Lecture 1

Introduction to optimization

Emil Gustavsson
Department of Mathematical Sciences
Chalmers University of Technology and Göteborg University
November 1, 2012

CHALMERS | GÖTEBORGS UNIVERSITET

What is optimization?

Optimization is the mathematical disciplin which is concerned with finding the maxima and minima of functions, possibly subject to constraints.

Basic notation

Some notations used in this course.

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

Notations

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

$\mathbf{x} \in \mathbb{R}^n$	: vector of decision variables, $x_j,\ j = 1, \ldots, n$,
$f : \mathbb{R}^n \rightarrow \mathbb{R} \cup \pm\infty$	: objective function,
$X \subseteq \mathbb{R}^n$	: ground set,
$g_i : \mathbb{R}^n \rightarrow \mathbb{R}$	: constraint function defining restrictions on $\mathbf{x}$,

Optimization problem

A general **optimization problem** then is to

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

(If we consider a maximization problem, we can change the sign of f. In this course, we only consider minimization problems.)

Examples I

Linear Programming (LP):

- Linear objective function $f(\mathbf{x}) = \mathbf{c}^T\mathbf{x} = \sum_{j=1}^n c_j x_j$,
- Affine constraint functions $g_i(\mathbf{x}) = \mathbf{a}_i^T\mathbf{x} - b_i,\ i \in \mathcal{I} \cup \mathcal{E}$
- 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.

Examples II

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):

- $X \subseteq \mathbb{Z}^n$, (in many cases $X \in \{0, 1\}^n$).

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.

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

What do we mean by solving the problem to

$$\underset{\mathbf{x} \in S}{\text{minimize}}\ f(\mathbf{x})?$$

Since no mathematical operation is involved, we need to define this notion properly.

Conventions, II

Let

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

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

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

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

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

where $S^* \subseteq S$ is nonempty if and only if the infinum value f^* is attained.

Conventions, III

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

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

Conventions, IV

If we have a optimization problem

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

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

Example, 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)

Example, 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 $f^* = 1$.

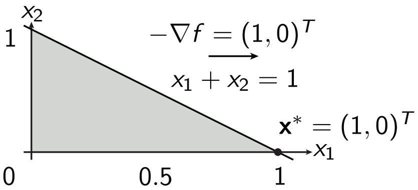
(Convex, quadratic, constrained optimization problem)

Example, 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 $\mathbf{x}^* = (1, 0)^T$ 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.

The McDonalds diet problem

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 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?

The McDonalds diet problem

What we have to our disposal is the following table.

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%	

Formulating the problem I

We define the sets

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

Define the parameters

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

and the decision variable

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

Formulating the problem II

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

(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.

Solution, I

The optimal solution is then

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

$$\begin{aligned} \text{Total cost} &= 118.47 \text{ kr} \\ \text{Total intake of calories} &= 3093.51 \text{ kcal} \end{aligned}$$

Solution, II

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

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

$$\begin{aligned} \text{Total cost} &= 150 \text{ kr} \\ \text{Total intake of calories} &= 3200 \text{ kcal} \end{aligned}$$

Going on a diet

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

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

Diet solution, I

The optimal solution is then

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

$$\begin{aligned} \text{Total cost} &= 251.01 \text{ kr} \\ \text{Total intake of calories} &= 2127.47 \text{ kcal} \end{aligned}$$

Diet solution, II

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

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

$$\begin{aligned} \text{Total cost} &= 270 \text{ kr} \\ \text{Total intake of calories} &= 2210 \text{ kcal} \end{aligned}$$

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 Convex duality

Lecture 8 Linear programming, I

Lecture 9 Linear programming, II

Lecture 10 Convex optimization

Lecture 11 Integer programming

Lecture 12 Nonlinear optimization methods, convex feasible sets

Lecture 13 Nonlinear optimization methods, general sets

Lecture 14 Overview of the course.