Modern Optimization Techniques

0. Overview

Lars Schmidt-Thieme

Information Systems and Machine Learning Lab (ISMLL)
Institute for Computer Science
University of Hildesheim, Germany

Outline

1. Linear Optimization

2. Optimization Problems

3. Application Areas

4. Classification of Optimization Problems

5. Overview of the Lecture

6. Organizational Stuff

Outline

1. Linear Optimization

2. Optimization Problems

3. Application Areas

4. Classification of Optimization Problems

5. Overview of the Lecture

6. Organizational Stuff

Optimization Problems

$$\text{find an } x \in \mathcal{X} \text{ with } f(x) \text{ maximal}$$

or for short

$$\arg\max_{x \in \mathcal{X}} f(x)$$

- $f : \mathbb{R}^N \to \mathbb{R}$: **objective function**
- $\mathcal{X} \subseteq \mathbb{R}^N$: **feasible area**, e.g., $\mathcal{X} := \mathbb{R}^N$
- $x \in \mathcal{X}$: **optimization variables** $x_1, x_2, \ldots, x_N$
- $x^* \in \arg\max_{x \in \mathcal{X}} f(x)$: **optima, solutions**

Linear Objectives Without Constraints

- Most simple case:
  - linear objective,
  - no constraints

  i.e.,

$$
\begin{aligned}
\max_{x} f(x) &:= cx, && x \in \mathbb{R} \\
\max_{x_1,x_2} f(x_1, x_2) &:= c_1x_1 + c_2x_2, && x_1, x_2 \in \mathbb{R} \\
\max_{x_1,x_2,x_3} f(x_1, x_2, x_3) &:= c_1x_1 + c_2x_2 + c_3x_3, && x_1, x_2, x_3 \in \mathbb{R} \\
\max f(x) &:= c^T x, && x, c \in \mathbb{R}^N
\end{aligned}
$$

- Q: Where can a linear function have its maximum?

Linear Objectives Without Constraints

- Most simple case:
  - linear objective,
  - no constraints

  i.e.,

$$
\begin{aligned}
\max_x f(x) &:= cx, & x \in \mathbb{R} \\
\max_{x_1, x_2} f(x_1, x_2) &:= c_1 x_1 + c_2 x_2, & x_1, x_2 \in \mathbb{R} \\
\max_{x_1, x_2, x_3} f(x_1, x_2, x_3) &:= c_1 x_1 + c_2 x_2 + c_3 x_3, & x_1, x_2, x_3 \in \mathbb{R} \\
\max f(x) &:= c^T x, & x, c \in \mathbb{R}^N
\end{aligned}
$$

- Q: Where can a linear function have its maximum?

- Linear functions are either constant or unbounded, thus maximizing them makes no sense.

Example

Plastic Cup Factory

A local family-owned plastic cup manufacturer wants to optimize their production mix in order to maximize their profit. They produce personalized beer mugs and champagne glasses. The profit on a case of beer mugs is $25 while the profit on a case of champagne glasses is $20. The cups are manufactured with a machine called a plastic extruder which feeds on plastic resins. Each case of beer mugs requires 20 lbs. of plastic resins to produce while champagne glasses require 12 lbs. per case. The daily supply of plastic resins is limited to at most 1800 pounds. About 15 cases of either product can be produced per hour. At the moment the family wants to limit their work day to 8 hours.

source: *https://sites.math.washington.edu/~burke/crs/407/notes/section1.pdf*

Example

product	resources materials	time	profit	amount
A	20	$1/15$	25	x_1
B	12	$1/15$	20	x_2
limit	≤ 1800	≤ 8	max.	

Example

product	resources materials	time	profit	amount
A	20	$1/15$	25	x_1
B	12	$1/15$	20	x_2
limit	≤ 1800	≤ 8	max.	

$$
\begin{aligned}
\max\ & 25x_1 + 20x_2 \\
\text{s.t. } & 20x_1 + 12x_2 \leq 1800 \\
& 1/15x_1 + 1/15x_2 \leq 8 \\
& x_1, x_2 \geq 0
\end{aligned}
$$

Example

product	resources materials	time	profit	amount
A	20	1/15	25	x_1
B	12	1/15	20	x_2
limit	≤ 1800	≤ 8	max.	

$$
\begin{aligned}
\max\ & 25x_1 + 20x_2 \\
\text{s.t. } & 20x_1 + 12x_2 \leq 1800 \\
& 1/15x_1 + 1/15x_2 \leq 8 \\
& x_1, x_2 \geq 0
\end{aligned}
$$

$$
\begin{aligned}
\max\ & c^T x \\
\text{s.t. } & Bx \leq b \\
& x \geq 0 \\
\text{with } & c, x \in \mathbb{R}^N \\
& B \in \mathbb{R}^{Q \times N} \\
& b \in \mathbb{R}^Q
\end{aligned}
$$

Linear Optimization Problems

- A problem

$$\begin{aligned} \max\ & c^T x \\ \text{s.t. } & Bx \leq b \\ & x \geq 0 \end{aligned}$$

$$\text{with } c, x \in \mathbb{R}^N, \quad B \in \mathbb{R}^{Q \times N}, \quad b \in \mathbb{R}^Q$$

  is called **linear optimization problem**.
  - linear objective $f(x) := c^T x$
  - $Bx \leq b$ are called **inequality constraints**
    - Q linear constraints
    - define the **feasible area** $\mathcal{X} := \{x \in \mathbb{R}^N \mid Bx \leq b, x \geq 0\}$
- most simple optimization problem
- linear optimization problems without constraints are unbounded
  - no optimum exists, problem makes no sense
- Q: Where can a linear optimization problem (with constraints) have its maximum?

