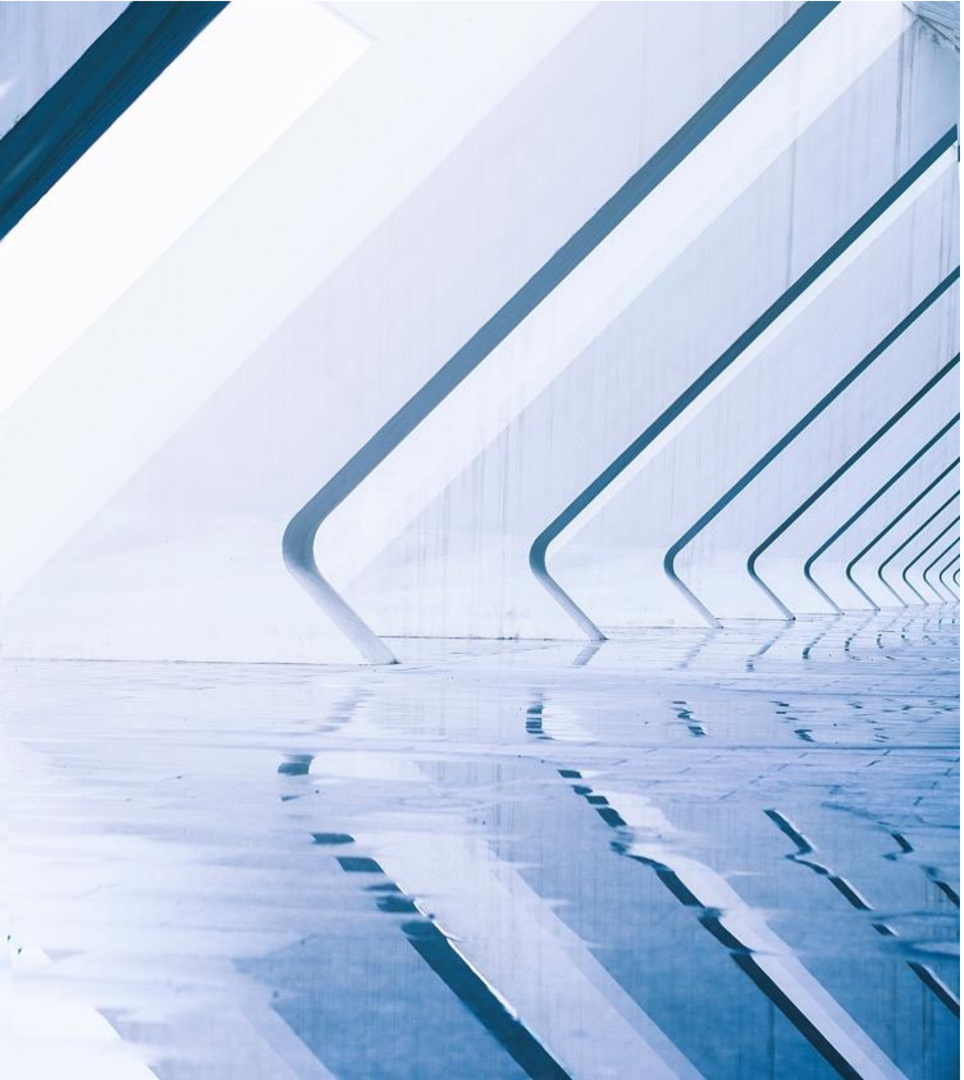


Theory and Basics

LPs and polyhedra

Timo Berthold

Agenda for this lecture

- Fundamentals about Mathematical Optimization
- LP history
- Polyhedral theory
- LP theory

Mathematical Optimization

Optimization Problem

$\rightarrow$ variables $\rightarrow$ solution $\rightarrow$ feasibility
$\rightarrow$ constraints $\rightarrow$ activity $\rightarrow$ feasibility
$\rightarrow$ objective function $\rightarrow$ value $\rightarrow$ quality

$\Rightarrow$ optimal solution $=$ feasible $+$ best possible objective value

Dual bound $\Leftarrow$ **GAP** $\Rightarrow$ **Primal bound**

If the gap is zero, we have a mathematical proof, that the incumbent solution is optimal.

Example: Primal-Dual Gap

- Example: What is the shortest path from
  - Zuse Institute Berlin
    - (where CO@Work should haven taken place)
  - TU Berlin
    - (the university hosting CO@Work)
- Primal Bound: 8.07 km driving route
  - Heuristic solution
  - Addressing a different objective
- Dual Bound: Bee line: 6.74 km
    - https://www.distance.to/
- Gap: $\frac{8.07-6.74}{8.07}$ = 16.48%

