Optimeringslära för F (SF1811) / Optimization (SF1841)

1. Information about the course

2. Examples of optimization problems

3. Introduction to linear programming

Information about the course

- Course in basic optimization theory and linear algebra
- This theory is applied in many areas:
  - Telecommunications, *e.g.* Base station placement
  - Signal processing, *e.g.* Speech recognition
  - Control theory, *e.g.* Trajectory planning
  - Economy, *e.g.* Portfolio optimization
  - Logistics, *e.g.* Production chain management
  - Medicine, *e.g.* Radiation therapy planning

Course Goals

The goal is that after the course you should be able to

- explain basic concepts in optimization, such as variables/objective-functions/constraints and convexity, and use them to formulate relevant optimization problems.
- apply and analyze the algorithms for linear, quadratic and nonlinear programming presented in the course.
- use linear algebra to describe, for example, feasible solutions/directions and determine convexity of functions.
- determine if a given solution can be optimal to a nonlinear program, with or without constraints, and use Newton's method and Lagrange relaxation to solve and analyze when applicable.

Teachers

- Per Enqvist (Email: penqvist@math.kth.se)
- Johan Markdahl (Email: markdahl@math.kth.se)
- Hildur Æsa Oddsdóttir (Email: haodd@mat.kth.se)

Course material

The following material is sold at the student office at the math. department, Lindstedtsvägen 25.

- Optimization, by Amol Sasane and Krister Svanberg.
- Exercises in Optimization

Other recommended literature:
Linear and Nonlinear Programming, second edition,
by Griva, Nash och Sofer.
It is encouraged to buy this book, especially since the same book is also used in the continuation courses
SF2812 Applied linear optimization
SF2822 Applied nonlinear optimization

Other material

- On some lectures I will use this kind of slides.
  When I do, I will post them on the home page after the class.
- They will *not* cover the material in the course.
- Other notes from lectures and exercise session will be posted on the homepage, as is, *i.e.* for your information, but they have not been proof read or define the course in any way.

The course home page

The home page of the course

http://www.math.kth.se/optsyst/grundutbildning/kurser/SF1811/

will be updated actively during the course and you can find

1. a preliminary schedule for the classes
2. old exams
3. reading instructions
4. the home assignments (will be posted soon)

Home assignments

Three voluntary assignments are available in the course.

Homeassignment 1: Linear problem, Matlab exercise

written report + possible oral

(2 bonus points for final exam)

Homeassignment 2: Optimization problem formulation exercise,

poster session "mandatory attendance 13/2"

(1 bonus point for final exam)

Homeassignment 3: Nonlinear problem, Matlab exercise

written report + possible oral

(2 bonus points for final exam)

Final exam

The maximum result at the exam is 50 points + 5 home assignment bonus points.

Preliminary grading limits:

Grade	A	B	C	D	E	FX
Points	43-55	38-42	33-37	28-32	25-27	23-24

- At the exam a sheet with formulas is handed out. No other aids are allowed, no calculators.
- The final examination is scheduled at wednesday March 13, 2013, at the time 8.00-13.00.
- Remember to sign up for the exam on "My Pages".

Course contents

1. Linear Optimization
2. Linear Algebra
3. Quadratic optimization
4. Nonlinear optimization without constraints
5. Nonlinear optimization with constraints

The General optimization problem

$$\left[\begin{array}{ll} \text{max} & f(x) \\ \text{such that} & x \in \mathcal{F}. \end{array}\right] \qquad (1)$$

$f(x)$ is called the objective function and specifies what should be maximized.

The feasible region to (1) is given, *e.g.*, by

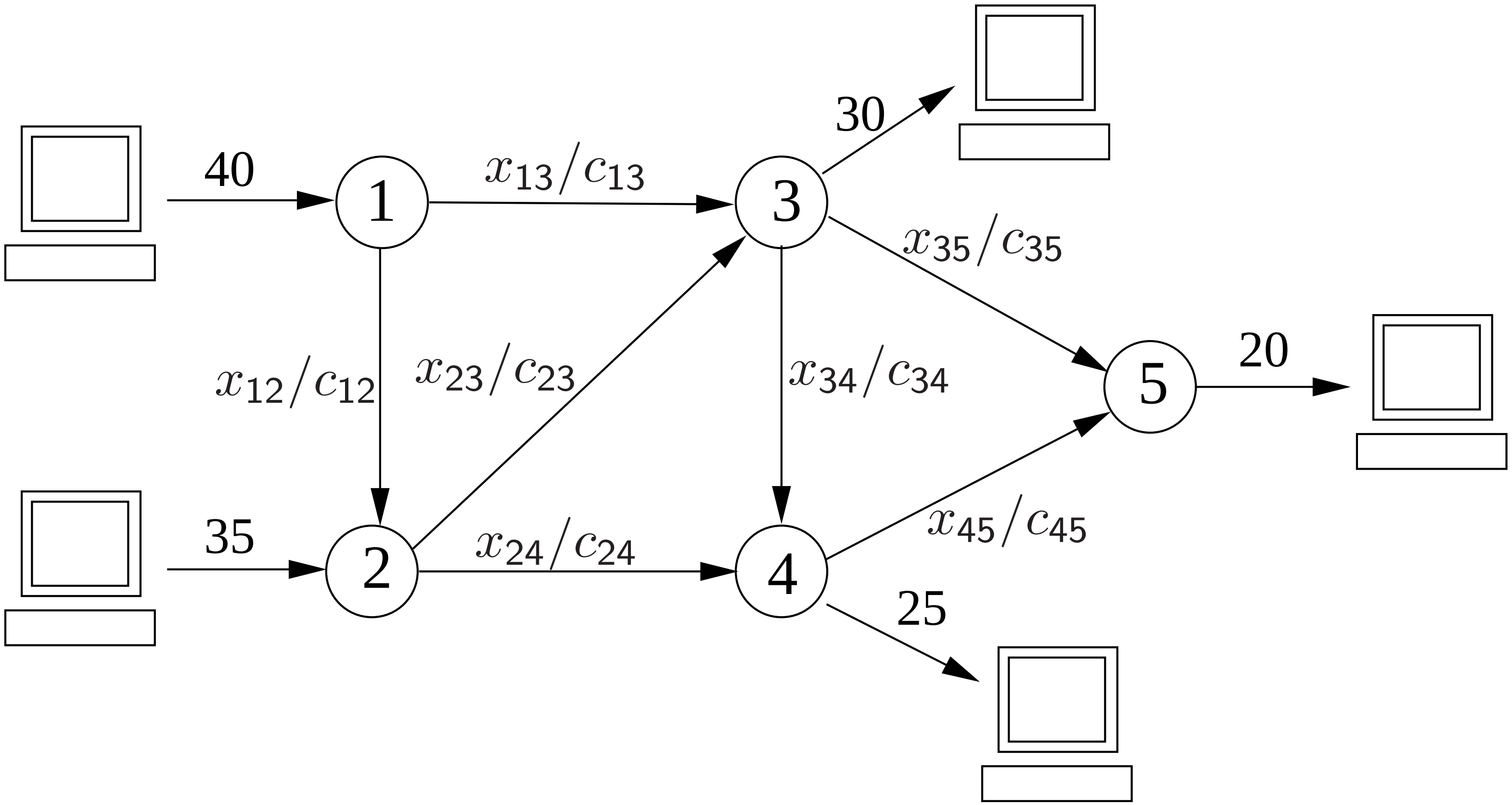
$$\mathcal{F} = \{x : g_1(x) = 0; g_2(x) \geq 0; \; x_1 \geq 0, x_2 \in \{0, 1\}\}$$

