# Optimization methods
## NOPT048

Jirka Fink
https://ktiml.mff.cuni.cz/~fink/

Department of Theoretical Computer Science and Mathematical Logic
Faculty of Mathematics and Physics
Charles University in Prague

Summer semester 2015/16
Last change on May 24, 2016

# Content

## Plan of the lecture

- Linear and integer optimization
- Convex sets and Minkowski-Weyl theorem
- Simplex methods
- Duality of linear programming
- Ellipsoid method
- Unimodularity
- Minimal weight maximal matching
- Matroid
- Cut and bound method

## General information

| | |
|---|---|
| E-mail | fink@ktiml.mff.cuni.cz |
| Homepage | http://ktiml.mff.cuni.cz/˜fink/ |
| Consultations | Individual schedule |

## Examination

- Tutorial conditions
  - Tests
  - Theretical homeworks
  - Practical homeworks
- Pass the exam

## Literature


- A. Schrijver, Theory of linear and integer programming, John Wiley, 1986
- W. J .Cook, W. H. Cunningham, W. R. Pulleyblank, A. Schrijver, Combinatorial Optimization, John Wiley, 1997
- J. Matoušek, B. Gärtner, Understanding and using linear programming, Springer, 2006.
- J. Matoušek Introduction to Discrete Geometry. ITI Series 2003-150, MFF UK, 2003

# Outline

# Optimization

## Mathematical optimization

is the selection of a best element (with regard to some criteria) from some set of available alternatives.

## Examples

- Minimize $x^2 + y^2$ where $(x, y) \in \mathbb{R}^2$
- Maximal matching in a graph
- Minimal spanning tree
- Shortest path between given two vertices

## Optimization problem

Given a set of solutions $M$ and an objective function $f : M \to \mathbb{R}$, optimization problem is finding a solution $x \in M$ with the maximal (or minimal) objective value $f(x)$ among all solutions of $M$.

## Duality between minimization and maximization

If $\min_{x \in M} f(x)$ exists, then also $\max_{x \in M} -f(x)$ exists and $-\min_{x \in M} f(x) = \max_{x \in M} -f(x)$.

# Notation: Vector and matrix

## Matrix

A matrix of type $m \times n$ is a rectangular array of $m$ rows and $n$ columns of real numbers. Matrices are written as $A$, $B$, $C$, etc.

## Vector

A vector is an $n$-tuple of real numbers. Vectors are written as $\boldsymbol{c}$, $\boldsymbol{x}$, $\boldsymbol{y}$, etc. Usually, vectors are column matrices of type $n \times 1$.

## Scalar

A scalar is a real number. Scalars are written as $a$, $b$, $c$, etc.

## Special vectors

$\mathbf{0}$ and $\mathbf{1}$ are vectors of zeros and ones, respectively.

## Transpose

The transpose of a matrix $A$ is matrix $A^{\mathrm{T}}$ created by reflecting $A$ over its main diagonal. The transpose of a column vector $\boldsymbol{x}$ is the row vector $\boldsymbol{x}^{\mathrm{T}}$.

# Notation: Matrix product

## Elements of a vector and a matrix

- The $i$-th element of a vector $\boldsymbol{x}$ is denoted by $\boldsymbol{x}_i$.
- The $(i, j)$-th element of a matrix $A$ is denoted by $A_{i,j}$.
- The $i$-th row of a matrix $A$ is denoted by $A_{i,\star}$.
- The $j$-th column of a matrix $A$ is denoted by $A_{\star,j}$.

## Dot product of vectors

The dot product (also called inner product or scalar product) of vectors $\boldsymbol{x}, \boldsymbol{y} \in \mathbb{R}^n$ is the scalar $\boldsymbol{x}^{\mathrm{T}}\boldsymbol{y} = \sum_{i=1}^{n} \boldsymbol{x}_i \boldsymbol{y}_i$.

## Product of a matrix and a vector

The product $A\boldsymbol{x}$ of a matrix $A \in \mathbb{R}^{m \times n}$ of type $m \times n$ and a vector $\boldsymbol{x} \in \mathbb{R}^n$ is a vector $\boldsymbol{y} \in \mathbb{R}^m$ such that $\boldsymbol{y}_i = A_{i,\star}\boldsymbol{x}$ for all $i = 1, \dots, m$.

## Product of two matrices

The product $AB$ of a matrix $A \in \mathbb{R}^{m \times n}$ and a matrix $B \in \mathbb{R}^{n \times k}$ a matrix $C \in \mathbb{R}^{m \times k}$ such that $C_{\star,j} = AB_{\star,j}$ for all $j = 1, \dots, k$.

# Notation: System of linear equations and inequalities

## Equality and inequality of two vectors

For vectors $\boldsymbol{x}, \boldsymbol{y} \in \mathbb{R}^n$ we denote

- $\boldsymbol{x} = \boldsymbol{y}$ if $\boldsymbol{x}_i = \boldsymbol{y}_i$ for every $i = 1, \ldots, n$ and
- $\boldsymbol{x} \leq \boldsymbol{y}$ if $\boldsymbol{x}_i \leq \boldsymbol{y}_i$ for every $i = 1, \ldots, n$.

## System of linear equations

Given a matrix $A \in \mathbb{R}^{m \times n}$ of type $m \times n$ and a vector $\boldsymbol{b} \in \mathbb{R}^m$, the formula $A\boldsymbol{x} = \boldsymbol{b}$ means a system of $m$ linear equations where $\boldsymbol{x}$ is a vector of $n$ real variables.

## System of linear inequalities

Given a matrix $A \in \mathbb{R}^{m \times n}$ of type and a vector $\boldsymbol{b} \in \mathbb{R}^m$, the formula $A\boldsymbol{x} \leq \boldsymbol{b}$ means a system of $m$ linear inequalities where $\boldsymbol{x}$ is a vector of $n$ real variables.

## Example: System of linear inequalities in two different notations

$$\begin{array}{rcrcrcr} 2\boldsymbol{x}_1 & + & \boldsymbol{x}_2 & + & \boldsymbol{x}_3 & \leq & 14 \\ 2\boldsymbol{x}_1 & + & 5\boldsymbol{x}_2 & + & 5\boldsymbol{x}_3 & \leq & 30 \end{array}$$

$$\begin{pmatrix} 2 & 1 & 1 \\ 2 & 5 & 5 \end{pmatrix} \begin{pmatrix} \boldsymbol{x}_1 \\ \boldsymbol{x}_2 \\ \boldsymbol{x}_3 \end{pmatrix} \leq \begin{pmatrix} 14 \\ 30 \end{pmatrix}$$

# Example of linear programming: Optimized diet

## Express using linear programming the following problem

Find the cheapest vegetable salad from carrots, white cabbage and cucumbers containing required amount the vitamins A and C and dietary fiber.

| Food | Carrot | White Cabbage | Cucumber | Required per meal |
|---|---|---|---|---|
| Vitamin A [mg/kg] | 35 | 0.5 | 0.5 | 0.5 mg |
| Vitamin C [mg/kg] | 60 | 300 | 10 | 15 mg |
| Dietary Fiber [g/kg] | 30 | 20 | 10 | 4 g |
| Price [EUR/kg] | 0.75 | 0.5 | 0.15 | |

## Formulation using linear programming

Let $\boldsymbol{x}_1$, $\boldsymbol{x}_2$ and $\boldsymbol{x}_3$ be real variables denoting the amount of carrots, white cabbage and cucumbers, respectively. The linear programming problem is

$$\begin{array}{lrcrcrcl}
\text{Minimize} & 0.75\boldsymbol{x}_1 & + & 0.5\boldsymbol{x}_2 & + & 0.15\boldsymbol{x}_3 & & \\
\text{subject to} & 35\boldsymbol{x}_1 & + & 0.5\boldsymbol{x}_2 & + & 0.5\boldsymbol{x}_3 & \geq & 0.5 \\
 & 60\boldsymbol{x}_1 & + & 300\boldsymbol{x}_2 & + & 10\boldsymbol{x}_3 & \geq & 15 \\
 & 30\boldsymbol{x}_1 & + & 20\boldsymbol{x}_2 & + & 10\boldsymbol{x}_3 & \geq & 4 \\
 & & & & & \boldsymbol{x}_1, \boldsymbol{x}_2, \boldsymbol{x}_3 & \geq & 0
\end{array}$$

# Example of linear programming: Network flow

## Network flow problem

Given direct graph $(V, E)$ with capacities $\boldsymbol{c} \in \mathbb{R}^E$ and a source $s \in V$ and a sink $t \in V$, find the maximal flow from $s$ to $t$ satisfying the flow conservation and capacity constrains.

## Formulation using linear programming

Variables: flow $\boldsymbol{f}_e$ for every edge $e \in E$

Capacity constrains: $\mathbf{0} \leq \boldsymbol{f} \leq \boldsymbol{c}$

Flow conservation: $\sum_{uv \in E} \boldsymbol{f}_{uv} = \sum_{vw \in E} \boldsymbol{f}_{vw}$ for every $v \in V \setminus \{s, t\}$

Objective function: Maximize $\sum_{sw \in E} \boldsymbol{f}_{sw} - \sum_{us \in E} \boldsymbol{f}_{us}$

# Example of integer linear programming: Vertex cover

## Vertex cover problem

Given undirected graph $(V, E)$, find the smallest set of vertices $U \subseteq V$ covering every edge of $E$; that is, $U \cup e \neq \emptyset$ for every $e \in E$.

## Formulation using integer linear programming

Variables: cover $\boldsymbol{x}_v \in \{0, 1\}$ for every vertex $v \in V$

Covering: $\boldsymbol{x}_u + \boldsymbol{x}_v \geq 1$ for every edge $uv \in E$

Objective function: Minimize $\mathbf{1}^{\mathrm{T}}\boldsymbol{x}$

## Canonical form

Linear programming problem in the canonical form is an optimization problem to find $\boldsymbol{x} \in \mathbb{R}^n$ which maximizes $\boldsymbol{c}^{\mathrm{T}}\boldsymbol{x}$ and satisfies $A\boldsymbol{x} \leq \boldsymbol{b}$ where $A \in \mathbb{R}^{m \times n}$ and $\boldsymbol{b} \in \mathbb{R}^m$.

## Equation form

Linear programming problem in the equation form is a problem to find $\boldsymbol{x} \in \mathbb{R}^n$ which maximizes $\boldsymbol{c}^{\mathrm{T}}\boldsymbol{x}$ and satisfies $A\boldsymbol{x} = \boldsymbol{b}$ and $\boldsymbol{x} \geq \boldsymbol{0}$ where $A \in \mathbb{R}^{m \times n}$ and $\boldsymbol{b} \in \mathbb{R}^m$.

## Conversions

- Every $\boldsymbol{x} \in \mathbb{R}^n$ satisfies $A\boldsymbol{x} = \boldsymbol{b}$ if and only if it satisfies $A\boldsymbol{x} \geq \boldsymbol{b}$ and $A\boldsymbol{x} \leq \boldsymbol{b}$.
- Every $\boldsymbol{x} \in \mathbb{R}^n$ satisfies $A\boldsymbol{x} \leq \boldsymbol{b}$ if and only if there exists $\boldsymbol{z} \in \mathbb{R}^m$ satisfying $A\boldsymbol{x} + \boldsymbol{z} = \boldsymbol{b}$ and $\boldsymbol{z} \geq \boldsymbol{0}$.
- Every occurrence of a variable $\boldsymbol{x}$ can be replaced by $\boldsymbol{x}^+ - \boldsymbol{x}^-$ when contains $\boldsymbol{x}^+, \boldsymbol{x}^- \geq 0$ are added.

## Example: Conversion from the canonical form into the equation form

- $\max \boldsymbol{c}^{\mathrm{T}}\boldsymbol{x}$ such that $A\boldsymbol{x} \leq \boldsymbol{b}$
- $\max \boldsymbol{c}^{\mathrm{T}}\boldsymbol{x}$ such that $A\boldsymbol{x} + \boldsymbol{z} = \boldsymbol{b}$ and $\boldsymbol{z} \geq 0$
- $\max \boldsymbol{c}^{\mathrm{T}}\boldsymbol{x}^+ - \boldsymbol{c}^{\mathrm{T}}\boldsymbol{x}^-$ such that $A\boldsymbol{x}^+ - A\boldsymbol{x}^- + \boldsymbol{z} = \boldsymbol{b}$ and $\boldsymbol{z}, \boldsymbol{x}^+, \boldsymbol{x}^- \geq 0$

# Related problems

## Integer linear programming

Integer linear programming problem is an optimization problem to find $\boldsymbol{x} \in \mathbb{Z}^n$ which maximizes $\boldsymbol{c}^{\mathrm{T}}\boldsymbol{x}$ and satisfies $A\boldsymbol{x} \leq \boldsymbol{b}$ where $A \in \mathbb{R}^{m \times n}$ and $\boldsymbol{b} \in \mathbb{R}^m$.

## Mix integer linear programming

Some variables are integer and others are real.

## Binary linear programming

Every variable is either 0 or 1. (1)

## Complexity

- A linear programming problem is efficiently solvable, both in theory and in practice.
- The classical algorithm for linear programming is the *Simplex method* which is fast in practice but it is not known whether it always run in polynomial time.
- Polynomial time algorithms the *ellipsoid* and the *interior point* methods.
- No strongly polynomial-time algorithms for linear programming is known.
- Integer linear programming is NP-hard.

1. Show that binary linear programming is a special case of integer linear programming.

## Basic terminology

- Number of variables: $n$
- Number of constrains: $m$
- Solution: $\boldsymbol{x}$
- Objective function: e.g. $\max \boldsymbol{c}^{\mathrm{T}}\boldsymbol{x}$
- Feasible solution: a solution satisfying all constrains, e.g. $A\boldsymbol{x} \geq \boldsymbol{b}$
- Optimal solution: a feasible solution maximizing $\boldsymbol{c}^{\mathrm{T}}\boldsymbol{x}$
- Infeasible problem: a problem having no feasible solution
- Unbounded problem: a problem having a feasible solution with arbitrary large value of given objective function
- Polyhedron: a set of points $\boldsymbol{x} \in \mathbb{R}^n$ satisfying $A\boldsymbol{x} \geq \boldsymbol{b}$ for some $A \in \mathbb{R}^{m \times n}$ and $\boldsymbol{b} \in \mathbb{R}^m$
- Polytope: a bounded polyhedron

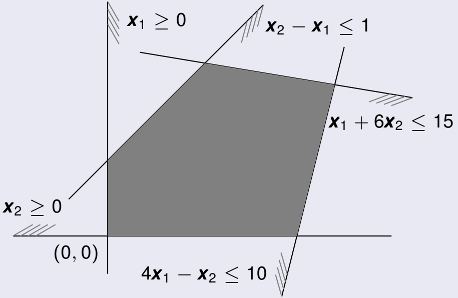
# Graphical method: Set of feasible solutions

## Example

Draw the set of all feasible solutions $(x_1, x_2)$ satisfying the following conditions.

$$\begin{array}{rcrcl} x_1 & + & 6x_2 & \leq & 15 \\ 4x_1 & - & x_2 & \leq & 10 \\ -x_1 & + & x_2 & \leq & 1 \\ & & x_1, x_2 & \geq & 0 \end{array}$$

## Solution



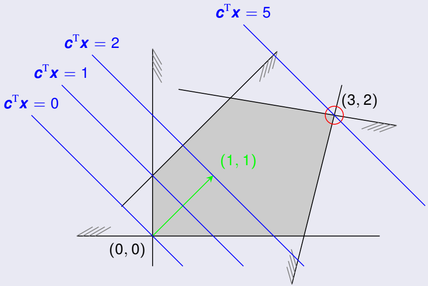
# Graphical method: Optimal solution

## Example

Find the optimal solution of the following problem.

$$
\begin{array}{lrcrcr}
\text{Maximize} & x_1 & + & x_2 & & \\
 & x_1 & + & 6x_2 & \leq & 15 \\
 & 4x_1 & - & x_2 & \leq & 10 \\
 & -x_1 & + & x_2 & \leq & 1 \\
 & & & x_1, x_2 & \geq & 0
\end{array}
$$

## Solution



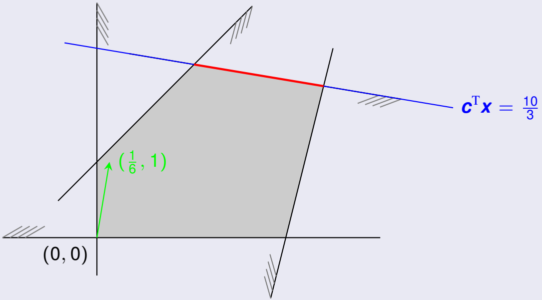
# Graphical method: Multiple optimal solutions

## Example

Find all optimal solutions of the following problem.

$$\begin{array}{lrcrcr} \text{Maximize} & \frac{1}{6}x_1 & + & x_2 & & \\ & x_1 & + & 6x_2 & \leq & 15 \\ & 4x_1 & - & x_2 & \leq & 10 \\ & -x_1 & + & x_2 & \leq & 1 \\ & & & x_1, x_2 & \geq & 0 \end{array}$$

## Solution

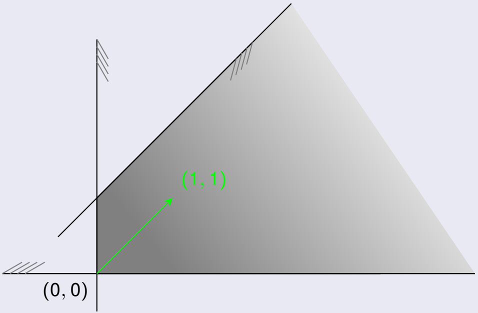


## Example

Show that the following problem is unbounded.

$$\begin{array}{lrcrcl} \text{Maximize} & x_1 & + & x_2 & & \\ & -x_1 & + & x_2 & \leq & 1 \\ & & & x_1, x_2 & \geq & 0 \end{array}$$

## Solution

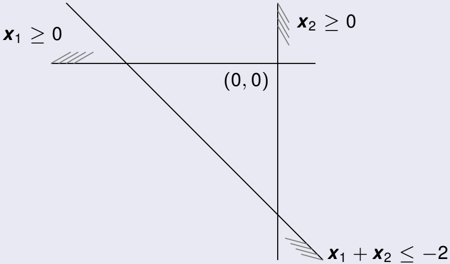


## Example

Show that the following problem has no feasible solution.

$$\begin{array}{lrcrcr} \text{Maximize} & \mathbf{x}_1 & + & \mathbf{x}_2 & & \\ & \mathbf{x}_1 & + & \mathbf{x}_2 & \leq & -2 \\ & & & \mathbf{x}_1, \mathbf{x}_2 & \geq & 0 \end{array}$$

## Solution



# Outline

# Linear and affine spaces in $\mathbb{R}^n$

## Definition

A set $L \subseteq \mathbb{R}^n$ is *linear* (also called a linear space) if

- $\mathbf{0} \in L$,
- $\boldsymbol{x} + \boldsymbol{y} \in L$ for every $\boldsymbol{x}, \boldsymbol{y} \in L$ and
- $\alpha \boldsymbol{x} \in L$ for every $\boldsymbol{x} \in L$ and $\alpha \in \mathbb{R}$.

## Definition

If $L \subseteq \mathbb{R}^n$ is a linear space and $\boldsymbol{a} \in R^n$ is a vector, then $L + \boldsymbol{a} = \{\boldsymbol{x} + \boldsymbol{a};\ \boldsymbol{x} \in L\}$ is called an *affine space*.

## Observation

An affine space $A \subseteq \mathbb{R}^n$ is linear if and only if $A$ contains the origin $\mathbf{0}$.

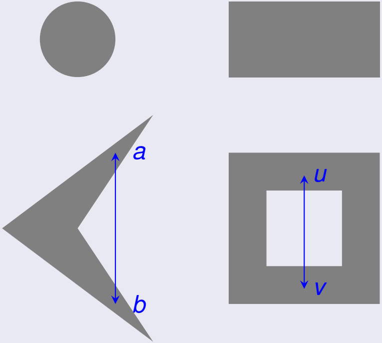
## Observation

If $A \subseteq \mathbb{R}^n$ is an affine space, then $A - \boldsymbol{x}$ is a linear space for for every $\boldsymbol{x} \in A$. Furthermore, all spaces $A - \boldsymbol{x}$ are the same for all $\boldsymbol{x} \in A$.

## Definition

A set $S \subseteq \mathbb{R}^n$ is *convex* if $S$ contains whole segment between every two points of $S$.

## Example



# Linear, affine and convex hulls

## Observation

- The intersection of arbitrary many linear spaces is also a linear space.
- The intersection of arbitrary many affine spaces is also an affine space.
- The intersection of arbitrary many convex sets is also a convex set.

