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 2017/18
Last change on May 22, 2018

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 https://ktiml.mff.cuni.cz/˜fink/

Consultations Individual schedule

Examination

- Tutorial conditions
  - Tests
  - Theoretical 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

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

$$
\begin{array}{lcccccccl}
 & \text{Carrot} & & \text{White cabbage} & & \text{Cucumber} & & & \\
\text{Minimize} & 0.75\boldsymbol{x}_1 & + & 0.5\boldsymbol{x}_2 & + & 0.15\boldsymbol{x}_3 & & & \text{Cost} \\
\text{subject to} & 35\boldsymbol{x}_1 & + & 0.5\boldsymbol{x}_2 & + & 0.5\boldsymbol{x}_3 & \geq & 0.5 & \text{Vitamin A} \\
 & 60\boldsymbol{x}_1 & + & 300\boldsymbol{x}_2 & + & 10\boldsymbol{x}_3 & \geq & 15 & \text{Vitamin C} \\
 & 30\boldsymbol{x}_1 & + & 20\boldsymbol{x}_2 & + & 10\boldsymbol{x}_3 & \geq & 4 & \text{Dietary fiber} \\
 & & & & & \boldsymbol{x}_1, \boldsymbol{x}_2, \boldsymbol{x}_3 & \geq & 0 &
\end{array}
$$

Matrix notation of the linear programming problem

Formulation using linear programming

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

Matrix notation

- Minimize

$$
\begin{pmatrix} 0.75 \\ 0.5 \\ 0.15 \end{pmatrix}^T \begin{pmatrix} \boldsymbol{x}_1 \\ \boldsymbol{x}_2 \\ \boldsymbol{x}_3 \end{pmatrix}
$$

- Subject to

$$
\begin{pmatrix} 35 & 0.5 & 0.5 \\ 60 & 300 & 10 \\ 30 & 20 & 10 \end{pmatrix} \begin{pmatrix} \boldsymbol{x}_1 \\ \boldsymbol{x}_2 \\ \boldsymbol{x}_3 \end{pmatrix} \geq \begin{pmatrix} 0.5 \\ 15 \\ 4 \end{pmatrix}
$$

- and $\boldsymbol{x}_1, \boldsymbol{x}_2, \boldsymbol{x}_3 \geq 0$

Notation: Vector and matrix

Scalar

A scalar is a real number. Scalars 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 $n \times 1$.

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.

Special vectors

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

Transpose

The transpose of a matrix A is matrix A^{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 $i = 1, \ldots, m$.

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 $j = 1, \ldots, k$.

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

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

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

Mathematical 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
$-\min_{x \in M} f(x) = \max_{x \in M} -f(x)$.

Linear programming problem

A linear program is the problem of maximizing (or minimizing) a given linear function over the set of all vectors that satisfy a given system of linear equations and inequalities.

Equation form: $\min \boldsymbol{c}^{\mathrm{T}}\boldsymbol{x}$ subject to $A\boldsymbol{x} = \boldsymbol{b}, \boldsymbol{x} \geq \mathbf{0}$

Canonical form: $\max \boldsymbol{c}^{\mathrm{T}}\boldsymbol{x}$ subject to $A\boldsymbol{x} \leq \boldsymbol{b}$,

where $\boldsymbol{c} \in \mathbb{R}^n$, $\boldsymbol{b} \in \mathbb{R}^m$, $A \in \mathbb{R}^{m \times n}$ a $\boldsymbol{x} \in \mathbb{R}^n$.

Conversion from the equation form to the canonical form

$\max -\boldsymbol{c}^{\mathrm{T}}\boldsymbol{x}$ subject to $A\boldsymbol{x} \leq \boldsymbol{b}, -A\boldsymbol{x} \leq -\boldsymbol{b}, -\boldsymbol{x} \leq \mathbf{0}$

Conversion from the canonical form to the equation form

$\min -\boldsymbol{c}^{\mathrm{T}}\boldsymbol{x}' + \boldsymbol{c}^{\mathrm{T}}\boldsymbol{x}''$ subject to $A\boldsymbol{x}' - A\boldsymbol{x}'' + I\boldsymbol{x}''' = \boldsymbol{b},\ \boldsymbol{x}', \boldsymbol{x}'', \boldsymbol{x}''' \geq \mathbf{0}$

Basic terminology

- Number of variables: n
- Number of constrains: m
- Solution: an arbritrary vector $\boldsymbol{x}$ of $\mathbb{R}^n$
- Objective function: a function to be minimized or maximized, e.g. $\max \boldsymbol{c}^{\mathrm{T}}\boldsymbol{x}$
- Feasible solution: a solution satisfying all constrains, e.g. $A\boldsymbol{x} \le \boldsymbol{b}$
- 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 $A\boldsymbol{x} \le \boldsymbol{b}$ for some $A \in \mathbb{R}^{m \times n}$ and $\boldsymbol{b} \in \mathbb{R}^m$
- Polytope: a bounded polyhedron

Example of linear programming: Network flow

Network flow problem

Given a 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 $\boldsymbol{x}_e$ for every edge $e \in E$

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

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

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

Matrix notation

- Add an auxiliary edge $\boldsymbol{x}_{ts}$ with a sufficiently large capacity $\boldsymbol{c}_{ts}$

Objective function: $\max \boldsymbol{x}_{ts}$

Flow conservation: $A\boldsymbol{x} = \mathbf{0}$ where A is the incidence matrix

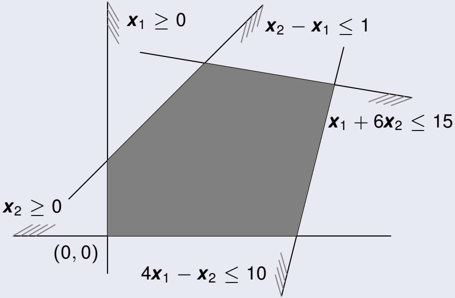
Capacity constrains: $\boldsymbol{x} \leq \boldsymbol{c}$ and $\boldsymbol{x} \geq 0$

Example

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

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

Solution



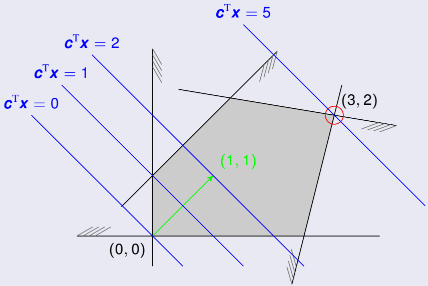
Graphical method: Optimal solution

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



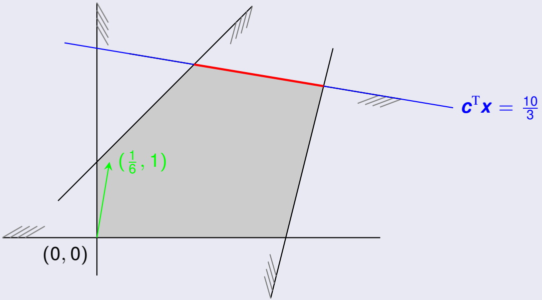
Graphical method: Multiple optimal solutions

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

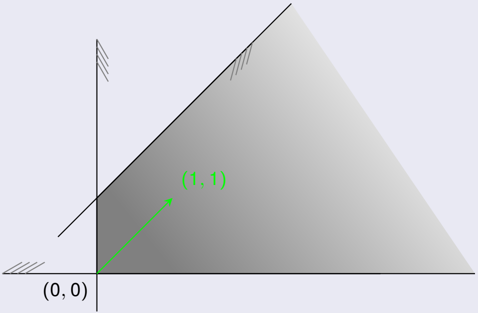


Example

Show that the following problem is unbounded.

$$\begin{array}{lrcrcl} \text{Maximize} & \boldsymbol{x}_1 & + & \boldsymbol{x}_2 & & \\ & -\boldsymbol{x}_1 & + & \boldsymbol{x}_2 & \leq & 1 \\ & & & \boldsymbol{x}_1, \boldsymbol{x}_2 & \geq & 0 \end{array}$$

Solution



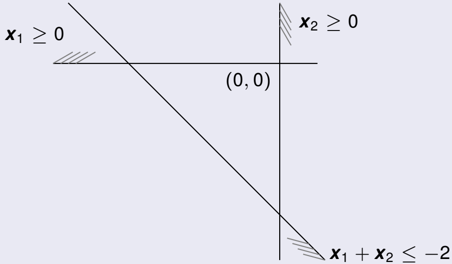
Graphical method: Infeasible problem

Example

Show that the following problem has no feasible solution.

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

Solution



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.

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 are *ellipsoid* and *interior point* methods.
- No strongly polynomial-time algorithms for linear programming is known.
- Integer linear programming is NP-hard.