- Each point $x \in \mathcal{F}$ is called feasible
- A point $\hat{x} \in \mathcal{F}$ is called optimal if $f(\hat{x}) \geq f(x)$, for all $x \in \mathcal{F}$.

Three examples of optimization problems

1. Network optimization
   - Linear programming problem
2. Linear regression (model fitting)
   - Quadratic optimization problem
3. Traffic control in communication systems
   - Nonlinear optimization problem with constraints

Network optimization



Data is sent from computers at nodes 1 and 2 to computers at nodes 3, 4, 5.

The cost for traffic in the link between node i and j is c_{ij} SEK/Kbyte.

The computer traffic from node i to node j is denoted x_{ij} (Kbyte).

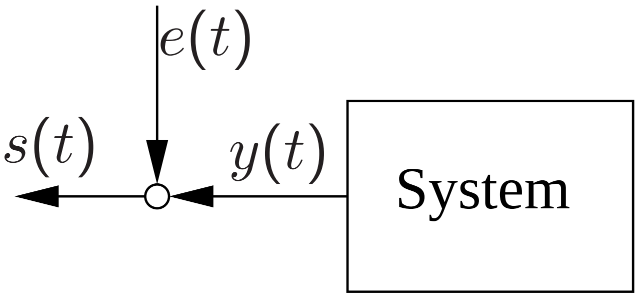
The aim is to minimize the total computer traffic cost.

Network optimization

Resulting optimization problem

$$
\begin{array}{ll}
\text{minimize} & \displaystyle\sum_{\text{all links}} c_{ij} x_{ij} \\
\text{such that} & x_{12} + x_{13} = 40 \\
& -x_{12} + x_{23} + x_{24} = 35 \\
& -x_{13} - x_{23} + x_{34} + x_{35} = -30 \\
& -x_{24} - x_{34} + x_{45} = -25 \\
& -x_{35} - x_{45} = -20 \\
& x_{ij} \geq 0, \text{ all flows in the links}
\end{array}
$$

Linear regression (model fitting)



Problem: Fit a linear regression model to measured data.

$$\text{Regression model}: \quad y(t) = \sum_{j=1}^{n} \alpha_j \psi_j(t)$$

- $\psi_j(t)$, $j = 1, \ldots, n$ are the regressors (known functions)
- $\alpha_j = j = 1, \ldots, n$ are the model parameters (to be determined)
- $e(t)$ measure noise (not known)
- $s(t)$ observations - measured data.

Linear Regression

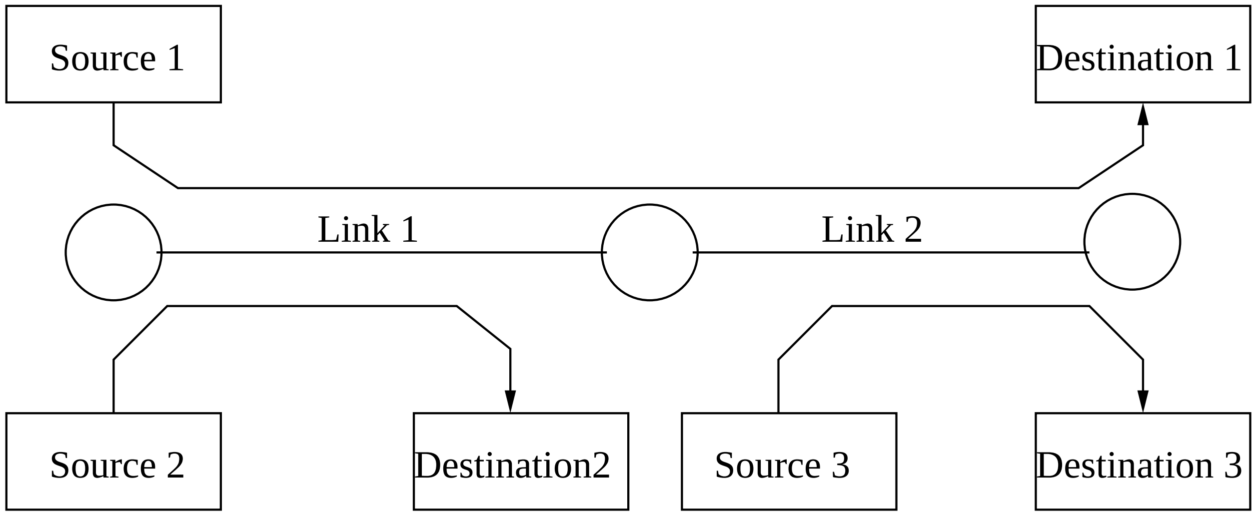
Idéa for model fitting: Minimize the quadratic sum of the prediction errors

$$\min_{\alpha_j} \quad \frac{1}{2}\sum_{i=1}^{m}(\sum_{j=1}^{n}\alpha_j\psi_j(t_i) - s(t_i))^2$$

- This is a least-squares problem.
- Special case of quadratic optimization.

