TMA947 / MMG621 — Nonlinear optimization

Lecture 1 — Introduction to optimization

Emil Gustavsson, Zuzana Nedělková

October 20, 2017

What is optimization?

Optimization is a mathematical discipline which is concerned with finding the minima or maxima of functions, possibly subject to constraints.

Basic notation

- Vectors are written with bold face, i.e., $\boldsymbol{x} \in \mathbb{R}^n$.
- Elements in a vector are written as $x_j,\ j = 1, \ldots, n$.
- All vectors are column vectors.
- The inner product of $\boldsymbol{a}$ and $\boldsymbol{b}$ is written as $\boldsymbol{a}^{\mathrm{T}}\boldsymbol{b} = \boldsymbol{b}^{\mathrm{T}}\boldsymbol{a} = \sum_{j=1}^{n} a_j b_j$.
- The norm $||\cdot||$ denotes the Euclidean norm, i.e., $||\boldsymbol{x}|| = \sqrt{\boldsymbol{x}^{\mathrm{T}}\boldsymbol{x}} = \sqrt{\sum_{j=1}^{n} x_j^2}$.
- We utilize vector inequalities, $\boldsymbol{a} \leq \boldsymbol{b}$, meaning that $a_j \leq b_j$, $j = 1, \ldots, n$.

Optimization problem formulation

In order to introduce a general optimization problem, we need to define the following:

$\boldsymbol{x} \in \mathbb{R}^n$	: vector of decision variables $x_j,\ j = 1, \ldots, n$,
$f : \mathbb{R}^n \to \mathbb{R} \cup \pm\infty$	: objective function,
$X \subseteq \mathbb{R}^n$	: ground set,
$g_i : \mathbb{R}^n \to \mathbb{R}$	: constraint function defining restriction on $\boldsymbol{x}$,
$g_i \geq 0,\ i \in \mathcal{I}$	: inequality constraints,
$g_i = 0,\ i \in \mathcal{E}$	: equality constraints.

A general **optimization problem** is to

$$
\begin{aligned}
\underset{\boldsymbol{x}}{\text{minimize}} \quad & f(\boldsymbol{x}), && \text{(1a)}\\
\text{subject to} \quad & g_i(\boldsymbol{x}) \leq 0, \quad i \in \mathcal{I}, && \text{(1b)}\\
& g_i(\boldsymbol{x}) = 0, \quad i \in \mathcal{E}, && \text{(1c)}\\
& \boldsymbol{x} \in X. && \text{(1d)}
\end{aligned}
$$

(If we consider a maximization problem, we change the sign of f to get a minimization problem.)

Classification of optimization problems

Linear Programming (LP):
- Linear objective function $f(\boldsymbol{x}) = \boldsymbol{c}^{\mathsf{T}}\boldsymbol{x} = \sum_{j=1}^{n} c_j x_j$,
- Affine constraint functions $g_i(\boldsymbol{x}) = \boldsymbol{a}_i{}^{\mathsf{T}}\boldsymbol{x} - b_i,\ i \in \mathcal{I} \cup \mathcal{E}$
- Ground set X defined by affine equalities/inequalities.

Nonlinear programming (NLP):
- Some functions $f, g_i,\ i \in \mathcal{I} \cup \mathcal{E}$ are nonlinear.

Unconstrained optimization:
- $\mathcal{I} \cup \mathcal{E} = \emptyset$,
- $X = \mathbb{R}^n$.

Constrained optimization:
- $\mathcal{I} \cup \mathcal{E} \neq \emptyset$, and/or
- $X \subset \mathbb{R}^n$.

Integer programming (IP):
- $X \subseteq \mathbb{Z}^n$ or $X \subseteq \{0,1\}^n$.

Convex programming (CP):
- $f, g_i,\ i \in \mathcal{I}$ are convex functions,
- $g_i,\ i \in \mathcal{E}$ are affine,
- X is closed and convex.

