MS-C2105 - Introduction to Optimization
Lecture 1

Fabricio Oliveira (with modifications by Harri Hakula)

Systems Analysis Laboratory
Department of Mathematics and Systems Analysis

Aalto University
School of Science

February 24, 2022

Outline of this lecture

Introduction

What is optimisation?

Mathematical programming and optimisation

Modelling real-world problems using optimisation

A first optimisation model

Production planning problems

Classification

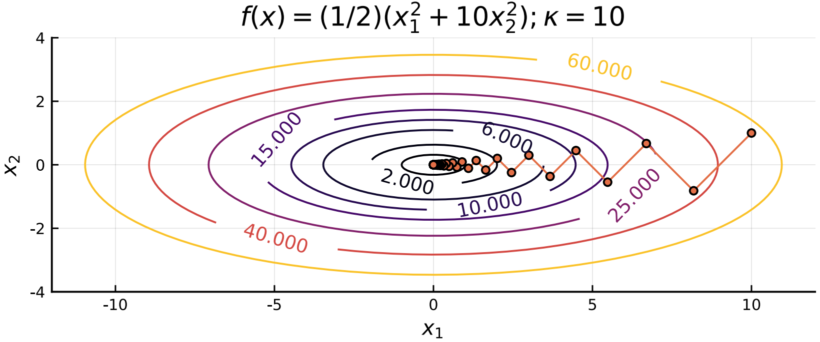
Reading: Taha: Chapter 1; Winston: Chapter 1

What is optimisation?

Discipline of applied mathematics. The idea is to search values for variables in a given domain that maximise/minimise function values.

Can be achieved by

- Analysing properties of functions/extreme points or
- Applying numerical methods.



What is optimisation?

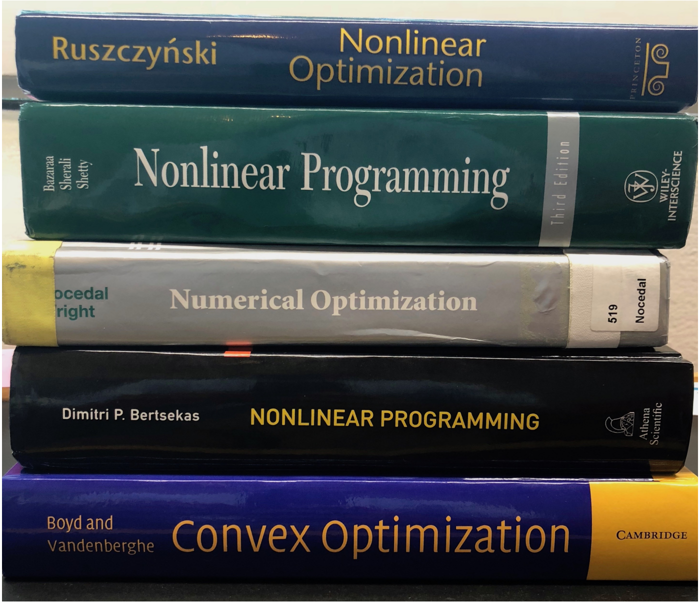
Optimisation has important applications in fields such as

- mathematical programming & operations research (OR) ;
- economics;
- statistics;
- machine learning & artificial intelligence...

Mathematical programming is a modelling paradigm that relies on optimisation to model decision processes:

- **variables** $\rightarrow$ decisions: business decisions, parameter definitions, settings, geometries, ...;
- **domain** $\rightarrow$ constraints: logic, design, engineering, ...;
- **function** $\rightarrow$ objective function: measurement of (decision) quality.

However, there is some confusion between terms optimisation/ programming. In this course, we focus on optimisation models.


Ruszczyński
Nonlinear Optimization
PRINCETON
Bazaraa Sherali Shetty
Nonlinear Programming
Third Edition
WILEY-INTERSCIENCE
Numerical Optimization
519
Nocedal
Dimitri P. Bertsekas
NONLINEAR PROGRAMMING
Athena Scientific
Boyd and Vandenberghe
Convex Optimization
CAMBRIDGE

Types of programming/ optimisation models

The simpler the assumptions are that define a type of problems, the better the methods to solve such problems.