Linear Optimization Problems

- A problem

$$\begin{aligned}\max\ & c^T x\\ \text{s.t. } & Bx \leq b\\ & x \geq 0\\ \text{with } & c, x \in \mathbb{R}^N, \quad B \in \mathbb{R}^{Q\times N}, \quad b \in \mathbb{R}^Q\end{aligned}$$

  is called **linear optimization problem**.
  - linear objective $f(x) := c^T x$
  - $Bx \leq b$ are called **inequality constraints**
    - Q linear constraints
    - define the **feasible area** $\mathcal{X} := \{x \in \mathbb{R}^N \mid Bx \leq b, x \geq 0\}$
- most simple optimization problem
- linear optimization problems without constraints are unbounded
  - no optimum exists, problem makes no sense
- the optimum always is on the **border of the feasible area**.

$$\partial \mathcal{X} = \{x \in \mathcal{X} \mid \exists q : (Bx)_q = b_q \text{ or } \exists n : x_n = 0\}$$

Slack Variables

- Introduce Q further variables $x_{N+1}, \ldots, x_{N+Q}$ to measure the slack of each constraint:

$$x_{N+1:N+Q} := b - Bx \geq 0$$

- each variable $x_n, n = 1{:}N + Q$ represents a constraint / a border of the feasible region:
  - $x_n, n = 1{:}N$ the constraint $x_n \geq 0$ and
  - $x_{N+q}, q = 1{:}Q$ the constraint $B_{q,.}^T x \leq b_q$
  - $x_n = 0$ means the constraint is sharp, i.e., x is on the respective border
- a linear objective with linear constraints assumes its maximum at a corner of the feasible region,
  - always N constraints are sharp

Simplex Tableau

- let's assume $b \geq 0$ and thus $x_{1:N} = 0$ is feasible.
  - otherwise solve $B(x + \tilde{x}) \leq b$, i.e., $Bx \leq b - B\tilde{x} =: b^{\text{new}}$ for a feasible $\tilde{x}$.

- start with

$$x_{1:N} := 0_N$$
$$x_{N+1:N+Q} := b - B\, x_{1:N} = b$$

then holds

$$\begin{pmatrix} B & I_{Q\times Q} \\ c^T & 0_Q \end{pmatrix} x_{1:N+Q} = \begin{pmatrix} b \\ 0 \end{pmatrix}$$

- coefficients can be collected in a matrix called **simplex tableau**:

$$T := \begin{pmatrix} B & I_{Q\times Q} & b \\ c^T & 0_Q & 0 \end{pmatrix}$$

Pivot Step

$$T := \begin{pmatrix} B & I_{Q\times Q} & b \\ c^T & 0_Q & 0 \end{pmatrix}$$

- if $c_n > 0$, we can increase the objective by increasing x_n
- but increasing x_n may also decrease some slacks x_{N+q}:
  for each $q = 1{:}Q$ check:
  - if $B_{q,n} > 0$, then we may increase x_n maximally by

$$\frac{b_q}{B_{q,n}}$$

  - thus choose

$$q := \arg\min \{\frac{b_q}{B_{q,n}} \mid q = 1{:}Q, B_{q,n} > 0\}$$

$$x_n := \frac{b_q}{B_{q,n}}, \quad x_{N+q} := 0$$

- make column n the q-th unit vector I_q (same as in Gaussian elimin.):
  - normalize row q s.t. the pivot cell gets 1: $T_{q,.} := T_{q,.}/T_{q,n}$
  - eliminate column n in all other rows: $T_{r,.} := T_{r,.} - T_{r,n}T_{q,.}\ \forall r \neq q$.

Stop and Solution

- stop once there is no positive c_n anymore.
- solution x^*:
  - non-zero x^*_n are those having unit vector I_q (for a $q \in 1{:}Q$) in column n of the final tableau
  - their value is just $T_{q,N+Q+1}$

Worked Example

$$\max c^T x = (5\ 4\ 3)^T x$$

$$\text{s.t. } Bx = \begin{pmatrix} 2 & 3 & 1 \\ 4 & 1 & 2 \\ 3 & 4 & 2 \end{pmatrix} x \leq b = \begin{pmatrix} 5 \\ 11 \\ 8 \end{pmatrix}$$

$$x \geq 0$$

Worked Example

$$\max c^T x = (5\ 4\ 3)^T x$$
$$\text{s.t. } Bx = \begin{pmatrix} 2 & 3 & 1 \\ 4 & 1 & 2 \\ 3 & 4 & 2 \end{pmatrix} x \leq b = \begin{pmatrix} 5 \\ 11 \\ 8 \end{pmatrix}$$
$$x \geq 0$$

$$T^{(0)} := \begin{pmatrix} B & I_{Q\times Q} & b \\ c^T & 0_Q & 0 \end{pmatrix} = \begin{pmatrix} 2 & 3 & 1 & 1 & 0 & 0 & 5 \\ 4 & 1 & 2 & 0 & 1 & 0 & 11 \\ 3 & 4 & 2 & 0 & 0 & 1 & 8 \\ 5 & 4 & 3 & 0 & 0 & 0 & 0 \end{pmatrix}$$

