in \mathbb{R}^{n\times n}$, $\boldsymbol{b}^TA \in \mathbb{R}^n$, and introducing $\boldsymbol{z} = A\boldsymbol{x}$:
$$\boldsymbol{x}^T A^T A\boldsymbol{x} = \boldsymbol{z}^T\boldsymbol{z} = \|\boldsymbol{z}\|^2 \geq 0, \quad \forall \boldsymbol{x} \in \mathbb{R}^n.$$
According to this last inequality, A^TA will be positive definite $\Leftrightarrow$ for all $\boldsymbol{x} \neq 0$ then $A\boldsymbol{x} \neq 0$, which is equivalent to $\mathrm{rank}(A) = n$.

Example of regression problem: Concrete mixing

Mix concrete using n different gravel sizes $s_1, s_2, \ldots, s_n$.

- The ideal mixture is given by $\boldsymbol{c} = (c_1, c_2, \ldots, c_n)$, where c_i $(0 \leq c_i \leq 1)$ is the fraction of size s_i in the mix, and $\sum_{i=1}^{n} c_i = 1$.
- Gravel mixtures come from m different mines: $M_1,\ldots,M_m$.
- The gravel composition at each mine M_j is given by $C_j = (c_1^j, \ldots, c_n^j)$ where $0 \leq c_i^j \leq 1$ for all $i = 1, \ldots, n$ and $\sum_{i=1}^{n} c_i^j = 1$

	s_1	$\ldots$	s_n	
M_1	c_1^1	$\ldots$	c_n^1	$x_1 =$ fraction from M_1 in the mix
$\vdots$	$\vdots$		$\vdots$	$\vdots$
M_m	c_1^m	$\ldots$	c_n^m	$x_m =$ fraction from M_m in the mix

- In the mix, the amount of grave with size k should be close to c_k.

Concrete mixing: mathematical formulation

Exercise 2. To be delivered before 27-IX-2021 as: `Ex02-YourSurname.pdf`.

Find the best possible approximation $\boldsymbol{x} = (x_1, ..., x_m)$ *of the ideal mixture,* $\boldsymbol{c} = (c_1, ..., c_n)$*, by using the material from the* m *mines.*

Show that the optimal mixture will be the point $\boldsymbol{x}$ such that:

$$\min \ (C\boldsymbol{x} - \boldsymbol{c})^T (C\boldsymbol{x} - \boldsymbol{c}),$$
$$s.t. \quad \sum_{j=1}^{m} x_j = 1, \quad \text{and} \quad x_j \geq 0,$$

where the matrix $C = (C_1, ..., C_m)$ has C_j as columns, and $\boldsymbol{c} = (c_1 \cdots c_n)^T$.

The infinite-dimensional optimization problem

There is a more general nonlinear optimization **infinite-dimensional** problem that can be written as follows:

Find the state function $x(t)$ and the control function $u(t)$ such that

$$
\begin{array}{lll}
\textit{minimize} & \mathcal{J} = \Phi(t_0, x_0, t_f, x_f) + \int_{t_0}^{t_f} F(t, x(t), u(t))dt, & \\
\textit{subject to} & t \in [t_0, t_f], & \\
 & \dot{x}(t) = f(t, x(t), u(t)), & \textit{(dynamic constraints)}, \\
 & b_L \leq b(t_0, x_0, t_f, x_f) \leq b_U, & \textit{(boundary conditions)}, \\
 & g_L \leq g(t, x(t), u(t)) \leq g_U, & \textit{(path constraints)}, \\
 & u_L \leq u(t) \leq u_U, & \textit{(control constraints)}.
\end{array}
$$

In the multi-objective optimization problem, the cost function $\mathcal{J}$ may involve several independent quantities, this is:

$$\vec{\mathcal{J}} = (\mathcal{J}_1 \cdots \mathcal{J}_M)^T.$$

The infinite-dimensional optimization problem

There are multiple solution approaches for the infinite-dimensional optimization problem. They are commonly divided into:

- Indirect methods. The initial problem is transformed into a Hamiltonian boundary-value problem that must be solved. These methods equire the derivation of the necessary conditions of optimality using calculus of variations.
- Direct methods. The original problem is first discretized and then re-written as a finite-dimensional nonlinear optimization problem (NLO).