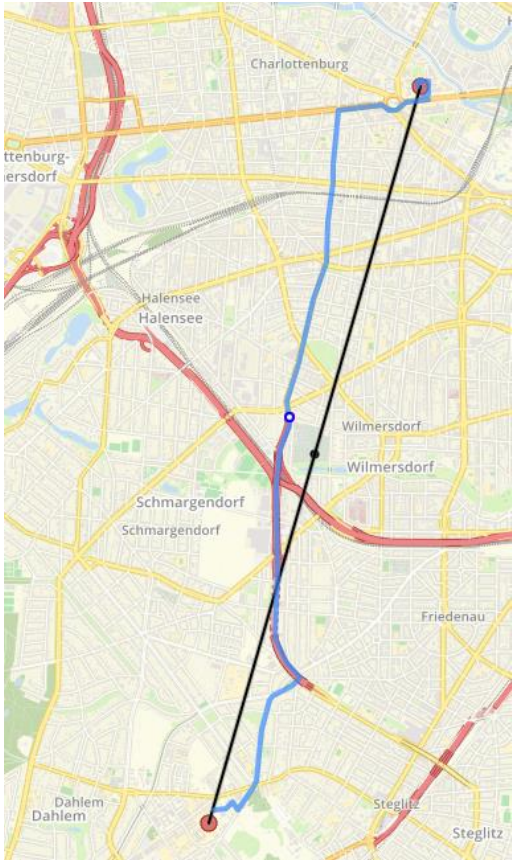


image source: distance.to

Linear Programming

Linear Program

Objective function:

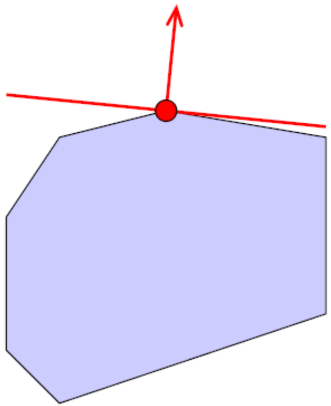
- ▷ linear function

Feasible set:

- ▷ described by linear constraints

Variable domains:

- ▷ real values



$$\begin{aligned} \min \quad & c^T x \\ s.t. \quad & Ax = b \\ & x \in R^n_{\geq 0} \end{aligned}$$

- ▷ convex set
- ▷ “basic” solutions

Various forms, all equivalent

- All representations can be converted into each other:
  - min to max: multiply objective vector c bei -1
  - Equation to inequality: $a_i^T x = b_i \rightarrow a_i^T x \leq b_i, -a_i^T x \leq -b_i$
  - $\leq$-inequality to $\geq$-inequality: multiply by -1
  - Inequality to equation: Introduce slack variable, $a_i^T x \leq b_i \rightarrow a_i^T x + s_i = b_i$
  - Unbounded variable to bounded: $x = x^+ - x^-, x \in \mathbb{R},\ x^+, x^- \in \mathbb{R}_{\geq 0}$
  - Bounded to unbounded: Consider bounds as constraints
- LP literature typically uses the standard form $\max \{c^T x \mid Ax = b, x \geq 0\}$
- MIP literature often uses inequalities for the constraints

Integer Programming

Integer Program

Objective function:

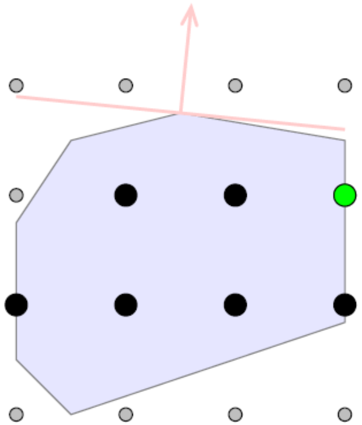
- ▷ linear function

Feasible set:

- ▷ described by linear constraints

Variable domains:

- ▷ integer values



$$\begin{aligned} \min \quad & c^T x \\ s.t. \quad & Ax \leq b \\ & x \in \mathbb{Z}_{\geq 0} \end{aligned}$$

- ▷ not even connected
- ▷ $\mathcal{NP}$-hard problem

Mixed-Integer Programming

Definition: MIP

Objective function:

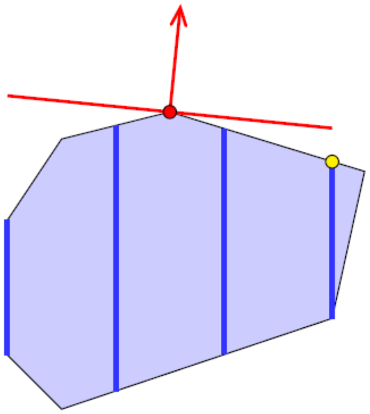
- ▷ linear function

Feasible set:

- ▷ described by linear constraints

Variable domains:

- ▷ integer or real values



$$\begin{aligned} \min \quad & c^T x \\ s.t. \quad & Ax \leq b \\ & x \in \mathbb{Z}^I \times \mathbb{R}^C \end{aligned}$$

- ▷ not even connected
- ▷ $\mathcal{NP}$-hard problem

LP history

LP History: Solving Linear Equation systems

- Egyptians and Babylonians considered about 2000 B.C. the solution of special linear equations. They described examples and did not formulate methods in today's style.
- What we call Gaussian elimination today has been explicitly described in Chinese "Nine Books of Arithmetic" which is a compendium written in the period 2010 B.C. to A.D. 9, but the methods were probably known even before that.
- Gauss, by the way, never described Gaussian elimination. He just used it and stated that the linear equations he used can be solved per eliminationem vulgarem

LP History: The first LP

- In 1827 Fourier described a variable elimination method for linear inequalities, today often called Fourier-Motzkin elimination (Motzkin 1936).
  - By adding one variable and one inequality, Fourier-Motzkin elimination can be turned into an LP solver.
- Who formulated the first LP?
  - The usual credit goes to George J. Stigler (1939)