Worked Example

$$T^{(0)} = \begin{pmatrix} 2 & 3 & 1 & 1 & 0 & 0 & 5 \\ 4 & 1 & 2 & 0 & 1 & 0 & 11 \\ 3 & 4 & 2 & 0 & 0 & 1 & 8 \\ 5 & 4 & 3 & 0 & 0 & 0 & 0 \end{pmatrix}$$

remember: Pivot (n, q):

$$T := \begin{pmatrix} B & I_{Q\times Q} & b \\ c & 0_Q & 0 \end{pmatrix}$$

$$n := \arg\max_{n:c_n>0} c_n, \quad q := \arg\min_{q=1:Q} \frac{b_q}{B_{q,n}}$$

Worked Example

$$T^{(0)} = \begin{pmatrix} 2 & 3 & 1 & 1 & 0 & 0 & 5 \\ 4 & 1 & 2 & 0 & 1 & 0 & 11 \\ 3 & 4 & 2 & 0 & 0 & 1 & 8 \\ 5 & 4 & 3 & 0 & 0 & 0 & 0 \end{pmatrix}$$

$$T^{(1)} = \begin{pmatrix} 1 & 3/2 & 1/2 & 1/2 & 0 & 0 & 5/2 \\ 0 & -5 & 0 & -2 & 1 & 0 & 1 \\ 0 & -1/2 & 1/2 & -3/2 & 0 & 1 & 1/2 \\ 0 & -7/2 & 1/2 & -5/2 & 0 & 0 & -25/2 \end{pmatrix}$$

Note: $T^{(0)}$ pivot $(1, 1)$

Worked Example

$$T^{(0)} = \begin{pmatrix} 2 & 3 & 1 & 1 & 0 & 0 & 5 \\ 4 & 1 & 2 & 0 & 1 & 0 & 11 \\ 3 & 4 & 2 & 0 & 0 & 1 & 8 \\ 5 & 4 & 3 & 0 & 0 & 0 & 0 \end{pmatrix}$$

$$T^{(1)} = \begin{pmatrix} 1 & 3/2 & 1/2 & 1/2 & 0 & 0 & 5/2 \\ 0 & -5 & 0 & -2 & 1 & 0 & 1 \\ 0 & -1/2 & 1/2 & -3/2 & 0 & 1 & 1/2 \\ 0 & -7/2 & 1/2 & -5/2 & 0 & 0 & -25/2 \end{pmatrix}$$

$$T^{(2)} = \begin{pmatrix} 1 & 2 & 0 & 2 & 0 & 11 & 2 \\ 0 & -5 & 0 & -2 & 1 & 0 & 1 \\ 0 & -1 & 1 & -3 & 0 & 22 & 1 \\ 0 & -3 & 0 & -1 & 0 & -11 & -13 \end{pmatrix}$$

Note: $T^{(0)}$ pivot $(1,1)$, $T^{(1)}$ pivot $(3,3)$.

Worked Example

$$T^{(0)} = \begin{pmatrix} 2 & 3 & 1 & 1 & 0 & 0 & 5 \\ 4 & 1 & 2 & 0 & 1 & 0 & 11 \\ 3 & 4 & 2 & 0 & 0 & 1 & 8 \\ 5 & 4 & 3 & 0 & 0 & 0 & 0 \end{pmatrix}$$

$$T^{(1)} = \begin{pmatrix} 1 & 3/2 & 1/2 & 1/2 & 0 & 0 & 5/2 \\ 0 & -5 & 0 & -2 & 1 & 0 & 1 \\ 0 & -1/2 & 1/2 & -3/2 & 0 & 1 & 1/2 \\ 0 & -7/2 & 1/2 & -5/2 & 0 & 0 & -25/2 \end{pmatrix}$$

$$T^{(2)} = \begin{pmatrix} 1 & 2 & 0 & 2 & 0 & 11 & 2 \\ 0 & -5 & 0 & -2 & 1 & 0 & 1 \\ 0 & -1 & 1 & -3 & 0 & 22 & 1 \\ 0 & -3 & 0 & -1 & 0 & -11 & -13 \end{pmatrix}$$

$$x^* = (2\ 0\ 1)^T \text{ with } c^T x^* = 13$$

Note: $T^{(0)}$ pivot $(1,1)$, $T^{(1)}$ pivot $(3,3)$.

Simplex Algorithm (for $x = 0_N$ feasible, i.e. $b \geq 0$)

max-simplex(c, B, b):

$$T := \begin{pmatrix} B & I_{Q\times Q} & b \\ c^T & 0_Q & 0 \end{pmatrix}$$

  $(n, q) :=$ find-pivot(T)
  while $(n, q) \neq (-1, -1)$:
    $T_{q,.} = T_{q,.} / T_{q,n}$
    for $r := 1{:}Q + 1, r \neq q$:
      $T_{r,.} := T_{r,.} - T_{r,n} T_{q,.}$
    $(n, q) :=$ find-pivot(T)

  $x := 0_N$
  for $n := 1{:}N$:
    if $\exists q \in 1{:}Q : T_{.,n} = I_q$:
      $x_n := T_{q,N+Q+1}$
  return x

find-pivot(T):
  Ns $:= \{n \in 1{:}N \mid T_{Q+1,n} > 0\}$
  if $\exists n \in \text{Ns} : T_{-(Q+1),n} \leq 0_Q$
    raise exception "problem unbounded"
  if Ns $= \emptyset$:
    return $(-1, -1)$
  $n := \arg\max_{n \in \text{Ns}} T_{Q+1,n}$
  $q := \text{argmin}_{q=1:Q, T_{q,n}>0} \frac{T_{q,N+Q+1}}{T_{q,n}}$
  return (n, q)

Note: $I_n := (\mathbb{I}(m = n))_{m\in 1:N} = (0 \dots 0\ 1\ 0 \dots 0)^T$ with a 1 at index n.