Some useful notation:

- $x \in \mathbb{R}^n$: vector of (decision) variables x_j, $j = 1, \dots, n$;
- $f : \mathbb{R}^n \to \mathbb{R} \cup \{\pm\infty\}$: objective function;
- $X \subseteq \mathbb{R}^n$: ground set (physical constraints);
- $g_i, h_i : \mathbb{R}^n \to \mathbb{R}$: constraint functions;
- $g_i(x) \leq 0$ for $i = 1, \dots, m$: inequality constraints;
- $h_i(x) = 0$ for $i = 1, \dots, l$: equality constraints.

Types of programming

Our goal will be to solve variations of the general problem P:

$$\begin{aligned}(P) \ : \ \text{min. } & f(x) \\ \text{s.t.: } & g_i(x) \leq 0, i = 1, \dots, m \\ & h_i(x) = 0, i = 1, \dots, l \\ & x \in X.\end{aligned}$$

- **Linear programming (LP):** linear $f(x) := c^\top x$ with $c \in \mathbb{R}^n$; constraint functions $g_i(x)$ and $h_i(x)$ are affine ($a_i^\top x - b_i$, with $a_i \in \mathbb{R}^n$, $b \in \mathbb{R}$); $X = \{x \in \mathbb{R}^n : x_j \geq 0, j = 1, \dots, n\}$.
- **Nonlinear programming (NLP):** some (or all) of the functions f, g_i or h_i are nonlinear;
- **(Mixed-)integer programming ((M)IP):** LP where (some of the) variables are binary (or integer). $X = \mathbb{R}^k \times \{0,1\}^{n-k}$
- **Mixed-integer nonlinear programming (MINLP):** MIP+NLP.

A first optimisation model

Let us start with a simple example to illustrate the optimisation modelling framework.

Consider the following: a carpenter makes tables and chairs using wood and her labour. The carpenter has a limited availability of labour and wood (see table below). What is the optimal weekly production of tables and chairs?

	Table	Chair	Available (per week)
Selling price ($)	800	600	-
Workload (h)	3	5	40
Wood (u)	7	4	60

A first optimisation model

Three key steps:

1. **Determine what needs to be decided** (*decision variables*)

   x_1 - amount of tables
   x_2 - amount of chairs

2. **How solutions are assessed** (*objective function*)

   Maximise revenue: max. $z = 800x_1 + 600x_2$

3. **The requirements that must be satisfied** (*constraints*)

   $3x_1 + 5x_2 \leq 40$ (available labour)
   $7x_1 + 4x_2 \leq 60$ (available wood)
   $x_1, x_2 \geq 0$

A first optimisation model

The complete model:

$$
\begin{aligned}
\text{max. } & z = 800x_1 + 600x_2 && \text{(profit)} \\
\text{s.t.: } & 3x_1 + 5x_2 \leq 40 && \text{(available labour)} \\
& 7x_1 + 4x_2 \leq 60 && \text{(available wood)} \\
& x_1, x_2 \geq 0
\end{aligned}
$$

Remarks: models are simplified representations of reality.
Simplifying assumptions in this example include:

- fractional number of chairs/ tables;
- no uncertainty;
- no production cost and no wastage of resources;
- perfect demand (all production is sold)...

A first optimisation model

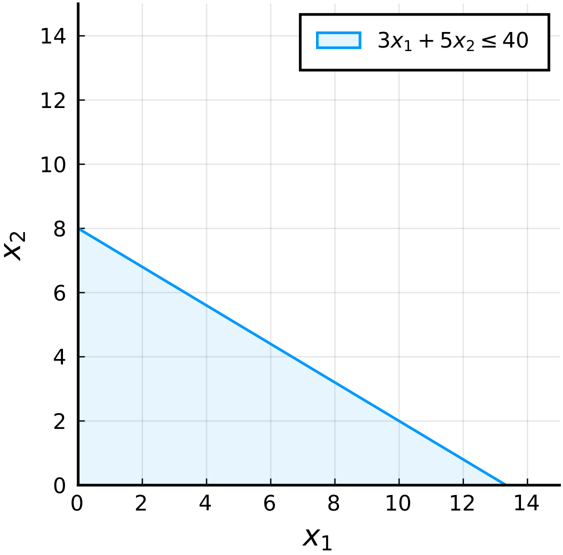
The most suitable optimisation method for solving an optimisation model depends on the model's mathematical properties.