Example of integer linear programming: Vertex cover

Vertex cover problem

Given an 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 $\sum_{v \in V} \boldsymbol{x}_v$

Matrix notation

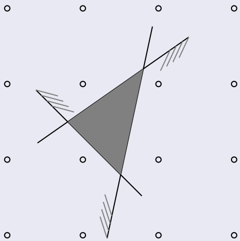
Variables: Cover $\boldsymbol{x} \in \{0, 1\}^V$ (i.e. $\mathbf{0} \leq \boldsymbol{x} \leq \mathbf{1}$ and $\boldsymbol{x} \in \mathbb{Z}^V$)

Covering: $A^{\mathrm{T}}\boldsymbol{x} \geq \mathbf{1}$ where A is the incidence matrix

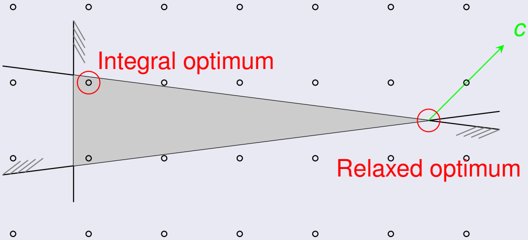
Objective function: Minimize $\mathbf{1}^{\mathrm{T}}\boldsymbol{x}$

Relation between optimal integer and relaxed solution

Non-empty polyhedron may not contain an integer solution



Integer feasible solution may not be obtained by rounding of a relaxed solution



Example: Ice cream production planning

Problem description

- An ice cream manufacturer needs to plan production of ice cream for next year
- The estimated demand of ice cream for month $i \in \{1, \ldots, n\}$ is $\boldsymbol{d}_i$ (in tons)
- Price for storing ice cream is a per ton and month
- Changing the production by 1 ton from month $i - 1$ to month i cost b
- Produced ice cream cannot be stored longer than one month
- The total cost has to be minimized

Example: Ice cream production planning

Solution

1. Variable $\boldsymbol{x}_i$ determines the amount of produced ice cream in month $i \in \{0, \ldots, n\}$
2. Variable $\boldsymbol{s}_i$ determines the amount of stored ice cream from month $i-1$ month i
3. The stored quantity is computed by $\boldsymbol{s}_i = \boldsymbol{s}_{i-1} + \boldsymbol{x}_i - \boldsymbol{d}_i$ for every $i \in \{1, \ldots, n\}$
4. Durability is ensured by $\boldsymbol{s}_i \le \boldsymbol{d}_i$ for all $i \in \{1, \ldots, n\}$
5. Non-negativity of the production and the storage $\boldsymbol{x}, \boldsymbol{s} \ge \mathbf{0}$
6. Objective function min $b\sum_{i=1}^n |\boldsymbol{x}_i - \boldsymbol{x}_{i-1}| + a\sum_{i=1}^n \boldsymbol{s}_i$ is non-linear
7. We introduce variables $\boldsymbol{y}_i$ for $i \in \{1, \ldots, n\}$ to avoid the absolute value
8. Linear programming problem formulation

$$
\begin{array}{lrcll}
\text{Minimize} & b\sum_{i=1}^n \boldsymbol{y}_i + a\sum_{i=1}^n \boldsymbol{s}_i & & & \\
\text{subject to} & \boldsymbol{s}_{i-1} - \boldsymbol{s}_i + \boldsymbol{x}_i & = & \boldsymbol{d}_i & \text{for } i \in \{1, \ldots, n\} \\
& \boldsymbol{s}_i & \le & \boldsymbol{d}_i & \text{for } i \in \{1, \ldots, n\} \\
& \boldsymbol{x}_i - \boldsymbol{x}_{i-1} - \boldsymbol{y}_i & \le & 0 & \text{for } i \in \{1, \ldots, n\} \\
& -\boldsymbol{x}_i + \boldsymbol{x}_{i-1} - \boldsymbol{y}_i & \le & 0 & \text{for } i \in \{1, \ldots, n\} \\
& \boldsymbol{x}, \boldsymbol{s}, \boldsymbol{y} & \ge & \mathbf{0} &
\end{array}
$$

9. We can bound the initial and final amount of ice cream $\boldsymbol{s}_0$ a $\boldsymbol{s}_n$
10. and also bound the production $\boldsymbol{x}_0$

Finding shortest paths from a vertex s in an oriented graph

Shortest path problem

Given an oriented graph (V, E) with length of edges $\mathbf{c} \in \mathbb{Z}^n$ and a starting vertex s, find the length of a shortest path from s to all vertices.

Linear programming problem

$$\begin{array}{llcll} \text{Maximize} & \sum_{u \in V} \mathbf{x}_u & & & \\ \text{subject to} & \mathbf{x}_v - \mathbf{x}_u & \leq & \mathbf{c}_{uv} & \text{for every edge } uv \\ & \mathbf{x}_s & = & 0 & \end{array}$$

Proof (optimal solution $\mathbf{x}_u^\star$ of LP gives the distance from s to u for $\forall u \in V$)

1. Let $\mathbf{y}_u$ be the length of a shortest path from s to u
2. It holds that $\mathbf{y} \geq \mathbf{x}^\star$
   - Let P be edges on the shortest path from s to z
   - $\mathbf{y}_z = \sum_{uv \in P} \mathbf{c}_{uv} \geq \sum_{uv \in P} \mathbf{x}_v^\star - \mathbf{x}_u^\star = \mathbf{x}_z^\star - \mathbf{y}_s^\star = \mathbf{x}_z^\star$
3. It holds that $\mathbf{y} = \mathbf{x}^\star$
   - For the sake of contradiction assume that $\mathbf{y} \neq \mathbf{x}^\star$
   - So $\mathbf{y} \geq \mathbf{x}^\star$ and $\sum_{u \in V} \mathbf{y}_u > \sum_{u \in V} \mathbf{x}_u^\star$
   - But $\mathbf{y}$ is a feasible solution and $\mathbf{x}^\star$ is an optimal solution

Outline

Linear space

Definition: Linear (vector) space

A set $(V, +, \cdot)$ is called a linear (vector) space over a field T if

- $+ : V \times V \rightarrow V$ i.e. V is closed under addition $+$
- $\cdot : T \times V \rightarrow V$ i.e. V is closed under multiplication by T
- $(V, +)$ is an Abelian group
- For every $\boldsymbol{x} \in V$ it holds that $1 \cdot \boldsymbol{x} = \boldsymbol{x}$ where $1 \in T$
- For every $a, b \in T$ and every $\boldsymbol{x} \in V$ it holds that $(ab) \cdot \boldsymbol{x} = a \cdot (b \cdot \boldsymbol{x})$
- For every $a, b \in T$ and every $\boldsymbol{x} \in V$ it holds that $(a + b) \cdot \boldsymbol{x} = a \cdot \boldsymbol{x} + b \cdot \boldsymbol{x}$
- For every $a \in T$ and every $\boldsymbol{x}, \boldsymbol{y} \in V$ it holds that $a \cdot (\boldsymbol{x} + \boldsymbol{y}) = a \cdot \boldsymbol{x} + a \cdot \boldsymbol{y}$

Observation

If V is a linear space and $L \subseteq V$, then L is a linear space if and only 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 $\alpha \in T$.

Observation

A non-empty set $V \subseteq \mathbb{R}^n$ is a linear space if and only if $\alpha \boldsymbol{x} + \beta \boldsymbol{y} \in V$ for all $\alpha, \beta \in \mathbb{R}$, $\boldsymbol{x}, \boldsymbol{y} \in V$.

Definition

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

Basic observations

- If $L \subseteq \mathbb{R}^n$ is an affine space, then $L + \boldsymbol{x}$ is an affine space for every $\boldsymbol{x} \in \mathbb{R}^n$.
- If $L \subseteq \mathbb{R}^n$ is an affine space, then $L - \boldsymbol{x}$ is a linear space for every $\boldsymbol{x} \in L$. ①
- If $L \subseteq \mathbb{R}^n$ is an affine space, then $L - \boldsymbol{x} = L - \boldsymbol{y}$ for every $\boldsymbol{x}, \boldsymbol{y} \in L$. ②
- An affine space $L \subseteq \mathbb{R}^n$ is linear if and only if L contains the origin $\mathbf{0}$. ③

System of linear equations

- The set of all solutions of $A\boldsymbol{x} = \mathbf{0}$ is a linear space and every linear space is the set of all solutions of $A\boldsymbol{x} = \mathbf{0}$ for some A. ④
- 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. ⑤

1. By definition, $L = V + \boldsymbol{a}$ for some linear space V and some vector $\boldsymbol{a} \in \mathbb{R}^n$. Observe that $L - \boldsymbol{x} = V + (\boldsymbol{a} - \boldsymbol{x})$ and we prove that $V + (\boldsymbol{a} - \boldsymbol{x}) = V$ which implies that $L - \boldsymbol{x}$ is a linear space. There exists $\boldsymbol{y} \in V$ such that $\boldsymbol{x} = \boldsymbol{y} + \boldsymbol{a}$. Hence, $\boldsymbol{a} - \boldsymbol{x} = \boldsymbol{a} - \boldsymbol{y} - \boldsymbol{a} = -\boldsymbol{y} \in V$. Since V is closed under addition, it follows that $V + (\boldsymbol{a} - \boldsymbol{x}) \subseteq V$. Similarly, $V - (\boldsymbol{a} - \boldsymbol{x}) \subseteq V$ which implies that $V \subseteq V + (\boldsymbol{a} - \boldsymbol{x})$. Hence, $V = V + (\boldsymbol{a} - \boldsymbol{x})$ and the statement follows.
2. We proved that $L = V + \boldsymbol{a}$ for some linear space $V \subseteq \mathbb{R}^n$ and some vector $\boldsymbol{a} \in \mathbb{R}^n$ and $L - \boldsymbol{x} = V + (\boldsymbol{a} - \boldsymbol{x}) = V$ for every $\boldsymbol{x} \in L$. So, $L - \boldsymbol{x} = V = L - \boldsymbol{y}$.
3. Every linear space must contain the origin by definition. For the opposite implication, we set $\boldsymbol{x} = \mathbf{0}$ and apply the previous statement.
4. If V is a linear space, then we can obtain rows of A from the basis of the orthogonal space of V.
5. If L is an affine space, then $L = V + \boldsymbol{a}$ for some vector space V and some vector $\boldsymbol{a}$ and there exists a matrix A such that $V = \{\boldsymbol{x};\ A\boldsymbol{x} = \mathbf{0}\}$. Hence, $V + \boldsymbol{a} = \{\boldsymbol{x} + \boldsymbol{a};\ A\boldsymbol{x} = \mathbf{0}\} = \{\boldsymbol{y};\ A\boldsymbol{y} - A\boldsymbol{a} = \mathbf{0}\} = \{\boldsymbol{y};\ A\boldsymbol{y} = \boldsymbol{b}\}$ where we substitute $\boldsymbol{x} + \boldsymbol{a} = \boldsymbol{y}$ and set $\boldsymbol{b} = A\boldsymbol{a}$.
   If $L = \{\boldsymbol{x};\ A\boldsymbol{x} = \boldsymbol{b}\}$ is non-empty, then let $\boldsymbol{y}$ be an arbitrary vertex of L. Furthermore, $L - \boldsymbol{y} = \{\boldsymbol{x} - \boldsymbol{y};\ A\boldsymbol{x} = \boldsymbol{b}\} = \{\boldsymbol{z};\ A\boldsymbol{y} + A\boldsymbol{z} = \boldsymbol{b}\} = \{\boldsymbol{z};\ A\boldsymbol{z} = \mathbf{0}\}$ is a linear space since $A\boldsymbol{y} = \boldsymbol{b}$.