$\min x1 + x2$	costs
$2x1 + x2 \geq 3$	protein
$x1 + 2x2 \geq 3$	carbohydrates
$x1 \geq 0$	potatoes
$x2 \geq 0$	beans

minimizing the cost of food

- Full example:
  - 77 foods, 9 nutritients
  - Stigler's heuristic solution was 0.7% from optimal

LP History: The first LP algorithms

- 1939 L. V. Kantorovitch (1912-1986): Foundations of linear programming
- 1947 G. B. Dantzig (1914-2005): Invention of the (primal) simplex algorithm
- 1954 C.E. Lemke & E.M.L. Beale: Dual simplex algorithm
- 1953 G.B. Dantzig, 1954 W. Orchard Hays, and 1954 G. B. Dantzig & W. Orchard Hays: Revised simplex algorithm

More on this later…

First computational study in 1953

COMPUTATIONAL EXPERIENCE IN SOLVING LINEAR PROGRAMS*

A. HOFFMAN, M. MANNOS, D. SOKOLOWSKY and N. WIEGMANN

1. Introduction. This paper is a discussion of three methods which have been employed to solve problems in linear programming, and a comparison of results which have been yielded by their use on the Standards Eastern Automatic Computer (SEAC) at the National Bureau of Standards.

LP history: commercial implementations

- The first commercial LP-Code was on the market in 1954 and available on an IBM CPC (card programmable calculator)
  - Record: 71 variables, 26 constraints, 8 h running time
- About 1960: LP became commercially viable, used largely by oil companies
- 1972: first commercial IP solver (almost 50 years ago)

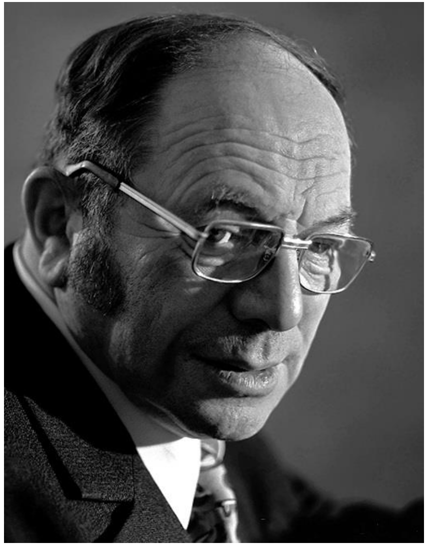
```
NIC 10424                                              Karl Kelley
NWG/RFC 345                                  University of Illinois
                                                      May 26, 1972

    INTEREST IN MIXED INTEGER PROGRAMMING (MPSX ON 360/91 AT CCN)

    MPSX is a newer version of the IBM project MPS, used for integer
programming.  From what I've been told, MPSX outperforms the previous
package.  In addition, it has available a feature of mixed integer
programming.
```

LP history: Nobel Prize Winners

- Leonid V. Kantorovich and Tjalling C. Koopmans received the Nobel Prize for Economics in 1975 for their work on „Optimal use of scarce resources“ – essentially for the foundation and economic interpretation of LP



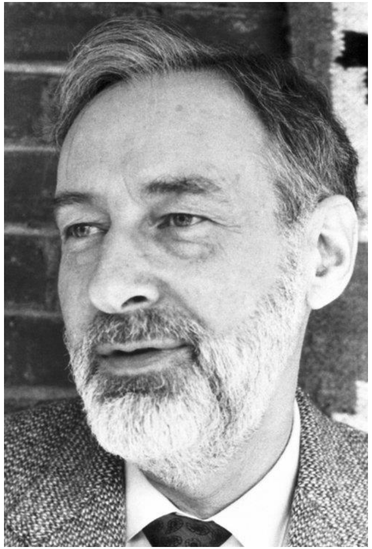


image source: Wikipedia

Quiz time

1. Dantzig‘s simplex algorithm was published in:
   a) 1947
   b) 1954
   c) 1958
2. The set of feasible solutions for a MIP is also called:
   a) a simplex set
   b) a mixed integer set
   c) a linear set
3. Who formulated the diet problem?
   a) George Stigler
   b) Leonid V. Kantorovich
   c) Jean B. J. Fourier

Quiz time

1. Dantzig's simplex algorithm was published in:
   a) 1947
   b) 1954
   c) 1958
2. The set of feasible solutions for a MIP is also called:
   a) a simplex set
   b) a mixed integer set
   c) a linear set
3. Who formulated the diet problem?
   a) George Stigler
   b) Leonid V. Kantorovich
   c) Jean B. J. Fourier

Polyhedral theory

LP & Polyhedra

For nice, interactive visualizations of the 120 regular convex polyhedra, check out: https://polyhedra.tessera.li/