- is the model linear?
- are there integer variables?
- are the nonlinear terms convex?
- are gradients available?

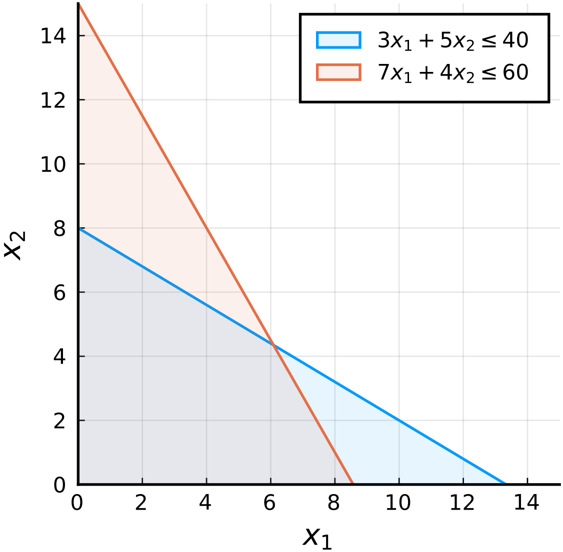
In this course, we will learn how to specify a suitable method for a model given its properties. For now, we will concentrate on (continuous) linear (optimisation) models.

Linear models have particular properties that can be exploited to devise an efficient optimisation method.

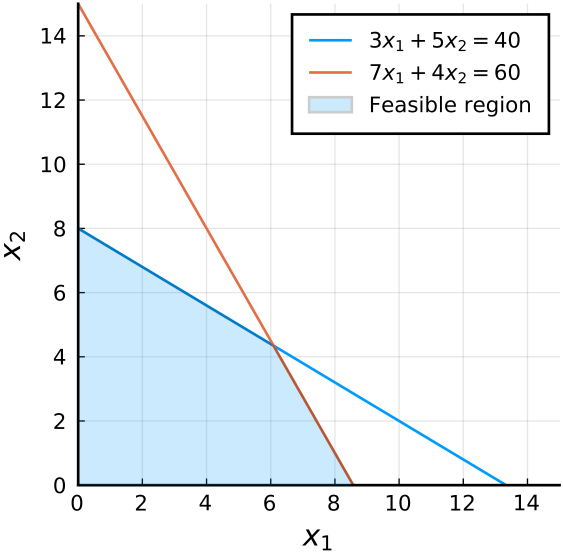
A first optimisation model - geometry of linear models



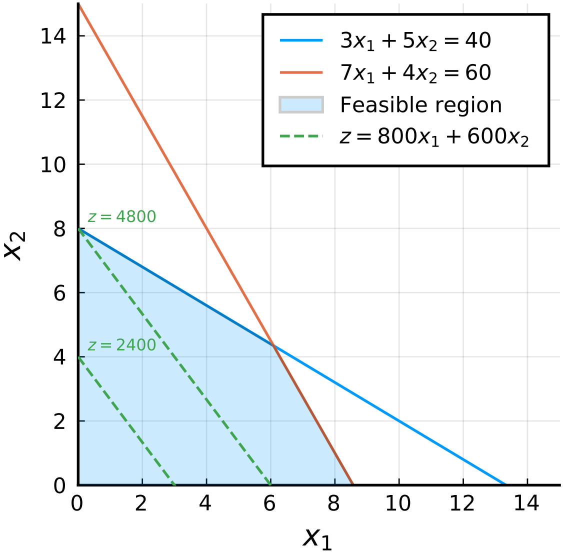
A first optimisation model - geometry of linear models