Conventions

Let $S = \{\boldsymbol{x} \in \mathbb{R}^n \mid g_i(\boldsymbol{x}) \leq 0,\ i \in \mathcal{I},\ g_i(\boldsymbol{x}) = 0,\ i \in \mathcal{E},\ \boldsymbol{x} \in X\}$ denote a feasible set.

What do we mean by solving the problem to $\underset{\boldsymbol{x} \in S}{\text{minimize}}\ f(\boldsymbol{x})$?

Let

$$f^* := \underset{\boldsymbol{x} \in S}{\text{infimum}}\ f(\boldsymbol{x})$$

denote the infimum value of f over the set S. If the value f^* is attained at some point $\boldsymbol{x}^*$ in S, we can write

$$f^* := \underset{\boldsymbol{x} \in S}{\text{minimum}}\ f(\boldsymbol{x}),$$

and have $f(\boldsymbol{x}^*) = f^*$. Another well-defined operator defines the set of minimal solutions to the problem

$$S^* := \arg \underset{\boldsymbol{x} \in S}{\text{minimum}}\ f(\boldsymbol{x}),$$

where $S^* \subseteq S$ is nonempty if and only if the infimum value f^* is attained at some point $\boldsymbol{x}^*$ in S.

Now we can define what we mean by the problem to $\underset{\boldsymbol{x} \in S}{\text{minimize}}\ f(\boldsymbol{x})$.

"to $\underset{\boldsymbol{x} \in S}{\text{minimize}}\ f(\boldsymbol{x})$" means "find f^* and an $\boldsymbol{x}^* \in S^*$"

If we have an optimization problem

$$P: \quad \underset{\boldsymbol{x} \in S}{\text{minimize}}\ f(\boldsymbol{x})$$

- A point $\boldsymbol{x}$ is **feasible** in problem P if $\boldsymbol{x} \in S$. The point is **infeasible** in problem P if $\boldsymbol{x} \notin S$.
- The problem P is feasible if there exist a $\boldsymbol{x} \in S$ and the problem P is infeasible if $S = \emptyset$.
- A point $\boldsymbol{x}^*$ is an **optimal solution** to P if $\boldsymbol{x}^* \in \arg \underset{\boldsymbol{x} \in S}{\text{minimum}}\ f(\boldsymbol{x})$.
- f^* is an **optimal value** to P if $f^* = \underset{\boldsymbol{x} \in S}{\text{minimum}}\ f(\boldsymbol{x})$.

Examples

I. Consider the problem to

$$\begin{aligned} \text{minimize} \quad & (x+1)^2, \\ \text{subject to} \quad & x \in \mathbb{R}, \end{aligned}$$

Easy problem, $(x+1)^2$ is convex, no constraints. Just solve $f'(x) = 0$, and get the optimal solution $x^* = -1$ and the optimal value $f^* = 0$.

(Convex, quadratic, unconstrained optimization problem)

II. A more complicated problem is to

$$\begin{aligned} \text{minimize} \quad & (x+1)^2, \\ \text{subject to} \quad & x \geq 0. \end{aligned}$$

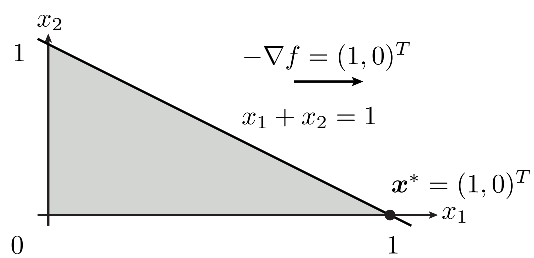
Now the "$f'(x) = 0$" trick does not work and we need to consider the boundary. We get the optimal solution $x^* = 0$ and the optimal value $f^* = 1$.

(Convex, quadratic, constrained optimization problem)

III. Consider the problem to

$$\begin{aligned} \text{minimize} \quad & -x_1, \\ \text{subject to} \quad & x_1 + x_2 \leq 1, \\ & x_1, x_2 \geq 0. \end{aligned}$$

We solve this graphically. So optimal solution is $\boldsymbol{x}^* = (1, 0)^T$ and the optimal value if $f^* = -1$.