Convex set

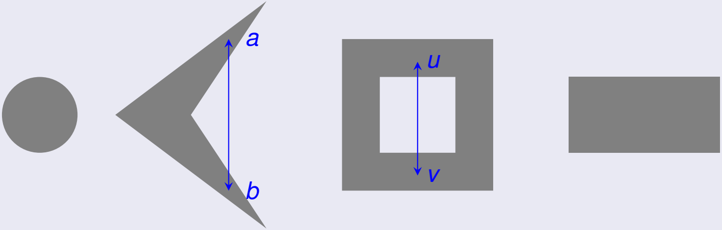
Observation (Exercise)

A set $S \subseteq \mathbb{R}^n$ is an affine space if and only if S contains whole line given every two points of S.

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 linear spaces is also a linear space. ①
- The non-empty intersection of affine spaces is an affine space. ②
- The intersection of convex sets is also a convex set. ③

Definition

Let $S \subseteq \mathbb{R}^n$ be an non-empty set.

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

Observation

Let $S \subseteq \mathbb{R}^n$ be an non-empty set.

- A set S is linear if and only if $S = \text{span}(S)$. ④
- A set S is affine if and only if $S = \text{aff}(S)$. ⑤
- A set S is convex if and only if $S = \text{conv}(S)$. ⑥
- $\text{span}(S) = \text{aff}(S \cup \{\mathbf{0}\})$

1. Use definition and logic.
2. Let L_i be affine space for i in an index set I and $L = \cap_{i \in I} L_i$ and $\boldsymbol{a} \in L$. We proved that $L - \boldsymbol{a} = \bigcap_{i \in I}(L_i - \boldsymbol{a})$ is a linear space which implies that L is an affine space.
3. Use definition and logic.
4. Similar as the convex version.
5. Similar as the convex version.
6. We proved that $\operatorname{conv}(S)$ is convex, so if $S = \operatorname{conv}(S)$, then S is convex. In order to prove that $S = \operatorname{conv}(S)$ if S is convex, we observe that $\operatorname{conv}(S) \subseteq S$ since $\operatorname{conv}(S) = \bigcap_{M \supseteq S, M \text{ convex}} M$ and S is included in this intersection. Similarly, $\operatorname{conv}(S) \supseteq S$ since every M in the intersection contains S.

Linear, affine and convex combinations

Definition

Let $\mathbf{v}_1, \ldots, \mathbf{v}_k$ be vectors of $\mathbb{R}^n$ where k is a positive integer.

- The sum $\sum_{i=1}^k \alpha_i \mathbf{v}_i$ is called a *linear combination* if $\alpha_1, \ldots, \alpha_k \in \mathbb{R}$.
- The sum $\sum_{i=1}^k \alpha_i \mathbf{v}_i$ is called an *affine combination* if $\alpha_1, \ldots, \alpha_k \in \mathbb{R}$, $\sum_{i=1}^k \alpha_i = 1$.
- The sum $\sum_{i=1}^k \alpha_i \mathbf{v}_i$ is called a *convex combination* if $\alpha_1, \ldots, \alpha_k \geq 0$ and $\sum_{i=1}^k \alpha_i = 1$.

Lemma

Let $S \subseteq \mathbb{R}^n$ be a non-empty set.

- The set of all linear combinations of S is a linear space. ①
- The set of all affine combinations of S is an affine space. ②
- The set of all convex combinations of S is a convex set. ③

Lemma

- A linear space S contains all linear combinations of S. ④
- An affine space S contains all affine combinations of S. ⑤
- A convex set S contains all convex combinations of S. ⑥

1. We have to verify that the set of all linear combinations has closure under addition and multiplication by scalars. In order to verify the closure under multiplication, let $\sum_{i=1}^k \alpha_i \boldsymbol{v}_i$ be a linear combination of S and $c \in \mathbb{R}$ be a scalar. Then, $c \sum_{i=1}^k \alpha_i \boldsymbol{v}_i = \sum_{i=1}^k (c\alpha_i) \boldsymbol{v}_i$ is a linear combination of of S. Similarly, the set of all linear combinations has closure under addition and it contains the origin.
2. Similar as the convex version: Show that S contains whole line defined by arbitrary pair of points of S.
3. Let $\sum_{i=1}^k \alpha_i \boldsymbol{u}_i$ and $\sum_{j=1}^l \beta_j \boldsymbol{v}_j$ be two convex combinations of S. In order to prove that the set of all convex combinations of S contains the line segment between $\sum_{i=1}^k \alpha_i \boldsymbol{u}_i$ and $\sum_{j=1}^l \beta_j \boldsymbol{v}_j$, let us consider $\gamma_1, \gamma_2 \geq 0$ such that $\gamma_1 + \gamma_2 = 1$. Then, $\gamma_1 \sum_{i=1}^k \alpha_i \boldsymbol{u}_i + \gamma_2 \sum_{j=1}^l \beta_j \boldsymbol{v}_j = \sum_{i=1}^k (\gamma_1 \alpha_i) \boldsymbol{u}_i + \sum_{j=1}^l (\gamma_2 \beta_j) \boldsymbol{v}_j$ is a convex combination of S since $(\gamma_1 \alpha_i), (\gamma_2 \beta_j) \geq 0$ and $\sum_{i=1}^k (\gamma_1 \alpha_i) + \sum_{j=1}^l (\gamma_2 \beta_j) = 1$.
4. Similar as the convex version.
5. Let $\sum_{i=1}^k \alpha_i \boldsymbol{v}_i$ be an affine combination of S. Since $S - \boldsymbol{v}_k$ is a linear space, the linear combination $\sum_{i=1}^k \alpha_i (\boldsymbol{v}_i - \boldsymbol{v}_k)$ of $S - \boldsymbol{v}_k$ belongs into $S - \boldsymbol{v}_k$. Hence, $\boldsymbol{v}_k + \sum_{i=1}^k \alpha_i (\boldsymbol{v}_i - \boldsymbol{v}_k) = \sum_{i=1}^k \alpha_i \boldsymbol{v}_i$ belongs to S.
6. We prove by induction on k that S contains every convex combination $\sum_{i=1}^k \alpha_i \boldsymbol{v}_i$ of S. The statement holds for $k \leq 2$ by the definition of a convex set. Let $\sum_{i=1}^k \alpha_i \boldsymbol{v}_i$ be a convex combination of k vectors of S and we assume that $\alpha_k < 1$, otherwise $\alpha_1 = \cdots = \alpha_{k-1} = 0$ so $\sum_{i=1}^k \alpha_i \boldsymbol{v}_i = \boldsymbol{v}_k \in S$. Hence, $\sum_{i=1}^k \alpha_i \boldsymbol{v}_i = (1 - \alpha_k) \sum_{i=1}^k \frac{\alpha_i}{1-\alpha_k} \boldsymbol{v}_i + \alpha_k \boldsymbol{v}_k = (1 - \alpha_k) \boldsymbol{y} + \alpha_k \boldsymbol{v}_k$ where we observe

that $\boldsymbol{y} := \sum_{i=1}^{k} \frac{\alpha_i}{1-\alpha_k} \boldsymbol{v}_i$ is a convex combination of $k-1$ vectors of S which by induction belongs to S. Furthermore, $(1-\alpha_k)\boldsymbol{y} + \alpha_k \boldsymbol{v}_k$ is a convex combination of S which by induction also belongs to S.

Linear, affine and convex combinations

Theorem

Let $S \subseteq \mathbb{R}^n$ be a non-empty set.

- The linear hull of a set S is the set of all linear combinations of S. ①
- The affine hull of a set S is the set of all affine combinations of S. ②
- The convex hull of a set S is the set of all convex combinations of S. ③

1. Similar as the convex version.
2. Similar as the convex version.
3. Let T be the set of all convex combinations of S. First, we prove that $\operatorname{conv}(S) \subseteq T$. The definition states that $\operatorname{conv}(S) = \bigcap_{M \supseteq S, M \text{ convex}} M$ and we proved that T is a convex set containing S, so T is included in this intersection which implies that $\operatorname{conv}(S)$ is a subset of T.
   In order to prove $\operatorname{conv}(S) \supseteq T$, we again consider the intersection $\operatorname{conv}(S) = \bigcap_{M \supseteq S, M \text{ convex}} M$. We proved that a convex set M contains all convex combinations of M which implies that if $M \supseteq S$ then M also contains all convex combinations of S. So, in this intersection every M contains T which implies that $\operatorname{conv}(S) \supseteq T$.