## Definition

- The *linear hull* $\text{span}(S)$ of $S \subseteq \mathbb{R}^n$ is the intersection of all linear sets containing $S$.
- The *affine hull* $\text{aff}(S)$ of $S \subseteq \mathbb{R}^n$ is the intersection of all affine sets containing $S$.
- The *convex hull* $\text{conv}(S)$ of $S \subseteq \mathbb{R}^n$ is the intersection of all convex sets containing $S$.

## Informally

The linear, the affine and the convex hull of a set $S \subseteq \mathbb{R}^n$ is the smallest (with respect to inclusion) linear, affine and convex set containing $S$, respectively.

## Observation

- A set $S \subseteq \mathbb{R}^n$ is linear if and only if $S = \text{span}(S)$.
- A set $S \subseteq \mathbb{R}^n$ is affine if and only if $S = \text{aff}(S)$.
- A set $S \subseteq \mathbb{R}^n$ is convex if and only if $S = \text{conv}(S)$.

## Definition

- The sum $\sum_{i=1}^{k} \alpha_i \boldsymbol{a}_i$ is called a *linear combination* of $S \subseteq \mathbb{R}^n$ if $k \in \mathbb{N}$, $\boldsymbol{a}_i \in S$ and $\alpha_i \in \mathbb{R}$ for $i = 1, \ldots, k$.
- The sum $\sum_{i=1}^{k} \alpha_i \boldsymbol{a}_i$ is called an *affine combination* of $S \subseteq \mathbb{R}^n$ if $k \in \mathbb{N}$, $\boldsymbol{a}_i \in S$, $\alpha_i \in \mathbb{R}$ and $\sum_{i=1}^{k} \alpha_i = 1$ for $i = 1, \ldots, k$.
- The sum $\sum_{i=1}^{k} \alpha_i \boldsymbol{a}_i$ is called a *convex combination* of $S \subseteq \mathbb{R}^n$ if $k \in \mathbb{N}$, $\boldsymbol{a}_i \in S$, $\alpha_i \geq 0$ and $\sum_{i=1}^{k} \alpha_i = 1$ for $i = 1, \ldots, k$.

## Theorem

- The linear hull of a set $S \subseteq \mathbb{R}^n$ is the set of all linear combinations of $S$.
- The affine hull of a set $S \subseteq \mathbb{R}^n$ is the set of all affine combinations of $S$.
- The convex hull of a set $S \subseteq \mathbb{R}^n$ is the set of all convex combinations of $S$.

# Convex hull and convex combinations

## Observation

The set of all convex combinations of a set $C \subseteq \mathbb{R}^n$ is convex. ①

## Observation

If $C \subseteq \mathbb{R}^n$ is a convex set and $X \subseteq C$, then $C$ contains all convex combinations of $X$. ②

## Theorem

The convex hull of a set $S \subseteq \mathbb{R}^n$ is the set of all convex combinations of $S$.

## Proof

- Let $Z$ be the set of all convex combinations of $S$.
- $\text{conv}(S) \subseteq Z$: Observe that $Z$ is a convex set containing $S$.
- $Z \subseteq \text{conv}(S)$: Observe that convex combinations of points of $S$ belong into $\text{conv}(S)$.

1. Let $\mathbf{a} = \sum \alpha_i \mathbf{a}_i$ and $\mathbf{b} = \sum \beta_i \mathbf{b}_i$ be convex combinations of $C$. The point $\mathbf{x} = \alpha \mathbf{a} + \beta \mathbf{b}$ on the segment between $\mathbf{a}$ and $\mathbf{b}$ is also a convex combination of $C$ since $\mathbf{x} = \sum \alpha \alpha_i \mathbf{a}_i + \sum \beta \beta_i \mathbf{b}_i$.
2. By induction by $k$, we prove for every $X \subseteq C$ that every convex combinations of $k$ points of $X$ belong into $C$. Let $\sum \alpha_i \mathbf{a}_i$ be a convex combination of points of $X$. WLOG $\alpha_i > 0$.
   For $k = 2$ the statement follows from the definition of convexity.
   For $k > 2$, let $\alpha' = \alpha_1 + \alpha_2$ and $\mathbf{a}' = \frac{\alpha_1}{\alpha'} \mathbf{a}_1 + \frac{\alpha_2}{\alpha'} \mathbf{a}_2$. Since $\mathbf{a}'$ is a point on the segment between $\mathbf{a}_1$ and $\mathbf{a}_2$ it follows that $\mathbf{a}' \in C$. Now, $\alpha' \mathbf{a}' + \sum_{i=3}^{k} \alpha_i \mathbf{a}_i = \sum \alpha_i \mathbf{a}_i$ is a convex combination of $k - 1$ points of $X \cup \{\mathbf{a}'\}$, so it is contained in $C$ by the induction hypotheses.

# Independence and base

## Definition

- A set of vectors $S \subseteq \mathbb{R}^n$ is *linearly independent* if no vector of $S$ is a linear combination of others.
- A set of vectors $S \subseteq \mathbb{R}^n$ is *affinely independent* if no vector of $S$ is an affine combination of others.

## Definition

- A set of vectors $B \subseteq R^n$ is a *(linear) base* of a linear space $S$ if vectors of $B$ are linearly independent and $\text{span}(B) = S$.
- A set of vectors $B \subseteq R^n$ is an *(affine) base* of an affine space $S$ if vectors of $B$ are affinely independent and $\text{aff}(B) = S$.

## Question

Is it possible to analogously define a convex independence and a convex base?

## Observation

- All linear bases of a linear space have the same cardinality.
- All affine bases of an affine space have the same cardinality.

## Observation

Vectors $\mathbf{x}_0, \ldots, \mathbf{x}_k$ are affinely independent if and only if vectors $\mathbf{x}_1 - \mathbf{x}_0, \ldots, \mathbf{x}_k - \mathbf{x}_0$ are linearly independent.

## Observation

Let $S$ be a linear space and $B \subseteq S \setminus \{\mathbf{0}\}$. Then, $B$ is a linear base of $S$ if and only if $B \cup \{\mathbf{0}\}$ is an affine base of $S$.

## Definition

- The *dimension* of a linear space is the cardinality of its linear base.
- The *dimension* of an affine space is the cardinality of its affine base minus one.
- The *dimension* $\dim(S)$ of a set $S \subseteq \mathbb{R}^n$ is the dimension of affine hull of $S$.

## Observation

- A set of vectors $S$ is linearly independent if and only if $\mathbf{0}$ is not a non-trivial linear combination of $S$.
- A set of vectors $S$ is affinely independent if and only if $\mathbf{0}$ is not a non-trivial combination $\sum \alpha_i \mathbf{a}_i$ of $S$ such that $\sum \alpha_i = 0$ and $\boldsymbol{\alpha} \neq \mathbf{0}$.

### Theorem (Carathéodory)

Let $S \subseteq \mathbb{R}^n$. Every point of conv($S$) is a convex combinations of affinely independent points of $S$. ①

### Corollary

Let $S \subseteq \mathbb{R}^n$ be a set of dimension $d$. Then, every point of conv($S$) is a convex combinations of at most $d + 1$ points of $S$.

1. Let $\boldsymbol{x} \in \text{conv}(S)$. Let $\boldsymbol{x} = \sum_{i=1}^{k} \alpha_i \boldsymbol{x}_i$ be a convex combination of points of $S$ with the smallest $k$. If $\boldsymbol{x}_1, \ldots, \boldsymbol{x}_k$ are affinely dependent, then there exists a combination $\mathbf{0} = \sum \beta_i \boldsymbol{x}_i$ such that $\sum \beta_i = 0$ and $\boldsymbol{\beta} \neq \mathbf{0}$. Since this combination is non-trivial, there exists $j$ such that $\beta_j > 0$ and $\frac{\alpha_j}{\beta_j}$ is minimal. Let $\gamma_i = \alpha_i - \frac{\alpha_j \beta_i}{\beta_j}$. Observe that
   - $\boldsymbol{x} = \sum_{i \neq j} \gamma_i \boldsymbol{x}_i$
   - $\sum_{i \neq j} \gamma_i = 1$
   - $\gamma_i \geq 0$ for all $i \neq j$

   which contradicts the minimality of $k$.

# System of linear equations and inequalities

## Definition

- A *hyperplane* is a set $\{\boldsymbol{x} \in \mathbb{R}^n;\ \boldsymbol{a}^{\mathrm{T}}\boldsymbol{x} = b\}$ where $\boldsymbol{a} \in \mathbb{R}^n \setminus \{\boldsymbol{0}\}$ and $b \in \mathbb{R}$.
- A *half-space* is a set $\{\boldsymbol{x} \in \mathbb{R}^n;\ \boldsymbol{a}^{\mathrm{T}}\boldsymbol{x} \leq b\}$ where $\boldsymbol{a} \in \mathbb{R}^n \setminus \{\boldsymbol{0}\}$ and $b \in \mathbb{R}$.
- A *polyhedron* is an intersection of finitely many half-spaces.
- A *polytope* is a bounded polyhedron.

## Observation

For every $\boldsymbol{a} \in \mathbb{R}^n$ and $b \in \mathbb{R}$, the set of all $\boldsymbol{x} \in \mathbb{R}^n$ satisfying $\boldsymbol{a}^{\mathrm{T}}\boldsymbol{x} \leq b$ is convex.

## Corollary

Every polyhedron $A\boldsymbol{x} \leq \boldsymbol{b}$ is convex.

## Definition

- A set $S \subseteq \mathbb{R}^n$ is *closed* if $S$ contains the limit of every converging sequence of points of $S$ .
- A set $S \subseteq \mathbb{R}^n$ is *bounded* if $\max\{||\boldsymbol{x}||;\ \boldsymbol{x} \in S\} < b$ for some $b \in \mathbb{R}$.
- A set $S \subseteq \mathbb{R}^n$ is *compact* if every sequence of points of $S$ contains a converging subsequence with limit in $S$.

## Theorem

A set $S \subseteq \mathbb{R}^n$ is compact if and only if $S$ is closed and bounded.

## Theorem

If $f : S \to \mathbb{R}$ is a continuous function on a compact set $S \subseteq \mathbb{R}^n$, then $S$ contains a point $\boldsymbol{x}$ maximizing $f$ over $S$; that is, $f(\boldsymbol{x}) \geq f(\boldsymbol{y})$ for every $\boldsymbol{y} \in S$.

## Infimum and supremum

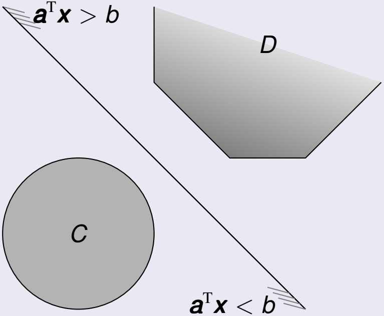
- Infimum of a set $S \subseteq \mathbb{R}$ is $\inf(S) = \max\{b \in \mathbb{R};\ b \leq x\ \forall x \in S\}$.
- Supremum of a set $S \subseteq \mathbb{R}$ is $\sup(S) = \min\{b \in \mathbb{R};\ b \geq x\ \forall x \in S\}$.
- $\inf(\emptyset) = \infty$ and $\sup(\emptyset) = -\infty$
- $\inf(S) = -\infty$ if $S$ has no lower bound

# Hyperplane separation theorem

## Theorem (strict version)

Let $C, D \subseteq \mathbb{R}^n$ be non-empty, closed, convex and disjoint sets and $C$ be bounded. Then, there exists a hyperplane $\boldsymbol{a}^{\mathrm{T}}\boldsymbol{x} = b$ which strictly separates $C$ and $D$; that is $C \subseteq \left\{\boldsymbol{x};\boldsymbol{a}^{\mathrm{T}}\boldsymbol{x} < b\right\}$ and $D \subseteq \left\{\boldsymbol{x};\boldsymbol{a}^{\mathrm{T}}\boldsymbol{x} > b\right\}$.

## Example



1. Find $\boldsymbol{c} \in C$ and $\boldsymbol{d} \in D$ with minimal distance $||\boldsymbol{d} - \boldsymbol{c}||$.
   1. Let $m = \inf\{||\boldsymbol{d} - \boldsymbol{c}||;\ \boldsymbol{c} \in C, \boldsymbol{d} \in D\}$.
   2. For every $n \in \mathbb{N}$ there exists $\boldsymbol{c}_n \in C$ and $\boldsymbol{d}_n \in D$ such that $||\boldsymbol{d}_n - \boldsymbol{c}_n|| \le m + \frac{1}{n}$.
   3. Since $C$ is compact, there exists a subsequence $\{\boldsymbol{c}_{k_n}\}_{n=1}^{\infty}$ converging to $\boldsymbol{c} \in C$.
   4. There exists $z \in \mathbb{R}$ such that for every $n \in \mathbb{N}$ the distance $||\boldsymbol{d}_n - \boldsymbol{c}||$ is at most $z$: $||\boldsymbol{d}_n - \boldsymbol{c}|| \le ||\boldsymbol{d}_n - \boldsymbol{c}_n|| + ||\boldsymbol{c}_n - \boldsymbol{c}|| \le m + 1 + \max\{||c' - c''||;\ c', c'' \in C\} = z$
   5. Since the set $D \cap \{\boldsymbol{x} \in \mathbb{R}^n;\ ||\boldsymbol{x} - \boldsymbol{c}|| \le z\}$ is compact, the sequence $\{\boldsymbol{d}_{k_n}\}_{n=1}^{\infty}$ has a subsequence $\{\boldsymbol{d}_{l_n}\}_{n=1}^{\infty}$ converging to $\boldsymbol{d} \in D$.
   6. Since $||\boldsymbol{d} - \boldsymbol{c}|| \le ||\boldsymbol{d} - \boldsymbol{d}_{l_n}|| + ||\boldsymbol{d}_{l_n} - \boldsymbol{c}_{l_n}|| + ||\boldsymbol{c}_{l_n} - \boldsymbol{c}|| \to m$, the distance $||\boldsymbol{d} - \boldsymbol{c}|| = m$ is minimal.
2. The required hyperplane is $\boldsymbol{a}^{\mathrm{T}}\boldsymbol{x} = b$ where $\boldsymbol{a} = \boldsymbol{d} - \boldsymbol{c}$ and $b = \frac{\boldsymbol{a}^{\mathrm{T}}\boldsymbol{c} + \boldsymbol{a}^{\mathrm{T}}\boldsymbol{d}}{2}$ since we prove that $\boldsymbol{a}^{\mathrm{T}}\boldsymbol{c}' \le \boldsymbol{a}^{\mathrm{T}}\boldsymbol{c} < b < \boldsymbol{a}^{\mathrm{T}}\boldsymbol{d} \le \boldsymbol{a}^{\mathrm{T}}\boldsymbol{d}'$ for every $\boldsymbol{c}' \in C$ and $\boldsymbol{d}' \in D$.
   1. In order to prove the most left inequality, let $\boldsymbol{c}' \in C$.
   2. Since $C$ is convex, $y = \boldsymbol{c} + \alpha(\boldsymbol{c}' - \boldsymbol{c}) \in C$ for every $0 \le \alpha \le 1$.
   3. From the minimality of the distance $||\boldsymbol{d} - \boldsymbol{c}||$ it follows that $||\boldsymbol{d} - y||^2 \ge ||\boldsymbol{d} - \boldsymbol{c}||^2$.
   4. $$\begin{aligned} (\boldsymbol{d} - \boldsymbol{c} - \alpha(\boldsymbol{c}' - \boldsymbol{c}))^{\mathrm{T}}(\boldsymbol{d} - \boldsymbol{c} - \alpha(\boldsymbol{c}' - \boldsymbol{c})) &\ge (\boldsymbol{d} - \boldsymbol{c})^{\mathrm{T}}(\boldsymbol{d} - \boldsymbol{c}) \\ \alpha^2(\boldsymbol{c}' - \boldsymbol{c})^{\mathrm{T}}(\boldsymbol{c}' - \boldsymbol{c}) - 2\alpha(\boldsymbol{d} - \boldsymbol{c})^{\mathrm{T}}(\boldsymbol{c}' - \boldsymbol{c}) &\ge 0 \\ \frac{\alpha}{2}||\boldsymbol{c}' - \boldsymbol{c}||^2 + \boldsymbol{a}^{\mathrm{T}}\boldsymbol{c} &\ge \boldsymbol{a}^{\mathrm{T}}\boldsymbol{c}' \end{aligned}$$
   5. Since the last inequality holds for arbitrarily small $\alpha > 0$, it follows that $\boldsymbol{a}^{\mathrm{T}}\boldsymbol{c} \ge \boldsymbol{a}^{\mathrm{T}}\boldsymbol{c}'$ holds.

## Corollary

The intersection of arbitrary many half-spaces is a closed convex set and every closed convex set is an intersection of (infinitely) many half-spaces.

## Observation

- The set of all solutions of $A\boldsymbol{x} = \boldsymbol{0}$ is a linear space and every linear space is the set of all solutions of $A\boldsymbol{x} = \boldsymbol{0}$ for some $A$.
- The set of all solutions of $A\boldsymbol{x} = \boldsymbol{b}$ is an affine space and every affine space is the set of all solutions of $A\boldsymbol{x} = \boldsymbol{b}$ for some $A$ and $\boldsymbol{b}$, assuming $A\boldsymbol{x} = \boldsymbol{b}$ is consistent. ①

## Definition

The set of all solutions of $A\boldsymbol{x} \leq \boldsymbol{b}$ is called a polyhedron.

1. Clearly, all solutions of $A\boldsymbol{x} = \mathbf{0}$ form a linear space $S$. For every solution $\boldsymbol{z}$ of $A\boldsymbol{x} = \boldsymbol{b}$ it holds that $S + \boldsymbol{z}$ is the affine space of all solutions of $A\boldsymbol{x} = \boldsymbol{b}$.
   Let $S$ be a linear space. Let rows of a matrix $A$ be a linear base of the orthogonal space to $S$. Then, $S$ are all solutions of $A\boldsymbol{x} = \mathbf{0}$. If $S + \boldsymbol{z}$ is an affine space and $\boldsymbol{b} = A\boldsymbol{z}$, then $S + \boldsymbol{z}$ are all solutions of $A\boldsymbol{x} = \boldsymbol{b}$.

# Faces of a polyhedron

## Definition

Let $P$ be a polyhedron. A half-space $\boldsymbol{\alpha}^{\mathrm{T}}\boldsymbol{x} \leq \beta$ is called a *supporting hyperplane* of $P$ if the inequality $\boldsymbol{\alpha}^{\mathrm{T}}\boldsymbol{x} \leq \beta$ holds for every $x \in P$ and the hyperplane $\boldsymbol{\alpha}^{\mathrm{T}}\boldsymbol{x} = \beta$ has a non-empty intersection with $P$.
The set of point in the intersetion $P \cap \left\{\boldsymbol{x};\ \boldsymbol{\alpha}^{\mathrm{T}}\boldsymbol{x} = \beta\right\}$ is called a *face* of $P$. By convention, the empty set and $P$ are also faces, and the other faces are *proper* faces.
①

## Definition

Let $P$ be a $d$-dimensional polyhedron.

- A 0-dimensional face of $P$ is called a *vertex* of $P$.
- A 1-dimensional face is of $P$ called an *edge* of $P$.
- A $(d-1)$-dimensional face of $P$ is called an *facet* of $P$.

## Observation

Let $P = \{x;\ A\boldsymbol{x} \leq \boldsymbol{b}\}$ of dimension $d$. Then for every row $i$, either

- $P \cap \{x;\ A_{i,\star}\boldsymbol{x} = \boldsymbol{b}_i\} = P$ or
- $P \cap \{x;\ A_{i,\star}\boldsymbol{x} = \boldsymbol{b}_i\} = \emptyset$ or
- $P \cap \{x;\ A_{i,\star}\boldsymbol{x} = \boldsymbol{b}_i\}$ is a proper face of dimension at most $d-1$.

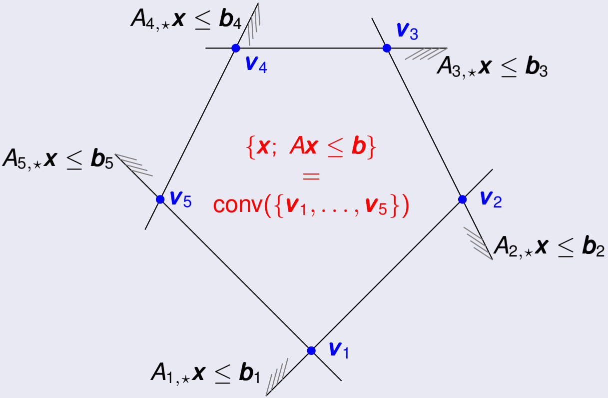
1. Observe, that every face of a polyhedron is also a polyhedron.

# Minkowski-Weyl

## Theorem (Minkowski-Weyl)

A set $S \subseteq \mathbb{R}^n$ is a polytope if and only if there exists a finite set $V \subseteq \mathbb{R}^n$ such that $S = \text{conv}(V)$.