The diet problem

As a first example of an real optimization problem, we consider the **diet problem** (first formulated by George Stigler).

> For a moderately active person, how much of each of a number of foods should be eaten on a daily basis so that the person's intake of nutrients will be at least equal to the recommended dietary allowances (RDAs), with the cost of the diet being minimal?

Good example to show

- how to model a real optimization problem,
- why a realistic model sometimes can be difficult to achieve.

We consider the case when the only allowed foods can be found at McDonalds.

> For a moderately active person, how much of each of a number of McDonald foods (see Table 1) should be eaten on a daily basis so that the person's intake of nutrients will be at least equal to the recommended dietary allowances (RDAs), with the cost of the diet being minimal?

Food	Calories	Carb	Protein	Vit A	Vit C	Calc	Iron	Cost
Big Mac	550 kcal	46g	25g	6%	2%	25%	25%	30kr
Cheeseburger	300 kcal	33g	15g	6%	2%	20%	15%	10kr
McChicken	360 kcal	40g	14g	0%	2%	10%	15%	35kr
McNuggets	280 kcal	18g	13g	0%	2%	2%	4%	40kr
Caesar Sallad	350 kcal	24g	23g	160%	35%	20%	10%	50kr
French Fries	380 kcal	48g	4g	0%	15%	2%	6%	20kr
Apple Pie	250 kcal	32g	2g	4%	25%	2%	6%	10kr
Coca-Cola	210 kcal	58g	0g	0%	0%	0%	0%	15kr
Milk	100 kcal	12g	8g	10%	4%	30%	8%	15kr
Orange Juice	150 kcal	30g	2g	0%	140%	2%	0%	15kr
RDA	2000 kcal	350g	55g	100%	100%	100%	100%	

Table 1: Given data for the diet problem

We define the **sets**

$$\begin{aligned} \text{Foods} &:= \{\text{Big Mac, Cheeseburger, McChicken, McNuggets, Caesar Sallad} \\ &\qquad \text{French Fried, Apple Pie, Coca-Cola, Milk, Orange Juice}\}, \\ \text{Nutrients} &:= \{\text{Calories, Carb, Protein, Vit A, Vit C, Calc, Iron.}\} \end{aligned}$$

Then we define the **parameters**

$$\begin{aligned} a_{ij} &= \text{Amount of nutrient } i \text{ in food } j,\ i \in \text{Nutrients},\ j \in \text{Foods}, \\ b_i &= \text{Recommended daily amount (RDA) of nutrient } i,\ i \in \text{Nutrients}, \\ c_j &= \text{Cost for food } j,\ j \in \text{Foods}, \end{aligned}$$

and the **decision variables**

$$x_j = \text{Amount of food } j \text{ we should eat each day},\ j \in \text{Foods}.$$

The **model** of the diet optimization problem is then to

$$\begin{aligned} \text{minimize} \quad & \sum_{j \in \text{Foods}} c_j x_j, && \text{(2a)} \\ \text{subject to} \quad & \sum_{j \in \text{Foods}} a_{ij} x_j \geq b_i, \quad i \in \text{Nutrients}, && \text{(2b)} \\ & x_j \geq 0, \quad j \in \text{Foods}. && \text{(2c)} \end{aligned}$$

(2a) We minimize the total cost, such that

(2b) we get enough of each nutrient, and such that

(2c) we don't sell anything to McDonalds.

The optimal solution is then

$$\boldsymbol{x} = \begin{pmatrix} x_{\text{Big Mac}} \\ x_{\text{Cheeseburger}} \\ x_{\text{McChicken}} \\ x_{\text{McNuggets}} \\ x_{\text{Caesar Sallad}} \\ x_{\text{French Fries}} \\ x_{\text{Apple Pie}} \\ x_{\text{Coca Cola}} \\ x_{\text{Milk}} \\ x_{\text{Orange Juice}} \end{pmatrix} = \begin{pmatrix} 0 \\ 7.48 \\ 0 \\ 0 \\ 0.27 \\ 0 \\ 3.03 \\ 0 \\ 0 \\ 0 \end{pmatrix}.$$

Total cost	= 118.47 kr.
Total intake of calories	= 3093.51 kcal.