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

Observation (Exercise)

- Vectors $\boldsymbol{v}_1, \ldots, \boldsymbol{v}_k \in \mathbb{R}^n$ are linearly dependent if and only if there exists a non-trivial combination $\alpha_1, \ldots, \alpha_k \in \mathbb{R}$ such that $\sum_{i=1}^k \alpha_i \boldsymbol{v}_i = \mathbf{0}$.
- Vectors $\boldsymbol{v}_1, \ldots, \boldsymbol{v}_k \in \mathbb{R}^n$ are affinely dependent if and only if there exists a non-trivial combination $\alpha_1, \ldots, \alpha_k \in \mathbb{R}$ such that $\sum_{i=1}^k \alpha_i \boldsymbol{v}_i = \mathbf{0}$ a $\sum_{i=1}^k \alpha_i = 0$.

Observation

- Vectors $\boldsymbol{v}_0, \ldots, \boldsymbol{v}_k \in \mathbb{R}^n$ are affinely independent if and only if vectors $\boldsymbol{v}_1 - \boldsymbol{v}_0, \ldots, \boldsymbol{v}_k - \boldsymbol{v}_0$ are linearly independent. ①
- Vectors $\boldsymbol{v}_1, \ldots, \boldsymbol{v}_k \in \mathbb{R}^n$ are linearly independent if and only if vectors $\mathbf{0}, \boldsymbol{v}_1, \ldots, \boldsymbol{v}_k$ are affinely independent. ②

1. If vectors $\boldsymbol{v}_1 - \boldsymbol{v}_0, \dots, \boldsymbol{v}_k - \boldsymbol{v}_0$ are linearly dependent, then there exists a non-trivial combination $\alpha_1, \dots, \alpha_k \in \mathbb{R}$ such that $\sum_{i=1}^k \alpha_i(\boldsymbol{v}_i - \boldsymbol{v}_0) = \mathbf{0}$. In this case, $\mathbf{0} = \sum_{i=1}^k \alpha_i(\boldsymbol{v}_i - \boldsymbol{v}_0) = \sum_{i=1}^k \alpha_i \boldsymbol{v}_i - \boldsymbol{v}_0 \sum_{i=1}^k \alpha_i = \sum_{i=0}^k \alpha_i \boldsymbol{v}_i$ is a non-trivial affine combination with $\sum_{i=0}^k \alpha_i = 0$ where $\alpha_0 = -\sum_{i=1}^k \alpha_i$.
   If $\boldsymbol{v}_0, \dots, \boldsymbol{v}_k \in \mathbb{R}^n$ are affinely dependent, then there exists a non-trivial combination $\alpha_0, \dots, \alpha_k \in \mathbb{R}$ such that $\sum_{i=0}^k \alpha_i \boldsymbol{v}_i = \mathbf{0}$ a $\sum_{i=0}^k \alpha_i = 0$. In this case, $\mathbf{0} = \sum_{i=0}^k \alpha_i \boldsymbol{v}_i = \alpha_0 \boldsymbol{v}_0 + \sum_{i=1}^k \alpha_i \boldsymbol{v}_i = \sum_{i=1}^k \alpha_i(\boldsymbol{v}_i - \boldsymbol{v}_0)$ is a non-trivial linear combination of vectors $\boldsymbol{v}_1 - \boldsymbol{v}_0, \dots, \boldsymbol{v}_k - \boldsymbol{v}_0$.
2. Use the previous observation with $\boldsymbol{v}_0 = \mathbf{0}$.

Basis

Definition

Let $B \subseteq \mathbb{R}^n$ and $S \subseteq \mathbb{R}^n$.

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

Observation

- All linear bases of a linear space have the same cardinality.
- All affine bases of an affine space have the same cardinality. ①

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.

1. For the sake of contradiction, let $\boldsymbol{a}_1, \dots, \boldsymbol{a}_k$ and $\boldsymbol{b}_1, \dots, \boldsymbol{b}_l$ be two basis of an affine space $L = V + \boldsymbol{x}$ where V a linear space and $l > k$. Then, $\boldsymbol{a}_1 - \boldsymbol{x}, \dots, \boldsymbol{a}_k - \boldsymbol{x}$ and $\boldsymbol{b}_1 - \boldsymbol{x}, \dots, \boldsymbol{b}_l - \boldsymbol{x}$ are two linearly independent sets of vectors of V. Hence, there exists i such that $\boldsymbol{a}_1 - \boldsymbol{x}, \dots, \boldsymbol{a}_k - \boldsymbol{x}, \boldsymbol{b}_i - \boldsymbol{x}$ are linearly independent, so $\boldsymbol{a}_1, \dots, \boldsymbol{a}_k, \boldsymbol{b}_i$ are affinely independent. Therefore, $\boldsymbol{b}_i$ cannot be obtained by an affine combination of $\boldsymbol{a}_1, \dots, \boldsymbol{a}_k$ and $\boldsymbol{b}_i \notin \operatorname{aff}(\boldsymbol{a}_1, \dots, \boldsymbol{a}_k)$ which contradicts the assumption that $\boldsymbol{a}_1, \dots, \boldsymbol{a}_k$ is a basis of L.

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 $\boldsymbol{x} \in \operatorname{conv}(S)$. 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.

Outline

Definition

- A *hyperplane* is a set $\{\boldsymbol{x} \in \mathbb{R}^n;\ \boldsymbol{a}^{\mathrm{T}}\boldsymbol{x} = b\}$ where $\boldsymbol{a} \in \mathbb{R}^n \setminus \{\mathbf{0}\}$ 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 $\boldsymbol{a} \in \mathbb{R}^n \setminus \{\mathbf{0}\}$ 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 $\{\boldsymbol{x};\ A\boldsymbol{x} \leq \boldsymbol{b}\}$ 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 there exists $b \in \mathbb{R}$ s.t. for every $\boldsymbol{x} \in S$ holds $||\boldsymbol{x}|| < b$.
- 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 $\boldsymbol{y} \in S$.

Definition

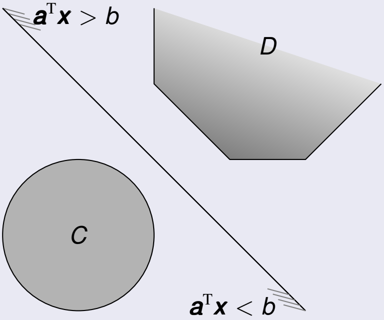
- Infimum of a set $S \subseteq \mathbb{R}$ is $\inf(S) = \max\{b \in \mathbb{R};\ b \leq x\ \forall x \in S\}$.
- Supremum of a set $S \subseteq \mathbb{R}$ is $\sup(S) = \min\{b \in \mathbb{R};\ b \geq x\ \forall x \in S\}$.
- $\inf(\emptyset) = \infty$ and $\sup(\emptyset) = -\infty$
- $\inf(S) = -\infty$ 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



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 $D \subseteq \left\{\boldsymbol{x}; \boldsymbol{a}^{\mathrm{T}}\boldsymbol{x} > b\right\}$.

Proof (overview)

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 $||\boldsymbol{d}_n - \boldsymbol{c}_n|| \leq m + \frac{1}{n}$.
   3. Since C is compact, there exists a subsequence $\left\{\boldsymbol{c}_{k_n}\right\}_{n=1}^{\infty}$ 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. ①
   5. Since the set $D \cap \{\boldsymbol{x} \in \mathbb{R}^n;\ ||\boldsymbol{x} - \boldsymbol{c}|| \leq z\}$ is compact, the sequence $\left\{\boldsymbol{d}_{k_n}\right\}_{n=1}^{\infty}$ has a subsequence $\left\{\boldsymbol{d}_{l_n}\right\}_{n=1}^{\infty}$ converging to $\boldsymbol{d} \in D$.
   6. Observe that the distance $||\boldsymbol{d} - \boldsymbol{c}||$ is m. ②
2. The required hyperplane is $\boldsymbol{a}^{\mathrm{T}}\boldsymbol{x} = b$ where $\boldsymbol{a} = \boldsymbol{d} - \boldsymbol{c}$ and $b = \frac{\boldsymbol{a}^{\mathrm{T}}\boldsymbol{c} + \boldsymbol{a}^{\mathrm{T}}\boldsymbol{d}}{2}$
   1. We prove that $\boldsymbol{a}^{\mathrm{T}}\boldsymbol{c}' \leq \boldsymbol{a}^{\mathrm{T}}\boldsymbol{c} < b < \boldsymbol{a}^{\mathrm{T}}\boldsymbol{d} \leq \boldsymbol{a}^{\mathrm{T}}\boldsymbol{d}'$ for every $\boldsymbol{c}' \in C$ and $\boldsymbol{d}' \in D$. ③
   2. Since C is convex, $y = \boldsymbol{c} + \alpha(\boldsymbol{c}' - \boldsymbol{c}) \in C$ for every $0 \leq \alpha \leq 1$.
   3. From the minimality of the distance $||\boldsymbol{d} - \boldsymbol{c}||$ it follows that $||\boldsymbol{d} - y||^2 \geq ||\boldsymbol{d} - \boldsymbol{c}||^2$.
   4. Using elementary operations observe that $\frac{\alpha}{2}||\boldsymbol{c}' - \boldsymbol{c}||^2 + \boldsymbol{a}^{\mathrm{T}}\boldsymbol{c} \geq \boldsymbol{a}^{\mathrm{T}}\boldsymbol{c}'$ ④
   5. which holds for arbitrarily small $\alpha > 0$, it follows that $\boldsymbol{a}^{\mathrm{T}}\boldsymbol{c} \geq \boldsymbol{a}^{\mathrm{T}}\boldsymbol{c}'$ holds.