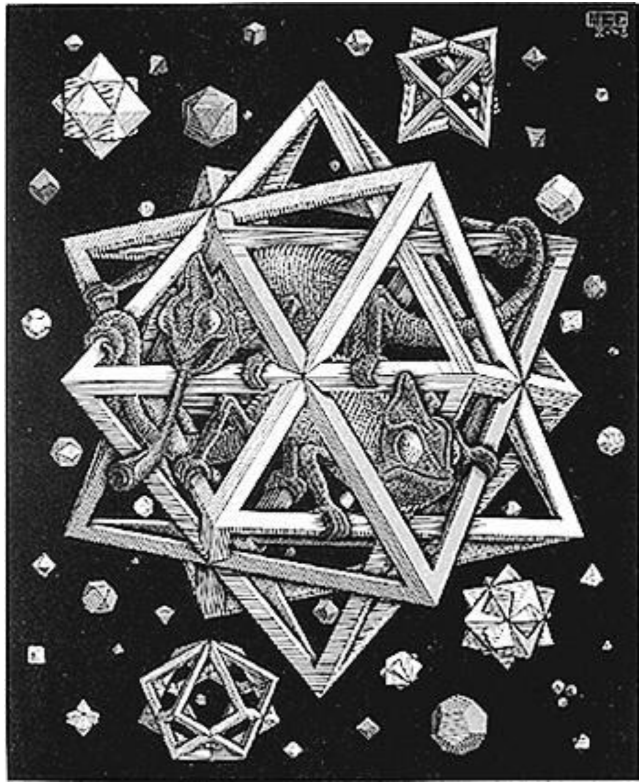


image source: Wikipedia

LP & Polyhedra

- Linear programming is optimizing over an n-dimensional variable vector
- polyhedron: intersection of finitely many halfspaces
  - $P = \{x \in \mathbb{R}^n \mid Ax \leq b\}$

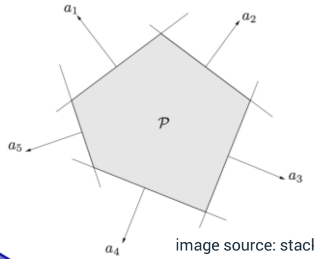


image source: stackexchange.com

- polytope: convex hull of finitely many points
  - $P = conv(V)$, V a finite set in $\mathbb{R}^n$.

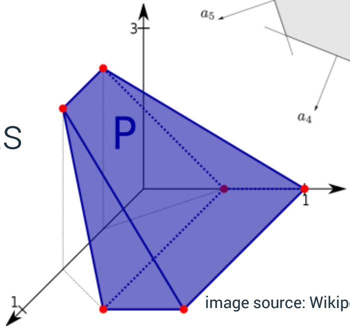


image source: Wikipedia

- convex polyhedral cone: conic combination (i.e., nonnegative linear combination) of finitely many points
  - $K = cone(E)$, E a finite set in $\mathbb{R}^n$.

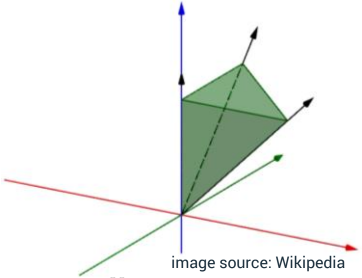
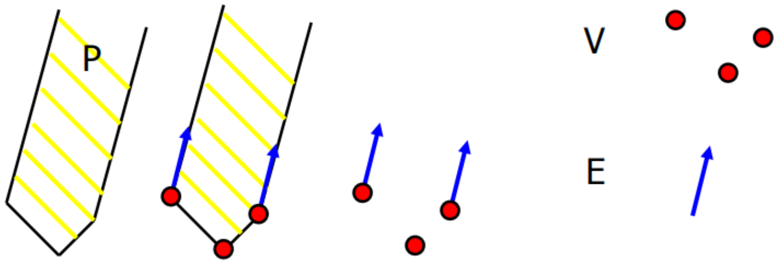


image source: Wikipedia

Representation of polyhedra

- Theorem: For a subset P of $\mathbb{R}^n$ the following are equivalent:
  1. P is a polyhedron.
  2. P is the intersection of finitely many halfspaces, i.e., there exist a matrix A and a vector b with $P = \{x \in \mathbb{R}^n \mid Ax \leq b\}$ (outer representation)
  3. P is the sum of a convex polytope and a finitely generated (polyhedral) cone, i.e., there exist finite sets V and E with $P = conv(V) + cone(E)$ (inner representation)



Important special cases

- $P = conv(0, \{e_i \mid i \in \{1, \dots, n\})$ $P = \{x \mid x \geq 0, \sum_{i=1}^{n} x_i \leq 1\}$
  - Simplex: n+1 points, n+1 inequalities

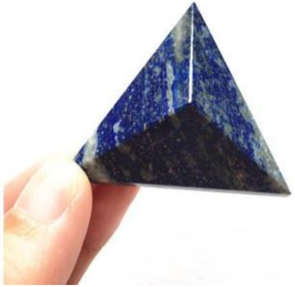


image source: aliexpress.com

- $P = conv(\{-e_i, e_i \mid i \in \{1, \dots, n\})$ $P = \{x \mid a^T x \leq 1, \forall\, a \in \{-1,1\}^n\}$
  - Cross polytope: 2n points, 2^n inequalities

image source: healingcrystals.com

- $P = conv(\{-1,1\}^n)$ $P = \{x \mid a^T x \leq 1, \forall\, a \in \{-e_i, e_i\},\ i \in \{1, \dots, n\}\}$
  - Cube: 2^n points, 2n inequalities

image source: naturshop.cz

Faces of polyhedra

An n-dimensional polyhedron has the following different faces:

- Vertex (0-dimensional)
- Edge (1-dimensional)
- ...
- Ridge = subfacet (n-2)-dimensional
- Facet (n-1)-dimensional

and each of them is a polyhedron itself!