## Illustration



$\Rightarrow$ Proof by induction on $d = \dim(S)$:

1. For $d = 0$, the size of $S$ is 0 or 1.
2. For $d > 0$, let $S = \{\boldsymbol{x};\ A\boldsymbol{x} \leq \boldsymbol{b}\}$ and $S_i = S \cap \{\boldsymbol{x};\ A_{i,\star}\boldsymbol{x} = \boldsymbol{b}_i\}$.
   Let $I$ be the set of rows $i$ such that $S_i$ is a proper face of $S$. Since $\dim(S_i) \leq \dim(S) - 1$ for all $i \in I$, the induction assumption implies that there exists a finite set $V_i \in \mathbb{R}^n$ such that $S_i = \operatorname{conv}(V_i)$.
   Let $V = \cup_{i \in I} V_i$. We prove that $\operatorname{conv}(V) = S$.
   - $\subseteq$ follows from $V_i \subseteq S_i \subseteq S$.
   - $\supseteq$ Let $\boldsymbol{x} \in S$. Let $L$ be a line containing $\boldsymbol{x}$.
     $S \cap L$ is a line segment with end-vertices $\boldsymbol{u}$ and $\boldsymbol{v}$.
     There exists $i, j \in I$ such that $A_{i,\star}\boldsymbol{u} = \boldsymbol{b}_i$ and $A_{j,\star}\boldsymbol{v} = \boldsymbol{b}_j$.
     Since $\boldsymbol{u} \in S_i$ and $\boldsymbol{v} \in S_j$, points $\boldsymbol{u}$ and $\boldsymbol{v}$ are convex combinations of $S$.
     Since $\boldsymbol{x}$ is a also a convex combination of $\boldsymbol{u}$ and $\boldsymbol{v}$, we have $\boldsymbol{x} \in \operatorname{conv}(S)$.

## Theorem (Minkowski-Weyl)

A set $S \subseteq \mathbb{R}^n$ is a polytope if and only if there exists a finite set $V \subseteq \mathbb{R}^n$ such that $S = \operatorname{conv}(V)$.

## Proof of the implication $\Leftarrow$ (main steps)

- Let $Q = \left\{ \binom{\boldsymbol{\alpha}}{\beta};\ \boldsymbol{\alpha} \in \mathbb{R}^n, \beta \in \mathbb{R}, -\mathbf{1} \le \boldsymbol{\alpha} \le 1, -1 \le \beta \le 1, \boldsymbol{\alpha}^{\mathrm{T}} \boldsymbol{v} \le \beta\ \forall \boldsymbol{v} \in V \right\}$.
- Observe that $\boldsymbol{\alpha}^{\mathrm{T}} \boldsymbol{v} \le \beta$ means the same as $\binom{\boldsymbol{v}}{-1}^{\mathrm{T}} \binom{\boldsymbol{\alpha}}{\beta} \le 0$.
- Since $Q$ is a polytope, there exists a finite set $W \subseteq \mathbb{R}^{n+1}$ such that $Q = \operatorname{conv}(W)$.
- We prove that $\operatorname{conv}(V) = \left\{ \boldsymbol{x} \in \mathbb{R}^n;\ \boldsymbol{\alpha}^{\mathrm{T}} \boldsymbol{x} \le \beta\ \forall \binom{\boldsymbol{\alpha}}{\beta} \in W \right\}$.

1. $\boldsymbol{x} \in \operatorname{conv}(V)$
2. $\boldsymbol{\alpha}^{\mathrm{T}} \boldsymbol{x} \le \beta\ \forall \binom{\boldsymbol{\alpha}}{\beta} \in Q_1$ where $Q_1 = \left\{ \binom{\boldsymbol{\alpha}}{\beta};\ \boldsymbol{\alpha}^{\mathrm{T}} \boldsymbol{v} \le \beta\ \forall \boldsymbol{v} \in \operatorname{conv}(V) \right\}$
3. $\boldsymbol{\alpha}^{\mathrm{T}} \boldsymbol{x} \le \beta\ \forall \binom{\boldsymbol{\alpha}}{\beta} \in Q_2$ where $Q_2 = \left\{ \binom{\boldsymbol{\alpha}}{\beta};\ \boldsymbol{\alpha}^{\mathrm{T}} \boldsymbol{v} \le \beta\ \forall \boldsymbol{v} \in V \right\}$
4. $\boldsymbol{\alpha}^{\mathrm{T}} \boldsymbol{x} \le \beta\ \forall \binom{\boldsymbol{\alpha}}{\beta} \in Q$
5. $\boldsymbol{\alpha}^{\mathrm{T}} \boldsymbol{x} \le \beta\ \forall \binom{\boldsymbol{\alpha}}{\beta} \in W$

(1) $\Rightarrow$ (2) $Q_1$ is the set of all conditions satisfied by all points of conv($V$).

(1) $\Leftarrow$ (2) Use the hyperplane separation theorem to separate $x \notin$ conv($V$) from conv($V$).

(2) $\Leftrightarrow$ (3) A condition $\boldsymbol{\alpha}^{\mathrm{T}}\boldsymbol{v} \leq \beta$ is satified by all $\boldsymbol{v} \in V$ if and only if the condition is satisfied by $\boldsymbol{v} \in$ conv($V$), so $Q_1 = Q_2$.

(3) $\Leftrightarrow$ (4) $\boldsymbol{\alpha}$ and $\beta$ in every condition $\boldsymbol{\alpha}^{\mathrm{T}}\boldsymbol{v} \leq \beta$ can be scaled so that $-\mathbf{1} \leq \boldsymbol{\alpha} \leq 1$ and $-1 \leq \beta \leq 1$ and the condition describe the same half-space.

(4) $\Leftrightarrow$ (5) Prove that if $\boldsymbol{\alpha}^{\mathrm{T}}\boldsymbol{x} \leq \beta$ holds for all conditions from $W$, then it also holds for all conditions from $Q =$ conv($W$).

### Observation

The intersection of two faces of a polyhedron $P$ is a face of $P$.

### Theorem

Let $P$ be a polyhedron and $V$ its vertices. Then, $\boldsymbol{x}$ is a vertex of $P$ if and only if $\boldsymbol{x} \notin \text{conv}(P \setminus \{\boldsymbol{x}\})$. Furthermore, if $P$ is bounded, then $P = \text{conv}(V)$. ①

### Observation (A face of a face is a face)

Let $F$ be a face of a polyhedron $P$ and let $E \subseteq F$. Then, $E$ is a face of $F$ if and only if $E$ is a face of $P$.

### Corollary

A set $F \subseteq \mathbb{R}^n$ is a face of a polyhedron $P = \{\boldsymbol{x} \in \mathbb{R}^n;\ A\boldsymbol{x} \leq \boldsymbol{b}\}$ if and only if $F$ is the set of all optimal solutions of the linear programming problem $\min\left\{\boldsymbol{c}^{\mathrm{T}}\boldsymbol{x};\ A\boldsymbol{x} \leq \boldsymbol{b}\right\}$ for some vector $\boldsymbol{c} \in \mathbb{R}^n$.

1. For simplicity, we prove this theorem only for bounded polyhedrons. Let $V_0$ be (inclusion) minimal set such that $P = \text{conv}(V_0)$. Let $V_e = \{\boldsymbol{x} \in P;\ \boldsymbol{x} \notin \text{conv}(P \setminus \{\boldsymbol{x}\})\}$. We prove that $V \subseteq V_e \subseteq V_0 \subseteq V$.

$V \subseteq V_e$: Let $\boldsymbol{z} \in V$ be a vertex. By definition, there exists a supporting hyperplane $\boldsymbol{c}^{\mathrm{T}}\boldsymbol{x} = t$ such that $P \cap \left\{\boldsymbol{x};\ \boldsymbol{c}^{\mathrm{T}}\boldsymbol{x} = t\right\} = \{\boldsymbol{z}\}$. Since $\boldsymbol{c}^{\mathrm{T}}\boldsymbol{x} < t$ for all $\boldsymbol{x} \in P \setminus \{\boldsymbol{z}\}$, it follows that $\boldsymbol{x} \in V_e$.

$V_e \subseteq V_0$: Let $\boldsymbol{z} \in V_e$. Since $\text{conv}(P \setminus \{\boldsymbol{z}\}) \neq P$, it follows that $\boldsymbol{z} \in V_0$.

$V_0 \subseteq V$: Let $\boldsymbol{z} \in V_0$ and $D = \text{conv}(V_0 \setminus \{\boldsymbol{z}\})$. From Minkovsky-Weil's theorem it follows that $V_0$ is finite and therefore, $D$ is compact. By the separation theorem, there exists a hyperplane $\boldsymbol{c}^{\mathrm{T}}\boldsymbol{x} = r$ separating $\{\boldsymbol{z}\}$ and $D$, that is $\boldsymbol{c}^{\mathrm{T}}\boldsymbol{x} < r < \boldsymbol{c}^{\mathrm{T}}\boldsymbol{z}$ for all $\boldsymbol{x} \in D$. Let $t = \boldsymbol{c}^{\mathrm{T}}\boldsymbol{z}$. Hence, $A = \left\{\boldsymbol{x};\ \boldsymbol{c}^{\mathrm{T}}\boldsymbol{x} = t\right\}$ is a supporting hyperplane of $P$.
We prove that $A \cap P = \{\boldsymbol{z}\}$. For contradiction, let $\boldsymbol{z}' \in P \cap A$ be a different from $\boldsymbol{z}$. Then, there exists a convex combination $\boldsymbol{z}' = \alpha_1\boldsymbol{x}_1 + \cdots + \alpha_k\boldsymbol{x}_k + \alpha_0\boldsymbol{z}$ of $V_0$. From $\boldsymbol{z} \neq \boldsymbol{z}'$ it follows that $\alpha_0 < 1$ and $\alpha_i > 0$ for some $i$. Since $\alpha_0\boldsymbol{c}^{\mathrm{T}}\boldsymbol{z} = t$ and $\alpha_i\boldsymbol{c}^{\mathrm{T}}\boldsymbol{x}_i < t$ and $\alpha_j\boldsymbol{c}^{\mathrm{T}}\boldsymbol{x}_j \leq t$, it holds that $\boldsymbol{c}^{\mathrm{T}}\boldsymbol{z}' < t$ which contradicts the assumption that $\boldsymbol{z}' \in A$.

# Minimal defining system of a polyhedron

## Definition

$P = \{\mathbf{x} \in \mathbb{R}^n;\ A'\mathbf{x} = \mathbf{b}',\ A''\mathbf{x} \leq \mathbf{b}''\}$ is a *minimal defining system* of a polyherdon $P$ if

- no condition can be removed and
- no inequality can be replaced by equality

without changing the polyhedron $P$.

## Observation

Let $\mathbf{z}$ be a point of a polyhedron $P = \{\mathbf{x} \in \mathbb{R}^n;\ A'\mathbf{x} = \mathbf{b}',\ A''\mathbf{x} \leq \mathbf{b}''\}$ such that $A''\mathbf{z} < \mathbf{b}''$. Then,

- $\dim(P) = n - \operatorname{rank}(A')$ and ①
- $\mathbf{z}$ does not belong in any proper face of $P$. ②

Furthermore, there exists such a point $\mathbf{z}$ in every minimal defining system of a polyhedron. ③

## Theorem

Let $P = \{\mathbf{x} \in \mathbb{R}^n;\ A'\mathbf{x} = \mathbf{b}',\ A''\mathbf{x} \leq \mathbf{b}''\}$ be a minimal defining system of a polyhedron $P$. Then, there exists a bijection between facets of $P$ and inequalities $A''\mathbf{x} \leq \mathbf{b}''$. ④

1. Let $L$ be the affine space defined by $A'\mathbf{x} = \mathbf{b}'$. Clearly, $\dim(P) \leq \dim(L) = n - \operatorname{rank}(A')$. Since $A''\mathbf{z} < \mathbf{b}''$, there exists $\epsilon > 0$ such that $P$ contains whole ball $B = \{\mathbf{x} \in L;\ ||x - z|| \leq \epsilon\}$. Since vectors of a base of the linear space $L - z$ can be scaled so that they belong into $B - z$, it follows that $\dim(P) \geq \dim(B) \geq \dim(L)$.
2. The point $\mathbf{z}$ cannot belong into any proper face of $P$ because a supporting hyperplane of such a face split the ball $B$.
3. For every row $i$ of $A''\mathbf{x} \leq \mathbf{b}''$ there exists $\mathbf{z}^i \in P$ such that $A''_{i,\star}\mathbf{z}^i < \mathbf{b}''_i$. Let $\mathbf{z} = \frac{1}{m''}\sum_{i=1}^{m''}\mathbf{z}^i$ be the center of gravity. Since $\mathbf{z}$ is a convex combination of points of $P$, point $\mathbf{z}$ belongs to $P$. From $A''_{i,\star}\mathbf{z}^i < \mathbf{b}''_i$, it follows that $A''_{i,\star}\mathbf{z} < \mathbf{b}''_i$, and therefore $A''\mathbf{z} < \mathbf{b}''$.
4. Let $R_i = \left\{\mathbf{x} \in \mathbb{R}^n;\ A''_{i,\star}\mathbf{x} = \mathbf{b}_i\right\}$ and $F_i = P \cap R_i$. From minimality if follows that $R_i$ is a supporting hyperplane, and therefore, $F_i$ is a face. Likewise in the previous observation, there exists $\mathbf{z} \in F_i$ satisfying $A''_{j,\star}\mathbf{z} < \mathbf{b}_j$ for all $j \neq i$ and so $\dim(F_i) = \dim(P) - 1$. Furthermore, $\mathbf{z} \notin F_j$ for all $j \neq i$, so $F_i \neq F_j$ for $j \neq i$. For contradiction, let $F$ be an another facet. There exists a facet $i$ such $F \subseteq F_i$, otherwise $\mathbf{z} = \frac{1}{m''}\sum_{i=1}^{m''}\mathbf{z}^i$ satisfies strictly all condition contradicting the assumption that $F$ is a proper facet. Since $F \neq F_i$, $F$ is a proper face of $F_i$ and so its dimension is at most $\dim(P) - 2$ contradicting the assumption that $F$ is a proper facet.

# Minimal defining system of a polyhedron

## Theorem

Let $P = \{\mathbf{x} \in \mathbb{R}^n;\ A'\mathbf{x} = \mathbf{b}',\ A''\mathbf{x} \le \mathbf{b}''\}$ be a minimal defining system of a polyhedron $P$. Then, there exists a bijection between facets of $P$ and inequalities $A''\mathbf{x} \le \mathbf{b}''$.

## Definition

A polyhedron $P \subseteq \mathbb{R}^n$ is of full-dimension if $\dim(P) = n$.

## Observation

If $P$ is a full-dimensional polyhedron, then $P$ has exactly one minimal defining system up-to multiplying conditions by constants. ①

## Corollary

Every proper face is an intersection of facets.

1. Affine space of dimension $n - 1$ is determined by a unique condition.

# Outline

# Notation

## Notation used in the Simplex method

- Linear programming problem in the equation form is a problem to find $\boldsymbol{x} \in \mathbb{R}^n$ which maximizes $\boldsymbol{c}^{\mathrm{T}}\boldsymbol{x}$ and satisfies $A\boldsymbol{x} = \boldsymbol{b}$ and $\boldsymbol{x} \geq \mathbf{0}$ where $A \in \mathbb{R}^{m \times n}$ and $\boldsymbol{b} \in \mathbb{R}^m$.
- We assume that rows of $A$ are linearly independent.
- For a subset $B \subseteq \{1, \ldots, n\}$, let $A_B$ be the matrix consisting of columns of $A$ whose indices belong to $B$.
- Similarly for vectors, $\boldsymbol{x}_B$ denotes the coordinates of $\boldsymbol{x}$ whose indices belong to $B$.
- The set $N = \{1, \ldots, n\} \setminus B$ denotes the remaining columns.

## Example

Consider $B = \{2, 4\}$. Then, $N = \{1, 3, 5\}$ and

$$A = \begin{pmatrix} 1 & 3 & 5 & 6 & 0 \\ 2 & 4 & 8 & 9 & 7 \end{pmatrix} \qquad A_B = \begin{pmatrix} 3 & 6 \\ 4 & 9 \end{pmatrix} \qquad A_N = \begin{pmatrix} 1 & 5 & 0 \\ 2 & 8 & 7 \end{pmatrix}$$

$$\boldsymbol{x}^{\mathrm{T}} = (3, 4, 6, 2, 7) \qquad \boldsymbol{x}_B^{\mathrm{T}} = (4, 2) \qquad \boldsymbol{x}_N^{\mathrm{T}} = (3, 6, 7)$$

Note that $A\boldsymbol{x} = A_B\boldsymbol{x}_B + A_N\boldsymbol{x}_N$.

# Basic feasible solutions

## Definitions

- A set of columns $B$ is a *base* if $A_B$ is a regular matrix.
- The *basic solution* $\boldsymbol{x}$ corresponding to a base $B$ is $\boldsymbol{x}_N = \mathbf{0}$ and $\boldsymbol{x}_B = A_B^{-1}\boldsymbol{b}$.
- A basic solution satisfying $\boldsymbol{x} \geq \mathbf{0}$ is called *basic feasible solution*.

## Observation

Basic feasible solutions are exactly vertices of the polyhedron $P = \{\boldsymbol{x};\ A\boldsymbol{x} = \boldsymbol{b}, \boldsymbol{x} \geq \mathbf{0}\}$.

## Lemma

A feasible solution $\boldsymbol{x}$ is basic if and only if the columns of the matrix $A_K$ are linearly independent where $K = \{j \in \{1, \dots, n\};\ \boldsymbol{x}_j > 0\}$.

# Example: Initial simplex tableau

## Canonical form

$$\begin{array}{lrcrcl}
\text{Maximize} & x_1 & + & x_2 & & \\
& -x_1 & + & x_2 & \leq & 1 \\
& x_1 & & & \leq & 3 \\
& & & x_2 & \leq & 2 \\
& & & x_1, x_2 & \geq & 0
\end{array}$$

## Equation form

$$\begin{array}{lrcrcrcrcrcl}
\text{Maximize} & x_1 & + & x_2 & & & & & & & & \\
& -x_1 & + & x_2 & + & x_3 & & & & & = & 1 \\
& x_1 & & & & & + & x_4 & & & = & 3 \\
& & & x_2 & & & & & + & x_5 & = & 2 \\
& & & & & & & & & x_1, x_2, x_3, x_4, x_5 & \geq & 0
\end{array}$$

## Simplex tableau

$$\begin{array}{rcrcrcr}
x_3 & = & 1 & + & x_1 & - & x_2 \\
x_4 & = & 3 & - & x_1 & & \\
x_5 & = & 2 & & & - & x_2 \\
\hline
z & = & & & x_1 & + & x_2
\end{array}$$

# Example: Initial simplex tableau

## Simplex tableau

$$
\begin{array}{rcrcrcr}
x_3 & = & 1 & + & x_1 & - & x_2 \\
x_4 & = & 3 & - & x_1 & & \\
x_5 & = & 2 & & & - & x_2 \\
\hline
z & = & & & x_1 & + & x_2
\end{array}
$$

## Initial basic feasible solution

- $B = \{3, 4, 5\}$, $N = \{1, 2\}$
- $x = (0, 0, 1, 3, 2)$

## Pivot

Two edges from the vertex $(0, 0, 1, 3, 2)$:

1. $(t, 0, 1 + t, 3 - t, 2)$ when $x_1$ is increased by $t$
2. $(0, r, 1 - r, 3, 2 - r)$ when $x_2$ is increased by $r$

These edges give feasible solutions for:

1. $t \leq 3$ since $x_3 = 1 + t \geq 0$ and $x_4 = 3 - t \geq 0$ and $x_5 = 2 \geq 0$
2. $r \leq 1$ since $x_3 = 1 - r \geq 0$ and $x_4 = 3 \geq 0$ and $x_5 = 2 - r \geq 0$

In both cases, the objective function is increasing. We choose $x_2$ as a pivot.