1. $||\boldsymbol{d}_n - \boldsymbol{c}|| \leq ||\boldsymbol{d}_n - \boldsymbol{c}_n|| + ||\boldsymbol{c}_n - \boldsymbol{c}|| \leq m + 1 + \max\{||c' - c''||;\ c', c'' \in C\} = z$
2. $||\boldsymbol{d} - \boldsymbol{c}|| \leq ||\boldsymbol{d} - \boldsymbol{d}_{l_n}|| + ||\boldsymbol{d}_{l_n} - \boldsymbol{c}_{l_n}|| + ||\boldsymbol{c}_{l_n} - \boldsymbol{c}|| \to m$
3. The inner two inequalities are obvious. We only prove the first inequality since the last one is analogous.
4. $$\begin{aligned} ||\boldsymbol{d} - y||^2 &\geq ||\boldsymbol{d} - \boldsymbol{c}||^2 \\ (\boldsymbol{d} - \boldsymbol{c} - \alpha(\boldsymbol{c}' - \boldsymbol{c}))^{\mathrm{T}}(\boldsymbol{d} - \boldsymbol{c} - \alpha(\boldsymbol{c}' - \boldsymbol{c})) &\geq (\boldsymbol{d} - \boldsymbol{c})^{\mathrm{T}}(\boldsymbol{d} - \boldsymbol{c}) \\ \alpha^2(\boldsymbol{c}' - \boldsymbol{c})^{\mathrm{T}}(\boldsymbol{c}' - \boldsymbol{c}) - 2\alpha(\boldsymbol{d} - \boldsymbol{c})^{\mathrm{T}}(\boldsymbol{c}' - \boldsymbol{c}) &\geq 0 \\ \frac{\alpha}{2}||\boldsymbol{c}' - \boldsymbol{c}||^2 + \boldsymbol{a}^{\mathrm{T}}\boldsymbol{c} &\geq \boldsymbol{a}^{\mathrm{T}}\boldsymbol{c}' \end{aligned}$$

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.

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

Minimal defining system of a polyhedron

Definition

$P = \{ \boldsymbol{x} \in \mathbb{R}^n;\ A'\boldsymbol{x} = \boldsymbol{b}',\ A''\boldsymbol{x} \le \boldsymbol{b}'' \}$ 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

Every polyhedron has a minimal defining system.

Lemma

Let $P = \{ \boldsymbol{x} \in \mathbb{R}^n;\ A'\boldsymbol{x} = \boldsymbol{b}',\ A''\boldsymbol{x} \le \boldsymbol{b}'' \}$ be a *minimal defining system* of a polyherdon P. Let $P' = \{ \boldsymbol{x} \in P;\ A''_{i,\star}\boldsymbol{x} = \boldsymbol{b}''_i \}$ for some row i of $A''\boldsymbol{x} \le \boldsymbol{b}''$. Then $\dim(P') < \dim(P)$.
1

Corollary

Let $P = \{x;\ A\boldsymbol{x} \le \boldsymbol{b}\}$ 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. There exists $x \in P \setminus P'$. Since $\operatorname{aff}(P') \subseteq \left\{\boldsymbol{x};\ A''_{i,\star}\boldsymbol{x} = \boldsymbol{b}''_i\right\}$, it follows that $x \notin \operatorname{aff}(P')$. Hence, $\dim(P') + 1 = \dim(P' \cup \{x\}) \leq \dim(P)$.

A point inside a polyhedron

Theorem

Let P be a non-empty polyhedron defined by a minimal system $\{\boldsymbol{x} \in \mathbb{R}^n;\ A'\boldsymbol{x} = \boldsymbol{b}',\ A''\boldsymbol{x} \le \boldsymbol{b}''\}$. Then,

1. there exists a point $\boldsymbol{z} \in P$ such that $A''\boldsymbol{z} < \boldsymbol{b}''$ and
2. $\dim(P) = n - \operatorname{rank}(A')$, and
3. and $\boldsymbol{z}$ does not belong in any proper face of P.

Proof

1. There exists a point $\boldsymbol{z} \in P$ such that $A''\boldsymbol{z} < \boldsymbol{b}''$.
   1. For every row i of $A''\boldsymbol{x} \le \boldsymbol{b}''$ there exists $\boldsymbol{z}^i \in P$ such that $A''_{i,\star}\boldsymbol{z}^i < \boldsymbol{b}''_i$.
   2. Let $\boldsymbol{z} = \frac{1}{m''} \sum_{i=1}^{m''} \boldsymbol{z}^i$ be the center of gravity.
   3. Since $\boldsymbol{z}$ is a convex combination of points of P, point $\boldsymbol{z}$ belongs to P and $A''\boldsymbol{z} < \boldsymbol{b}''$.
2. $\dim(P) = n - \operatorname{rank}(A')$
   1. Let L be the affine space defined by $A'\boldsymbol{x} = \boldsymbol{b}'$.
   2. There exists $\epsilon > 0$ such that P contains whole ball $B = \{\boldsymbol{x} \in L;\ ||x - z|| \le \epsilon\}$.
   3. Vectors of a base of the linear space $L - z$ can be scaled so that they belong into $B - z$.
   4. $\dim(L) \ge \dim(P) \ge \dim(B) \ge \dim(L) = n - \operatorname{rank}(A')$.
3. The point $\boldsymbol{z}$ does not belong in any proper face of p.
   1. The point $\boldsymbol{z}$ cannot belong into any proper face of P because a supporting hyperplane of such a face split the ball B.

A bijection between faces and inequalities

Theorem

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

Proof

1. Let $R_i = \{\boldsymbol{x};\ A''_{i,\star}\boldsymbol{x} = \boldsymbol{b}_i\}$ and $F_i = P \cap R_i$.
2. From minimality if follows that R_i is a supporting hyperplane, and therefore, F_i is a face.
3. There exists a point $\boldsymbol{y}^i \in F_i$ satisfying $A''_{j,\star}\boldsymbol{y}^i < \boldsymbol{b}''_j$ for all $j \neq i$. ①
4. So $\dim(F_i) = \dim(P) - 1$ and F_i is a facet.
5. Furthermore, $\boldsymbol{y}^i \notin F_j$ for all $j \neq i$, so $F_i \neq F_j$ for $j \neq i$.
6. For contradiction, let F be an another facet.
7. There exists a facet i such $F \subseteq F_i$. ②
8. 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.

1. From minimality it follows that there exists $\boldsymbol{x}$ satifying all conditions of P except $A''_{i,\star}\boldsymbol{x} < \boldsymbol{b}''_i$. Let $\boldsymbol{z}$ be a point from the previous theorem. A point $\boldsymbol{y}^i$ can be obtained as a convex combination of $\boldsymbol{x}$ and $\boldsymbol{z}$.
2. Otherwise $\frac{1}{m''}\sum_{i=1}^{m''}\boldsymbol{y}^i$ satisfies strictly all condition contradicting the assumption that F is a proper facet.

Fully dimensional polyhedron

Definition

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

Corollary

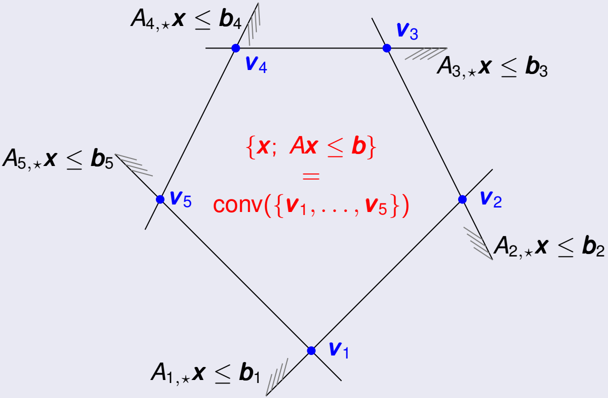
If P is a full-dimensional polyhedron, then P has exactly one minimal defining system up-to multiplying conditions by constants. 1

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

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



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 $S = \text{conv}(V)$.

Proof of the implication $\Rightarrow$ (main steps) by induction on $\dim(S)$

For $\dim(S) = 0$ the size of S is 1 and the statement holds. Assume that $\dim(S) > 0$.

1. Let $S = \left\{\boldsymbol{x} \in \mathbb{R}^n;\ A'\boldsymbol{x} = \boldsymbol{b}',\ A''\boldsymbol{x} \le \boldsymbol{b}''\right\}$ be a minimal defining system.
2. Let $S_i = \left\{\boldsymbol{x} \in S;\ A''_{i,\star}\boldsymbol{x} = \boldsymbol{b}''_i\right\}$ where i is a row of $A''\boldsymbol{x} \le \boldsymbol{b}''$.
3. Since $\dim(S_i) < \dim(S)$, there exists a finite set $V_i \subseteq \mathbb{R}^n$ such that $S_i = \text{conv}(V_i)$.
4. Let $V = \bigcup_i V_i$. We prove that $\text{conv}(V) = S$.
   - $\subseteq$ Follows from $V_i \subseteq S_i \subseteq S$ and convexity of 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 V_i and V_j, resp.
     Since $\boldsymbol{x}$ is a also a convex combination of $\boldsymbol{u}$ and $\boldsymbol{v}$, we have $\boldsymbol{x} \in \text{conv}(V)$.