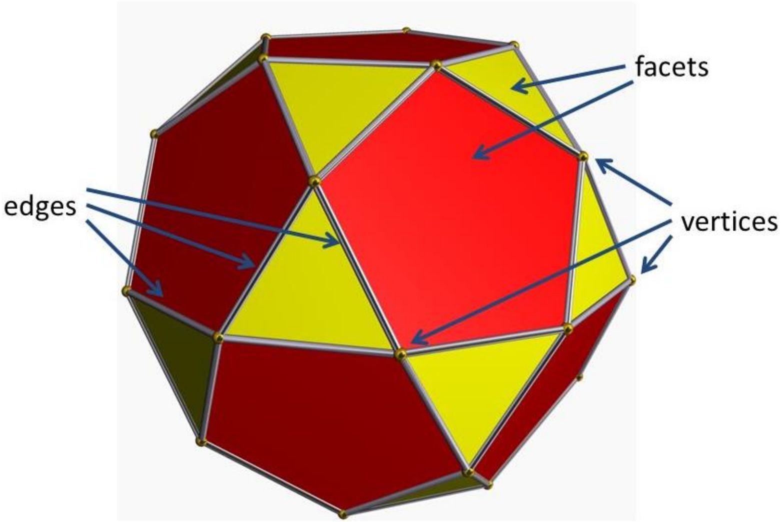


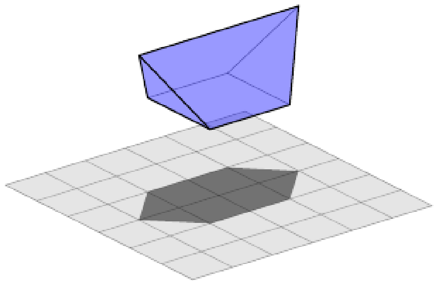
image source: Carmelly Griffith

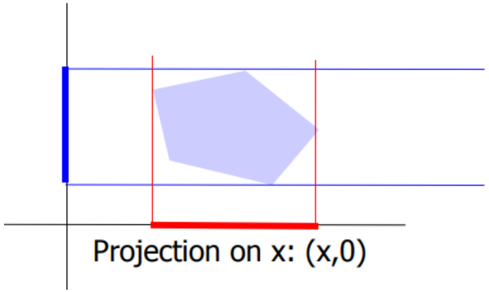
Farkas Lemma

- The Farkas-Lemma (1908):
- A polyhedron defined by an inequality system $Ax \leq b$ is empty, if and only if there is a vector y such that $y \geq 0, y^T A = 0^T, y^T b < 0$
- Hence, we can reformulate $Ax \leq b$ to a wrong statement
- Theorem of Alternatives
- Foundation of all important higher-level LP theory: Duality theorems, complementary slackness, proof of LP optimality, etc.
- Example: $x \leq 1, -x \leq -2 \Rightarrow A = \begin{pmatrix} 1 \\ -1 \end{pmatrix}, \mathrm{b} = \begin{pmatrix} 1 \\ -2 \end{pmatrix}$. Choose $\mathrm{y} = \begin{pmatrix} 1 \\ 1 \end{pmatrix}$.
  - $y^T A = 1 \cdot 1 + 1 \cdot (-1) = 0$ and $y^T b = 1 \cdot 1 + 1 \cdot (-2) = -1 < 0$
  - We added the two inequalities and got a wrong statement

Fourier-Motzkin Elimination

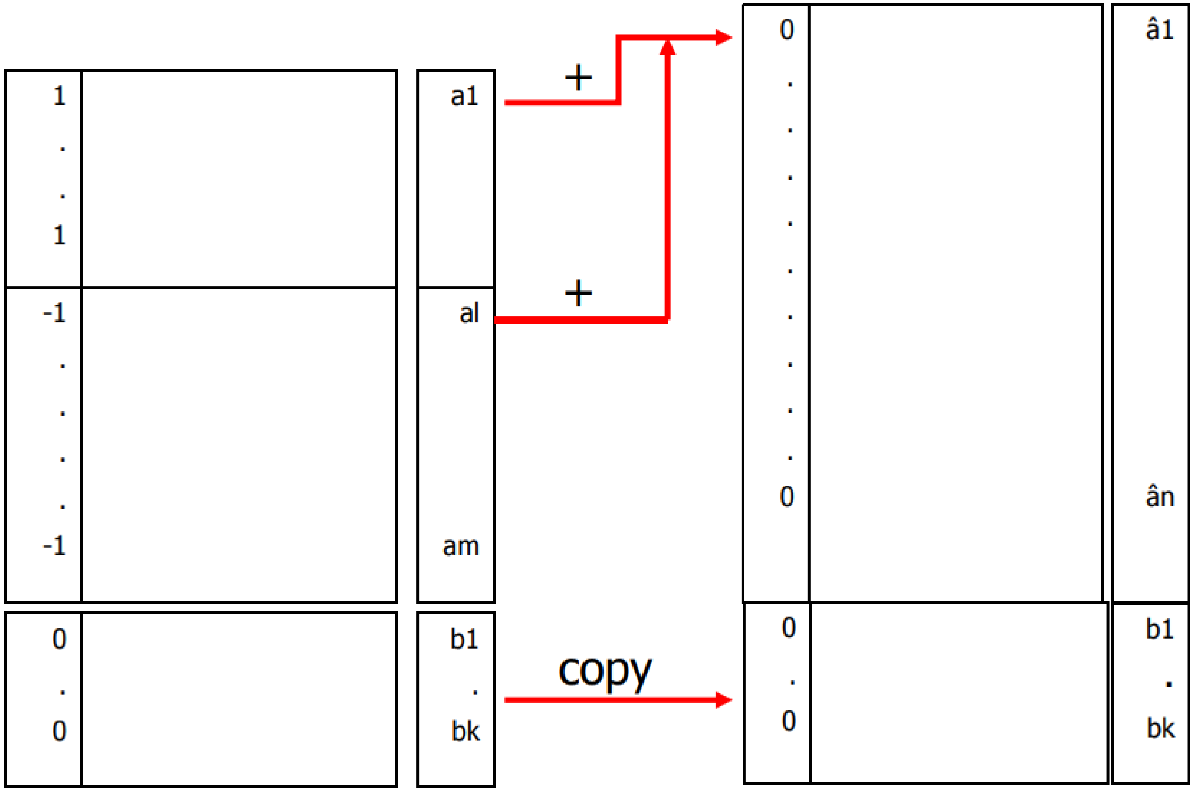
- Fourier, 1827, rediscovered by Motzkin, 1936
- Method: successive projection of a polyhedron in n-dimensional space into a vector space of dimension n-1 by elimination of one variable.