# Example: Pivot step

## Simplex tableau

$$\begin{array}{rcrcrcr} x_3 & = & 1 & + & x_1 & - & x_2 \\ x_4 & = & 3 & - & x_1 & & \\ x_5 & = & 2 & & & - & x_2 \\ \hline z & = & & & x_1 & + & x_2 \end{array}$$

## Basis

- Original basis $B = \{3, 4, 5\}$
- $x_2$ enters the basis (by our choice).
- $(0, r, 1 - r, 3, 2 - r)$ is feasible for $r \leq 1$ since $x_3 = 1 - r \geq 0$.
- Therefore, $x_3$ leaves the basis.
- New base $B = \{2, 4, 5\}$

## New simplex tableau

$$\begin{array}{rcrcrcr} x_2 & = & 1 & + & x_1 & - & x_3 \\ x_4 & = & 3 & - & x_1 & & \\ x_5 & = & 1 & - & x_1 & + & x_3 \\ \hline z & = & 1 & + & 2x_1 & - & x_3 \end{array}$$

# Example: Next step

## Simplex tableau

$$\begin{array}{rcrcrcr} x_2 & = & 1 & + & x_1 & - & x_3 \\ x_4 & = & 3 & - & x_1 & & \\ x_5 & = & 1 & - & \color{red}{x_1} & + & x_3 \\ \hline z & = & 1 & + & 2x_1 & - & x_3 \end{array}$$

## Next pivot

- Basis $B = \{2, 4, 5\}$ with a basis feasible solution $(0, 1, 0, 3, 1)$.
- This vertex has two incident edges but only one increases the objective function.
- The edge increasing objective function is $(t, 1 + t, 0, 3 - t, 1 - t)$.
- Feasible solutions for $x_2 = 1 + t \geq 0$ and $x_4 = 3 - t \geq 0$ and $x_5 = 1 - t \geq 0$.
- Therefore, $x_1$ enters the basis and $x_5$ leaves the basis.

## New simplex tableau

$$\begin{array}{rcrcrcr} x_1 & = & 1 & + & x_3 & - & x_5 \\ x_2 & = & 2 & & & - & x_5 \\ x_4 & = & 2 & - & x_3 & + & x_5 \\ \hline z & = & 3 & + & x_3 & - & 2x_5 \end{array}$$

# Example: Last step

## Simplex tableau

$$\begin{array}{rcrcrcr} x_1 & = & 1 & + & x_3 & - & x_5 \\ x_2 & = & 2 & & & - & x_5 \\ x_4 & = & 2 & - & {\color{red}x_3} & + & x_5 \\ \hline z & = & 3 & + & x_3 & - & 2x_5 \end{array}$$

## Next pivot

- Basis $B = \{1, 2, 4\}$ with a basis feasible solution $(1, 2, 0, 2, 0)$.
- This vertex has two incident edges but only one increases the objective function.
- The edge increasing objective function is $(1 + t, 2, t, 2 - t, 0)$.
- Feasible solutions for $x_1 = 1 + t \geq 0$ and $x_2 = 2 \geq 0$ and $x_4 = 2 - t \geq 0$.
- Therefore, $x_3$ enters the basis and $x_4$ leaves the basis.

## New simplex tableau

$$\begin{array}{rcrcrcr} x_1 & = & 3 & - & x_4 & & \\ x_2 & = & 2 & & & - & x_5 \\ x_3 & = & 2 & - & x_4 & + & x_5 \\ \hline z & = & 5 & - & x_4 & - & x_5 \end{array}$$

# Example: Optimal solution

## Simplex tableau

$$\begin{array}{rcrcrcr} x_1 & = & 3 & - & x_4 & & \\ x_2 & = & 2 & & & - & x_5 \\ x_3 & = & 2 & - & x_4 & + & x_5 \\ \hline z & = & 5 & - & x_4 & - & x_5 \end{array}$$

## No other pivot

- Basis $B = \{1, 2, 3\}$ with a basis feasible solution $(3, 2, 2, 0, 0)$.
- This vertex has two incident edges but no one increases the objective function.
- We have an optimal solution.

## Why this is an optimal solution?

- Consider an arbitrary feasible solution $\tilde{\boldsymbol{y}}$.
- The value of objective function is $\tilde{z} = 5 - \tilde{\boldsymbol{y}}_4 - \tilde{\boldsymbol{y}}_5$.
- Since $\tilde{\boldsymbol{y}}_4, \tilde{\boldsymbol{y}}_5 \geq 0$, the objective value is $\tilde{z} = 5 - \tilde{\boldsymbol{y}}_4 - \tilde{\boldsymbol{y}}_5 \leq 5 = z$.

# Simplex tableau in general

## Definition

A simplex tableau determined by a feasible basis $B$ is a system of $m+1$ linear equations in variables $\boldsymbol{x}_1, \ldots, \boldsymbol{x}_n$, and $z$ that has the same set of solutions as the system $A\boldsymbol{x} = \boldsymbol{b}$, $z = \boldsymbol{c}^{\mathrm{T}}\boldsymbol{x}$, and in matrix notation looks as follows:

$$\begin{array}{cccccc} \boldsymbol{x}_B & = & \boldsymbol{p} & + & Q\boldsymbol{x}_N \\ \hline z & = & z_0 & + & \boldsymbol{r}^{\mathrm{T}}\boldsymbol{x}_N \end{array}$$

where $\boldsymbol{x}_B$ is the vector of the basis variables, $\boldsymbol{x}_N$ is the vector on non-basis variables, $\boldsymbol{p} \in \mathbb{R}^m$, $\boldsymbol{r} \in \mathbb{R}^{n-m}$, $Q$ is an $m \times (n-m)$ matrix, and $z_0 \in \mathbb{R}$.

## Observation

For each basis $B$ there exists exactly one simplex tableau, and it is given by

- $Q = -A_B^{-1}A_N$
- $\boldsymbol{p} = A_B^{-1}\boldsymbol{b}$
- $z_0 = \boldsymbol{c}_B^{\mathrm{T}}A_B^{-1}\boldsymbol{b}$
- $r = \boldsymbol{c}_n^{\mathrm{T}} - (\boldsymbol{c}_B^{\mathrm{T}}A_B^{-1}A_N)^{\mathrm{T}}$

# Properties of a simplex tableau

## Simplex tableau in general

$$\begin{array}{ccccc} \boldsymbol{x}_B & = & \boldsymbol{p} & + & Q\boldsymbol{x}_N \\ \hline z & = & z_0 & + & \boldsymbol{r}^{\mathrm{T}}\boldsymbol{x}_N \end{array}$$

## Observation

Basis $B$ is feasible if and only if $\boldsymbol{p} \geq \mathbf{0}$.

## Observation

The solution corresponding to a basis $B$ is optimal if and only if $\boldsymbol{r} \leq 0$.

## Observation

If a linear programming problem in the equation form is feasible and bounded, then it has an optimal basis solution.

# Pivot step

## Simplex tableau in general

$$\begin{array}{rcl} \mathbf{x}_B & = & \mathbf{p} + Q\mathbf{x}_N \\ \hline z & = & z_0 + \mathbf{r}^{\mathrm{T}}\mathbf{x}_N \end{array}$$

## Find a pivot

- If $\mathbf{r} \leq \mathbf{0}$, then we have an optimal solution.
- Otherwise, choose an arbitrary entering variable $\mathbf{x}_v$ such that $\mathbf{r}_v > 0$.
- If $Q_{\star,v} \geq \mathbf{0}$, then the corresponding edge is unbounded and the problem is also unbounded.
- Otherwise, find a leaving variable $\mathbf{x}_u$ which limits the increment of the entering variable most strictly, i.e. $Q_{u,v} < 0$ and $-\frac{\mathbf{p}_u}{Q_{u,v}}$ is minimal.

## Update the simplex tableau

Gaussian elimination. Postponed for a tutorial.

## Pivot rules

**Largest coefficient** Choose an improving variable with the largest coefficient.

**Largest increase** Choose an improving variable that leads to the largest absolute improvement in $z$.

**Steepest edge** Choose an improving variable whose entering into the basis moves the current basic feasible solution in a direction closest to the direction of the vector $c$, i.e.

$$\frac{\boldsymbol{c}^{\mathrm{T}}(\boldsymbol{x}_{new} - \boldsymbol{x}_{old})}{||\boldsymbol{x}_{new} - \boldsymbol{x}_{old}||}$$

**Bland's rule** Choose an improving variable with the smallest index, and if there are several possibilities of the leaving variable, also take the one with the smallest index.

**Random edge** Select the entering variable uniformly at random among all improving variables.

# Initial feasible basis

## Equation form

Maximize $\mathbf{c}^{\mathrm{T}}\mathbf{x}$ such that $A\mathbf{x} = \mathbf{b}$ and $\mathbf{x} \geq 0$.

## Auxiliary linear program

We introduce variables $\mathbf{x}_{n+1}, \dots, \mathbf{x}_{n+m}$ and solve an auxiliary linear program:
Maximize $-\mathbf{x}_{n+1} \cdots - \mathbf{x}_{n+m}$ such that $(A|I)\mathbf{x} = \mathbf{b}$ and $\mathbf{x} \geq 0$.

## Observation

The original linear program has a feasible solution if and only if an optimal solution of the auxiliary linear program satisfies $\mathbf{x}_{n+1} = \cdots = \mathbf{x}_{n+m} = 0$.

# Complexity

## Degeneracy

- Different basis may correspond to the same solution. ①
- The simplex method may loop forever between these basis.
- Bland's or lexicographic rules prevent visiting the same basis twice.

## The number of visited vertices

- The total number of vertices is finite since the number of basis is finite.
- The objective value of visited vertices is increasing, so every vertex is visited at most once. ②
- The number of visited vertices may be exponential, e.g. the Klee-Minty cube. ③
- Practical linear programming problems in equation forms with $m$ equations typically need between $2m$ and $3m$ pivot steps to solve.

## Open problem

Is there a pivot rule which guarantees a polynomial number of steps?

1. For example, the apex of the 3-dimensional $k$-side pyramid belongs to $k$ faces, so there are $\binom{k}{3}$ basis determining the apex.
2. In degeneracy, the simplex method stay in the same vertex; and when the vertex is left, it is not visited again.
3. The Klee-Minty cube is a "deformed" $n$-dimensional cube with $2n$ facets and $2^n$ vertices. The Dantzig's original pivot rule (largest coefficient) visits all vertices of this cube.

# Outline

# Duality of linear programming: Example

## Find an upper bound for the following problem

$$\begin{array}{lrcrcr} \text{Maximize} & 2x_1 & + & 3x_2 & & \\ \text{subject to} & 4x_1 & + & 8x_2 & \leq & 12 \\ & 2x_1 & + & x_2 & \leq & 3 \\ & 3x_1 & + & 2x_2 & \leq & 4 \\ & & & x_1, x_2 & \geq & 0 \end{array}$$

## Simple estimates

- $2x_1 + 3x_2 \leq 4x_1 + 8x_2 \leq 12$ ①
- $2x_1 + 3x_2 \leq \frac{1}{2}(4x_1 + 8x_2) \leq 6$ ②
- $2x_1 + 3x_2 = \frac{1}{3}(4x_1 + 8x_2 + 2x_1 + x_2) \leq 5$ ③

## What is the best combination of conditions?

Every non-negative linear combination of inequalities which gives an inequality $d_1x_1 + d_2x_2 \leq h$ with $d_1 \geq 2$ and $d_2 \geq 3$ provides the upper bound $2x_1 + 3x_2 \leq d_1x_1 + d_2x_2 \leq h$.

1. The first condition
2. A half of the first condition
3. A third of the sum of the first and the second conditions

# Duality of linear programming: Example

## Find an upper bound for the following problem

$$\begin{array}{lrcrcl} \text{Maximize} & 2x_1 & + & 3x_2 & & \\ \text{subject to} & 4x_1 & + & 8x_2 & \leq & 12 \\ & 2x_1 & + & x_2 & \leq & 3 \\ & 3x_1 & + & 2x_2 & \leq & 4 \\ & & & x_1, x_2 & \geq & 0 \end{array}$$

## Non-negative combination of inequalities with coefficients $y_1$, $y_2$ and $y_3$

$(4y_1 + 2y_2 + 3y_3)x_1 + (8y_1 + y_2 + 2y_3)x_2 \leq 12y_1 + 3y_2 + 4y_3$ where

- $d_1 = 4y_1 + 2y_2 + 3y_3 \geq 2$
- $d_2 = 8y_1 + y_2 + 2y_3 \geq 3$
- $h = 12y_1 + 2y_2 + 4y_3$ to be minimized

## Dual program ①

$$\begin{array}{lrcrcrcl} \text{Minimize} & 12y_1 & + & 2y_2 & + & 4y_3 & & \\ \text{subject to} & 4y_1 & + & 2y_2 & + & 3y_3 & \geq & 2 \\ & 8y_1 & + & y_2 & + & 2y_3 & \geq & 3 \\ & & & & & y_1, y_2, y_3 & \geq & 0 \end{array}$$

1. The primal optimal solution is $\boldsymbol{x}^{\mathrm{T}} = (\frac{1}{2}, \frac{5}{4})$ and the dual solution is $\boldsymbol{y}^{\mathrm{T}} = (\frac{5}{16}, 0, \frac{1}{4})$, both with the same objective value 4.75.

# Duality of linear programming: General

## Primal linear program

Maximize $\boldsymbol{c}^{\mathrm{T}}\boldsymbol{x}$ subject to $A\boldsymbol{x} \leq \boldsymbol{b}$ and $\boldsymbol{x} \geq \mathbf{0}$

## Dual linear program

Minimize $\boldsymbol{b}^{\mathrm{T}}\boldsymbol{y}$ subject to $A^{\mathrm{T}}\boldsymbol{y} \geq \boldsymbol{c}$ and $\boldsymbol{y} \geq \mathbf{0}$

## Weak duality theorem

For every primal feasible solution $\boldsymbol{x}$ and dual feasible solution $\boldsymbol{y}$ hold $\boldsymbol{c}^{\mathrm{T}}\boldsymbol{x} \leq \boldsymbol{b}^{\mathrm{T}}\boldsymbol{y}$.

## Corollary

If one program is unbounded, then the other one is infeasible.

## Duality theorem

Exactly one of the following possibilities occurs

1. Neither primal nor dual has a feasible solution
2. Primal is unbounded and dual is infeasible
3. Primal is infeasible and dual is unbounded
4. There are feasible solutions $\boldsymbol{x}$ and $\boldsymbol{y}$ such that $\boldsymbol{c}^{\mathrm{T}}\boldsymbol{x} = \boldsymbol{b}^{\mathrm{T}}\boldsymbol{y}$

# Dualization

## Every linear programming problem has its dual, e.g.

- Maximize $\mathbf{c}^{\mathrm{T}}\mathbf{x}$ subject to $A\mathbf{x} \geq \mathbf{b}$ and $\mathbf{x} \geq \mathbf{0}$
- Maximize $\mathbf{c}^{\mathrm{T}}\mathbf{x}$ subject to $-A\mathbf{x} \leq -\mathbf{b}$ and $\mathbf{x} \geq \mathbf{0}$
- Minimize $-\mathbf{b}^{\mathrm{T}}\mathbf{y}$ subject to $-A^{\mathrm{T}}\mathbf{y} \geq \mathbf{c}$ and $\mathbf{y} \geq \mathbf{0}$
- Minimize $\mathbf{b}^{\mathrm{T}}\mathbf{y}$ subject to $A^{\mathrm{T}}\mathbf{y} \geq \mathbf{c}$ and $\mathbf{y} \leq \mathbf{0}$

## A dual of a dual problem is the (original) primal problem

- Minimize $\mathbf{b}^{\mathrm{T}}\mathbf{y}$ subject to $A^{\mathrm{T}}\mathbf{y} \geq \mathbf{c}$ and $\mathbf{y} \geq \mathbf{0}$
- -Maximize $-\mathbf{b}^{\mathrm{T}}\mathbf{y}$ subject to $A^{\mathrm{T}}\mathbf{y} \geq \mathbf{c}$ and $\mathbf{y} \geq \mathbf{0}$
- -Minimize $\mathbf{c}^{\mathrm{T}}\mathbf{x}$ subject to $A\mathbf{x} \geq -\mathbf{b}$ and $\mathbf{x} \leq \mathbf{0}$
- -Minimize $-\mathbf{c}^{\mathrm{T}}\mathbf{x}$ subject to $-A\mathbf{x} \geq -\mathbf{b}$ and $\mathbf{x} \geq \mathbf{0}$
- Maximize $\mathbf{c}^{\mathrm{T}}\mathbf{x}$ subject to $A\mathbf{x} \leq \mathbf{b}$ and $\mathbf{x} \geq \mathbf{0}$

# Dualization: General rules

| | Primal linear program | Dual linear program |
|---|---|---|
| Variables | $\boldsymbol{x}_1, \ldots, \boldsymbol{x}_n$ | $\boldsymbol{y}_1, \ldots, \boldsymbol{y}_m$ |
| Matrix | $A$ | $A^{\mathrm{T}}$ |
| Right-hand side | $\boldsymbol{b}$ | $\boldsymbol{c}$ |
| Objective function | $\max \boldsymbol{c}^{\mathrm{T}}\boldsymbol{x}$ | $\min \boldsymbol{b}^{\mathrm{T}}\boldsymbol{y}$ |
| Constraints | $i$-the constraint has $\leq$ | $\boldsymbol{y}_i \geq 0$ |
| | $i$-the constraint has $\geq$ | $\boldsymbol{y}_i \leq 0$ |
| | $i$-the constraint has $=$ | $\boldsymbol{y}_i \in \mathbb{R}$ |
| | $\boldsymbol{x}_j \geq 0$ | $j$-th constraint has $\geq$ |
| | $\boldsymbol{x}_j \leq 0$ | $j$-th constraint has $\leq$ |
| | $\boldsymbol{x}_j \in \mathbb{R}$ | $j$-th constraint has $=$ |

## Feasibility versus optimality

Finding a feasible solution of a linear program is computationally as difficult as finding an optimal solution.

## Using duality

The optimal solutions of linear programs

- Primal: Maximize $\mathbf{c}^\mathrm{T}\mathbf{x}$ subject to $A\mathbf{x} \leq \mathbf{b}$ and $\mathbf{x} \geq \mathbf{0}$
- Dual: Minimize $\mathbf{b}^\mathrm{T}\mathbf{y}$ subject to $A^\mathrm{T}\mathbf{y} \geq \mathbf{c}$ and $\mathbf{y} \geq \mathbf{0}$

are exactly feasible solutions satisfying

$$\begin{aligned} A\mathbf{x} &\leq \mathbf{b} \\ A^\mathrm{T}\mathbf{y} &\geq \mathbf{c} \\ \mathbf{c}^\mathrm{T}\mathbf{x} &\geq \mathbf{b}^\mathrm{T}\mathbf{y} \\ \mathbf{x}, \mathbf{y} &\geq \mathbf{0} \end{aligned}$$

# Complementary slackness

## Theorem

Feasible solutions $\boldsymbol{x}$ and $\boldsymbol{y}$ of linear programs

- Primal: Maximize $\boldsymbol{c}^{\mathrm{T}}\boldsymbol{x}$ subject to $A\boldsymbol{x} \leq \boldsymbol{b}$ and $\boldsymbol{x} \geq \boldsymbol{0}$
- Dual: Minimize $\boldsymbol{b}^{\mathrm{T}}\boldsymbol{y}$ subject to $A^{\mathrm{T}}\boldsymbol{y} \geq \boldsymbol{c}$ and $\boldsymbol{y} \geq \boldsymbol{0}$

are optimal if and only if

- $\boldsymbol{x}_i = 0$ or $A^{\mathrm{T}}_{i,\star}\boldsymbol{y} = \boldsymbol{c}_i$ for every $i = 1, \ldots, n$ and
- $\boldsymbol{y}_j = 0$ or $\mathrm{A}_{j,\star}\boldsymbol{x} = \boldsymbol{b}_j$ for every $j = 1, \ldots, m$.

## Proof

$$\boldsymbol{c}^{\mathrm{T}}\boldsymbol{x} = \sum_{i=1}^{n} \boldsymbol{c}_i\boldsymbol{x}_i \leq \sum_{i=1}^{n} (\boldsymbol{y}^{\mathrm{T}}A_{\star,i})\boldsymbol{x}_i = \boldsymbol{y}^{\mathrm{T}}A\boldsymbol{x} = \sum_{j=1}^{m} \boldsymbol{y}_j(A_{j,\star}\boldsymbol{x}) \leq \sum_{j=1}^{m} \boldsymbol{y}_j\boldsymbol{b}_j = \boldsymbol{b}^{\mathrm{T}}\boldsymbol{y}$$

# Fourier–Motzkin elimination: Example

## Goal: Find a feasible solution

$$\begin{array}{rcrcrcr} 2x & - & 5y & + & 4z & \leq & 10 \\ 3x & - & 6y & + & 3z & \leq & 9 \\ 5x & + & 10y & - & z & \leq & 15 \\ -x & + & 5y & - & 2z & \leq & -7 \\ -3x & + & 2y & + & 6z & \leq & 12 \end{array}$$

## Express the variable $x$ in each condition

$$\begin{array}{rcrcrcr} x & \leq & 5 & + & \frac{5}{2}y & - & 2z \\ x & \leq & 3 & + & 2y & - & z \\ x & \leq & 3 & - & 2y & + & \frac{1}{5}z \\ x & \geq & 7 & + & 5y & - & 2z \\ x & \geq & -4 & + & \frac{2}{3}y & + & 2z \end{array}$$

## Eliminate the variable $x$

The original system has a feasible solution if and only if there exist $y$ and $z$ satisfying

$$\max\left\{7 + 5y - 2z, -4 + \frac{2}{3}y + 2z\right\} \leq \min\left\{5 + \frac{5}{2}y - 2z, 3 + 2y - z, 3 - 2y + \frac{1}{5}z\right\}$$

# Fourier–Motzkin elimination: Example

## Rewrite into a system of inequalities

Real numbers $x$ and $y$ satisfy

$$\max\left\{7 + 5y - 2z, -4 + \frac{2}{3}y + 2z\right\} \le \min\left\{5 + \frac{5}{2}y - 2z, 3 + 2y - z, 3 - 2y + \frac{1}{5}z\right\}$$

if and only they satisfy

$$\begin{array}{rcrcrcrcrcrcr}
7 & + & 5y & - & 2z & \le & 5 & + & \frac{5}{2}y & - & 2z \\
7 & + & 5y & - & 2z & \le & 3 & + & 2y & - & z \\
7 & + & 5y & - & 2z & \le & 3 & - & 2y & + & \frac{1}{5}z \\
-4 & + & \frac{2}{3}y & + & 2z & \le & 5 & + & \frac{5}{2}y & - & 2z \\
-4 & + & \frac{2}{3}y & + & 2z & \le & 3 & + & 2y & - & z \\
-4 & + & \frac{2}{3}y & + & 2z & \le & 3 & - & 2y & + & \frac{1}{5}z
\end{array}$$

## Next steps

Eliminate variables $y$ and $z$ in a similar way.