Traffic control in communication systems

Consider a communication network consisting of two links. Three sources sends data over the network to three different destinations.



- Source 1 uses both links.
- Source 2 uses link 1.
- Source 3 uses link 2.

- Link 1 has capacity 2 (normalized entity [amount of data/sek])
- Link 2 has capacity 1
- The three sources sends data with the speeds x_r, $r = 1, 2, 3$.
- The three sources has each the utility function $U_r(x)$, $r = 1, 2, 3$. A common choice of utility function is $U_r(x_r) = w_r \log(x_r)$.

For an efficient and fair distribution of the available capacity, the data speeds are chosen according to the following optimization criterium

$$\begin{aligned} \text{maximize} \quad & U_1(x_1) + U_2(x_2) + U_3(x_3) \\ \text{such that} \quad & x_1 + x_2 \leq 2 \\ & x_1 + x_3 \leq 1 \\ & x_1 \geq 0,\ x_2 \geq 0,\ x_3 \geq 0 \end{aligned}$$

Introduction to linear programming (LP)

1. Formulation of a prototype problem
2. Feasible solution and optimality
3. Graphic illustration of feasibility and optimality
4. Introduction to the simplex method
5. Some highlights from LP theory

Formulation of the prototype problem

- A factory produces two products P_1 och P_2
- The machines M_1 and M_2 both produces products P_1 and P_2
- The production capacity for the machines are a_{ij} [day/unit]:

M_1	$a_{11} = \frac{1}{40}$	$a_{12} = \frac{1}{60}$
M_2	$a_{21} = \frac{1}{50}$	$a_{22} = \frac{1}{50}$
	P_1	P_2

- The price for the products are [SEK/unit]:

product	price
P_1	$c_1 = 200$
P_2	$c_2 = 400$

- AIM: Determine the production volume to maximize the profit.

Let

- x_k number of units P_k produced per day, $k = 1, 2$.

The maximal profit is given by the solution to the following optimization problem:

$$
\begin{aligned}
\text{maximize} \quad & 200x_1 + 400x_2 \\
\text{such that} \quad & \tfrac{1}{40}x_1 + \tfrac{1}{60}x_2 \leq 1 \\
& \tfrac{1}{50}x_1 + \tfrac{1}{50}x_2 \leq 1 \\
& x_k \geq 0,\ k = 1, 2
\end{aligned}
$$

Compact formulation:

$$
\begin{aligned}
\max \quad & c^T x \\
\text{s.t.} \quad & Ax \le b, \\
& x \ge 0
\end{aligned} \tag{2}
$$

where

$$
A = \begin{bmatrix} \frac{1}{40} & \frac{1}{60} \\ \frac{1}{50} & \frac{1}{50} \end{bmatrix}, \qquad b = \begin{bmatrix} 1 \\ 1 \end{bmatrix}, \qquad c = \begin{bmatrix} 200 \\ 400 \end{bmatrix}
$$

The optimization problem (2) is a linear programming problem:

- Linear objective function $c^T x$
- Linear constraints $Ax \le b$ and $x \ge 0$ (componentwise inequalities)

Feasible solution and optimality

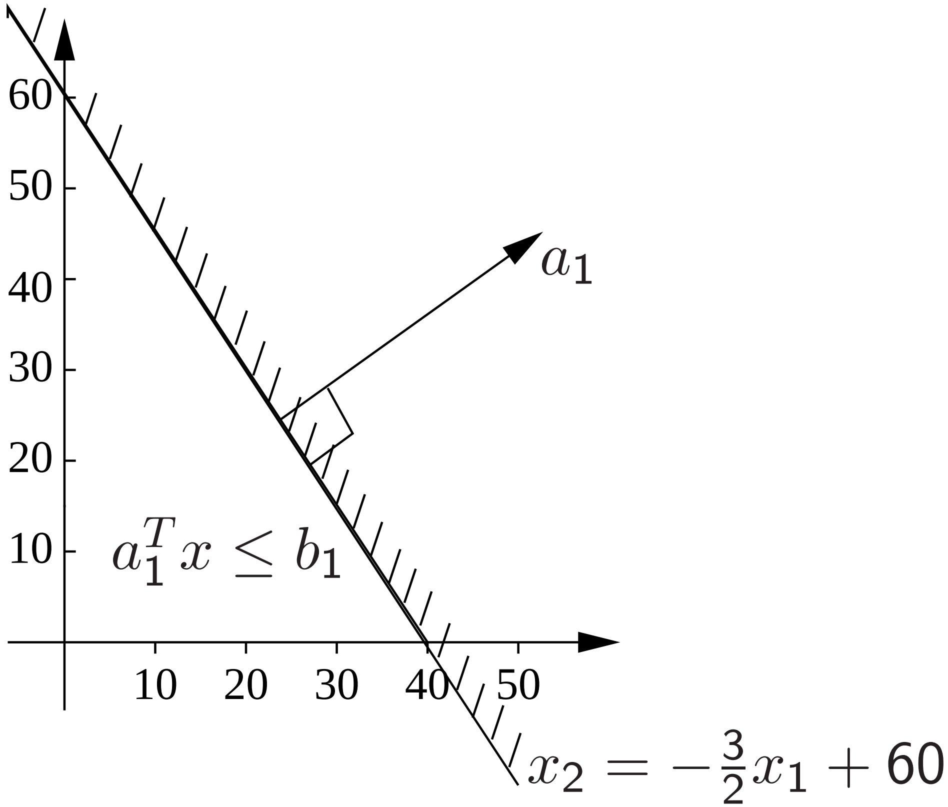
The feasible region to (2) is given by

$$\mathcal{F} = \{x : Ax \leq b;\ x \geq 0\}$$

- Each point $x \in \mathcal{F}$ is called feasible
- A point $\hat{x} \in \mathcal{F}$ is called optimal if $c^T\hat{x} \geq c^T x$, for all $x \in \mathcal{F}$.