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 $S = \operatorname{conv}(V)$.

Lemma

A condition $\boldsymbol{\alpha}^{\mathrm{T}}\boldsymbol{v} \le \beta$ is satisfied by all points $\boldsymbol{v} \in V$ if and only if the condition is satisfied by all points $\boldsymbol{v} \in \operatorname{conv}(V)$.

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

1. Let $Q = \left\{ \binom{\boldsymbol{\alpha}}{\beta};\ \boldsymbol{\alpha} \in \mathbb{R}^n, \beta \in \mathbb{R}, -\mathbf{1} \le \boldsymbol{\alpha} \le 1, -1 \le \beta \le 1, \boldsymbol{\alpha}^{\mathrm{T}}\boldsymbol{v} \le \beta\ \forall \boldsymbol{v} \in V \right\}$. ①
2. Since Q is a polytope, there exists a finite set $W \subseteq \mathbb{R}^{n+1}$ s.t. $Q = \operatorname{conv}(W)$. ②
3. Let $Y = \left\{ \boldsymbol{x} \in \mathbb{R}^n;\ \boldsymbol{\alpha}^{\mathrm{T}}\boldsymbol{x} \le \beta\ \forall \binom{\boldsymbol{\alpha}}{\beta} \in W \right\}$ and we prove that $\operatorname{conv}(V) = Y$.

   $\subseteq$ From $V \subseteq Y$ it follows that $\operatorname{conv}(V) \subseteq Y$. ③
   $\supseteq$ We prove that $\boldsymbol{x} \notin \operatorname{conv}(V) \Rightarrow \boldsymbol{x} \notin Y$.

   - There exists $\boldsymbol{\alpha} \in \mathbb{R}^n, \beta \in \mathbb{R}$ s.t. $\boldsymbol{\alpha}^{\mathrm{T}}\boldsymbol{x} > \beta$ and $\forall \boldsymbol{v} \in V : \boldsymbol{\alpha}^{\mathrm{T}}\boldsymbol{v} \le \beta$ ④
   - Assume that $-\mathbf{1} \le \boldsymbol{\alpha} \le 1, -1 \le \beta \le 1$. ⑤
   - Observe that $\binom{\boldsymbol{\alpha}}{\beta} \in Q$ and $\boldsymbol{x}$ fails at least one condition of Q.
   - Hence, $\boldsymbol{x}$ fails at least one condition of W. ⑥

1. Observe that $\boldsymbol{\alpha}^{\mathrm{T}}\boldsymbol{v} \le \beta$ means the same as $\binom{\boldsymbol{v}}{-1}^{\mathrm{T}}\binom{\boldsymbol{\alpha}}{\beta} \le 0$. Therefore, Q is described by $|V| + 2n + 2$ inequalities. Furthermore, conditions $-\mathbf{1} \le \boldsymbol{\alpha} \le 1$ and $-1 \le \beta \le 1$ implies that Q is bounded.
2. Here we use the implication $\Rightarrow$ of Minkovski-Weyl theorem which we already proved.
3. Every point of V satifies all conditions of Q since Q contains only conditions satisfied by all points of V. Since $W \subseteq \mathrm{conv}(W) = Q$, it follows that every point of V satifies all conditions of W. Hence, $V \subseteq Y$. Since Y is convex, the inclusion $\mathrm{conv}(V) \subseteq Y$.
4. Apply Hyperplane separation theorem on sets Q and $\{x\}$.
5. Scale the vector $\binom{\boldsymbol{\alpha}}{\beta}$ so that it fit into this box.
6. Use lemma.

Faces

Theorem

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

Proof (only for bounded polyhedrons)

- Let V_0 be (inclusion) minimal set such that $P = \operatorname{conv}(V_0)$.
- Let $V_e = \{\boldsymbol{x} \in P;\ \boldsymbol{x} \notin \operatorname{conv}(P \setminus \{\boldsymbol{x}\})\}$.
- We prove that $V = V_e = V_0$. 1

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

$V_0 \subseteq V$: Let $\boldsymbol{z} \in V_0$ and $D = \operatorname{conv}(V_0 \setminus \{\boldsymbol{z}\})$. From Minkovski-Weyl'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 $\boldsymbol{z}' \in A$.

Faces

Theorem (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.

Observation (Exercise)

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

Observation (Exercise)

A non-empty 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 a linear programming problem $\min\left\{\boldsymbol{c}^{\mathrm{T}}\boldsymbol{x};\ A\boldsymbol{x} \leq \boldsymbol{b}\right\}$ for some vector $\boldsymbol{c} \in \mathbb{R}^n$.

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 $\boldsymbol{x} \geq \boldsymbol{0}$ 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 $B \subseteq \{1, \dots, n\}$, 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 $N = \{1, \dots, n\} \setminus B$ 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 $A\boldsymbol{x} = A_B\boldsymbol{x}_B + A_N\boldsymbol{x}_N$.

Definitions

Consider the equation form $A\mathbf{x} = \mathbf{b}$ and $\mathbf{x} \geq \mathbf{0}$ with n variables and $\mathrm{rank}(A) = m$ rows.

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

Lemma

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

Observation

Basic feasible solutions are exactly vertices of the polyhedron
$P = \{\mathbf{x};\ A\mathbf{x} = \mathbf{b}, \mathbf{x} \geq \mathbf{0}\}$. ② ③

1. Remember that non-basic variables are always equal to zero.
2. If $\boldsymbol{x}$ is a basic feasible solution and B is the corresponding basis, then $\boldsymbol{x}_N = \mathbf{0}$ and so $K \subseteq B$ which implies that columns of A_K are also linearly independent.
   If columns of A_K are linearly independent, then we can extend K into B by adding columns of A so that columns of A_B are linearly independent which implies that B is a basis of $\boldsymbol{x}$.
3. Note that basic variables can also be zero. In this case, the basis B corresponding to a basic solution $\boldsymbol{x}$ may not be unique since there may be many ways to extend K into a basis B. This is called degeneracy.

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

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

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

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

Pivot

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

1. $(t, 0, 1 + t, 3 - t, 2)$ when $\boldsymbol{x}_1$ is increased by t
2. $(0, r, 1 - r, 3, 2 - r)$ when $\boldsymbol{x}_2$ is increased by r

These edges give feasible solutions for:

1. $t \leq 3$ since $\boldsymbol{x}_3 = 1 + t \geq 0$ and $\boldsymbol{x}_4 = 3 - t \geq 0$ and $\boldsymbol{x}_5 = 2 \geq 0$
2. $r \leq 1$ since $\boldsymbol{x}_3 = 1 - r \geq 0$ and $\boldsymbol{x}_4 = 3 \geq 0$ and $\boldsymbol{x}_5 = 2 - r \geq 0$

In both cases, the objective function is increasing. We choose $\boldsymbol{x}_2$ as a pivot.

Example: Pivot step

Simplex tableau

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

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 basis $B = \{2,4,5\}$

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

$$
\begin{array}{rclllll}
\boldsymbol{x}_2 & = & 1 & + & \boldsymbol{x}_1 & - & \boldsymbol{x}_3 \\
\boldsymbol{x}_4 & = & 3 & - & \boldsymbol{x}_1 & & \\
\boldsymbol{x}_5 & = & 1 & - & \color{red}{\boldsymbol{x}_1} & + & \boldsymbol{x}_3 \\
\hline
z & = & 1 & + & 2\boldsymbol{x}_1 & - & \boldsymbol{x}_3
\end{array}
$$

Next pivot

- Basis $B = \{2, 4, 5\}$ with a basic 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 $\boldsymbol{x}_2 = 1 + t \geq 0$ and $\boldsymbol{x}_4 = 3 - t \geq 0$ and $\boldsymbol{x}_5 = 1 - t \geq 0$.
- Therefore, $\boldsymbol{x}_1$ enters the basis and $\boldsymbol{x}_5$ leaves the basis.

New simplex tableau

$$
\begin{array}{rclllll}
\boldsymbol{x}_1 & = & 1 & + & \boldsymbol{x}_3 & - & \boldsymbol{x}_5 \\
\boldsymbol{x}_2 & = & 2 & & & - & \boldsymbol{x}_5 \\
\boldsymbol{x}_4 & = & 2 & - & \boldsymbol{x}_3 & + & \boldsymbol{x}_5 \\
\hline
z & = & 3 & + & \boldsymbol{x}_3 & - & 2\boldsymbol{x}_5
\end{array}
$$

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 basic 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

$$
\begin{array}{rcrcrcr}
\boldsymbol{x}_1 & = & 3 & - & \boldsymbol{x}_4 & & \\
\boldsymbol{x}_2 & = & 2 & & & - & \boldsymbol{x}_5 \\
\boldsymbol{x}_3 & = & 2 & - & \boldsymbol{x}_4 & + & \boldsymbol{x}_5 \\
\hline
z & = & 5 & - & \boldsymbol{x}_4 & - & \boldsymbol{x}_5
\end{array}
$$

No other pivot

- Basis $B = \{1, 2, 3\}$ with a basic 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 $\tilde{z} = 5 - \tilde{\boldsymbol{y}}_4 - \tilde{\boldsymbol{y}}_5 \leq 5 = z$.