# Fourier–Motzkin elimination: In general

## Observation

Let $A\boldsymbol{x} \leq \boldsymbol{b}$ be a system with $n \geq 1$ variables and $m$ inequalities. There is a system $A'\boldsymbol{x}' \leq \boldsymbol{b}'$ with $n-1$ variables and at most $\max\left\{m, m^2/4\right\}$ inequalities, with the following properties:

1. $A\boldsymbol{x} \leq \boldsymbol{b}$ has a solution if and only if $A'\boldsymbol{x}' \leq \boldsymbol{b}'$ has a solution, and
2. each inequality of $A'\boldsymbol{x}' \leq \boldsymbol{b}'$ is a positive linear combination of some inequalities from $A\boldsymbol{x} \leq \boldsymbol{b}$.

## Proof

1. WLOG: $A_{i,1} \in \{-1, 0, 1\}$ for all $i = 1, \ldots, n$
2. Let $C = \{i;\ A_{i,1} = 1\}$, $F = \{i;\ A_{i,1} = -1\}$ and $L = \{i;\ A_{i,1} = 0\}$
3. Let $A'\boldsymbol{x}' \leq \boldsymbol{b}'$ be the system of $n-1$ variables and $|C| \cdot |F| + |L|$ inequalities

$$\begin{array}{lrcll} j \in C, k \in F: & (A_{j,\star} + A_{k,\star})\boldsymbol{x} & \leq & \boldsymbol{b}_j + \boldsymbol{b}_k & (1) \\ l \in L: & A_{l,\star}\boldsymbol{x} & \leq & \boldsymbol{b}_l & (2) \end{array}$$

4. Assuming $A'\boldsymbol{x}' \leq \boldsymbol{b}'$ has a solution $\boldsymbol{x}'$, we find a solution $\boldsymbol{x}$ of $A\boldsymbol{x} \leq \boldsymbol{b}$:
   - (1) is equivalent to $A'_{k,\star}\boldsymbol{x}' - \boldsymbol{b}_k \leq \boldsymbol{b}_j - A'_{j,\star}\boldsymbol{x}'$ for all $j \in C, k \in F$,
   - which is equivalent to $\max_{k \in F}\left\{A'_{k,\star}\boldsymbol{x}' - \boldsymbol{b}_k\right\} \leq \min_{j \in C}\left\{\boldsymbol{b}_j - A'_{j,\star}\boldsymbol{x}'\right\}$
   - Choose $\boldsymbol{x}_1$ between these bounds and $\boldsymbol{x} = (\boldsymbol{x}_1, \boldsymbol{x}')$ satisfies $A\boldsymbol{x} \leq \boldsymbol{b}$

# Farkas lemma

## Definition

A cone generated by vectors $\boldsymbol{a}_1, \dots, \boldsymbol{a}_n \in \mathbb{R}^m$ is the set of all non-negative combinations of $\boldsymbol{a}_1, \dots, \boldsymbol{a}_n$, i.e. $\left\{\sum_{i=1}^n \alpha_i \boldsymbol{a}_i;\ \alpha_1, \dots, \alpha_n \geq 0\right\}$.

## Proposition (Farkas lemma geometrically)

Let $\boldsymbol{a}_1, \dots, \boldsymbol{a}_n, \boldsymbol{b} \in \mathbb{R}^m$. Then exactly one of the following two possibilities occurs:

1. The point $\boldsymbol{b}$ lies in the cone generated by $\boldsymbol{a}_1, \dots, \boldsymbol{a}_n$.
2. There exists a hyperplane $h = \left\{\boldsymbol{x} \in \mathbb{R}^m;\ \boldsymbol{y}^{\mathrm{T}}\boldsymbol{x} = 0\right\}$ containing $\mathbf{0}$ for some $\boldsymbol{y} \in \mathbb{R}^m$ separating $\boldsymbol{a}_1, \dots, \boldsymbol{a}_n$ and $\boldsymbol{b}$, i.e. $\boldsymbol{y}^{\mathrm{T}}\boldsymbol{a}_i \geq 0$ for all $i = 1, \dots, n$ and $\boldsymbol{y}^{\mathrm{T}}\boldsymbol{b} < 0$.

## Proposition (Farkas lemma)

Let $A \in R^{m \times n}$ and $\boldsymbol{b} \in \mathbb{R}^m$. Then exactly one of the following two possibilities occurs:

1. There exists a vector $\boldsymbol{x} \in \mathbb{R}^n$ satisfying $A\boldsymbol{x} = \boldsymbol{b}$ and $\boldsymbol{x} \geq \mathbf{0}$.
2. There exists a vector $\boldsymbol{y} \in \mathbb{R}^m$ satisfying $\boldsymbol{y}^{\mathrm{T}}A \geq \mathbf{0}$ and $\boldsymbol{y}^{\mathrm{T}}b < \mathbf{0}$.

## Proposition (Farkas lemma)

Let $A \in R^{m \times n}$ and $\boldsymbol{b} \in \mathbb{R}^m$. The following statements hold.

1. The system $A\boldsymbol{x} = \boldsymbol{b}$ has a non-negative solution $\boldsymbol{x} \in \mathbb{R}^n$ if and only if every $\boldsymbol{y} \in \mathbb{R}^m$ with $\boldsymbol{y}^{\mathrm{T}} A \geq \mathbf{0}^{\mathrm{T}}$ satisfies $\boldsymbol{y}^{\mathrm{T}} \boldsymbol{b} \geq 0$.
2. The system $A\boldsymbol{x} \leq \boldsymbol{b}$ has a non-negative solution $\boldsymbol{x} \in \mathbb{R}^n$ if and only if every non-negative $\boldsymbol{y} \in \mathbb{R}^m$ with $\boldsymbol{y}^{\mathrm{T}} A \geq \mathbf{0}^{\mathrm{T}}$ satisfies $\boldsymbol{y}^{\mathrm{T}} \boldsymbol{b} \geq 0$.
3. The system $A\boldsymbol{x} \leq \boldsymbol{b}$ has a solution $\boldsymbol{x} \in \mathbb{R}^n$ if and only if every non-negative $\boldsymbol{y} \in \mathbb{R}^m$ with $\boldsymbol{y}^{\mathrm{T}} A = \mathbf{0}^{\mathrm{T}}$ satisfies $\boldsymbol{y}^{\mathrm{T}} \boldsymbol{b} \geq 0$.

## Overview of the proof of duality

Fourier–Motzkin elimination

$\Downarrow$

Farkas lemma, 3rd version

$\Downarrow$

Farkas lemma, 2nd version

$\Downarrow$

Duality of linear programming

# Farkas lemma

## Proposition (Farkas lemma, 3rd version)

Let $A \in R^{m \times n}$ and $\boldsymbol{b} \in \mathbb{R}^m$. Then, the system $A\boldsymbol{x} \leq \boldsymbol{b}$ has a solution $\boldsymbol{x} \in \mathbb{R}^n$ if and only if every non-negative $\boldsymbol{y} \in \mathbb{R}^m$ with $\boldsymbol{y}^{\mathrm{T}} A = \mathbf{0}^{\mathrm{T}}$ satisfies $\boldsymbol{y}^{\mathrm{T}} \boldsymbol{b} \geq 0$.

## Proof

$\Rightarrow$ If $\boldsymbol{x}$ satisfies $A\boldsymbol{x} \leq \boldsymbol{b}$ and $\boldsymbol{y} \geq \mathbf{0}$ satifies $\boldsymbol{y}^{\mathrm{T}} A = \mathbf{0}^{\mathrm{T}}$, then $\boldsymbol{y}^{\mathrm{T}} \boldsymbol{b} \geq \boldsymbol{y}^{\mathrm{T}} A\boldsymbol{x} \geq \mathbf{0}^{\mathrm{T}} \boldsymbol{x} = \mathbf{0}$

$\Leftarrow$ If $A\boldsymbol{x} \leq \boldsymbol{b}$ has no solution, the find $\boldsymbol{y} \geq \mathbf{0}$, $\boldsymbol{y}^{\mathrm{T}} A = \mathbf{0}^{\mathrm{T}}$, $\boldsymbol{y}^{\mathrm{T}} \boldsymbol{b} < 0$ by the induction on $n$

$n = 0$
- The system $A\boldsymbol{x} \leq \boldsymbol{b}$ equals to $\mathbf{0} \leq \boldsymbol{b}$ which is infeasible, so $b_i < 0$ for some $i$
- Choose $\boldsymbol{y} = e_i$ (the $i$-th unit vector)

$n > 0$
- Using Fourier–Motzkin elimination we obtain an infeasible system $A'\boldsymbol{x}' \leq \boldsymbol{b}'$
- There exists a non-negative matrix $M$ such that $(\mathbf{0}|A') = MA$ and $\boldsymbol{b}' = M\boldsymbol{b}$
- By induction, there exists $\boldsymbol{y}' \geq 0$, $\boldsymbol{y}'^{\mathrm{T}} A' = \mathbf{0}^{\mathrm{T}}$, $\boldsymbol{y}'^{\mathrm{T}} \boldsymbol{b}' < 0$
- We verify that $\boldsymbol{y} = M^{\mathrm{T}} \boldsymbol{y}'$ satifies all requirements of the induction

  $\boldsymbol{y} = M^{\mathrm{T}} \boldsymbol{y}' \geq \mathbf{0}$

  $\boldsymbol{y}^{\mathrm{T}} A = (M^{\mathrm{T}} \boldsymbol{y}')^{\mathrm{T}} A = \boldsymbol{y}'^{\mathrm{T}} MA = \boldsymbol{y}'^{\mathrm{T}} (\mathbf{0}|A') = \mathbf{0}^{\mathrm{T}}$

  $\boldsymbol{y}^{\mathrm{T}} \boldsymbol{b} = (M^{\mathrm{T}} \boldsymbol{y}')^{\mathrm{T}} \boldsymbol{b} = \boldsymbol{y}'^{\mathrm{T}} M\boldsymbol{b} = \boldsymbol{y}'^{\mathrm{T}} \boldsymbol{b}' < \mathbf{0}^{\mathrm{T}}$

# Farkas lemma

## Proposition (Farkas lemma, 3rd version)

Let $A \in R^{m \times n}$ and $\boldsymbol{b} \in \mathbb{R}^m$. Then, the system $A\boldsymbol{x} \leq \boldsymbol{b}$ has a solution $\boldsymbol{x} \in \mathbb{R}^n$ if and only if every non-negative $\boldsymbol{y} \in \mathbb{R}^m$ with $\boldsymbol{y}^{\mathrm{T}} A = \mathbf{0}^{\mathrm{T}}$ satisfies $\boldsymbol{y}^{\mathrm{T}} \boldsymbol{b} \geq 0$.

## Proposition (Farkas lemma, 2nd version)

Let $A \in R^{m \times n}$ and $\boldsymbol{b} \in \mathbb{R}^m$. The system $A\boldsymbol{x} \leq \boldsymbol{b}$ has a non-negative solution $\boldsymbol{x} \in \mathbb{R}^n$ if and only if every non-negative $\boldsymbol{y} \in \mathbb{R}^m$ with $\boldsymbol{y}^{\mathrm{T}} A \geq \mathbf{0}^{\mathrm{T}}$ satisfies $\boldsymbol{y}^{\mathrm{T}} \boldsymbol{b} \geq 0$.

## Proof of the 2nd version using the 3rd version

The following statements are equivalent

1. $A\boldsymbol{x} \leq \boldsymbol{b}$, $\boldsymbol{x} \geq \mathbf{0}$ has a solution
2. $\binom{A}{-I}\boldsymbol{x} \leq \binom{\boldsymbol{b}}{\mathbf{0}}$ has a solution
3. Every $\boldsymbol{y}, \boldsymbol{y}' \geq \mathbf{0}$ with $\binom{\boldsymbol{y}}{\boldsymbol{y}'}^{\mathrm{T}} \binom{A}{-I} = \mathbf{0}^{\mathrm{T}}$ satisfies $\binom{\boldsymbol{y}}{\boldsymbol{y}'}^{\mathrm{T}} \binom{\boldsymbol{b}}{\mathbf{0}} \geq 0$
4. Every $\boldsymbol{y}, \boldsymbol{y}' \geq \mathbf{0}$ with $\boldsymbol{y}^{\mathrm{T}} A = \boldsymbol{y}'$ satisfies $\boldsymbol{y}^{\mathrm{T}} \boldsymbol{b} \geq 0$
5. Every $\boldsymbol{y} \geq \mathbf{0}$ with $\boldsymbol{y}^{\mathrm{T}} A \geq \mathbf{0}$ satisfies $\boldsymbol{y}^{\mathrm{T}} \boldsymbol{b} \geq 0$

# Proof of the duality of linear programming

## Proposition (Farkas lemma, 2nd version)

Let $A \in R^{m \times n}$ and $\boldsymbol{b} \in \mathbb{R}^m$. The system $A\boldsymbol{x} \leq \boldsymbol{b}$ has a non-negative solution if and only if every non-negative $\boldsymbol{y} \in \mathbb{R}^m$ with $\boldsymbol{y}^{\mathrm{T}} A \geq \mathbf{0}^{\mathrm{T}}$ satisfies $\boldsymbol{y}^{\mathrm{T}} \boldsymbol{b} \geq 0$.

## Duality

- Primal: Maximize $\boldsymbol{c}^{\mathrm{T}} \boldsymbol{x}$ subject to $A\boldsymbol{x} \leq \boldsymbol{b}$ and $\boldsymbol{x} \geq \mathbf{0}$
- Dual: Minimize $\boldsymbol{b}^{\mathrm{T}} \boldsymbol{y}$ subject to $A^{\mathrm{T}} \boldsymbol{y} \geq \boldsymbol{c}$ and $\boldsymbol{y} \geq \mathbf{0}$

If the primal problem has an optimal solution $\boldsymbol{x}^\star$, then the dual problem has an optimal solution $\boldsymbol{y}^\star$ and $\boldsymbol{c}^{\mathrm{T}} \boldsymbol{x}^\star = \boldsymbol{b}^{\mathrm{T}} \boldsymbol{y}^\star$.

## Proof of duality using Farkas lemma

1. Let $\boldsymbol{x}^\star$ be an optimal solution of the primal problem and $\gamma = \boldsymbol{c}^{\mathrm{T}} \boldsymbol{x}^\star$
2. $\epsilon > 0$ iff $A\boldsymbol{x} \leq \boldsymbol{b}$ and $\boldsymbol{x} \geq \mathbf{0}$ and $\boldsymbol{c}^{\mathrm{T}} \boldsymbol{x} \geq \gamma + \epsilon$ is infeasible
3. $\epsilon > 0$ iff $\binom{A}{-\boldsymbol{c}^{\mathrm{T}}} \boldsymbol{x} \leq \binom{\boldsymbol{b}}{-\gamma-\epsilon}$ and $\boldsymbol{x} \geq \mathbf{0}$ is infeasible
4. $\epsilon > 0$ iff $\boldsymbol{u}, z \geq 0$ and $\binom{\boldsymbol{u}}{z}^{\mathrm{T}} \binom{A}{-\boldsymbol{c}^{\mathrm{T}}} \geq \mathbf{0}^{\mathrm{T}}$ and $\binom{\boldsymbol{u}}{z}^{\mathrm{T}} \binom{\boldsymbol{b}}{-\gamma-\epsilon} < 0$ is feasible
5. $\epsilon > 0$ iff $\boldsymbol{u}, z \geq 0$ and $A^{\mathrm{T}} \boldsymbol{u} \geq z\boldsymbol{c}$ and $\boldsymbol{b}^{\mathrm{T}} \boldsymbol{u} < z(\gamma + \epsilon)$ is feasible

# Proof of the duality of linear programming

## Duality

- Primal: Maximize $\boldsymbol{c}^{\mathrm{T}}\boldsymbol{x}$ subject to $A\boldsymbol{x} \leq \boldsymbol{b}$ and $\boldsymbol{x} \geq \mathbf{0}$
- Dual: Minimize $\boldsymbol{b}^{\mathrm{T}}\boldsymbol{y}$ subject to $A^{\mathrm{T}}\boldsymbol{y} \geq \boldsymbol{c}$ and $\boldsymbol{y} \geq \mathbf{0}$

If the primal problem has an optimal solution $\boldsymbol{x}^\star$, then the dual problem has an optimal solution $\boldsymbol{y}^\star$ and $\boldsymbol{c}^{\mathrm{T}}\boldsymbol{x}^\star = \boldsymbol{b}^{\mathrm{T}}\boldsymbol{y}^\star$.

## Proof of duality using Farkas lemma (continue)

1. Let $\boldsymbol{x}^\star$ be an optimal solution of the primal problem and $\gamma = \boldsymbol{c}^{\mathrm{T}}\boldsymbol{x}^\star$
2. $\epsilon > 0$ iff $\boldsymbol{u}, z \geq 0$ and $A^{\mathrm{T}}\boldsymbol{u} \geq z\boldsymbol{c}$ and $\boldsymbol{b}^{\mathrm{T}}\boldsymbol{u} < z(\gamma + \epsilon)$ is feasible
3. For $\epsilon > 0$, there exists $\boldsymbol{u}', z' \geq 0$ with $A^{\mathrm{T}}\boldsymbol{u}' \geq z'\boldsymbol{c}$ and $\boldsymbol{b}^{\mathrm{T}}\boldsymbol{u}' < z'(\gamma + \epsilon)$
4. For $\epsilon = 0$ it holds that $\boldsymbol{u}', z' \geq 0$ and $A^{\mathrm{T}}\boldsymbol{u}' \geq z'\boldsymbol{c}$ so $\boldsymbol{b}^{\mathrm{T}}\boldsymbol{u}' \geq z'\gamma$
5. Since $z'\gamma \leq \boldsymbol{b}^{\mathrm{T}}\boldsymbol{u}' < z'(\gamma + \epsilon)$ and $z' \geq 0$ it follows that $z' > 0$
6. Let $\boldsymbol{v} = \frac{1}{z'}\boldsymbol{u}'$
7. Since $A^{\mathrm{T}}\boldsymbol{v} \geq \boldsymbol{c}$ and $\boldsymbol{v} \geq \mathbf{0}$, the dual solution $\boldsymbol{v}$ is feasible
8. Since the dual is feasible and bounded, there exists an optimal dual solution $\boldsymbol{y}^\star$
9. Hence, $\boldsymbol{b}^{\mathrm{T}}\boldsymbol{y}^\star < \gamma + \epsilon$ for every $\epsilon > 0$, and so $\boldsymbol{b}^{\mathrm{T}}\boldsymbol{y}^\star \leq \gamma$
10. From the weak duality theorem it follows that $\boldsymbol{b}^{\mathrm{T}}\boldsymbol{y}^\star = \boldsymbol{c}^{\mathrm{T}}\boldsymbol{x}^\star$

# Outline

# Integer linear programming

## Integer linear programming

Integer linear programming problem is an optimization problem to find $\boldsymbol{x} \in \mathbb{Z}^n$ which maximizes $\boldsymbol{c}^{\mathrm{T}}\boldsymbol{x}$ and satisfies $A\boldsymbol{x} \leq \boldsymbol{b}$ where $A \in \mathbb{R}^{m \times n}$ and $\boldsymbol{b} \in \mathbb{R}^m$.

## Mix integer linear programming

Some variables are integer and others are real.