Outline

1. Linear Optimization

2. Optimization Problems

3. Application Areas

4. Classification of Optimization Problems

5. Overview of the Lecture

6. Organizational Stuff

Optimization Problems

An **unconstrained optimization problem** has the form:

$$\text{minimize} \quad f(x)$$

where

- $f : \mathbb{R}^N \to \mathbb{R}$
- a minimal x^* exists
  - we often will denote its minimal value by $p^* := f(x^*)$

Optimization Problems — A simple example

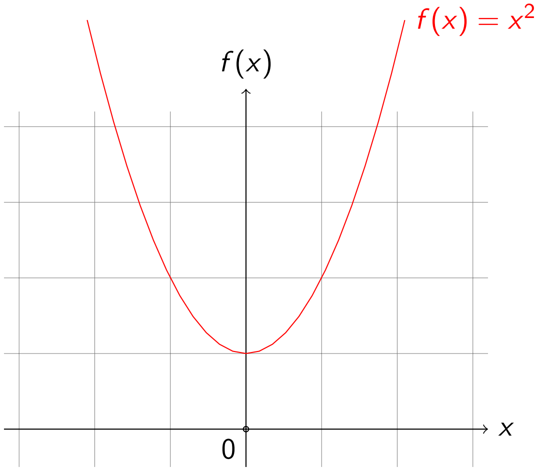
Say we have $f(x) = x^2 + 1$:

$$\text{minimize} \quad x^2 + 1$$

Optimization Problems — A simple example

Say we have $f(x) = x^2 + 1$:

$$\text{minimize} \quad x^2 + 1$$



Optimization Problems — A simple example

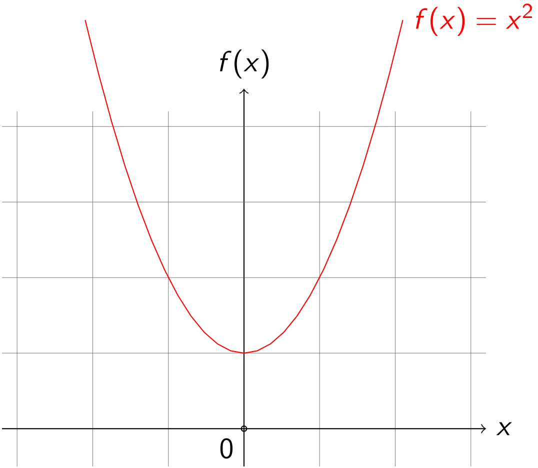
Say we have $f(x) = x^2 + 1$:

$$\text{minimize} \quad x^2 + 1$$

$$\frac{\partial f(x)}{\partial x} \overset{!}{=} 0$$

$$2x = 0$$

$$x = 0$$



Optimization Problems — A simple example

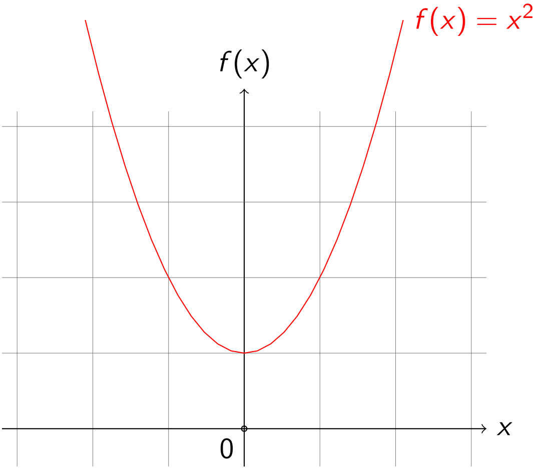
Say we have $f(x) = x^2 + 1$:

$$\text{minimize} \quad x^2 + 1$$

$$\frac{\partial f(x)}{\partial x} \stackrel{!}{=} 0$$

$$2x = 0$$

$$x = 0$$

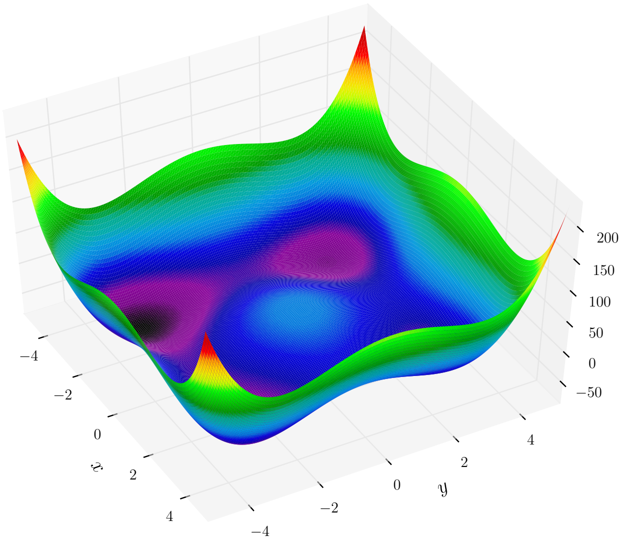


So:

$$x^* = 0$$

$$p^* = f(x^*) = 0^2 + 1 = 1$$

Optimization Problems



Optimization Problems — Constraints

A **constrained optimization problem** has the form:

$$\begin{array}{rll} \text{minimize} & f(x) & \\ \text{subject to} & g_p(x) = 0, & p = 1, \ldots, P \\ & h_q(x) \leq 0, & q = 1, \ldots, Q \end{array}$$

where

- $f : \mathbb{R}^N \to \mathbb{R}$
- $g_1, \ldots, g_P : \mathbb{R}^N \to \mathbb{R}$
- $h_1, \ldots, h_Q : \mathbb{R}^N \to \mathbb{R}$
- a minimal x^* exists