- Can check whether a polyhedron is nonempty
- Can be used to prove Farkas Lemma
- Can be used for Linear Programming

One step of FME



From inner to outer representation

- Each elimination step might square the number of rows, hence in total $O(m^{2^n})$
- Fourier-Motzkin essentially the best-known method for polyhedral transformations:
  - Let a polyhedron be given via: P = conv(V)+cone(E)
  - Goal: Find a representation of P in the form $P = \{x \in \mathbb{R}^n \mid Ax \leq b\}$
  - Idea: Write P as follows : $P = \{x, y, z \in \mathbb{R}^d \mid x = Vy + Ez, \sum y_i = 1, y \geq 0, z \geq 0\}$
  - Eliminate y and z
- With some tricks, FME can be reduced to single-exponential running time
  - which is already best possible for cube/cross polytope

LP theory

The dual LP

$$\min\{c^T x \mid Ax \geq b, x \geq 0\}$$

- How do we get a proof of solution quality?
- Easy case: One of our constraints underestimates the objective function
  - E.g., if $c^T x = 2x_1 + x_2$ and one constraint $x_1 + x_2 \geq 3$, then also $\min c^T x \geq 3$
- Observation 1: Constraints are invariant to scaling by positive numbers
- Observation 2: The sum of two valid constraints is a valid constraint
  - Consequence: Conic combinations of constraints are valid
- Task: Find a conic combination of all constraints that is a maximal underestimator for our objective
- This is an LP. The dual LP

The dual LP

$$\min\{c^T x \mid Ax \geq b, x \geq 0\}$$

- Task: Find a conic combination of all constraints that is a maximal underestimator for our objective

The dual LP

$$\min\{c^T x \mid Ax \geq b, x \geq 0\}$$

- Task: Find a conic combination of all constraints that is a maximal underestimator for our objective
- Conic combination: $y \geq 0$

The dual LP

$$\min\{c^T x \mid Ax \geq b, x \geq 0\}$$

- Task: Find a conic combination of all constraints that is a maximal underestimator for our objective
- Conic combination: $y \geq 0$
- combination of all constraints...: $\mathrm{y}^{\mathrm{T}} Ax \geq y^T \mathrm{b}$

The dual LP

$$\min\{c^T x \mid Ax \geq b, x \geq 0\}$$

- Task: Find a conic combination of all constraints that is a maximal underestimator for our objective
- Conic combination: $y \geq 0$
- combination of all constraints...: $\mathrm{y}^{\mathrm{T}} Ax \geq y^T \mathrm{b}$
- ...that is an (...)underestimator for our objective: $\mathrm{y}^T A \leq c$

The dual LP

$$\min\{c^T x \mid Ax \geq b, x \geq 0\}$$

- Task: Find a conic combination of all constraints that is a maximal underestimator for our objective
- Conic combination: $y \geq 0$
- combination of all constraints...: $\mathrm{y}^{\mathrm{T}} Ax \geq y^T \mathrm{b}$
- ...that is an (...)underestimator for our objective: $\mathrm{y}^T A \leq c$
- ...and that is a maximal underestimator : $\max \ y^T b$

The dual LP

$$\min\{c^T x \mid Ax \geq b, x \geq 0\}$$

- Task: Find a conic combination of all constraints that is a maximal underestimator for our objective
- Conic combination: $y \geq 0$
- combination of all constraints...: $\mathrm{y}^{\mathrm{T}} A x \geq y^T \mathrm{b}$
- ...that is an (...)underestimator for our objective: $\mathrm{y}^T A \leq c$
- ...and that is a maximal underestimator : $\max\ y^T b$
- This is an LP. The dual LP!

$$\max\{y^T b \mid \mathrm{y}^{\mathrm{T}} A \leq c, y \geq 0\}$$

Dual LP: Scheme

primal: $\min\{c^T x \mid Ax \geq b, x \geq 0\}$ and dual: $\max\{y^T b \mid y^T A \leq c, y \geq 0\}$

- Each constraint in the primal LP becomes a variable in the dual LP
- Each variable in the primal LP becomes a constraint in the dual LP
- The objective direction is inversed – maximize in the primal becomes minimize in the dual and vice-versa
- Inequality constraints become ≥-bounds
  - Consequently, equations become free variables

By this scheme: Easy to see that the dual of the dual is the primal

Weak duality theorem

$$\max\{y^T b \mid y^T A \leq c, y \geq 0\} \leq \min\{c^T x \mid Ax \geq b, x \geq 0\}$$

- Trivial to proof:
- $y^T b \leq y^T(Ax) = (y^T A)x \leq c^T x$
- q.e.d.