## Relaxed problem and solution

- Given a (mix) integer linear programming problem, the corresponding relaxed problem is the linear programming problem where all integral constraints $\boldsymbol{x}_i \in \mathbb{Z}$ are relaxed; that is, replaced by $\boldsymbol{x}_i \in \mathbb{R}$.
- Relaxed solution is a feasible solution of the relaxed problem.
- Optimal relaxed solution is the optimal feasible solution of the relaxed problem.

## Observation

Let $\boldsymbol{x}^\star$ be an integral optimal solution and $\boldsymbol{x}^r$ be a relaxed optimal solution. Then, $\boldsymbol{c}^{\mathrm{T}}\boldsymbol{x}^r \geq \boldsymbol{c}^{\mathrm{T}}\boldsymbol{x}^\star$.

# Branch and bound

## Branch

Consider a mix integer linear programming problem
$\max\{\boldsymbol{x} \in \mathbb{R}^n;\ A\boldsymbol{x} \leq \boldsymbol{b},\ \boldsymbol{x}_i \in \mathbb{Z},\ i \in I\}$ where $I$ is a set of integral variables.

- Let $\boldsymbol{x}^r$ be the optimal relaxed solution.
- If $\boldsymbol{x}_i^r \in \mathbb{Z}$ for all $i \in I$, then $\boldsymbol{x}^r$ is an optimal solution.
- Otherwise, choose $j \in I$ such that $\boldsymbol{x}_j^r \notin \mathbb{Z}$ and
- recursively solve two subproblems
  - $\max\left\{\boldsymbol{x} \in \mathbb{R}^n;\ A\boldsymbol{x} \leq \boldsymbol{b},\ \boldsymbol{x}_j \leq \left\lfloor \boldsymbol{x}_j^r \right\rfloor,\ \boldsymbol{x}_i \in \mathbb{Z},\ i \in I\right\}$ and
  - $\max\left\{\boldsymbol{x} \in \mathbb{R}^n;\ A\boldsymbol{x} \leq \boldsymbol{b},\ \boldsymbol{x}_j \geq \left\lceil \boldsymbol{x}_j^r \right\rceil,\ \boldsymbol{x}_i \in \mathbb{Z},\ i \in I\right\}$.
- The optimal solution of the original problem is the better one of subproblems.

## Bound

Let $\boldsymbol{x}'$ be an integral feasible solution and $\boldsymbol{x}^r$ be an optimal relaxed solution of a subproblem. If $\boldsymbol{c}^{\mathrm{T}}\boldsymbol{x}' \geq \boldsymbol{c}^{\mathrm{T}}\boldsymbol{x}^r$, then the subproblem does not contain better integral feasible solution than $\boldsymbol{x}'$.

## Observation

If the polyhedron $\{\boldsymbol{x} \in \mathbb{R}^n;\ A\boldsymbol{x} \leq \boldsymbol{b}\}$ is bounded, then the Brand and bound algorithm finds an optimal solution of the mix integer linear programming problem.

# Rational and integral polyhedrons

## Definition: Rational polyhedron

A polyhedron is called rational if it is defined by a rational linear system, that is $A \in \mathbb{Q}^{m \times n}$ and $\boldsymbol{b} \in \mathbb{Q}^m$.

## Exercise

Every vertex of a rational polyhedron is rational.

## Definition: Integral polyhedron

A rational polyhedron is called integral if every non-empty face contains an integral point.

## Observation

Let $P$ be a rational polyhedron which has a vertex. Then, $P$ is integral if and only if every vertex of $P$ is integral.

## Theorem

A rational polytope $P$ is integral if and only if for all integral vector $\boldsymbol{c}$ the optimal value of $\max \left\{\boldsymbol{c}^{\mathrm{T}}\boldsymbol{x};\ \boldsymbol{x} \in P\right\}$ is an integer.

# Rational and integral polyhedrons

## Theorem

A rational polytope $P$ is integral if and only if for all integral vector $\boldsymbol{c}$ the optimal value of $\max\left\{\boldsymbol{c}^{\mathrm{T}}\boldsymbol{x};\ \boldsymbol{x} \in P\right\}$ is an integer.

## Proof

- $\Rightarrow$ Every vertex of $P$ is integral, so optimal values are integrals.
- $\Leftarrow$ Let $\boldsymbol{v}$ be a vertex of $P$. We prove that $\boldsymbol{v}_1$ is an integer.
  1. Let $\boldsymbol{c}$ be an integer vector such that $\boldsymbol{v}$ is the only optimal solution.
  2. Since we can scale the vector $\boldsymbol{c}$, we assume that $\boldsymbol{c}^{\mathrm{T}}\boldsymbol{v} > \boldsymbol{c}^{\mathrm{T}}\boldsymbol{u} + \boldsymbol{u}_1 - \boldsymbol{v}_1$ for all others vertices $u$ of $P$.
  3. Let $\boldsymbol{d} = \boldsymbol{c} + e_1$.
  4. Observe that $\boldsymbol{v}$ is an optimal of solution of $\max\left\{\boldsymbol{d}^{\mathrm{T}}\boldsymbol{x};\ \boldsymbol{x} \in P\right\}$.
  5. Hence, $\boldsymbol{v}_1 = \boldsymbol{d}^{\mathrm{T}}\boldsymbol{v} - \boldsymbol{c}^{\mathrm{T}}\boldsymbol{v}$ is an integer.

# Gomory-Chvátal cutting plane: Example

## Interger linear programming problem

$$\begin{array}{lrcrcr} \text{Maximize} & & & x_2 & & \\ \text{subject to} & 2x_1 & + & 3x_2 & \leq & 27 \\ & 2x_1 & - & 2x_2 & \leq & 7 \\ & -2x_1 & - & 6x_2 & \leq & -11 \\ & -6x_1 & + & 8x_2 & \leq & 21 \\ & \multicolumn{5}{l}{x_1, x_2 \in \mathbb{Z}} \end{array}$$

## Relaxed problem

Optimal relaxed solution is $(\frac{9}{2}, 6)^{\mathrm{T}}$.

## Cutting plane 1

$$\begin{array}{lrcrcr} \text{The last inequality} & -3x_1 & + & 4x_2 & \leq & \frac{21}{2} \\ \text{Every feasible } x \in \mathbb{Z}^2 \text{ satisfies} & -3x_1 & + & 4x_2 & \leq & 10 \end{array}$$

## Cutting plane 2

$$\begin{array}{lrcrcr} \text{Cutting plane 1} & -6x_1 & + & 8x_2 & \leq & 20 \\ \text{The first inequality} & 6x_1 & + & 9x_2 & \leq & 81 \\ \text{Sum} & & & 17x_2 & \leq & 101 \\ \text{Every feasible } x \in \mathbb{Z}^2 \text{ satisfies} & & & x_2 & \leq & 5 \end{array}$$

# Gomory-Chvátal cutting plane proof

## System of inequalities

Consider a system $P = \{\boldsymbol{x};\ A\boldsymbol{x} \le \boldsymbol{b}\}$ with $n$ variables and $m$ inequalities.

## Definition: Gomory-Chvátal cutting plane

- Consider a non-negative linear combination of inequalities $\boldsymbol{y} \in \mathbb{R}^m$
- Let $\boldsymbol{c} = \boldsymbol{y}^{\mathrm{T}} A$ and $d = \boldsymbol{y}^{\mathrm{T}} \boldsymbol{b}$
- Every point $\boldsymbol{x} \in P$ satifies $\boldsymbol{c}^{\mathrm{T}}\boldsymbol{x} \le d$
- Furthermore, if $\boldsymbol{c}$ is integral, every integral point $\boldsymbol{x}$ satisfies $\boldsymbol{c}^{\mathrm{T}}\boldsymbol{x} \le \lfloor d \rfloor$
- The inequality $\boldsymbol{c}^{\mathrm{T}}\boldsymbol{x} \le \lfloor d \rfloor$ is called a Gomory-Chvátal cutting plane

## Definition: Gomory-Chvátal cutting plane proof

A cutting plane proof of an inequality $\boldsymbol{w}^{\mathrm{T}}\boldsymbol{x} \le t$ is a sequence of inequalities $a_{m+k}^{\mathrm{T}}\boldsymbol{x} \le b_{m+k}$ where $k = 1, \ldots, M$ such that

- for each $k = 1, \ldots, M$ the inequality $a_{m+k}^{\mathrm{T}}\boldsymbol{x} \le b_{m+k}$ is a cutting plane derived from the system $a_i^{\mathrm{T}}\boldsymbol{x} \le b_i$ for $i = 1, \ldots, m+k-1$ and
- $\boldsymbol{w}^{\mathrm{T}}\boldsymbol{x} \le t$ is the last inequality $a_{m+M}^{\mathrm{T}}\boldsymbol{x} \le b_{m+M}$.

# Gomory-Chvátal cutting plane: Theorems

## Theorem: Existence of a cutting plane proof for every valid inequality

Let $P = \{\boldsymbol{x};\ A\boldsymbol{x} \le \boldsymbol{b}\}$ be a rational polytope and let $\boldsymbol{w}^{\mathrm{T}}\boldsymbol{x} \le t$ be an inequality with $\boldsymbol{w}^{\mathrm{T}}$ intergal satisfied by all integral vectors in $P$. Then there exists a cutting plane proof of $\boldsymbol{w}^{\mathrm{T}}\boldsymbol{x} \le t'$ from $A\boldsymbol{x} \le \boldsymbol{b}$ for some $t' \le t$.

## Theorem: Cutting plane proof for $\mathbf{0}^{\mathrm{T}}\boldsymbol{x} \le -1$ in polytopes without integral point

Let $P = \{\boldsymbol{x};\ A\boldsymbol{x} \le \boldsymbol{b}\}$ be a rational polytope that contains no integral point. Then there exists a cutting plane proof of $\mathbf{0}^{\mathrm{T}}\boldsymbol{x} \le -1$ from $A\boldsymbol{x} \le \boldsymbol{b}$.

## Lemma

Let $F$ be a face of a rational polytope $P$. If $\boldsymbol{c}^{\mathrm{T}}\boldsymbol{x} \le \lfloor d \rfloor$ is a cutting plane for $F$, then there exists a cutting plane $\boldsymbol{c}'^{\mathrm{T}}\boldsymbol{x} \le d'$ such that

$$F \cap \left\{\boldsymbol{x};\ \boldsymbol{c}'^{\mathrm{T}}\boldsymbol{x} \le \lfloor d' \rfloor\right\} = F \cap \left\{\boldsymbol{x};\ \boldsymbol{c}^{\mathrm{T}}\boldsymbol{x} \le \lfloor d \rfloor\right\}.$$

# Gomory-Chvátal cutting plane: Proof of the lemma

## Lemma

Let $F$ be a face of a rational polytope $P$. If $\boldsymbol{c}^{\mathrm{T}}\boldsymbol{x} \le \lfloor d \rfloor$ is a cutting plane for $F$, then there exists a cutting plane $\boldsymbol{c}'^{\mathrm{T}}\boldsymbol{x} \le d'$ such that

$$F \cap \left\{\boldsymbol{x};\ \boldsymbol{c}'^{\mathrm{T}}\boldsymbol{x} \le \lfloor d' \rfloor\right\} = F \cap \left\{\boldsymbol{x};\ \boldsymbol{c}^{\mathrm{T}}\boldsymbol{x} \le \lfloor d \rfloor\right\}.$$

## Proof

1. Let $P = \{\boldsymbol{x};\ A'\boldsymbol{x} \le \boldsymbol{b}', A''\boldsymbol{x} \le \boldsymbol{b}''\}$ and $F = \{\boldsymbol{x};\ A'\boldsymbol{x} \le \boldsymbol{b}', A''\boldsymbol{x} = \boldsymbol{b}''\}$ where $A''$ and $b''$ are integral
2. Assume $d = \max\{\boldsymbol{c}^{\mathrm{T}}\boldsymbol{x};\ \boldsymbol{x} \in F\}$
3. By Farkas' lemma, there exists vectors $\boldsymbol{y}' \ge \boldsymbol{0}$ and $\boldsymbol{y}''$ such that $\boldsymbol{y}'^{\mathrm{T}}A' + \boldsymbol{y}''^{\mathrm{T}}A'' = \boldsymbol{c}^{\mathrm{T}}$ and $\boldsymbol{y}'^{\mathrm{T}}b' + \boldsymbol{y}''^{\mathrm{T}}b'' = d$
4. $\boldsymbol{c}' = \boldsymbol{c} - \lfloor \boldsymbol{y}'' \rfloor^{\mathrm{T}} A'' = \boldsymbol{y}'^{\mathrm{T}}A' + (\boldsymbol{y}'' - \lfloor \boldsymbol{y}'' \rfloor)^{\mathrm{T}} A''$
   $d' = d - \lfloor \boldsymbol{y}'' \rfloor^{\mathrm{T}} b'' = \boldsymbol{y}'^{\mathrm{T}}b' + (\boldsymbol{y}'' - \lfloor \boldsymbol{y}'' \rfloor)^{\mathrm{T}} b''$
5. Since $\boldsymbol{y}'$ and $(\boldsymbol{y}'' - \lfloor \boldsymbol{y}'' \rfloor)^{\mathrm{T}}$ are non-negative, $\boldsymbol{c}'^{\mathrm{T}}\boldsymbol{x} \le d'$ is a valid inequality for $P$
6. Hence, $F \cap \{\boldsymbol{x};\ \boldsymbol{c}'^{\mathrm{T}}\boldsymbol{x} \le \lfloor d' \rfloor\} = F \cap \{\boldsymbol{x};\ \boldsymbol{c}'^{\mathrm{T}}\boldsymbol{x} \le \lfloor d' \rfloor, \lfloor \boldsymbol{y}''^{\mathrm{T}} \rfloor A''\boldsymbol{x} = \lfloor \boldsymbol{y}''^{\mathrm{T}} \rfloor \boldsymbol{b}''\} = F \cap \{\boldsymbol{x};\ \boldsymbol{c}^{\mathrm{T}}\boldsymbol{x} \le \lfloor d \rfloor\}$.

## Theorem: Cutting plane proof for $\mathbf{0}^{\mathrm{T}}\boldsymbol{x} \leq -1$ in polytopes without integral point

Let $P = \{\boldsymbol{x};\ A\boldsymbol{x} \leq \boldsymbol{b}\}$ be a rational polytope that contains no integral point. Then there exists a cutting plane proof of $\mathbf{0}^{\mathrm{T}}\boldsymbol{x} \leq -1$ from $A\boldsymbol{x} \leq \boldsymbol{b}$.

## Proof

Induction by $\dim(P)$. Trivial for $\dim(P) = 0$. Assume $\dim(P) \geq 1$.

1. Let $\boldsymbol{w}^{\mathrm{T}}\boldsymbol{x} \leq l$ induces a proper face of $P$ and $\bar{P} = \left\{\boldsymbol{x} \in P;\ \boldsymbol{w}^{\mathrm{T}}\boldsymbol{x} \leq \lfloor l \rfloor\right\}$
2. We derive $\mathbf{0}^{\mathrm{T}}\boldsymbol{x} \leq -1$ from $A\boldsymbol{x} \leq \boldsymbol{b}$, $\boldsymbol{w}^{\mathrm{T}}\boldsymbol{x} \leq \lfloor l \rfloor$ by the following two cases
   - If $\bar{P} = \emptyset$, we use Farkas' lemma
   - If $\bar{P} \neq \emptyset$, let $F = \left\{\boldsymbol{x} \in P;\ \boldsymbol{w}^{\mathrm{T}}\boldsymbol{x} = \lfloor l \rfloor\right\}$
     - Since $\dim(F) < \dim(P)$, there exists a cutting plane proof of $\mathbf{0}^{\mathrm{T}}\boldsymbol{x} \leq -1$ from $A\boldsymbol{x} \leq b$, $\boldsymbol{w}^{\mathrm{T}}\boldsymbol{x} = \lfloor l \rfloor$
     - By lemma, there exists a cutting plane proof of $\boldsymbol{c}^{\mathrm{T}}\boldsymbol{x} \leq \lfloor d \rfloor$ such that $\bar{P} \cap \left\{\boldsymbol{x};\ \boldsymbol{c}^{\mathrm{T}}\boldsymbol{x} \leq \lfloor d \rfloor, \boldsymbol{w}^{\mathrm{T}}\boldsymbol{x} = \lfloor l \rfloor\right\} = \emptyset$
     - Applying these sequence of cuts to $\bar{P}$, we obtain $\boldsymbol{w}^{\mathrm{T}}\boldsymbol{x} \leq \lfloor l \rfloor - 1$
     - Repeat these steps on $\bar{P} = \left\{\boldsymbol{x} \in P;\ \boldsymbol{w}^{\mathrm{T}}\boldsymbol{x} \leq \lfloor l \rfloor - 1\right\}$
   - The number of repetitions is finite since $P$ is bounded

## Theorem: Existence of a cutting plane proof for every valid inequality

Let $P = \{\boldsymbol{x};\ A\boldsymbol{x} \le \boldsymbol{b}\}$ be a rational polytope and let $\boldsymbol{w}^{\mathrm{T}}\boldsymbol{x} \le t$ be an inequality with $\boldsymbol{w}^{\mathrm{T}}$ integral satisfied by all integral vectors in $P$. Then there exists a cutting plane proof of $\boldsymbol{w}^{\mathrm{T}}\boldsymbol{x} \le t'$ from $A\boldsymbol{x} \le \boldsymbol{b}$ for some $t' \le t$.

## Proof

Let $l = \max\left\{\boldsymbol{w}^{\mathrm{T}}\boldsymbol{x};\ \boldsymbol{x} \in P\right\}$ and $\bar{P} = \left\{\boldsymbol{x} \in P;\ \boldsymbol{w}^{\mathrm{T}}\boldsymbol{x} \le \lfloor l \rfloor\right\}$

- If $P$ contains no integer point, then there exists a cutting plane proof of $\mathbf{0}^{\mathrm{T}}\boldsymbol{x} \le -1$ and $\boldsymbol{w}^{\mathrm{T}}\boldsymbol{x} \le t'$ for some $t' \le t$
- If $P$ contains an integral point, then:
  1. If $\lfloor l \rfloor \le t$, we are finished, so we suppose not
  2. $F = \left\{\boldsymbol{x} \in \bar{P} : \boldsymbol{w}^{\mathrm{T}}\boldsymbol{x} = \lfloor l \rfloor\right\}$ is a face of $\bar{P}$
  3. Since $F$ has no integral point, we derive $\mathbf{0}^{\mathrm{T}}\boldsymbol{x} \le -1$ from $A\boldsymbol{x} \le \boldsymbol{b}$, $\boldsymbol{w}^{\mathrm{T}}\boldsymbol{x} = \lfloor l \rfloor$
  4. By lemma, there exists a cutting plane proof of $\boldsymbol{c}^{\mathrm{T}}\boldsymbol{x} \le \lfloor d \rfloor$ from $A\boldsymbol{x} \le \boldsymbol{b}$, $\boldsymbol{w}^{\mathrm{T}}\boldsymbol{x} \le \lfloor l \rfloor$ such that $\bar{P} \cap \left\{\boldsymbol{x};\ \boldsymbol{c}^{\mathrm{T}}\boldsymbol{x} \le \lfloor d \rfloor, \boldsymbol{w}^{\mathrm{T}}\boldsymbol{x} = \lfloor l \rfloor\right\} = \emptyset$
  5. We apply this sequence of cuts to $\bar{P}$ to obtain cutting plane $\boldsymbol{w}^{\mathrm{T}}\boldsymbol{x} \le \lfloor l \rfloor - 1$
  6. Now, we continue until we derive $\boldsymbol{w}^{\mathrm{T}}\boldsymbol{x} \le t'$ for some $t' \le t$

# Total unimodularity

## Questions

- How to recognise whether a polytope $P = \{\boldsymbol{x}; A\boldsymbol{x} \le \boldsymbol{b}\}$ is integral?
- When $P$ is integral for every integral vector $\boldsymbol{b}$?

## Proposition

Let $A \in \mathbb{R}^{m \times m}$ be an integral and regular matrix. Then, $A^{-1}b$ is integral for every integral vector $\boldsymbol{b} \in \mathbb{R}^m$ if and only if $\det(A) \in \{1, -1\}$.

## Proof

$\Leftarrow$ By Cramer's rule, $A^{-1}$ is integral, so $A^{-1}b$ is integral for every integral $\boldsymbol{b}$

$\Rightarrow$

- $A^{-1}_{\star,i} = A^{-1}e_i$ is integral for every $i = 1, \dots, m$
- Since $A$ and $A^{-1}$ are integral, also $\det(A)$ and $\det(A^{-1})$ are both integers
- From $1 = \det(A) \cdot \det(A^{-1})$ it follows that $\det(A) = \det(A^{-1}) \in \{1, -1\}$

# Unimodular matrix

## Definition

A full row rank matrix $A$ is unimodular if $A$ is integral and each basis of $A$ has determinant $\pm 1$.