Graphical illustration of feasibility and optimality

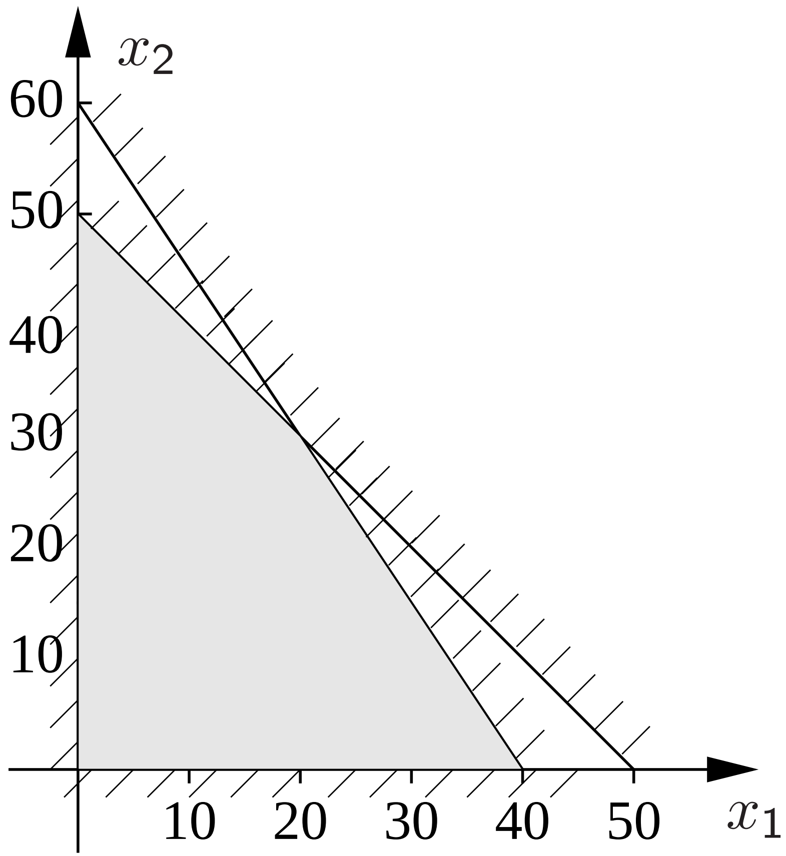
An inequality of the form $a_k^T x \leq b_k$, $k = 1, 2$ corresponds to a halfplane



Here

$$a_1^T = \begin{bmatrix} \frac{1}{40} & \frac{1}{60} \end{bmatrix}, \qquad b_1 = 1$$

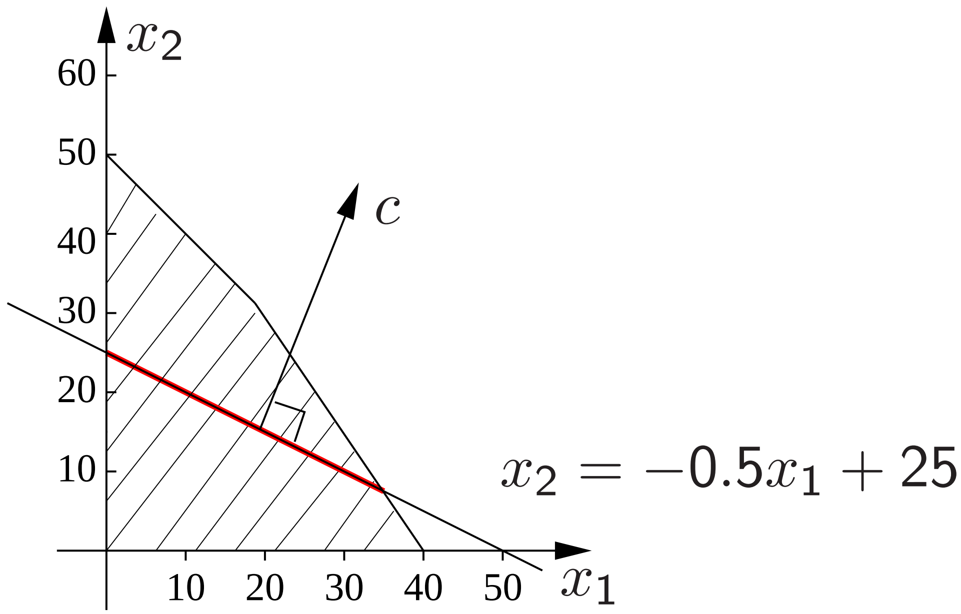
The feasible region for the prototype problem is a polyhedron with four corners