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:

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

where $\boldsymbol{x}_B$ is the vector of the basic variables, $\boldsymbol{x}_N$ is the vector on non-basic 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

- $Q = -A_B^{-1}A_N$
- $\boldsymbol{p} = A_B^{-1}\boldsymbol{b}$ ①
- $z_0 = \boldsymbol{c}_B^{\mathrm{T}}A_B^{-1}\boldsymbol{b}$
- $r = \boldsymbol{c}_N - (\boldsymbol{c}_B^{\mathrm{T}}A_B^{-1}A_N)^{\mathrm{T}}$ ②

1. Since a matrix A_B is regular, we can multiply an equation $A_B\boldsymbol{x}_B + A_N\boldsymbol{x}_N = b$ by A_B^{-1} to obtain $\boldsymbol{x}_B = A_B^{-1}\boldsymbol{b} - A_B^{-1}A_N\boldsymbol{x}_N$, so $Q = -A_B^{-1}A_N$ and $\boldsymbol{p} = A_B^{-1}\boldsymbol{b}$.
2. The objective function is $\boldsymbol{c}_B^{\mathrm{T}}\boldsymbol{x}_B + \boldsymbol{c}_N^{\mathrm{T}}\boldsymbol{x}_N = \boldsymbol{c}_B^{\mathrm{T}}(A_B^{-1}\boldsymbol{b} - A_B^{-1}A_N\boldsymbol{x}_N) + \boldsymbol{c}_N^{\mathrm{T}}\boldsymbol{x}_N = \boldsymbol{c}_B^{\mathrm{T}}A_B^{-1}\boldsymbol{b} + (\boldsymbol{c}_N^{\mathrm{T}} - \boldsymbol{c}_B^{\mathrm{T}}A_B^{-1}A_N)\boldsymbol{x}_N$, so $z_0 = \boldsymbol{c}_B^{\mathrm{T}}A_B^{-1}\boldsymbol{b}$ and $r = \boldsymbol{c}_N - (\boldsymbol{c}_B^{\mathrm{T}}A_B^{-1}A_N)^{\mathrm{T}}$.

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 $\boldsymbol{r} \leq 0$. ①

Observation

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

1. The oposite implication may not hold for a degenerated optimal basis.

Simplex tableau in general

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

Find a pivot

- If $\boldsymbol{r} \leq \mathbf{0}$, then we have an optimal solution.
- Otherwise, choose an arbitrary entering variable $\boldsymbol{x}_v$ such that $\boldsymbol{r}_v > 0$.
- If $Q_{\star,v} \geq \mathbf{0}$, then the corresponding edge is unbounded and the problem is also unbounded. ①
- Otherwise, find a leaving variable $\boldsymbol{x}_u$ which limits the increment of the entering variable most strictly, i.e. $Q_{u,v} < 0$ and $-\frac{\boldsymbol{p}_u}{Q_{u,v}}$ is minimal.

Update the simplex tableau

Gaussian elimination: Express $\boldsymbol{x}_v$ from the row $\boldsymbol{x}_u = \boldsymbol{p}_u + Q_{u,\star}\boldsymbol{x}_N$ and substitute $\boldsymbol{x}_v$ using the obtained formula.

1. Consider the following edge: $\boldsymbol{x}_v = t$, remaining nonbasic variables are 0, and $\boldsymbol{x}_B = p + Q_{\star,v} t$. All solutions on this edge are feasible for $t \geq 0$ since $\boldsymbol{x} \geq \boldsymbol{0}$. For the objective value, $\boldsymbol{c}^{\mathrm{T}} \boldsymbol{x} = z_0 + \boldsymbol{r}^{\mathrm{T}} \boldsymbol{x}_N = z_0 + \boldsymbol{r}_v t \to \infty$ as $t \to \infty$, so the objective function is unbounded.

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 $\mathbf{x} \geq \mathbf{0}$.

Auxiliary linear program

- Multiply every row j with $\mathbf{b}_j < 0$ by -1. ①
- Introduce new variables $\mathbf{y} \in \mathbb{R}^m$ and solve an auxiliary linear program: Maximize $-\mathbf{1}^{\mathrm{T}}\mathbf{y}$ such that $A\mathbf{x} + I\mathbf{y} = \mathbf{b}$ and $\mathbf{x} \geq \mathbf{0}$, $\mathbf{y} \geq \mathbf{0}$.
- An initial basis contains variables $\mathbf{y}$ and an initial tableau is

$$\begin{array}{ccccc} \mathbf{y} & = & \mathbf{b} & + & A\mathbf{x} \\ \hline z & = & -\mathbf{1}^{\mathrm{T}}\mathbf{b} & + & (\mathbf{1}^{\mathrm{T}}A)\mathbf{x} \end{array}$$

- Whenever a variable of $\mathbf{y}$ become nonbasic, it can be removed from a tableau.
- When all variables of $\mathbf{y}$ are removed, express the original objective function $\mathbf{c}^{\mathrm{T}}\mathbf{x}$ using nonbasic variables and solve the problem.

Observation

The original linear program has a feasible solution if and only if an optimal solution of the auxiliary linear program satisfies $\mathbf{y} = \mathbf{0}$.

1. Now, assume that $\boldsymbol{b} \geq 0$.

Complexity

Degeneracy

- Different bases may correspond to the same solution. ①
- The simplex method may loop forever between these bases.
- 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 bases 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 $\binom{k}{3}$ bases 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

$$\begin{array}{lrcrcr} \text{Maximize} & 2\boldsymbol{x}_1 & + & 3\boldsymbol{x}_2 & & \\ \text{subject to} & 4\boldsymbol{x}_1 & + & 8\boldsymbol{x}_2 & \leq & 12 \\ & 2\boldsymbol{x}_1 & + & \boldsymbol{x}_2 & \leq & 3 \\ & 3\boldsymbol{x}_1 & + & 2\boldsymbol{x}_2 & \leq & 4 \\ & & & \boldsymbol{x}_1, \boldsymbol{x}_2 & \geq & 0 \end{array}$$

Simple estimates

- $2\boldsymbol{x}_1 + 3\boldsymbol{x}_2 \leq 4\boldsymbol{x}_1 + 8\boldsymbol{x}_2 \leq 12$ ①
- $2\boldsymbol{x}_1 + 3\boldsymbol{x}_2 \leq \frac{1}{2}(4\boldsymbol{x}_1 + 8\boldsymbol{x}_2) \leq 6$ ②
- $2\boldsymbol{x}_1 + 3\boldsymbol{x}_2 = \frac{1}{3}(4\boldsymbol{x}_1 + 8\boldsymbol{x}_2 + 2\boldsymbol{x}_1 + \boldsymbol{x}_2) \leq 5$ ③

What is the best combination of conditions?

Every non-negative linear combination of inequalities which gives an inequality $\boldsymbol{d}_1\boldsymbol{x}_1 + \boldsymbol{d}_2\boldsymbol{x}_2 \leq h$ with $d_1 \geq 2$ and $d_2 \geq 3$ provides the upper bound $2\boldsymbol{x}_1 + 3\boldsymbol{x}_2 \leq \boldsymbol{d}_1\boldsymbol{x}_1 + \boldsymbol{d}_2\boldsymbol{x}_2 \leq h$.

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

Consider a non-negative combination $\boldsymbol{y}$ of inequalities

$$\begin{array}{lrcrcrl}
\text{Maximize} & 2\boldsymbol{x}_1 & + & 3\boldsymbol{x}_2 & & & \\
\text{subject to} & 4\boldsymbol{x}_1 & + & 8\boldsymbol{x}_2 & \leq & 12 & \color{red}{/ \cdot \boldsymbol{y}_1} \\
 & 2\boldsymbol{x}_1 & + & \boldsymbol{x}_2 & \leq & 3 & \color{red}{/ \cdot \boldsymbol{y}_2} \\
 & 3\boldsymbol{x}_1 & + & 2\boldsymbol{x}_2 & \leq & 4 & \color{red}{/ \cdot \boldsymbol{y}_3} \\
 & & & \boldsymbol{x}_1, \boldsymbol{x}_2 & \geq & 0 &
\end{array}$$

Observations

- Every feasible solution $\boldsymbol{x}$ and non-negative combination $\boldsymbol{y}$ satisfies $(4\boldsymbol{y}_1 + 2\boldsymbol{y}_2 + 3\boldsymbol{y}_3)\boldsymbol{x}_1 + (8\boldsymbol{y}_1 + \boldsymbol{y}_2 + 2\boldsymbol{y}_3)\boldsymbol{x}_2 \leq 12\boldsymbol{y}_1 + 3\boldsymbol{y}_2 + 4\boldsymbol{y}_3$.
- If $4\boldsymbol{y}_1 + 2\boldsymbol{y}_2 + 3\boldsymbol{y}_3 \geq 2$ and $8\boldsymbol{y}_1 + \boldsymbol{y}_2 + 2\boldsymbol{y}_3 \geq 3$, then $12\boldsymbol{y}_1 + 2\boldsymbol{y}_2 + 4\boldsymbol{y}_3$ is an upper for the objective function.