Optimization Problems — Vocabulary

$$\begin{aligned} \text{minimize} \quad & f(x) \\ \text{subject to} \quad & g_p(x) = 0, \quad p = 1, \dots, P \\ & h_q(x) \leq 0, \quad q = 1, \dots, Q \end{aligned}$$

where

- $f : \mathbb{R}^N \rightarrow \mathbb{R}$ is the **objective function**
- $x \in \mathbb{R}^N$ is the **optimization variable**
- $g_p : \mathbb{R}^N \rightarrow \mathbb{R}, p = 1 : P$ are the **equality constraint functions**
  - usually g_p are linear: $g_p(x) := A_{p,.}x - a_p$, $A \in \mathbb{R}^{P \times N}, a \in \mathbb{R}^P$
- $h_q : \mathbb{R}^N \rightarrow \mathbb{R}, q = 1 : Q$ are the **inequality constraint functions**

Objective Domain, Feasible Area and Constraints

- often the objective domain will be all of $\mathbb{R}^N$ and the constraints define the feasible area:

$$\operatorname{dom} f = \mathbb{R}^N,$$
$$\mathcal{X} := \{x \in \mathbb{R}^N \mid g(x) = 0, h(x) \leq 0\}$$

- sometimes the objective domain will be a (simple, convex) open subset of $\mathbb{R}^N$ (e.g., all vectors with all positive entries) and the constraints refine the feasible area:

$$\text{e.g.,} \quad \operatorname{dom} f = \{x \in \mathbb{R}^N \mid x_n > 0 \ \forall n = 1{:}N\}$$
$$\mathcal{X} := \{x \in \operatorname{dom} f \mid g(x) = 0, h(x) \leq 0\}$$

Outline

1. Linear Optimization

2. Optimization Problems

3. Application Areas

4. Classification of Optimization Problems

5. Overview of the Lecture

6. Organizational Stuff

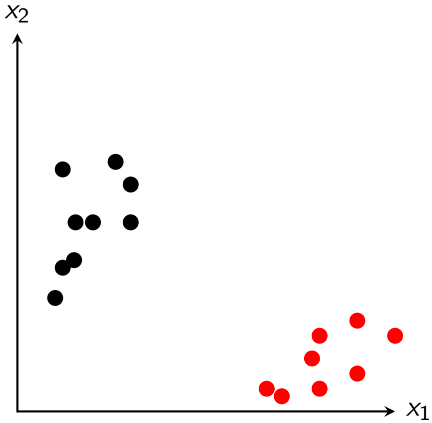
What is optimization good for?

The optimization problem is an abstraction of the problem of making the best possible choice of a vector in $\mathbb{R}^N$ from a set of candidate choices

- Machine Learning
- Logistics
- Computer Vision
- Decision Making
- Scheduling
- ...