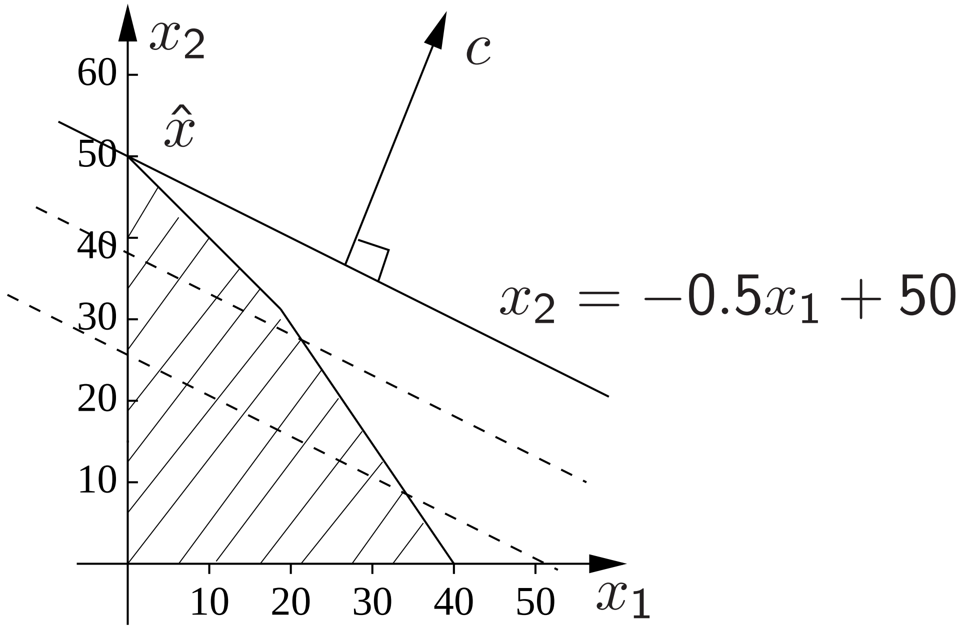


The objective function can be illustrated by iso-cost lines $c^T x = z$.

In the prototype problem it is given by

$$x_2 = -\frac{c_1}{c_2} x_1 + \frac{z}{c_2} = -0.5 x_1 + \frac{z}{400}$$





- We see that the optimum is achieved in the corner $\hat{x} = (0, 50)$
- The optimal value is $\hat{z} = c^T\hat{x} = 0 \cdot 200 + 50 \cdot 400 = 20\,000$.

This is a version of Theorem 4.7 in the book:

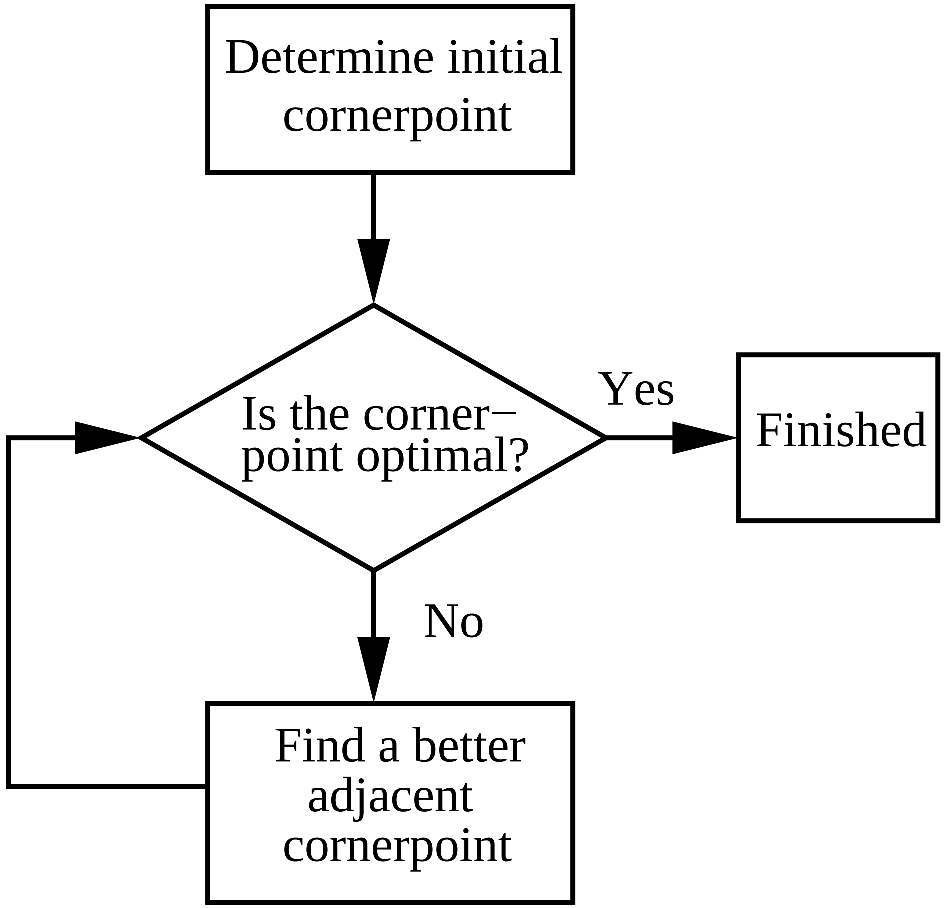
Theorem *If there is an optimal solution to*

$$\begin{aligned} \text{maximize} \quad & c^T x \\ \text{s.t.} \quad & Ax \leq b, \\ & x \geq 0 \end{aligned}$$

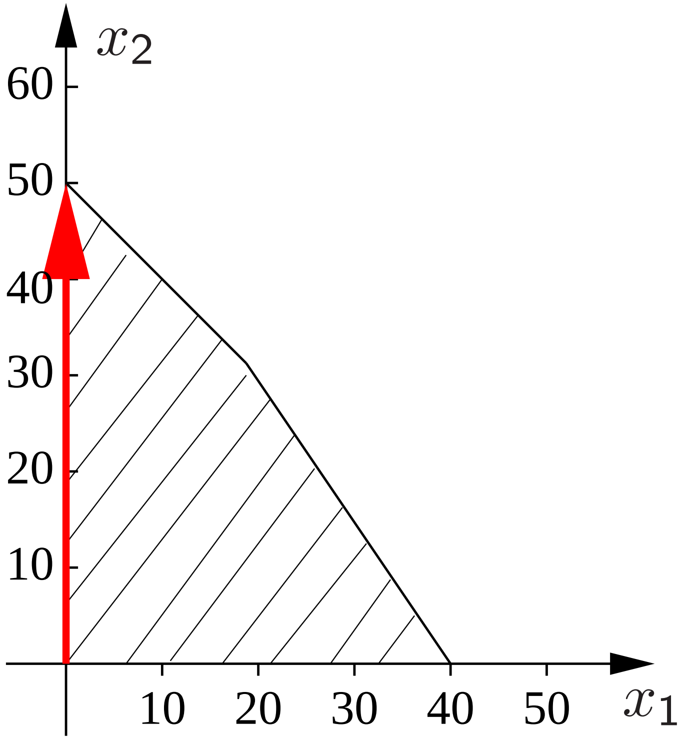
then, there is an optimal corner-solution.