## Theorem

Let $A \in \mathbb{R}^{m \times n}$ be an integral full row rank matrix. Then, the polyhedron $P = \{\boldsymbol{x};\ A\boldsymbol{x} = \boldsymbol{b}, \boldsymbol{x} \geq \boldsymbol{0}\}$ is integral for every integral vector $\boldsymbol{b}$ if and only if $A$ is unimodular.

## Proof

$\Leftarrow$

- Let $\boldsymbol{b}$ be an integral vector and let $\boldsymbol{x}'$ be a vertex of $P$
- Columns of $A$ corresponding to non-zero components of $\boldsymbol{x}'$ are linearly independent and we extend these columns into a basis $A_B$
- Hence, $\boldsymbol{x}'_B = A_B^{-1}\boldsymbol{b}$ is integral and $\boldsymbol{x}'_N = \boldsymbol{0}$

$\Rightarrow$

- We prove that $A_B^{-1}\boldsymbol{v}$ is integral for every base $B$ and integral vector $\boldsymbol{v}$
- Let $\boldsymbol{y}$ be integral vector such that $\boldsymbol{y} + A_B^{-1}\boldsymbol{v} \geq 0$
- Let $\boldsymbol{b} = A_B(\boldsymbol{y} + A_B^{-1}\boldsymbol{v}) = A_B\boldsymbol{y} + \boldsymbol{v}$ which is integral
- Let $\boldsymbol{z}_B = \boldsymbol{y} + B^{-1}\boldsymbol{v}$ and $\boldsymbol{z}_N = \boldsymbol{0}$
- From $A\boldsymbol{z} = A_B(\boldsymbol{y} + B^{-1}\boldsymbol{v}) = \boldsymbol{b}$ and $\boldsymbol{z} \geq \boldsymbol{0}$, it follows that $\boldsymbol{z} \in P$ and $\boldsymbol{z}$ is a vertex of $P$
- Hence, $A_B^{-1}\boldsymbol{v} = \boldsymbol{z}_B - \boldsymbol{y}$ is integral

# Totally unimodular matrix

## Definition

A matrix is totally unimodular if all of its square submatrices have determinant 0, 1 or $-1$.

## Exercise

Prove that every element of a totally unimodular matrix is 0, 1 or $-1$.
Find a matrix $A \in \{0, 1, -1\}^{m \times n}$ which is not totally unimodular.

## Exercise

Prove that $A$ is totally unimodular if and only if $(A|I)$ is unimodular.

# Totally unimodular matrix

## Theorem: Hoffman-Kruskal

Let $A \in \mathbb{Z}^{m \times n}$ and $P = \{\boldsymbol{x};\ A\boldsymbol{x} \le \boldsymbol{b}, \boldsymbol{x} \ge \boldsymbol{0}\}$. The polyhedron $P$ is integral for every integral $\boldsymbol{b}$ if and only if $A$ is totally unimodular.

## Proof

Adding slack variables, we observe that the following statements are equivalent.

1. $\{\boldsymbol{x};\ A\boldsymbol{x} \le \boldsymbol{b}, \boldsymbol{x} \ge \boldsymbol{0}\}$ is integral for every integral $\boldsymbol{b}$
2. $\{\boldsymbol{x};\ (A|I)\boldsymbol{z} = \boldsymbol{b}, \boldsymbol{z} \ge \boldsymbol{0}\}$ is integral for every integral $\boldsymbol{b}$
3. $(A|I)$ is unimodular
4. $A$ is totally unimodular

## Theorem

Let $A$ be an totally unimodular matrix and let $\boldsymbol{b}$ be an integral vector. Then, The polyhedron defined by $A\boldsymbol{x} \le \boldsymbol{b}$ is integral.

## Proof

- Let $F = \left\{\boldsymbol{x};\ A'\boldsymbol{x} \le \boldsymbol{b}', A''\boldsymbol{x} = \boldsymbol{b}''\right\}$ be a minimal face where $A''$ has full row rank
- Let $B$ be a basis of $A''$
- Then, $\boldsymbol{x}_B = A_B''^{-1}\boldsymbol{b}''$ and $\boldsymbol{x}_N = \boldsymbol{0}$ is an integral point in $F$

# Totally unimodular matrix: Application

## Observation

Let $A$ be a matrix of 0, 1 and $-1$ where every column has at most one $+1$ and at most one $-1$. Then, $A$ is totally unimodular.

## Proof

By the induction on $k$ prove that every $k \times k$ submatrix $N$ has determinant 0, $+1$ or $-1$

$k = 1$ Trivial

$k > 1$
- If $N$ has a column with at most one non-zero element, then we expand this column and use induction
- If $N$ has exactly one $+1$ and $-1$ in every column, then the sum of all rows is $\mathbf{0}$, so $N$ is singular

## Corollary

The incidence matrix of an oriented graph is totally unimodular.

## Observation: Other totally unimodular (TU) matrices

$A$ is TU iff $A^{\mathrm{T}}$ is TU iff $(A|I)$ is TU iff $(A|A)$ is TU iff $(A|-A)$ is TU

# Network flow

## Definition: Network flow

Let $G = (V, E)$ be an oriented graph with non-negative capacities of edges $c \in \mathbb{R}^E$. A network flow in $G$ is a vector $f \in \mathbb{R}^E$ such that

**Conservation:** $\sum_{uv \in E} f_{uv} = \sum_{vu \in E} f_{vu}$ for every vertex $v \in V$

**Capacity:** $0 \leq f \leq c$

The network flow problem is the optimization problem of finding a flow $f$ in $G$ that maximize $f_{ts}$ on a given edge $ts \in E$.

## Theorem

The polytope of network flow is integral for every integral $c$.

## Proof

1. Let $A$ be the incidence matrix of $G$
2. $A$ is totally unimodular
3. $(A| - A)$ and $(A| - A|I)$ are totally unimodular
4. $\left\{ f;\ \begin{pmatrix} A \\ -A \\ I \end{pmatrix} f \leq \begin{pmatrix} 0 \\ 0 \\ c \end{pmatrix}, f \geq \mathbf{0} \right\}$ is an integral polytope

# Duality of the network flow problem

## Primal: Network flow

Maximize $f_{ts}$ subject to $Af = \mathbf{0}$, $f \leq c$ and $f \geq \mathbf{0}$.

## Primal dual

Minimize $\boldsymbol{cz}$ subject to $A^{\mathrm{T}}\boldsymbol{y} + \boldsymbol{z} \geq e_{ts}$, that is $-\boldsymbol{y}_u + \boldsymbol{y}_v + \boldsymbol{z}_{uv} \geq 0$ for $uv \neq ts$ and $-\boldsymbol{y}_t + \boldsymbol{y}_s \geq 1$ assuming $f(ts)$ is unbounded.

## Observation

Dual problem has an integral optimal solution.

## Theorem

The dual problem is the minimal cut problem where $Z = \{uv \in E;\ z_{uv} = 1\}$ are cut edges and $U = \{u \in V;\ \boldsymbol{y}_u > \boldsymbol{y}_t\}$ is partition of vertices.

# Outline

# Augmenting paths

## Definitions

Let $M \subseteq E$ a matching of a graph $G = (V, E)$.

- A vertex $v \in V$ is $M$*-covered* if some edge of $M$ is incident with $v$.
- A vertex $v \in V$ is $M$*-exposed* if $v$ is not $M$-coveder.
- A path $P$ is $M$*-alternating* if its edges are alternately in and not in $M$.
- An $M$-alternating path is $M$*-augmenting* if both end-vertices are $M$-exposed.

## Augmenting path theorem of matchings

A matching $M$ in a graph $G = (V, E)$ is maximum if and only if there is no $M$-augmenting path.

## Proof

$\Rightarrow$ Every $M$-augmenting path increases the size of $M$

$\Leftarrow$ Let $N$ be a matching such that $|N| > |M|$ and we find an $M$-augmenting path

1. The graph $(V, N \cup M)$ contains a component $K$ which has more $N$ edges than $M$ edges
2. $K$ has at least two vertices $u$ and $v$ which are $N$-covered and $M$-exposed
3. Verteces $u$ and $v$ are joined by a path $P$ in $K$
4. Observe that $P$ is $M$-augmenting

# Tutte-Berge Formula

## Definition

Let def($G$) be the number of exposed edges by a maximal-size matching in $G = (V, E)$.

## Definition

Let $oc(G)$ be the number of odd components of a graph $G$.

## Observation

For every $A \subseteq V$ it holds that $\text{def}(G) \geq \text{oc}(G \setminus A) - |A|$.

## Theorem: Tutte-Berge Formula

$\text{def}(G) = \min \{\text{oc}(G \setminus A) - |A|;\ A \subseteq V\}$

## Proof

- $\geq$ Follows from the previous observation.
- $\leq$ An algorithm presented later.

## Tutte's matching theorem

A graph $G$ has a perfect matching if and only if $\text{oc}(G \setminus A) \leq |A|$ for every $A \subseteq V$.

## Construction of an $M$-alternating tree $T$ on vertices $A\dot\cup B$

**Init:** $A = \emptyset$ and $B = \{r\}$ where $r$ is an $M$-exposed root

**Step:** Let $uv \in E$ such that $u \in B$, $v \notin A \cup B$ and $vz \in M$ for some $z \in V$
Add $v$ to $A$ and $z$ to $B$

## Properties

- $r$ is the only $M$-exposed vertex of $T$
- For every $v$ of $T$, the path in $T$ from $v$ to $r$ is $M$-alternating

## Definition

$M$-alternation path $T$ is frustrated if every edge of $G$ having one ege in $B$ has the other end in $A$

## Observation

If $G$ has a matching $M$ and a frustrated $M$-alternating tree, then $G$ has no perfect matching.

## Proof

$B$ are single vertex components of $G \setminus A$, so $\text{oc}(G \setminus A) \geq |B| > |A|$

## Use $uv \in E$ to extend $T$

Input: A matching $M$ of a graph $G$, an $M$-alternating tree $T$, edge $uv \in E$ such that $u \in B$ and $v \notin A \cup B$ and $v$ is $M$-covered

Action: Let $vz \in M$ and extend $T$ by edges $uv$ and $vz$

## Use $uv \in E$ to augment $M$

Input: A matching $M$ of a graph $G$, an $M$-alternating tree $T$ with root $r$, edge $uv \in E$ such that $u \in B$ and $v \notin A \cup B$ and $v$ is $M$-exposed

Action: Let $P$ be the path obtained by attaching $uv$ to the path from $r$ to $v$ in $T$. Replace $M$ by $M \triangle E(P)$.

# Perfect matchings algorithm in a non-weighted bipartite graph

## Algorithm

Init: $M = \emptyset$ and $T = (\{r\}, \emptyset)$ where $r$ is an arbitrary vertex
**while** *there exists* $uv \in E$ *with* $u \in B(T)$ *and* $v \notin V(T)$ **do**
    **if** $v$ *is* $M$*-exposed* **then**
        Use $uv$ to augment $M$
        **if** *there is no* $M$*-exposed node in* $G$ **then**
            **return** $M$
        **else**
            Replace $T$ by $(\{r\}, \emptyset)$ where $r$ is an $M$-exposed vertex
    **else**
        Use $uv$ to extend $T$
**return** $G$ has no perfect matching since $T$ is a frustrated $M$-alternating path

## Theorem

The algorithm decides whether a given bipartite graph $G$ has a perfect matching and find one if exists. The algorithm calls $O(n)$ augmenting operations and $O(n^2)$ extending operations.

# Perfect matchings in bipartite graphs

## Minimal-weight perfect matching

Let $G$ be a graph with weights $\mathbf{c} \geq \mathbf{0}$ on edges. The minimal-weight perfect matching problem is minimizing $\mathbf{cx}$ subject to $A\mathbf{x} = \mathbf{1}$ and $\mathbf{x} \in \{0, 1\}^E$ where $A$ is the incidence matrix.

## Observation

The incidence matrix $A$ of a bipartite graph $G$ is totally unimodular.

## Proof

By the induction on $k$ prove that every $k \times k$ submatrix $N$ has determinant 0, $+1$ or $-1$

$k = 1$ Trivial

$k > 1$
- If $N$ has a column or a row with at most one non-zero element, then we expand this column and use induction
- Otherwise, the subgraph of edges corresponing to rows of $N$ contains a cycle and rows corresponing to edges of a cycle are linearly dependent.

## Theorem

If $A$ is an incidence matrix of a bipartite graph, then $\{\mathbf{x};\ A\mathbf{x} = \mathbf{1},\ \mathbf{x} \geq \mathbf{0}\}$ is integral.

# Duality and complementary slackness of perfect matchings

## Primal: relaxed perfect matching

Minimize $\boldsymbol{c}^{\mathrm{T}}\boldsymbol{x}$ subject to $A\boldsymbol{x} = \mathbf{1}$ and $\boldsymbol{x} \geq \mathbf{0}$.

## Dual

Maximize $\mathbf{1}\boldsymbol{y}$ subject to $A^{\mathrm{T}}\boldsymbol{y} \leq \boldsymbol{c}$ and $\boldsymbol{y} \in \mathbb{R}^E$, that is $\boldsymbol{y}_u + \boldsymbol{y}_v \leq \boldsymbol{c}_{uv}$.

## Idea of primal-dual algorithms

If we find a primal and a dual feasible solutions satisfying the complementary slackness, then solutions are optimal (relaxed) solutions.

## Definition

- An edge $uv \in E$ is called *tight* if $\boldsymbol{y}_u + \boldsymbol{y}_v = \boldsymbol{c}_{uv}$.
- Let $E_{\boldsymbol{y}}$ be the set of a tight edges of the dual solution $\boldsymbol{y}$.
- Let $M_{\boldsymbol{x}} = \{uv \in E;\ \boldsymbol{x}_{uv} = 1\}$ be the set of matching edge of the primal solution $\boldsymbol{x}$.

## Complementary slackness

$\boldsymbol{x}_{uv} = 0$ or $\boldsymbol{y}_u + \boldsymbol{y}_v = \boldsymbol{c}_{uv}$ for every edge $uv \in E$, that is $M_{\boldsymbol{x}} \subseteq E_{\boldsymbol{y}}$

# Weighted perfect matchings in a bipartite graph: Overview

## Complementary slackness

$\boldsymbol{x}_{uv} = 0$ or $\boldsymbol{y}_u + \boldsymbol{y}_v = \boldsymbol{c}_{uv}$ for every edge $uv \in E$, that is $M_{\boldsymbol{x}} \subseteq E_{\boldsymbol{y}}$

## Invariants

- Dual solution is feasible, that is $\boldsymbol{y}_u + \boldsymbol{y}_v \leq \boldsymbol{c}_{uv}$
- Every matching edge is tight
- $\boldsymbol{x} \in \{0,1\}^E$ and $M_{\boldsymbol{x}} = \{uv \in E;\ \boldsymbol{x}_{uv} = 1\}$ form a matching

## Initial solution satisfying invariants

$\boldsymbol{x} = \mathbf{0}$ and $\boldsymbol{y} = \mathbf{0}$

## Lemma: optimality

If $M_{\boldsymbol{x}}$ is a perfect matching, then $M_{\boldsymbol{x}}$ is a perfect matching with the minimal weight.

## Idea of the algorithm

- If there exists an $M_{\boldsymbol{x}}$-augmenting path $P$ in $(V, E_{\boldsymbol{y}})$, then $M_{\boldsymbol{x}} \triangle P$ is a new matching.
- Otherwise, update the dual solution $\boldsymbol{y}$ to enlarge $E_{\boldsymbol{y}}$.

# Minimal weight perfect matchings algorithm in a bipartite graph

## Algorithm

Init: $\boldsymbol{y} = \mathbf{0}$ and $M = \emptyset$ and $T = (\{r\}, \emptyset)$ where $r$ is an arbitrary vertex
**Loop**
  Find a perfect matching $M$ in $(V, E_{\boldsymbol{y}})$ or flustrated $M$-alternating tree
  **if** *$M$ is a perfect matching of $G$* **then**
    **return** Perfect matching $M$
  $\epsilon = \min\{c_{uv} - \boldsymbol{y}_u - \boldsymbol{y}_v;\ u, v \in E, u \in B(T), v \notin T\}$
  **if** $\epsilon = \infty$ **then**
    **return** Dual problem is unbounded, so there is no perfect matching
  $\boldsymbol{y}_u := \boldsymbol{y}_u + \epsilon$ for all $u \in B$
  $\boldsymbol{y}_v := \boldsymbol{y}_v - \epsilon$ for all $v \in A$

## Theorem

The algorithm decides whether a given bipartite graph $G$ has a perfect matching and a minimal-weight perfect matching if exists. The algorithm calls $O(n)$ augmenting operations and $O(n^2)$ extending operations and $O(n^2)$ dual changes.

# Shrinking odd circuits

## Definition

Let $C$ be an odd circuit in $G$. The graph $G \times C$ has vertices $(V(G) \setminus V(C)) \cup \{c'\}$ where $c'$ is a new vertex and edges

- $E(G)$ with both end-vertices in $V(G) \setminus V(C)$ and
- and $uc'$ for every edge $uv$ with $u \notin V(C)$ and $v \in V(C)$.

Edges $E(C)$ are removed.

## Proposition

Let $C$ be an odd circuit of $G$ and $M'$ be a matching $G \times C$. Then, there exists a matching $M$ of $G$ such that $M \subseteq M' \cup E(C)$ and the number of $M'$-exposed nodes of $G$ is the same as the number of $M'$-exposed nodes in $G \times C$.

## Corollary

$\text{def}(G) \leq \text{def}(G \times C)$

## Exercise

Find a graph $G$ with odd circuit $C$ such that $\text{def}(G) < \text{def}(G \times C)$.

## Use $uv$ to shrink and update $M'$ and $T$

**Input:** A matching $M'$ of a graph $G'$, an $M'$-alternating tree $T$, edge $uv \in E'$ such that $u, v \in B$

**Action:** Let $C$ be the circuit formed by $uv$ together with the path in $T$ from $u$ to $v$. Replace $G'$ by $G' \times C$, $M'$ by $M' \setminus E(C)$ and $T$ by the tree having edge-set $E(T) \setminus E(C)$.

## Observation

Let $G'$ be a graph obtained from $G$ by a sequence of odd-circuit shrinkings. Let $M'$ be matching of $G'$ and let $T$ be an $M'$ alternating tree of $G'$ such that all vertices of $A$ are original vertices of $G$. If $T$ is frustrated, then $G$ has no perfect matching.

# Perfect matchings algorithm in a non-weighted graph

## Algorithm

Init: $M' = M = \emptyset$, $G' = G$ and $T = (\{r\}, \emptyset)$ where $r$ is an arbitrary vertex
**while** *there exists* $uv \in E'$ *with* $u \in B$ *and* $v \notin A$ **do**
    **if** $v \notin T$ *is* $M'$*-exposed* **then**
        Use $uv$ to augment $M'$
        Extend $M'$ to a matching $G$
        Replace $M'$ by $M$ and $G'$ by $G$
        **if** *there is no* $M'$*-exposed node in* $G'$ **then**
            **return** Perfect matching $M$
        **else**
            Replace $T$ by $(\{r\}, \emptyset)$ where $r$ is an $M'$-exposed vertex
    **else if** $v \notin T$ *is* $M'$*-covered* **then**
        Use $uv$ to extend $T$
    **else if** $v \in B$ **then**
        Use $uv$ to shrink and update $M'$ and $T$
**return** $G$ has no perfect matching since $T$ is a frustrated $M$-alternating path

# Minimum-Weight perfect matchings in general graphs

## Observation

Let $M$ be a perfect matching of $G$ and $D$ be an odd set of vertices of $G$. Then there exists at least one edge $uv \in M$ between $D$ and $V \setminus D$.

## Linear programming for Minimum-Weight perfect matchings in general graphs

$$\begin{array}{lrcll} \text{Minimize} & \boldsymbol{c}\boldsymbol{x} & & & \\ \text{subject to} & \delta^u \boldsymbol{x} & = & 1 & \text{for all } u \in V \\ & \delta^D \boldsymbol{x} & \geq & 1 & \text{for all } D \in \mathcal{C} \\ & \boldsymbol{x} & \geq & \boldsymbol{0} & \end{array}$$

Where $\delta^D \in \{0,1\}^E$ is a vector such that $\delta^D_{uv} = 1$ if $|uv \cap D| = 1$ and $\delta^w = \delta^{\{w\}}$ and $\mathcal{C}$ is the set of all odd-size subsets of $V$.

## Exercise

Find a cutting plane proof of all odd-subset conditions.

## Theorem

Let $G$ be a graph and $\boldsymbol{c} \in \mathbb{R}^E$. Then $G$ has a perfect matching if and only if the LP problem is feasible. Moreover, if $G$ has a perfect matching, the minimum weight of a perfect matching is equal to the optimal value of the LP problem.