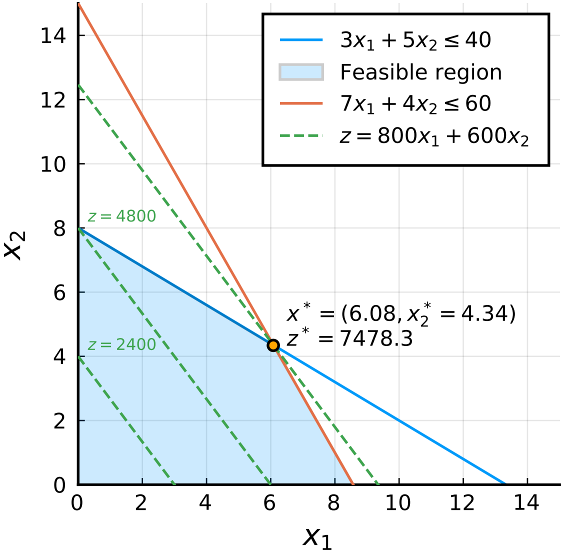
A first optimisation model - geometry of linear models



A first optimisation model - geometry of linear models



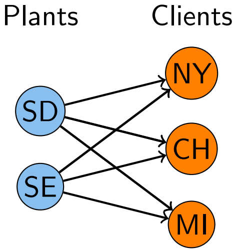
A first optimisation model - geometry of linear models



Real-world model: production planning problems (OR)

Problem statement:

- plan production and distribution;
- transportation cost proportional to distance travelled;
- factories have a capacity limit;
- clients have known demands.



	Clients			
Factory	NY	Chicago	Miami	Capacity
Seattle	2.5	1.7	1.8	350
San Diego	3.5	1.8	1.4	600
Demands	325	300	275	-

Table: Problem data: unit transportation costs, demands and capacities

Real-world model: production planning problems (OR)

Let $i \in I = \{\text{Seattle}, \text{San Diego}\}$ be the index set representing factories. Similarly, let $j \in J = \{\text{New York}, \text{Chicago}, \text{Miami}\}$.

Three key steps:

1. **Determine what needs to be decided** (*decision variables*)

   x_{ij} be the amount produced in factory i and sent to client j.

2. **How solutions are assessed** (*objective function*)

   Minimise total distribution cost:

   min. $z = 2.5x_{11} + 1.7x_{12} + 1.8x_{13} + 3.5x_{21} + 1.9x_{22} + 1.4x_{23}$,

   which can be more compactly expressed as

   $$\text{min. } z = \sum_{i \in I} \sum_{j \in J} c_{ij} x_{ij}$$

   where c_{ij} is the unit transportation cost from i to j.

Real-world model: production planning problems (OR)

3. **The requirements that must be satisfied** (*constraints*)

$$x_{11} + x_{12} + x_{13} \leq 350 \text{ (capacity limit Seattle)}$$
$$x_{21} + x_{22} + x_{23} \leq 600 \text{ (capacity limit San Diego)}$$
$$x_{11} + x_{21} \geq 325 \text{ (demand in New York)}$$
$$x_{12} + x_{22} \geq 300 \text{ (demand in Chicago)}$$
$$x_{13} + x_{23} \geq 275 \text{ (demand in Miami).}$$

These constraints can be expressed in the more compact form

$$\sum_{j \in J} x_{ij} \leq C_i, \forall i \in I$$
$$\sum_{i \in I} x_{ij} \geq D_j, \forall j \in J,$$

where C_i is the production capacity of factory i and D_j is the demand of client j.