- The Theorem shows that it is enough to look among the corner-points to find an optimal point.
- The concept "corner-point" will later be given a formal algebraic definition.

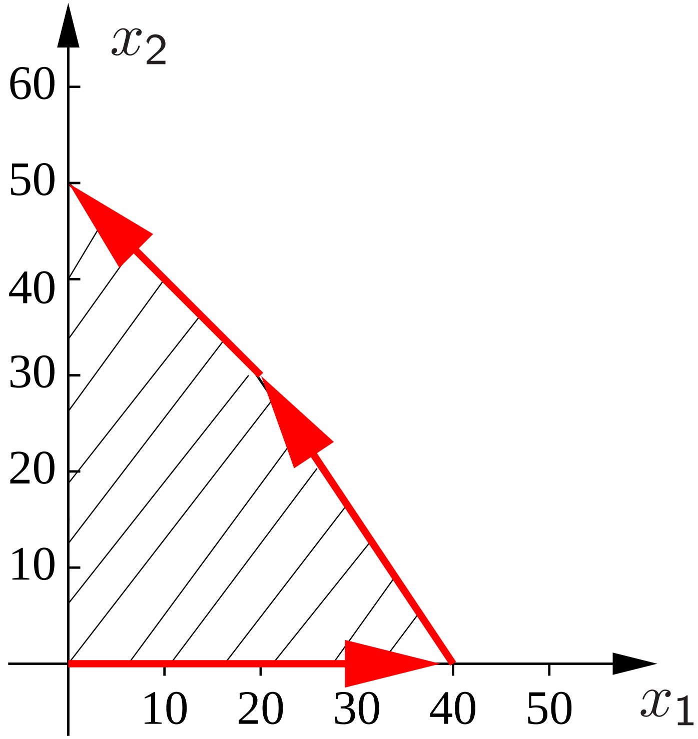
Introduction to the simplex method



The algorithm starts in a corner-point and checks if the objective function decrease when we move along an edge to another corner-point.



Alternative 1

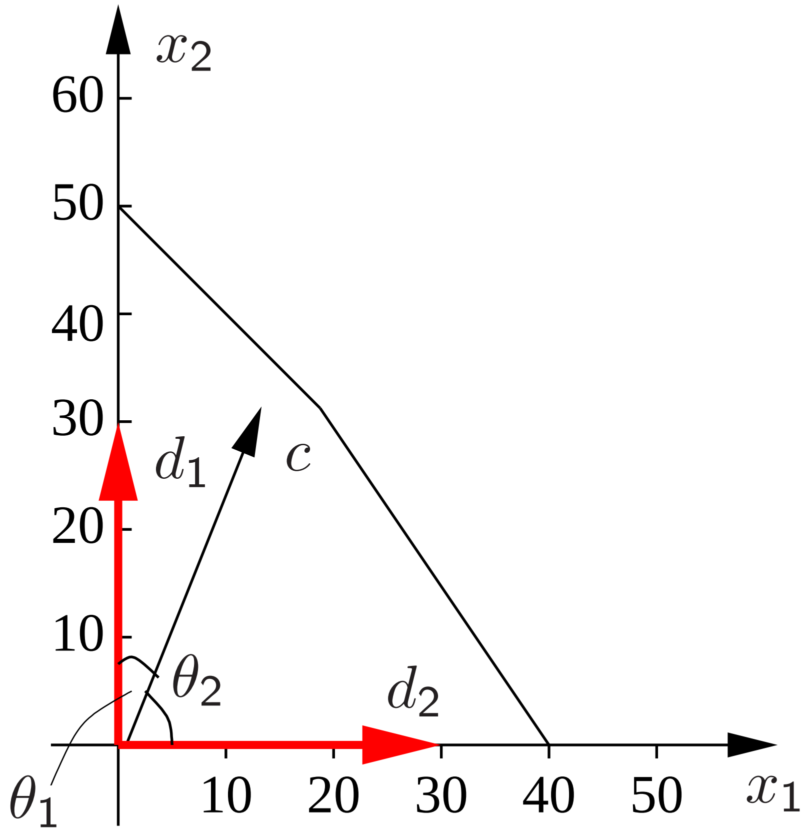


Alternative 2

- How do you choose the shortest way ?

- There exists no efficient algorithm that is guarranteed to find the shortest way to the optimum.
- The Simplex algorithm usually finds a short way to the optimum.
  - It is based on choosing what (locally) seems to be the best way.
- How do you know when you have found the optimum ?