Weak duality theorem

$$\max\{y^T b |\ y^T A \leq c, y \geq 0\} \leq \min\{c^T x \mid Ax \geq b, x \geq 0\}$$

- Trivial to proof:
- $y^T b \leq y^T(Ax) = (y^T A)x \leq c^T x$
- q.e.d.

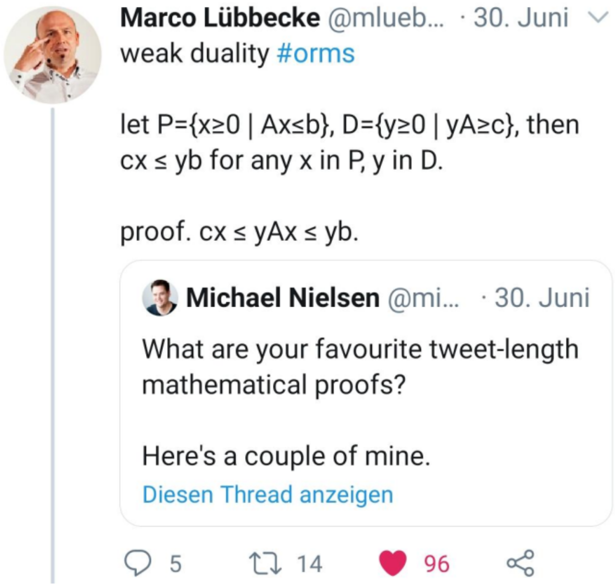


image source: Twitter

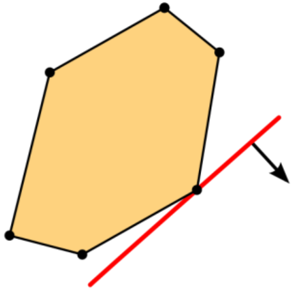
Strong duality theorem

- The most important and influential theorem in Optimization
- Primal has a finite optimum if and only if dual has a finite optimum
- $\min\{c^T x \mid Ax \geq b, x \geq 0\} = \max\{y^T b \mid y^T A \leq c, y \geq 0\}$
- A relation of this type is called min-max result
- Proof is not straight-forward, uses weak duality and Farkas lemma

- Three possibilities: finite (and equal) optima, unbounded and infeasible, both infeasible

Complementary slackness

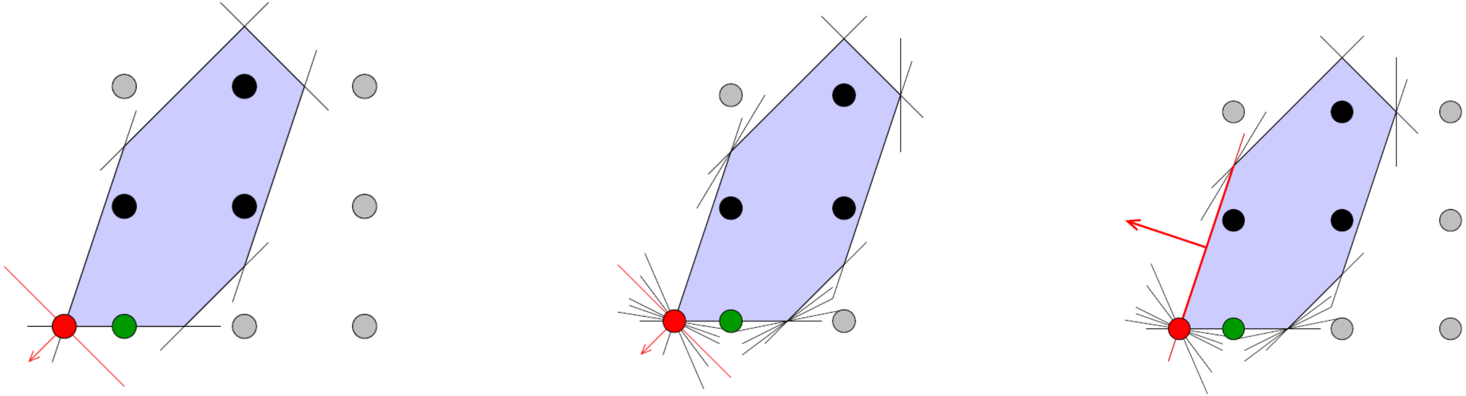
- $\min\{c^T x \mid Ax \geq b, x \geq 0\}$ = $\max\{y^T b| \, y^T A \leq c, y \geq 0\}$
- At an optimal solution pair (x,y), for all i: either $y_i = 0$ or $\sum_{j=1}^{n} a_{ij} \, x_j = b_i$ (or both)
- Analogously, for all j either $x_j = 0$ or $\sum_{j=1}^{n} a_{ij} \, y_i = c_j$
- Proof: By weak duality
- „to construct an optimality proof, we can only use constraints that are tight at the optimal point"



