Optimization

Màster de Fonaments de Ciència de Dades

Lecture 1. Introduction to optimization

Introduction

- **What is Optimization?** Given a system or process, find the best solution possibly subject to 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.

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,
  - ChatGPT,...
- **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 $\mathbb{R}^2$.

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

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.

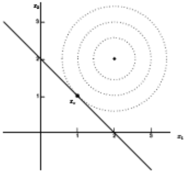
The set where we must look for the solution, the feasible set, is the entire two-dimensional space $\mathbb{R}^2$.

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

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

First introductory examples

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



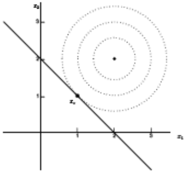
The mathematical model can be written as

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

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



The mathematical model can be written as

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

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

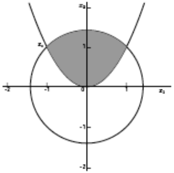
The solution is $(x^*, y^*)^T = (1, 1)^T$.

First introductory examples

Problem. *Find the point such that:*

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

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

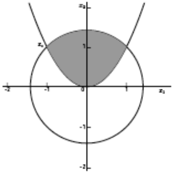


First introductory examples

Problem. *Find the point such that:*

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

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



The solution 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 and methods.**

Some classical optimization problems - I

1. Dido's (or isoperimetric) problem. Among all closed planar 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 minimal.
3. Snell's law of refraction. Given two points A and B on different sides 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 minimal.
   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.

Some classical optimization problems - II

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
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. The Brachistochrone. Let two points A and B be given in a vertical plane. Find the curve connecting A to B such that a particle moving along this curve starting from A with zero velocity reaches B in the shortest time under its own gravity.

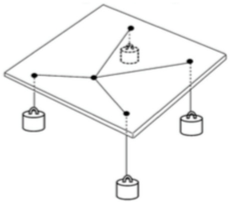
Some classical optimization problems - III

Exercise 0. (The Fermat point of a set of points) (See Campus Virtual for the instructions)
Given set of points y_1,...,y_m in the plane, find a point x^* whose sum of weighted distances to the given set of points is minimized.

The mathematical model of this problem is:

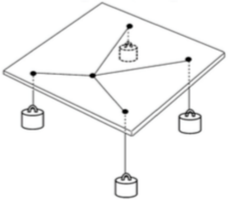
$$\min \sum_{i=1}^{m} w_i \|x^* - y_i\|, \quad \textit{subject to } x^* \in \mathbb{R}^2,$$

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).*

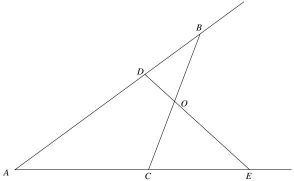


2. *Is the optimal solution always unique?*
3. *Show that an optimal solution minimizes the potential energy of the mechanical model defined as $\sum_{i=1}^{m} w_i h_i$, where h_i is the height of the ith weight measured from some reference level.*

Some classical optimization problems - IV

Exercise 1. (Smallest area problem)

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

General setup

Definition:

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

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

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:

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

  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:**

- ...

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 use the calculus of variations (Pontryagin's Maximum principle).
- Direct methods. The original problem is first discretized and then rewritten as a finite-dimensional nonlinear optimization problem (NLO).