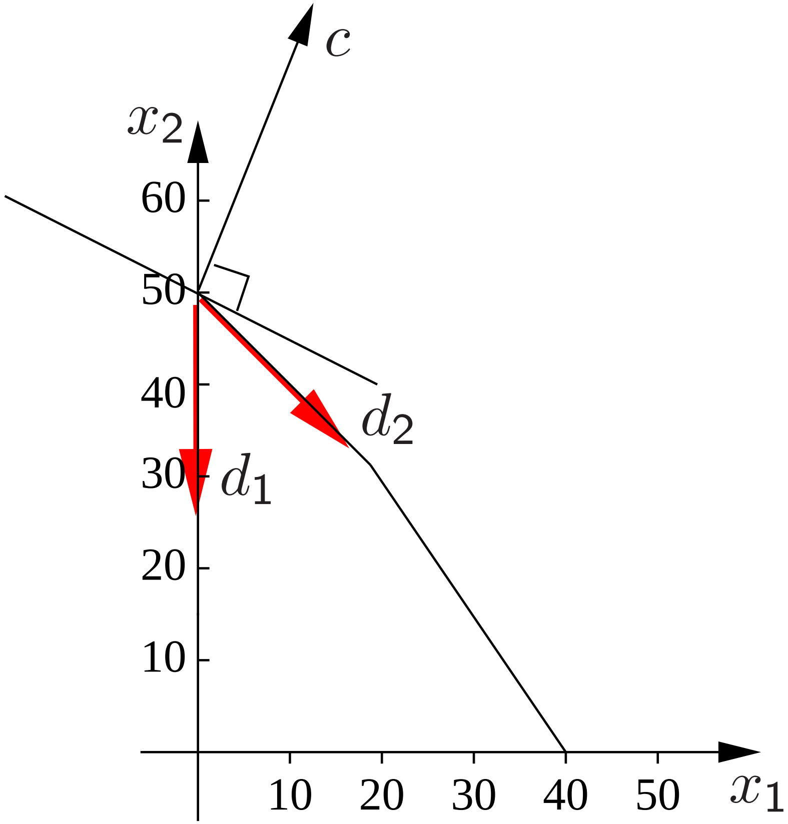
The choice of direction in the Simplex algorithm



It is locally better to follow the edge with unit direction d_1, than to follow the edge with unit direction d_2.

This follows since $c^T d_1 = \|c\| \cdot \underbrace{\|d_1\|}_{=1} \cos(\theta_1) > \|c\| \cdot \underbrace{\|d_2\|}_{=1} \cos(\theta_2) = c^T d_2$.

Stopping criterium



The optimal corner point is reached when

$$c^T d_1 \leq 0$$
$$c^T d_2 \leq 0$$

Highlights in the next lectures

- The Simplex algorithm:
  - The geometrical arguments are good for intuition, but needs to be replaced by linear algebra for numerical implementation
  - Corner-points and edge directions are determined as solutions of linear equation systems.
- Duality:
  - To every linear optimization problem, there exists a dual linear optimization problem
  - The dual provides new insight and can be used to derive optimality conditions.
- Applications: The diet problem, transportation problem, inventory planning, network optimization

Comment:

The prototype example is taken from
"An Introduction to Linear Programming and the Simplex Algorithm"
A www course that can be found at
www.isye.gatech.edu/~spyros/LP/LP.html

LP-problem on standard form

$$\begin{aligned} \text{minimize} \quad & \sum_{j=1}^{n} c_j x_j \\ \text{s.t.} \quad & \sum_{j=1}^{n} a_{ij} x_j = b_i, \; i = 1, \ldots, m \\ & x_j \geq 0, \; j = 1, \ldots, n \end{aligned} \qquad = \qquad \begin{aligned} \text{minimize} \quad & \mathbf{c}^\mathsf{T} \mathbf{x} \\ \text{s.t.} \quad & \mathbf{A}\mathbf{x} = \mathbf{b} \\ & \mathbf{x} \geq 0 \end{aligned}$$

where

$$\mathbf{A} = \begin{bmatrix} a_{11} & \ldots & a_{1n} \\ \vdots & & \vdots \\ a_{m1} & \ldots & a_{mn} \end{bmatrix}, \quad \mathbf{b} = \begin{bmatrix} b_1 \\ \vdots \\ b_m \end{bmatrix}, \quad \mathbf{c} = \begin{bmatrix} c_1 \\ \vdots \\ c_n \end{bmatrix}, \quad \mathbf{x} = \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix}$$

Assumptions

- It is assumed that $\mathbf{A}$ has linearly independent rows.
  - Otherwise the problem can be reduced by removing rows.
- We assume that $n > m$, which means that the linear constraint is underdetermined. If $n \leq m$ there is either
  - a unique solution
  - no solution - hence the optimization problem is not interesting
- All linear optimization problems can be transformed to standard form.
- We will derive a simplex algorithm for LP on standard form.

Rewriting a LP on standard form

"Tricks" for transforming an arbitrary LP-problem into standard form.