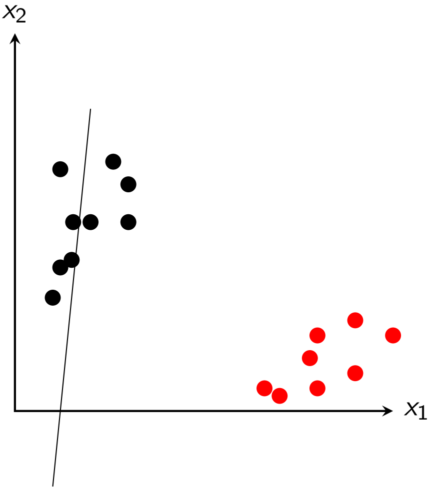
Application Areas — Machine Learning

Task: Classification



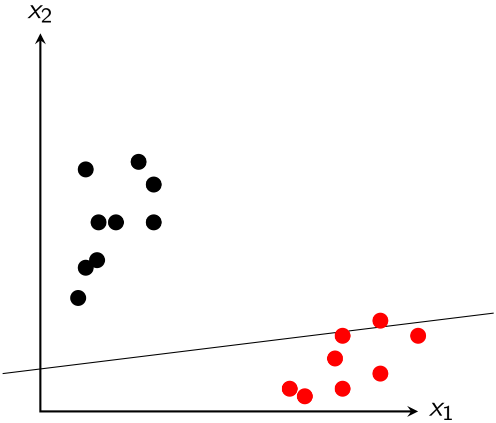
Application Areas — Machine Learning

Task: Classification



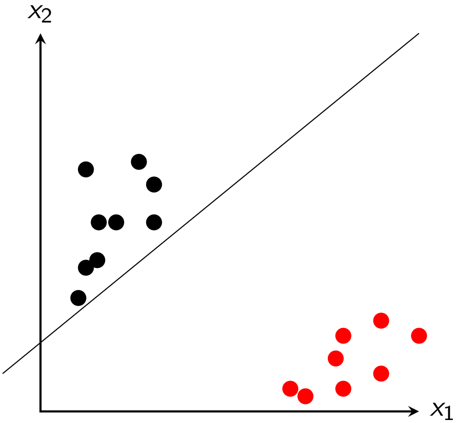
Application Areas — Machine Learning

Task: Classification



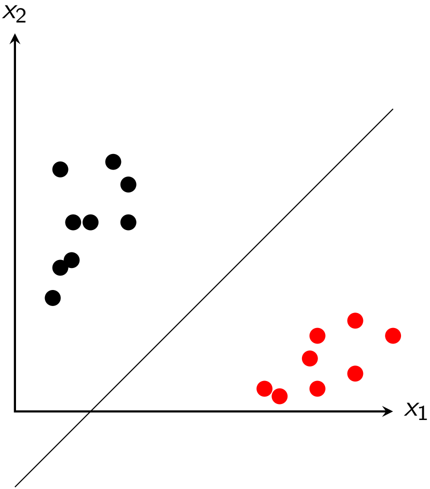
Application Areas — Machine Learning

Task: Classification



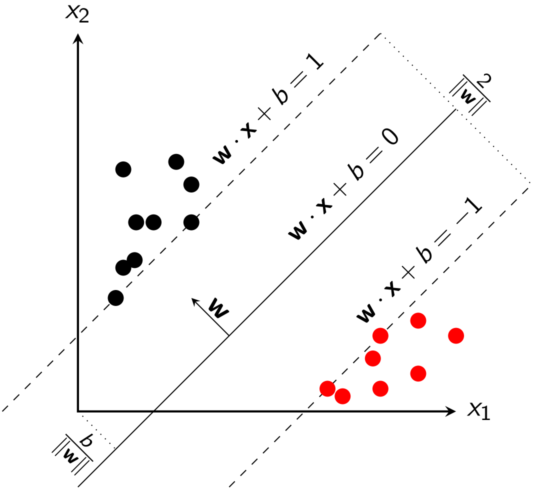
Application Areas — Machine Learning

Task: Classification



Application Areas — Machine Learning

Task: Classification



Application Areas — Logistics

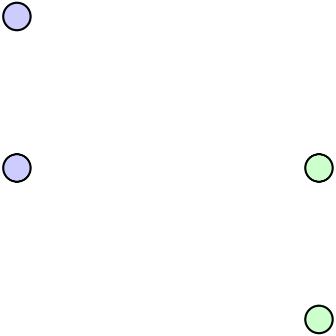
Suppose we have:

Application Areas — Logistics

Suppose we have:

- Factories

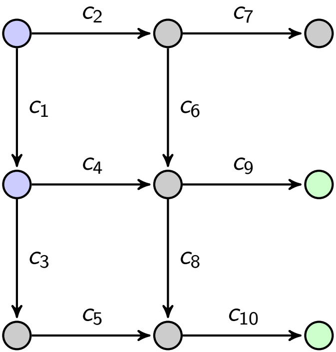
Application Areas — Logistics



Suppose we have:

- Factories
- Warehouses

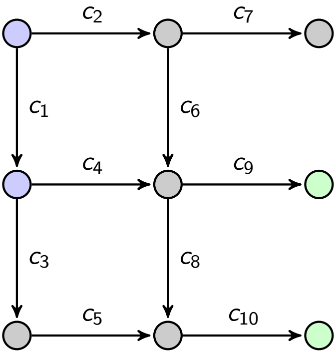
Application Areas — Logistics



Suppose we have:

- Factories
- Warehouses
- Roads with costs associated to them

Application Areas — Logistics



Suppose we have:

- Factories
- Warehouses
- Roads with costs associated to them

Determine how many products to ship from each factory to each warehouse to minimize shipping cost while meeting warehouse demands and not exceeding factory supplies

Outline

1. Linear Optimization

2. Optimization Problems

3. Application Areas

4. Classification of Optimization Problems

5. Overview of the Lecture

6. Organizational Stuff

Classification

There are many different ways to group mathematical optimization problems:

1. Linear vs. Non-linear
2. Convex vs. Non-convex
3. Constrained vs. Unconstrained

1. Linear vs. Non-Linear Problems

A function $f : \mathbb{R}^N \rightarrow \mathbb{R}$ is **linear** if it satistfies

$$f(\alpha x + \beta y) = \alpha f(x) + \beta f(y)$$

for all

- $x, y \in \mathbb{R}^N$
- $\alpha, \beta \in \mathbb{R}$

1. Linear vs. Non-Linear Problems

A function $f : \mathbb{R}^N \rightarrow \mathbb{R}$ is **linear** if it satistfies

$$f(\alpha x + \beta y) = \alpha f(x) + \beta f(y)$$

for all

- $x, y \in \mathbb{R}^N$
- $\alpha, \beta \in \mathbb{R}$

An unconstrained optimization problem

$$\text{minimize} \quad f(x)$$

with f being linear is unbounded (thus makes no sense).

Convex Functions

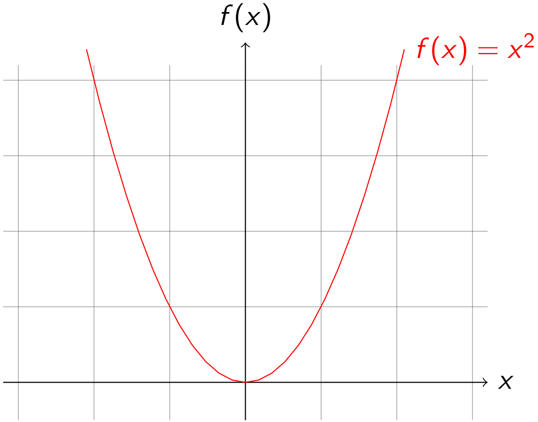
A function $f : \mathbb{R}^N \to \mathbb{R}$ is **convex** if it satisfies

$$f(\alpha x + \beta y) \leq \alpha f(x) + \beta f(y)$$

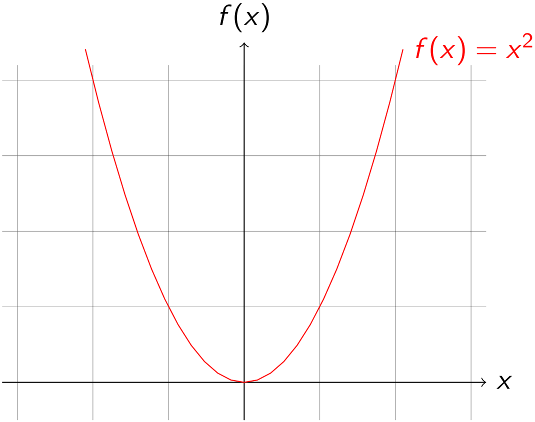
for all

- $x, y \in \mathbb{R}^N$
- $\alpha, \beta \in \mathbb{R}$
- $\alpha + \beta = 1$, $\alpha \geq 0$, $\beta \geq 0$