Real-world model: production planning problems (OR)

The complete model:

$$\begin{aligned} \text{min. } & z = 2.5x_{11} + 1.7x_{12} + 1.8x_{13} + 3.5x_{21} + 1.9x_{22} + 1.4x_{23} \\ \text{s.t.: } & x_{11} + x_{12} + x_{13} \leq 350,\ x_{21} + x_{22} + x_{23} \leq 600 \\ & x_{11} + x_{21} \geq 325,\ x_{12} + x_{22} \geq 300,\ x_{13} + x_{23} \geq 275 \\ & x_{11}, \dots, x_{23} \geq 0. \end{aligned}$$

Or, more compactly, in the so called algebraic (symbolic) form

$$\begin{aligned} \text{min. } & z = \sum_{i \in I} \sum_{j \in J} c_{ij} x_{ij} \\ \text{s.t.: } & \sum_{j \in J} x_{ij} \leq C_i, \forall i \in I \\ & \sum_{i \in I} x_{ij} \geq D_j, \forall j \in J \\ & x_{ij} \geq 0, \forall i \in I, \forall j \in J. \end{aligned}$$

Real-world model: classification problem (ML)

Suppose we are given a data set $D \subset \mathbb{R}^n$ that can be separated into two disjunct sets in $\mathbb{R}^n$: $I^- = \{x_1, \dots, x_N\}$ and $I^+ = \{x_1, \dots, x_M\}$.

Each element $x_i \in D$ is an observation of a given set of features; belonging to either I^- or I^+ defines a classification.

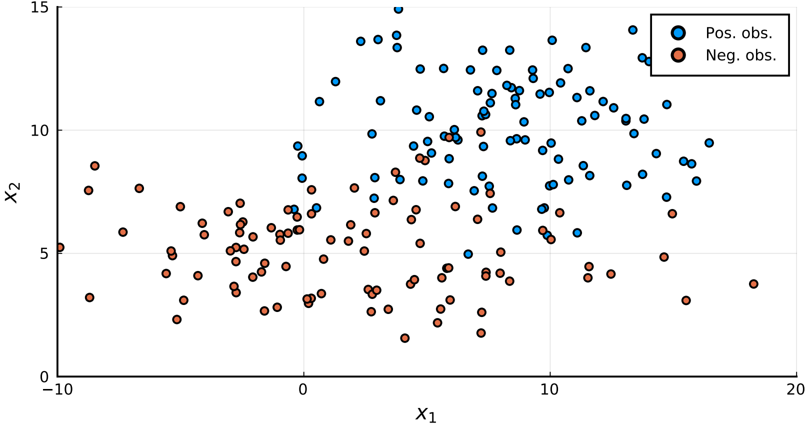
Our task is to obtain a function $f : \mathbb{R}^n \to \mathbb{R}$ from a given family of functions such that

$$f(x_i) < 0, i \in I^- \text{ and } f(x_i) > 0, i \in I^+.$$

f is selected as a linear classifier, i.e., $f(x_i) = a^\top x_i - b$, in which we try to set optimal a and b considering the classification error.

The best possible classifier is that which minimises misclassification.

Real-world model: classification problem (ML)



Real-world model: classification problem (ML)

Let us define the following error measures:

$$e^-(x_i \in I^-; a, b) = \begin{cases} 0, \text{ if } a^\top x_i - b \leq 0, \\ a^\top x_i - b, \text{ if } a^\top x_i - b > 0. \end{cases}$$

$$e^+(x_i \in I^+; a, b) = \begin{cases} 0, \text{ if } a^\top x_i - b \geq 0, \\ b - a^\top x_i, \text{ if } a^\top x_i - b < 0. \end{cases}$$

Using slack variables $\{u_i\}_{i=1}^M$ and $\{v_i\}_{i=1}^N$ to represent e^- and e^+, respectively, the optimal classifier is obtained from solving

$$\begin{aligned} (LC) \ : \ \text{min.} \ & \sum_{i=1}^M u_i + \sum_{i=1}^N v_i \\ \text{s.t.: } & a^\top x_i - b - u_i \leq 0, i = 1, \dots, M \\ & a^\top x_i - b + v_i \geq 0, i = 1, \dots, N \\ & a \in \mathbb{R}^n, b \in \mathbb{R} \\ & u_i \geq 0, i = 1, \dots, M; v_i \geq 0, i = 1, \dots, N. \end{aligned}$$

Real-world model: classification problem (ML)