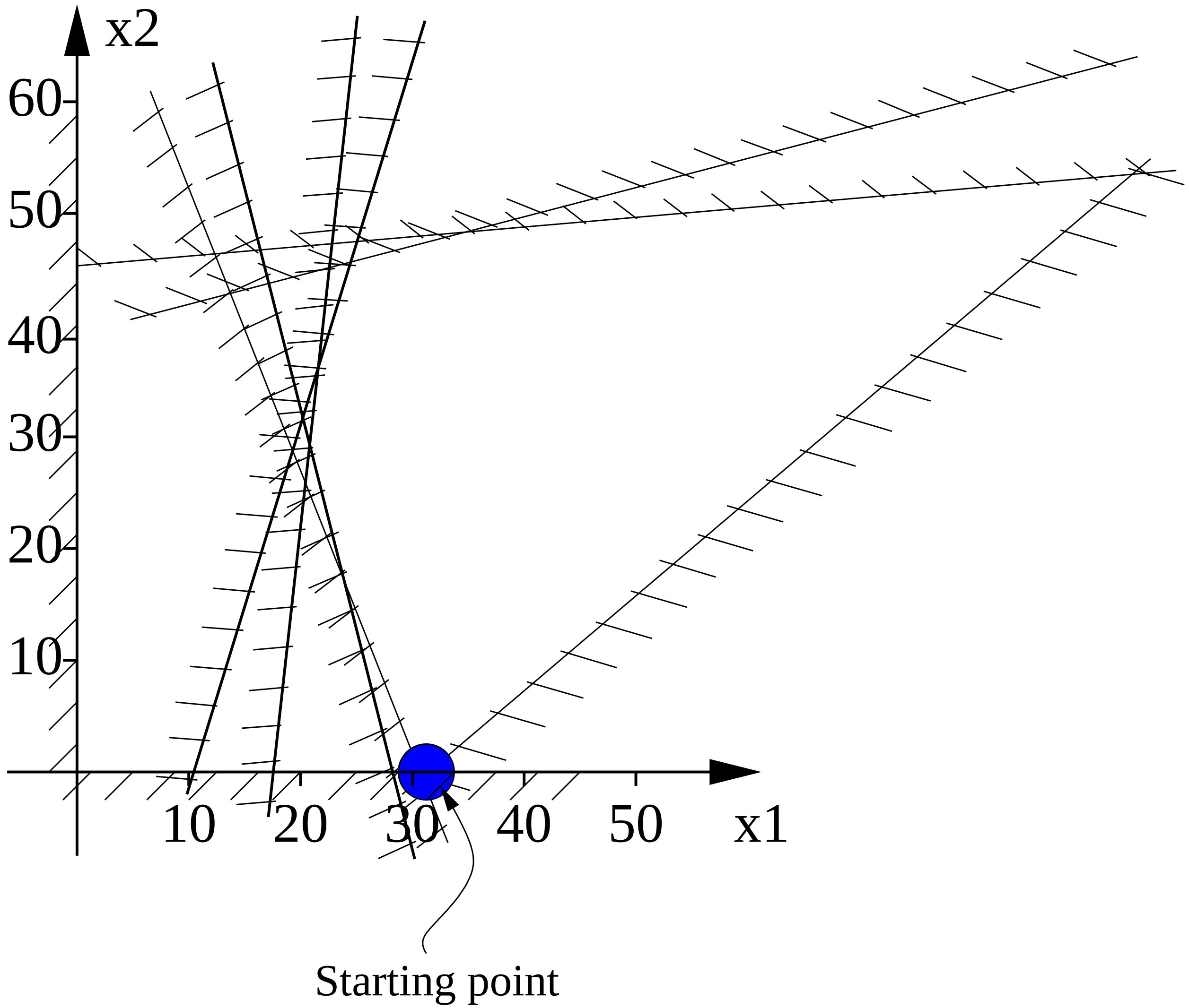
- $\max \mathbf{c}^\mathsf{T}\mathbf{x} = -\min(-\mathbf{c})^\mathsf{T}\mathbf{x}$
- Introduce slack- or surplus-variables if there are inequality constraints
  - $\displaystyle\sum_{j=1}^{n} a_{ij}x_j \leq b_i$ is replaced by $\displaystyle\sum_{j=1}^{n} a_{ij}x_j + x_{n+1} = b_i$,
    where $x_{n+1} \geq 0$ are so called *slack-variables*.
  - $\displaystyle\sum_{j=1}^{n} a_{ij}x_j \geq b_i$ is replaced by $\displaystyle\sum_{j=1}^{n} a_{ij}x_j - x_{n+1} = b_i$,
    where $x_{n+1} \geq 0$ are so called *surplus-variables*.
- A variable x_k with no limit on the sign, can be replaced with the difference $x_k = x'_k - x''_k$, where $x'_k, x''_k \geq 0$.

Some common forms of LP-problems

Form	Primal	Dual
Standard form	minimize $\mathbf{c}^\mathsf{T}\mathbf{x}$ s.t. $\mathbf{A}\mathbf{x} = \mathbf{b}$ $\mathbf{x} \geq 0$	maximize $\mathbf{b}^\mathsf{T}\mathbf{y}$ s.t. $\mathbf{A}^\mathsf{T}\mathbf{y} \leq \mathbf{c}$
Canonical form	minimize $\mathbf{c}^\mathsf{T}\mathbf{x}$ s.t. $\mathbf{A}\mathbf{x} \geq \mathbf{b}$ $\mathbf{x} \geq 0$	maximize $\mathbf{b}^\mathsf{T}\mathbf{y}$ s.t. $\mathbf{A}^\mathsf{T}\mathbf{y} \leq \mathbf{c}$ $\mathbf{y} \geq 0$
Pedagogical form	minimize $\mathbf{c}^\mathsf{T}\mathbf{x}$ s.t. $\mathbf{A}\mathbf{x} \geq \mathbf{b}$	maximize $\mathbf{b}^\mathsf{T}\mathbf{y}$ s.t. $\mathbf{A}^\mathsf{T}\mathbf{y} = \mathbf{c}$ $\mathbf{y} \geq 0$

Reading instructions

- Todays lecture: Chapter 1,2 and 3.
- Next lecture: Chapter 4 and parts of 5.



Given the starting point, if x_2 is maximized, where is the global optimum and which path would the Simplex algorithm choose ?

Standard form

Given the optimization problem

$$\begin{aligned} \text{maximize} \quad & x_1 \\ \text{s.t.} \quad & x_1 \leq x_2, \\ & x_2 + 3x_3 = 2, \\ & x_1 \geq 0, x_3 \geq 0. \end{aligned}$$

Write it on standard form, $\min c^T x$, s.t. $Ax = b$ and $x \geq 0$.

Basic solutions

In LP problems the equation system $Ax = b$ usually has an uncountable number of solutions.

Write $A = [a_1\ a_2\ \cdots\ a_n]$,
where $a_1, \ldots, a_n$ are n column vectors of dimension m $(< n)$.

Choosing m of the columns, $\beta_1, \cdots, \beta_m$, that are linearly independent we get a reduced equation system

$$A_\beta x_\beta = [a_{\beta_1}\ \cdots\ a_{\beta_m}]\,[x_{\beta_1}\ \cdots\ x_{\beta_m}]^T = b$$

which has a unique solution x_β. The x created from x_β with zeros in the positions not corresponding to a β_k is called a *basic solution*.
It will correspond to a corner-point.

If $x \geq 0$ it is called a *feasible basic solution*. More about this next time.