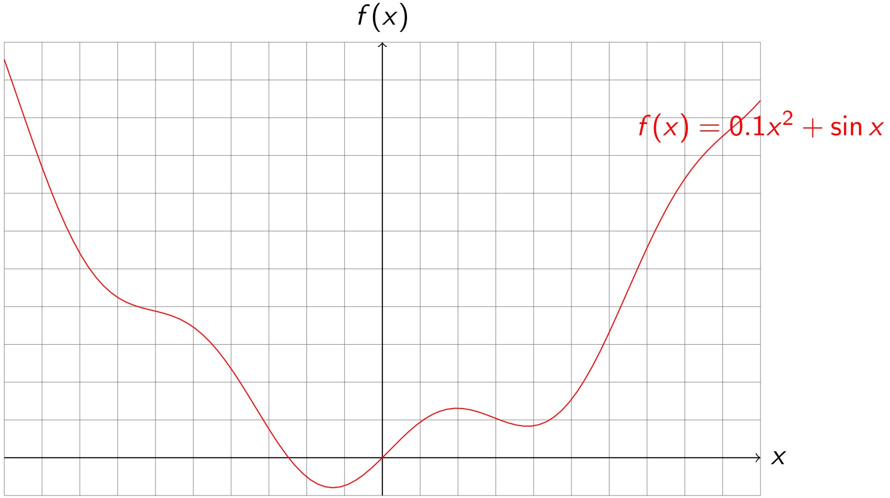
A convex function?



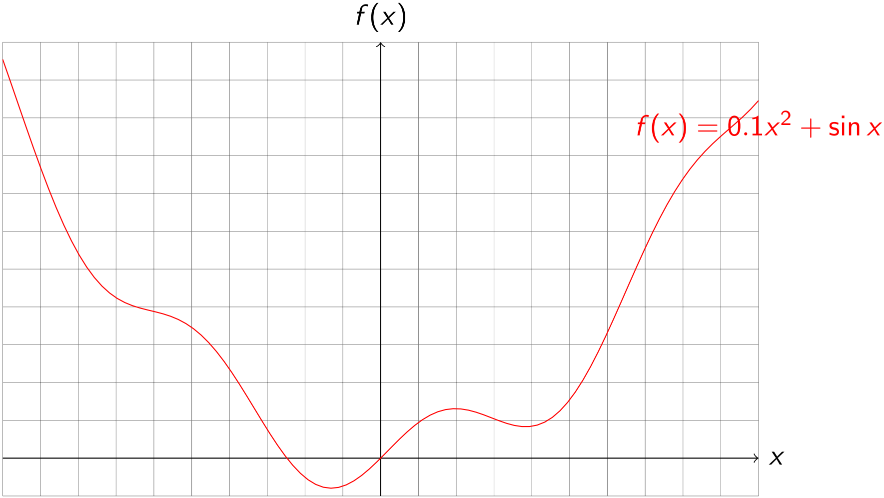
A convex function? Yes



A convex function?



A convex function? No



2. Convex vs. Non-Convex Optimization Problem

An (unconstrained!) optimization problem

$$\text{minimize} \quad f(x)$$

is said to be **convex** if

- the objective function f is convex.

3. Constrained vs. Unconstrained Problems

An **unconstrained optimization problem** has only

- the objective function f

$$\text{minimize} \quad f(x)$$

3. Constrained vs. Unconstrained Problems

An **unconstrained optimization problem** has only

- the objective function f

$$\text{minimize} \quad f(x)$$

A **constrained optimization problem** has besides

- objective function f
- the **equality constraint functions** $g_1, \ldots g_P$ and/or
- the **inequality constraint functions** $h_1, \ldots h_Q$

$$\begin{aligned} \text{minimize} \quad & f(x) \\ \text{subject to} \quad & g_p(x) = 0, \quad p = 1, \ldots, P \\ & h_q(x) \leq 0, \quad q = 1, \ldots, Q \end{aligned}$$

Linear vs. Non-Linear Constrained Problems

A constrained optimization problem

$$
\begin{aligned}
\text{minimize} \quad & f(x) \\
\text{subject to} \quad & g_p(x) = 0, \quad p = 1, \ldots, P \\
& h_q(x) \leq 0, \quad q = 1, \ldots, Q
\end{aligned}
$$

is said to be **linear** if

- the objective function f,
- the equality constraints $g_1, \ldots g_P$ and
- the inequality constraints $h_1, \ldots h_Q$ are linear.

Linear vs. Non-Linear Constrained Problems

A linear constrained optimization problem can be written as

$$\begin{aligned} \text{minimize} \quad & f(x) := c^T x \\ \text{subject to} \quad & g(x) := Ax - a = 0 \\ & h(x) := Bx - b \leq 0 \end{aligned}$$

with

- a vector $c \in \mathbb{R}^N$,
- a matrix $A \in \mathbb{R}^{P \times N}$, a vector $a \in \mathbb{R}^P$ and
- a matrix $B \in \mathbb{R}^{Q \times N}$, a vector $b \in \mathbb{R}^Q$.