# Minimum-Weight perfect matchings in general graphs: Duality

## Primal

$$\begin{array}{llcll} \text{Minimize} & \mathbf{c}\mathbf{x} & & & \\ \text{subject to} & \delta^u \mathbf{x} & = & 1 & \text{for all } u \in V \\ & \delta^D \mathbf{x} & \geq & 1 & \text{for all } D \in \mathcal{C} \\ & \mathbf{x} & \geq & \mathbf{0} & \end{array}$$

## Dual

$$\begin{array}{llcll} \text{Maximize} & \sum_{v \in V} \mathbf{y}_v + \sum_{D \in \mathcal{C}} \mathbf{z}_D & & & \\ \text{subject to} & \mathbf{y}_u + \mathbf{y}_v + \sum_{uv \in D \in \mathcal{C}} \mathbf{z}_D & \leq & \mathbf{c}_{uv} & \text{for all } uv \in E \\ & \mathbf{z} & \geq & \mathbf{0} & \end{array}$$

## Notation: Reduced cost

$\bar{\mathbf{c}}_{uv} := \mathbf{c}_{uv} - \mathbf{y}_u - \mathbf{y}_v - \sum_{uv \in D \in \mathcal{C}} \mathbf{z}_D$

An edge $e$ is tight if $\bar{\mathbf{c}}_e = 0$

## Complementary slackness

- $\mathbf{x}_e > 0$ implies $e$ is tight for all $e \in E$
- $\mathbf{z}_D > 0$ implies $\delta^D \mathbf{x} = 1$ for all $D \in \mathcal{C}$

# Minimum-Weight perfect matchings in general graphs: Change of $y$

## Updates weights and dual solution when shrinking a circuit $C$

Replace $\boldsymbol{c}'_{uv}$ by $\boldsymbol{c}'_{uv} - \boldsymbol{y}'_v$ for $u \in C$ and $v \notin C$ and set $\boldsymbol{y}'_{c'} = 0$ for the new vertex $c'$. Note that the reduced cost is unchanged.

## Expand $c'$ into circuit $C$

- Set $\boldsymbol{z}'_C = \boldsymbol{y}'_{c'}$
- Replace $\boldsymbol{c}'_{uv}$ by $\boldsymbol{c}'_{uv} + \boldsymbol{y}'_v$ for $u \in C$ and $v \notin C$
- Update $M'$ and $T$

## Change of $\boldsymbol{y}$ and $\boldsymbol{z}$ on a frustrated tree

**Input:** A graph $G'$ with weights $\boldsymbol{c}'$, a feasible dual solution $\boldsymbol{y}'$, a matching $M'$ of tight edges of $G'$ and an $M'$-alternating tree $T$ of tight edges of $G'$.

**Action:**

- $\epsilon_1 = \min \left\{\bar{\boldsymbol{c}}_e{}';\ e \text{ joins a vertex in } B \text{ and a vertex not in } T\right\}$
- $\epsilon_2 = \min \left\{\bar{\boldsymbol{c}}_e{}'/2;\ e \text{ joins two vertices of } B\right\}$
- $\epsilon_3 = \min \{\boldsymbol{y}'_v;\ v \in A \text{ and } v \text{ is a pseudonode of } G\}$
- $\epsilon = \min \{\epsilon_1, \epsilon_2, \epsilon_3\}$
- Replace $\boldsymbol{y}'_v$ by $\boldsymbol{y}'_v + \epsilon$ for all $v \in B$
- Replace $\boldsymbol{y}'_v$ by $\boldsymbol{y}'_v - \epsilon$ for all $v \in A$

# Minimal weight perfect matchings algorithm in a general graph

## Algorithm

Init: $M' = M = \emptyset$, $G' = G$ and $T = (\{r\}, \emptyset)$ where $r$ is an arbitrary vertex

**Loop**

- **if** *there exists* $uv \in E_{\boldsymbol{y}}$, $u \in B$, $v \notin E(T)$, *$v$ is $M'$-exposed* **then**
  - Use $uv$ to augment $M'$
  - **if** *there is no $M'$-exposed node* **then**
    - **return** *extended $M'$ to a perfect matching $G$*
  - **else**
    - Replace $T$ by $(\{r\}, \emptyset)$ where $r$ is an $M'$-exposed vertex
- **else if** *there exists* $uv \in E_{\boldsymbol{y}}$, $u \in B$, $v \notin E(T)$, *$v$ is $M'$-covered* **then**
  - Use $uv$ to extend $T'$
- **else if** *there exists* $uv \in E_{\boldsymbol{y}}$, $u, v \in B$ **then**
  - Use $uv$ to shrink and update $M'$, $T'$, $c'$
- **else if** *there exists a pseudonode* $v \in A$ *with* $\boldsymbol{y}_v = 0$ **then**
  - Expand $v$ and update $M'$, $T$, and $c'$
- **else**
  - Change $\boldsymbol{y}$
  - **if** $\epsilon = \infty$ **then**
    - **return** *$G$ has no perfect matching*

# Maximum-weight (general) matching

## Reduction to perfect matching problem

Let $G$ be a graph with non-negative weights $\boldsymbol{c}$.

- Let $G_1$ and $G_2$ be two copies of $G$
- Let $P$ be a perfect matching between $G_1$ and $G_2$ joining copied vertices
- Let $G^\star$ be a graph of vertices $V(G_1) \cup V(G_2)$ and edges $E(G_1) \cup E(G_2) \cup P$
- For $e \in E(G_1) \cup E(G_2)$ let $\boldsymbol{c}^\star(e)$ be the weight of the original edge $e$ on $G$
- For $e \in P$ let $\boldsymbol{c}^\star(e) = 0$

## Theorem

The maximal weight of a perfect matching in $G^\star$ equals twice the maximal weight of a matching in $G$.

## Note

For maximal-size matching, use weights $\boldsymbol{c} = \mathbf{1}$.

## Tutte's matching theorem

A graph $G$ has a perfect matching if and only if $\text{oc}(G \setminus A) \leq |A|$ for every $A \subseteq V$.

# Outline

## Problem

Determine whether a given fully-dimensional convex compact set $Z \subseteq \mathbb{R}^n$ (e.g. a polytope) is non-empty and find a point in $Z$ if exists.

## Separation oracle

Separation oracle determines whether a point $s$ belongs into $Z$. If $s \notin Z$, the oracle finds a hyperplane that separates $s$ and $Z$.

## Inputs

- Radius $R > 0$ of a ball $B(0, R)$ containing $Z$
- Radius $\epsilon > 0$ such that $Z$ contains $B(s, \epsilon)$ for some point $s$ if $Z$ is non-empty
- Separation oracle

# Ellipsoid method

## Idea

Consider an ellipsoid $E$ containing $Z$. In every step, reduce the volume of $E$ using an hyperplane provided by the oracle.

## Algorithm

Init: $s = \mathbf{0}$, $E = B(s, R)$
**Loop**
    **if** *volume of E is smaller than volume of* $B(0, \epsilon)$ **then**
        **return** *$Z$ is empty*
    Call the oracle
    **if** $s \in Z$ **then**
        **return** *$s$ is a point of $Z$*
    Update $s$ and $Z$ using the separation hyperplane fount by oracle

# Ellipsoid

## Definition: Ball

The ball in the centre $\boldsymbol{s} \in \mathbb{R}^n$ and radius $R \geq 0$ is $B(\boldsymbol{s}, R) = \{\boldsymbol{x} \in \mathbb{R}^n;\ ||\boldsymbol{x} - \boldsymbol{s}|| \leq R\}$.

## Definition

Ellipsoid $E$ is an affine transformation of the unit ball $B(\boldsymbol{0}, 1)$. That is,
$E = \{M\boldsymbol{x} + \boldsymbol{s};\ \boldsymbol{x} \in B(\boldsymbol{0}, 1)\}$ where $M$ is a regular matrix and $s$ is the centre of $E$.

## Notation

$$\begin{aligned} E &= \left\{\boldsymbol{y} \in \mathbb{R}^n;\ M^{-1}(\boldsymbol{y} - \boldsymbol{s}) \in B(\boldsymbol{0}, 1)\right\} \\ &= \left\{\boldsymbol{y} \in \mathbb{R}^n;\ (\boldsymbol{y} - \boldsymbol{s})^{\mathrm{T}}(M^{-1})^{\mathrm{T}} M^{-1}(\boldsymbol{y} - \boldsymbol{s}) \leq 1\right\} \\ &= \left\{\boldsymbol{y} \in \mathbb{R}^n;\ (\boldsymbol{y} - \boldsymbol{s})^{\mathrm{T}} Q^{-1}(\boldsymbol{y} - \boldsymbol{s}) \leq 1\right\} \end{aligned}$$

where $Q = MM^{\mathrm{T}}$ is a positive definite matrix

# Ellipsoid method: update of the ellipsoid

## Separation hyperplane

Consider a hyperplane $\boldsymbol{a}^{\mathrm{T}}\boldsymbol{x} = b$ such that $\boldsymbol{a}^{\mathrm{T}}\boldsymbol{s} \geq b$ and $Z \subseteq \left\{\boldsymbol{x};\ \boldsymbol{a}^{\mathrm{T}}\boldsymbol{x} \leq b\right\}$.
For simplicity, assume that the hyperplane contains $\boldsymbol{s}$, that is $\boldsymbol{a}^{\mathrm{T}}\boldsymbol{s} = b$.

## Update formulas (without proof)

$$\boldsymbol{s}' = \boldsymbol{s} - \frac{1}{n+1}\frac{Q\boldsymbol{a}}{\sqrt{\boldsymbol{a}^{\mathrm{T}}Q\boldsymbol{a}}}$$

$$Q' = \frac{n^2}{n^2-1}\left(Q - \frac{2}{n+1}\frac{Q\boldsymbol{a}\boldsymbol{a}^{\mathrm{T}}Q}{\boldsymbol{a}^{\mathrm{T}}Q\boldsymbol{a}}\right)$$

## Reduce of the volume (without proof)

$$\frac{\text{volume}(E')}{\text{volume}(E)} \leq e^{-\frac{1}{2n+2}}$$

## Corollary

The number of steps of the Ellipsoid method is at most $\left\lceil n(2n+2)\ln\frac{R}{\epsilon}\right\rceil$.

# Ellipsoid method: Estimation of radii for rational polytopes

## Largest coefficient of $A$ and $\boldsymbol{b}$

Let $L$ be the maximal absolute value of all coefficients of $A$ and $b$.

## Estimation of $R$

We find $R'$ such that $||\boldsymbol{x}||_\infty \leq R'$ for all $\boldsymbol{x}$ satisfying $A\boldsymbol{x} \leq \boldsymbol{b}$:

- Consider a vertex of the polytope satisfying a subsystem $A'\boldsymbol{x} = \boldsymbol{b}'$
- Cramer's rule: $\boldsymbol{x}_i = \frac{\det A'_i}{\det A'}$
- $|\det(A'_i)| \leq n!L^n$ using the definition of determinant
- $|\det(A')| \geq 1$ since $A'$ is integral and regular

From the choice $R' = n!L^n$, it follows that $\log(R) = O(n^2 \log(n) \log(L))$

## Estimation of $\epsilon$ (without proof)

A non-empty rational fully-dimensional polytope contains a ball with radius $\epsilon$ where $\log \frac{1}{\epsilon} = O(poly(n, m, \log L))$.

## Complexity of Ellipsoid method

Time complexity of Ellipsoid method is polynomial in the length of binary encoding of $A$ and $\boldsymbol{b}$.

# Strongly polynomial algorithm for linear programming

## Ellipsoid method is not strongly polynomial (without proof)

For every $M$ there exists a linear program with 2 variables and 2 constrains such that the ellipsoid method executes at least $M$ mathematical operations.

## Open problem

Decide whether there exist an algorithm for linear programming which is polynomial in the number of variables and constrains.

# Outline

# Minimum vertex cover problem

## Definition

A vertex cover in a graph $G = (V, E)$ is a set of vertices $S$ such that every edge of $E$ has at least one end vertex in $S$. Finding a minimal-size vertex cover is the minimum vertex cover problem.

## Integer linear programming formulation

$$\begin{array}{lll} \text{Minimize} & \sum_{v \in V} \boldsymbol{x}_v & \\ \text{subject to} & \boldsymbol{x}_u + \boldsymbol{x}_v \geq 1 & \text{for all } uv \in E \\ & \boldsymbol{x}_v \in \{0, 1\} & \text{for all } v \in V \end{array}$$

## Relaxed problem

$$\begin{array}{lll} \text{Minimize} & \sum_{v \in V} \boldsymbol{x}_v & \\ \text{subject to} & \boldsymbol{x}_u + \boldsymbol{x}_v \geq 1 & \text{for all } uv \in E \\ & 0 \leq \boldsymbol{x}_v \leq 1 & \text{for all } v \in V \end{array}$$

# Approximation algorithm for vertex cover problem

## Algorithm

- Let $\mathbf{x}^\star$ the optimal relaxed solution
- Let $S_{LP} = \left\{v \in V;\ \mathbf{x}^\star_v \geq \frac{1}{2}\right\}$

## Observation

$S_{LP}$ is a vertex cover.

## Observation

Let $S_{OPT}$ be the minimal vertex cover. Then $\frac{|S_{LP}|}{|S_{OPT}|} \leq 2$.

## Proof

- Since $\mathbf{x}^\star$ is the optimal relaxed solution, $\sum_{v \in V} \mathbf{x}^\star_v \leq |S_{OPT}|$
- From the rounding rule, it follows that $|S_{LP}| \leq 2 \sum_{v \in V} \mathbf{x}^\star_v$
- Hence, $|S_{LP}| \leq 2 \sum_{v \in V} \mathbf{x}^\star_v \leq 2|S_{OPT}|$

# Maximum independent set problem

## Definition

An independent set in a graph $G = (V, E)$ is a set of vertices $S$ such that every edge of $E$ has at most one end vertex in $S$. Finding a maximal-size independent is the maximal independent problem.

## Integer linear programming formulation

$$\begin{array}{lll} \text{Minimize} & \sum_{v \in V} \boldsymbol{x}_v & \\ \text{subject to} & \boldsymbol{x}_u + \boldsymbol{x}_v \leq 1 & \text{for all } uv \in E \\ & \boldsymbol{x}_v \in \{0, 1\} & \text{for all } v \in V \end{array}$$

## Relaxed problem

$$\begin{array}{lll} \text{Minimize} & \sum_{v \in V} \boldsymbol{x}_v & \\ \text{subject to} & \boldsymbol{x}_u + \boldsymbol{x}_v \leq 1 & \text{for all } uv \in E \\ & 0 \leq \boldsymbol{x}_v \leq 1 & \text{for all } v \in V \end{array}$$

# Maximum independent set problem

## Relaxed solution

The relaxed solution $\boldsymbol{x}_v = \frac{1}{2}$ for all $v \in V$ is feasible, so the optimal relaxed solution is at least $\frac{n}{2}$.

## Optimal integer solution

The maximal independent set of a complete graph $K_n$ is a single vertex.

## Conclusion

In general, an optimal integer solution can be far from an optimal relaxed solution and cannot be obtained by a simple rounding.

## Inapproximability of the minimmum independent set problem

Unless $P = NP$, for every $C$ there is no polynomial-time approximation algorithm for the maximum independent set with the approximation error at most $C$.

# Outline

# Greedy algorithm for spanning tree problem

## Definition

A subtree $(V, J)$ of a connected graph $(V, E)$ is called a spanning tree.
Maximum-weight spanning tree is a problem to find the spanning of maximum weight.

## Greedy algorithm for finding a non-weighted spanning tree

Init: $J = \emptyset$
**while** *there exists an edge* $e \notin J$ *such that* $J \cup \{e\}$ *is a forest* **do**
    Choose an arbitrary such $e$
    Replace $J$ by $J \cup \{e\}$

## Greedy algorithm for finding a maximal-weight spanning tree

Init: $J = \emptyset$
**while** *there exists an edge* $e \notin J$ *such that* $J \cup \{e\}$ *is a forest* **do**
    Choose such $e$ with maximum weight
    Replace $J$ by $J \cup \{e\}$

# General greedy algorithm

## Family of subsets

Consider a finite set $S$ with weights $c : S \rightarrow \mathbb{R}$ and a family of subsets $\mathcal{I} \subseteq 2^S$ called independent. Our problem is to find $A \in \mathcal{I}$

- with maximum cardinality or
- with maximum weight.

## When the following algorithm finds the maximal subset?

Init: $J = \emptyset$
**while** *there exists an element* $e \in S \setminus J$ *such that* $J \cup \{e\} \in \mathcal{I}$ **do**
    Choose such $e$ (with maximum weight)
    Replace $J$ by $J \cup \{e\}$

## Examples

- ✓ Spanning tree
- × Matching
- × Independent set of vertices

# Matroid

## Definition

A pair $(S, \mathcal{I})$ where $S$ is a finite set and $\mathcal{I} \subseteq 2^{\mathcal{I}}$ is called a matroid if

(M0) $\emptyset \in \mathcal{I}$

(M1) If $J' \subseteq J \in \mathcal{I}$, then $J' \in \mathcal{I}$

(M2) For every $A \subseteq S$, every maximal independent subset of $A$ has the same cardinality

The cardinality of maximal independent subset of $A$ is callled rank $r(A)$.

## Examples

- Forest matroid: $S$ are edges of a graph and every forest is independent
- Linear matroid: $S$ are vectors of a linear space and $\mathcal{I}$ contains linearly independent vectors
- Uniform matroid: $\mathcal{I} = \{J \subseteq S;\ |J| \leq k\}$ for some $k$

## Theorem

Let $(S, \mathcal{I})$ satisfies (M0) and (M1). Then the Greedy algorithm finds an optimal independent set of every $c \in \mathbb{R}^S$ if and only if $(S, \mathcal{I})$ is a matroid.

## Theorem

Let $(S, \mathcal{I})$ satisfies (M0) and (M1). Then the Greedy algorithm finds an optimal independent set of every $c \in \mathbb{R}^S$ if and only if $(S, \mathcal{I})$ is a matroid.

## Proof

$\Rightarrow$

- For contradiction, consider $J \subseteq A \subseteq S$ such that $J \in \mathcal{I}$ is inclusion-maximal subset of $A$ which is not cardinality-maximal
- Let $c$ be a characteristic vector of $A$
- The Greedy algorithm may find $J$ although it is not the maximal-weight independent set

$\Leftarrow$

- Let $J = \{e_1, \dots, e_m\}$ be fount by the Greedy algorithm
- Let $J' = \{q_1, \dots, q_l\}$ be an optimal solution
- Let $k$ be the least index with $c(q_k) > c(e_k)$
- Let $A = \{e_1, \dots, e_{k-1}, q_1, \dots, q_k\}$
- $\{e_1, \dots, e_{k-1}\} \subseteq J$ and $\{q_1, \dots, q_k\} \subseteq J'$ are independent by (M1)
- $\{e_1, \dots, e_{k-1}, q_i\}$ is dependent for every $q_i \in \{q_1, \dots, q_k\} \setminus \{e_1, \dots, e_{k-1}\}$ since the Greedy algorithm does not choose $q_i$ in the $k$-th step
- Sets $A$, $\{e_1, \dots, e_{k-1}\}$ and $\{q_1, \dots, q_k\}$ contradict (M2)

# Complexity of the Greedy algorithm

## Inefficient

Enumerating whole $\mathcal{I}$ is inefficient, e.g. providing all forests in the input.

## Oracula

The input contains $S$ and $c$ and an oracula which decides whether a given $A \subseteq S$ is independent.

## Complexity

Complexity is determined in the size of $S$ and the number of calls of oracula.

# Equivalent definitions of a matroid

## Theorem

A set system $(S, \mathcal{I})$ is a matroid if and only if

(I0) $\emptyset \in \mathcal{I}$

(I1) If $J' \subseteq J \in \mathcal{I}$, then $J' \in \mathcal{I}$

(I2) For every $A, B \in \mathcal{I}$ with $|A| > |B|$ there exists $e \in A \setminus B$ such that $B \cup \{e\} \in \mathcal{I}$

## Definition

A circuit of a set system $(M, \mathcal{I})$ is a minimal dependent set.

## Observation

Let $(S, \mathcal{I})$ be a matroid, let $J \in \mathcal{I}$ and $e \in S$. Then $J \cup \{e\}$ contains at most one circuit.

## Theorem

A set $C$ of subsets of $S$ is the set of circuits of a matroid if and only if

(C0) $\emptyset \notin C$

(C1) If $C_1, C_2 \in C$ and $C_1 \subseteq C_2$, then $C_1 = C_2$

(C2) If $C_1, C_2 \in C$, $C_1 \neq C_2$ and $e \in C_1 \cap C_2$, then there exists $C' \in C$, $C \subseteq (C_1 \cup C_2) \setminus \{e\}$