If we add the constraint that x_j should be integer, the solution is

$$\boldsymbol{x} = \begin{pmatrix} x_{\text{Big Mac}} \\ x_{\text{Cheeseburger}} \\ x_{\text{McChicken}} \\ x_{\text{McNuggets}} \\ x_{\text{Caesar Sallad}} \\ x_{\text{French Fries}} \\ x_{\text{Apple Pie}} \\ x_{\text{Coca Cola}} \\ x_{\text{Milk}} \\ x_{\text{Orange Juice}} \end{pmatrix} = \begin{pmatrix} 0 \\ 7 \\ 0 \\ 0 \\ 1 \\ 0 \\ 3 \\ 0 \\ 0 \\ 0 \end{pmatrix}.$$

Total cost	= 150 kr.
Total intake of calories	= 3200 kcal.

Now consider going on a diet, meaning that we would like to eat as few calories as possible. We reformulate our model to

$$\begin{aligned} \text{minimize} \quad & \sum_{j \in \text{Foods}} a_{\text{Calories},j} x_j, && (3a) \\ \text{subject to} \quad & \sum_{j \in \text{Foods}} a_{ij} x_j \geq b_i, \quad i \in \text{Nutrients} \setminus \{\text{Calories}\}, && (3b) \\ & x_j \geq 0, \quad j \in \text{Foods}. && (3c) \end{aligned}$$

The optimal solution is then

$$\boldsymbol{x} = \begin{pmatrix} x_{\text{Big Mac}} \\ x_{\text{Cheeseburger}} \\ x_{\text{McChicken}} \\ x_{\text{McNuggets}} \\ x_{\text{Caesar Sallad}} \\ x_{\text{French Fries}} \\ x_{\text{Apple Pie}} \\ x_{\text{Coca Cola}} \\ x_{\text{Milk}} \\ x_{\text{Orange Juice}} \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 3.96 \\ 12.41 \\ 0.36 \end{pmatrix}.$$

Total cost	= 251.01 kr.
Total intake of calories	= 2127.47 kcal.

If we add the constraint that x_j should be integer, the solution is

$$\boldsymbol{x} = \begin{pmatrix} x_{\text{Big Mac}} \\ x_{\text{Cheeseburger}} \\ x_{\text{McChicken}} \\ x_{\text{McNuggets}} \\ x_{\text{Caesar Sallad}} \\ x_{\text{French Fries}} \\ x_{\text{Apple Pie}} \\ x_{\text{Coca Cola}} \\ x_{\text{Milk}} \\ x_{\text{Orange Juice}} \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 1 \\ 11 \\ 6 \end{pmatrix}.$$

Total cost	= 270 kr.
Total intake of calories	= 2210 kcal.

The real diet problem

When first studied by the Stigler, the problem concerned the US military and had 77 different foods in the model. He didn't managed to solve the problem to optimality, but almost. The near optimal diet was

- Wheat flour
- Evaporated milk
- Cabbage
- Spinach
- Dried navy beans

at a cost of \$0.1 a day in 1939 US dollars.

Course material

Lecture 1	Define and model optimization problems, classification
Lecture 2	Convexity of sets, functions, optimization problems
Lecture 3	Optimality conditions, introduction
Lecture 4	Unconstrained optimization, methods, classification.
Lecture 5	Optimality conditions, continued
Lecture 6	The Karush-Kuhn-Tucker conditions
Lecture 7	Lagrangian duality
Lecture 8	Linear programming, introduction
Lecture 9	Linear programming, continued
Lecture 10	Duality in linear programming
Lecture 11	Convex optimization
Lecture 12	Integer programming
Lecture 13	Nonlinear optimization methods, convex feasible sets
Lecture 14	Nonlinear optimization methods, general sets
Lecture 15	Overview of the course