- Interpretation of dual variables as shadow prices: How much would the objective increase, if we relaxed the constraint?
- There is a solution, s.t. $y_i > 0 \Leftrightarrow \sum_{j=1}^{n} a_{ij} \, x_j = b_i$ (strong complementary slackness)

Primal and dual degenaracy

- naïvely: one optimal solution, determined by n constraints
- at a second thought: one optimal solution, k > n tight constraints
- but really: an optimal polyhedron
- Can be exploited in some ways. But mostly a burden

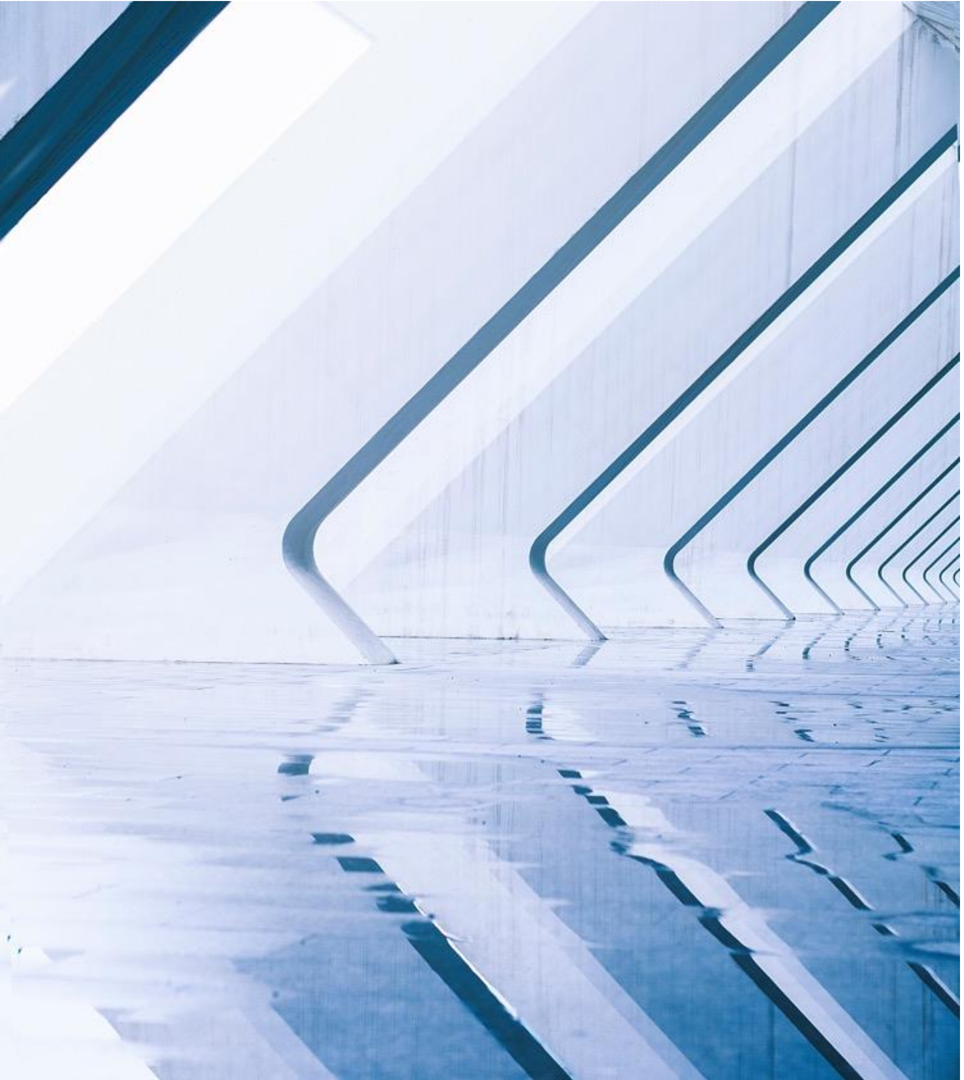


Quiz time

1. If a dual variable has a nonzero value in an optimal primal-dual solution:
   a) the corresponding primal variable is nonzero as well
   b) the corresponding primal constraint is nonzero as well
   c) the corresponding primal constraint is tight
2. Describing a polyhedron as $P = \{x \in \mathbb{R}^n \mid Ax \leq b\}$ is called the
   a) Standard form
   b) Outer representation
   c) Inner representation
3. The dual of $\min\{c^T x \mid Ax \geq b, x \geq 0\}$ is
   a) $\min\{y^T b \mid y^T A = c, y \leq 0\}$
   b) $\max\{y^T b \mid y^T A = c, y \geq 0\}$
   c) $\max\{y^T b \mid y^T A \leq c, y \geq 0\}$

Quiz time

1. If a dual variable has a nonzero value in an optimal primal-dual solution:
   a) the corresponding primal variable is nonzero as well
   b) the corresponding primal constraint is nonzero as well
   c) the corresponding primal constraint is tight
2. Describing a polyhedron as $P = \{x \in \mathbb{R}^n \mid Ax \leq b\}$ is called the
   a) Standard form
   b) Outer representation
   c) Inner representation
3. The dual of $\min\{c^T x \mid Ax \geq b, x \geq 0\}$ is
   a) $\min\{y^T b \mid y^T A = c, y \leq 0\}$
   b) $\max\{y^T b \mid y^T A = c, y \geq 0\}$
   c) $\max\{y^T b \mid y^T A \leq c, y \geq 0\}$



Thank You!

Timo Berthold