Convex vs. Non-Convex Constrained Problems

A constrained optimization problem

$$\begin{aligned} \text{minimize} \quad & f(x) \\ \text{subject to} \quad & g_p(x) = 0, \quad p = 1, \ldots, P \\ & h_q(x) \leq 0, \quad q = 1, \ldots, Q \end{aligned}$$

is said to be **convex** if

- the objective function f and
- the inequality constraints $h_1, \ldots h_Q$ are convex and
- the equality constraints $g_1, \ldots g_P$ are even **linear**.

Types of Optimization Problems

	linear	convex	non-convex
unconstrained	—	Convex Optimization	Global Optimization
constrained	Linear Programming	Convex Optimization	Global Optimization

Outline

1. Linear Optimization

2. Optimization Problems

3. Application Areas

4. Classification of Optimization Problems

5. Overview of the Lecture

6. Organizational Stuff

Syllabus

Mon. 07.11.	(1)	0. Overview
		1. Theory
Mon. 14.11.	(2)	1. Convex Sets and Functions
		2. Unconstrained Optimization
Mon. 21.11.	(3)	2.1 Gradient Descent
Mon. 28.11.	(4)	2.2 Stochastic Gradient Descent
Mon. 05.12.	(5)	2.3 Newton's Method
Mon. 12.12.	(6)	2.4 Quasi-Newton Methods
Mon. 19.12.	(7)	2.5 Subgradient Methods
	—	— *Christmas Break* —
Mon. 09.01.	(8)	2.6 Coordinate Descent
		3. Equality Constrained Optimization
Mon. 16.01.	(9)	3.1 Duality
Mon. 23.01.	(10)	3.2 Methods
		4. Inequality Constrained Optimization
Mon. 30.01.	(11)	4.1 Primal Methods
Mon. 06.01.	(12)	4.2 Barrier and Penalty Methods
Mon. 13.02.	(13)	4.3 Cutting Plane Methods

Outline

1. Linear Optimization

2. Optimization Problems

3. Application Areas

4. Classification of Optimization Problems

5. Overview of the Lecture

6. Organizational Stuff

Exercises and Tutorials (1/2)

- weekly sheet with 2 exercises
  - handed out **each Monday** 11:59 PM online via Learnweb
  - 1^{st} sheet will be handed out today (on 7.11).
- Solutions to the exercises are submitted online via Learnweb
  - until next **Saturday 11:59 PM**
  - 1^{st} sheet will be due on Saturday 12.11
- Exercises will be corrected.
- You will get feedback on your submission in the tutorial.

Exercises and Tutorials (2/2)

- Tutorials:
  - **Tue, 8am - 10am (Deepalakshmi Murugesu)** and
  - **Wed, 2pm - 4pm (Vijaya Krishna Yalavarthi)**

  starting this week (Wed. only for all).

- Successful participation in the tutorial gives up to 10% bonus points for the exam.
  - group submissions are allowed but do not contribute to the bonus points.
  - Plagiarism is illegal and usually leads to expulsion from the program.
    - about plagiarism see `https://en.wikipedia.org/wiki/Plagiarism`

Exams and credit points

- There will be a written exam at the end of the term
  - 2h, 4 problems, open book
- The course gives 6 ECTS (2+2 SWS)
- The course can be used in
  - Data Analytics MSc. (mandatory)
  - IMIT and AINF MSc. / Informatik / Gebiet KI & ML (elective)
  - Wirtschaftsinformatik MSc / Business Intelligence (elective)

Some books

- Stephen Boyd, Lieven Vandenberghe (2004):
  Convex Optimization, Cambridge University Press.

- David G. Luenberger, Yinyu Ye (2008; 3rd):
  Linear and Nonlinear Programming, Springer.

- Jorge Nocedal, Steven Wright (2006):
  Numerical Optimization, Springer.

- Igor Griva, Stephen G. Nash, Ariela Sofer (2009):
  Linear and nonlinear optimization, SIAM.

- Dimitri P. Bertsekas (2016; 3rd):
  Nonlinear Programming, Athena Scientific.

Summary

- **Optimization problems** are described by
  - an **objective function** f of real vectors $x \in \mathbb{R}^N$,
  - optionally, a list of **equality constraints** g, and
  - optionally, a list of **inequality constraints** h.
- Optimization problems are called **linear**,
  if their objective and both types of contraints are linear.
- Optimization problems are called **convex**,
  if their objective and inequality constraints are convex
  and their equality constraints even are linear.
- Linear optimization problems are the most simple,
  convex optimization problems the next simple problem class.
- Linear optimization problems can be solved by the **simplex algorithm**.

Further Readings

- to review linear optimization:
  - Luenberger and Ye, 2008, ch. 2 and 3.
- general introduction to convex optimization:
  - Boyd and Vandenberghe, 2004, ch. 1.
  - Luenberger and Ye, 2008, ch. 1.
  - Nocedal and Wright, 2006, ch. 1.
  - Griva, Nash, and Sofer, 2009, ch. 1.

Acknowledgement: An earlier version of the slides has been written by Lucas Rego Drumond (ISMLL).

References

Boyd, Stephen and Lieven Vandenberghe (2004). *Convex Optimization*. Cambridge University Press.

Griva, Igor, Stephen G. Nash, and Ariela Sofer (2009). *Linear and Nonlinear Optimization*. Society for Industrial and Applied Mathematics.

Luenberger, David G. and Yinyu Ye (2008). *Linear and Nonlinear Programming*. Springer.

Nocedal, Jorge and Stephen J. Wright (2006). *Numerical Optimization*. Springer Science+ Business Media.