Dual program ①

$$\begin{array}{lrcrcrcl}
\text{Minimize} & 12\boldsymbol{y}_1 & + & 2\boldsymbol{y}_2 & + & 4\boldsymbol{y}_3 & & \\
\text{subject to} & 4\boldsymbol{y}_1 & + & 2\boldsymbol{y}_2 & + & 3\boldsymbol{y}_3 & \geq & 2 \\
 & 8\boldsymbol{y}_1 & + & \boldsymbol{y}_2 & + & 2\boldsymbol{y}_3 & \geq & 3 \\
 & & & & & \boldsymbol{y}_1, \boldsymbol{y}_2, \boldsymbol{y}_3 & \geq & 0
\end{array}$$

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

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

- Minimize $\boldsymbol{b}^{\mathrm{T}}\boldsymbol{y}$ subject to $A^{\mathrm{T}}\boldsymbol{y} \geq \boldsymbol{c}$ and $\boldsymbol{y} \geq \mathbf{0}$ — Dual program
- -Maximize $-\boldsymbol{b}^{\mathrm{T}}\boldsymbol{y}$ subject to $A^{\mathrm{T}}\boldsymbol{y} \geq \boldsymbol{c}$ and $\boldsymbol{y} \geq \mathbf{0}$ — Equivalent formulation
- -Minimize $\boldsymbol{c}^{\mathrm{T}}\boldsymbol{x}$ subject to $A\boldsymbol{x} \geq -\boldsymbol{b}$ and $\boldsymbol{x} \leq \mathbf{0}$ — Dual of the dual program
- -Minimize $-\boldsymbol{c}^{\mathrm{T}}\boldsymbol{x}$ subject to $-A\boldsymbol{x} \geq -\boldsymbol{b}$ and $\boldsymbol{x} \geq \mathbf{0}$ — Simplified formulation
- Maximize $\boldsymbol{c}^{\mathrm{T}}\boldsymbol{x}$ subject to $A\boldsymbol{x} \leq \boldsymbol{b}$ and $\boldsymbol{x} \geq \mathbf{0}$ — The original primal program

	Maximizing program	Minimizing program
Variables	$x_1, \dots, x_n$	$y_1, \dots, y_m$
Matrix	A	A^{T}
Right-hand side	b	c
Objective function	$\max c^{\mathrm{T}} x$	$\min b^{\mathrm{T}} y$
Constraints	i-th constraint has $\leq$	$y_i \geq 0$
	i-th constraint has $\geq$	$y_i \leq 0$
	i-th constraint has $=$	$y_i \in \mathbb{R}$
	$x_j \geq 0$	j-th constraint has $\geq$
	$x_j \leq 0$	j-th constraint has $\leq$
	$x_j \in \mathbb{R}$	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

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

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

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

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

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 y and z 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

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

Overview

- Eliminate the variable y, find a feasible evaluation of z a and compute y a x.
- In every step, we eliminate one variable; however, the number of conditions may increase quadratically.
- If we start with m conditions, then after n eliminations the number of conditions is up to $4(m/4)^{2^n}$.

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

Proposition (Farkas lemma geometrically)

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

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

Farkas lemma

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

Observation (Exercise)

Variants of Farkas lemma are equivalent.

Proposition (Farkas lemma, 3rd version)

Let $A \in R^{m \times n}$ and $\boldsymbol{b} \in \mathbb{R}^m$. Then, the system $A\boldsymbol{x} \le \boldsymbol{b}$ 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 $\boldsymbol{y}^\mathrm{T}\boldsymbol{b} \ge 0$.

Proof (overview)

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

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

$n = 0$
- The system $A\boldsymbol{x} \le \boldsymbol{b}$ equals to $\mathbf{0} \le \boldsymbol{b}$ 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 $A'\boldsymbol{x}' \le \boldsymbol{b}'$
- There exists a non-negative matrix M such that $(\mathbf{0}|A') = MA$ and $\boldsymbol{b}' = M\boldsymbol{b}$
- By induction, there exists $\boldsymbol{y}' \ge 0$, $\boldsymbol{y}'^\mathrm{T} A' = \mathbf{0}^\mathrm{T}$, $\boldsymbol{y}'^\mathrm{T}\boldsymbol{b}' < 0$
- We verify that $\boldsymbol{y} = M^\mathrm{T}\boldsymbol{y}'$ satisfies all requirements of the induction

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

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

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

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 $A\boldsymbol{x} \le \boldsymbol{b}$ has a non-negative solution if and only if every non-negative $\boldsymbol{y} \in \mathbb{R}^m$ with $\boldsymbol{y}^{\mathrm{T}} A \ge \mathbf{0}^{\mathrm{T}}$ satisfies $\boldsymbol{y}^{\mathrm{T}}\boldsymbol{b} \ge 0$.

Duality

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

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 $A\boldsymbol{x} \le \boldsymbol{b}$ and $\boldsymbol{x} \ge \mathbf{0}$ and $\boldsymbol{c}^{\mathrm{T}}\boldsymbol{x} \ge \gamma + \epsilon$ is infeasible
3. $\epsilon > 0$ iff $\binom{A}{-\boldsymbol{c}^{\mathrm{T}}}\boldsymbol{x} \le \binom{\boldsymbol{b}}{-\gamma-\epsilon}$ and $\boldsymbol{x} \ge \mathbf{0}$ is infeasible
4. $\epsilon > 0$ iff $\boldsymbol{u}, z \ge 0$ and $\binom{\boldsymbol{u}}{z}^{\mathrm{T}}\binom{A}{-\boldsymbol{c}^{\mathrm{T}}} \ge \mathbf{0}^{\mathrm{T}}$ and $\binom{\boldsymbol{u}}{z}^{\mathrm{T}}\binom{\boldsymbol{b}}{-\gamma-\epsilon} < 0$ is feasible
5. $\epsilon > 0$ iff $\boldsymbol{u}, z \ge 0$ and $A^{\mathrm{T}}\boldsymbol{u} \ge z\boldsymbol{c}$ 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

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

Update formulas (without proof)

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

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

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 ϵ (without proof)

A non-empty rational fully-dimensional polytope contains a ball with radius ϵ 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

Matching problems

Perfect matching problem

Input: Graph (V, E)

Output: Perfect matching $M \subseteq E$ if it exists

Minimum weight perfect matching problem

Input: Graph (V, E) and weights $\boldsymbol{c}_e \geq 0$ on edges $e \in E$ ①

Output: Perfect matching $M \subseteq E$ minimizing the weight $\sum_{e \in M} \boldsymbol{c}_e$

Overview

1. Tools: Augmenting paths, Tutte-Berge formula, alternating trees
2. Perfect matching in bipartite graphs without weights
3. Minimum weight perfect matching in bipartite graphs
4. Tool: Shrinking odd circuits
5. Perfect matching in general graphs without weights
6. Minimum weight perfect matching in general graphs
7. Maximum weight matching

1. In the perfect matching problem, we can add a constant to weights of all edges without changing the set of all optimal perfect matchings. Therefore, if some edge has a negative weight, we can add a sufficiently large constant to all weights to ensure non-negativity of $\boldsymbol{c}$.

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

Definitions

- Let def(G) be the number of exposed vertices by a maximum size matching in G.
- Let $oc(G)$ be the number of odd components of a graph G. ①

Observations

- $\text{def}(G) \geq \text{oc}(G)$
- For every $A \subseteq V$ it holds that $\text{def}(G) \geq \text{oc}(G \setminus A) - |A|$. ②

Tutte's matching theorem

A graph G has a perfect matching if and only if $\text{oc}(G \setminus A) \leq |A|$ for every $A \subseteq V$. ③

Theorem: Tutte-Berge Formula (without proof)

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

1. A component of a graph is odd if it has odd number of vertices.
2. Every odd component has at least one exposed vertex.
3. $\Rightarrow$ If a graph G has a perfect matching, then $\mathrm{def}(G) = 0$, so from the previous observation it follows that $\mathrm{oc}(G \setminus A) \leq |A|$.

   $\Leftarrow$ We will present an algorithmic proof which finds a perfect matching or a subset $A \subseteq V$ such that $\mathrm{oc}(G \setminus A) > |A|$.

Building an alternating tree

Initialization of M-alternating tree T on vertices $A \dot\cup B$

$T = A = \emptyset$ and $B = \{r\}$ where r is an M-exposed root. ①

Use $uv \in E$ to extend T

Input: An 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 $\{uv, vz\}$ and A by v and B by z.

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.
- $|B| = |A| + 1$

Use $uv \in E$ to augment M

Input: An 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 u in T. Replace M by $M \triangle E(P)$.

1. An M-alternating tree T with the root r on vertices A and B is a tree obtained from this initialization by applying the following operation extend.

Frustrated tree

Definition

M-alternating tree T is M-frustrated if every edge of G having one end vertex in B has the other end vertex in A. ①

Observation

If a bipartite graph G has a matching M and an frustrated M-alternating tree, then G has no perfect matching. ② ③

1. That is, an M-alternating tree is frustrated if neither operation extend nor augment can be applied. Note that in bipartite graphs, there is no edge between vertices of B.
2. B are single vertex components in the graph $G \setminus A$. Therefore, $\text{oc}(G \setminus A) \geq |B| > |A|$.
3. This proves that Tutte's matching theorem for bipartite graphs: From every M-exposed vertex r we build an M-alternating tree T such that T can be used to augment M to cover r or T is frustrated.

Algorithm for perfect matching problem in a bipartite graph

Algorithm

Init: $M := \emptyset$

while *G contains an M-exposed vertex r* ① **do**

  $A := \emptyset$ and $B = \{r\}$ `# Build an` M`-alternating tree from` r.

  while *there exists $uv \in E$ with $u \in B$ and $v \notin A \cup B$* **do**

    if *v is M-covered* **then**

      Use uv to extend T

    else

      Use uv to augment M

      break `# Terminate the inner loop.`

  if *r is still M-exposed* ② **then**

    return There is no perfect matching `#` T `is a frustrated tree.`

return Perfect matching M

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.

1. Actually, it suffices to once iterate over all vertices.
2. That is, the augmentation was no applied.

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 $\boldsymbol{x} \geq \mathbf{0}$ where A is the incidence matrix.

Dual

Maximize $\mathbf{1}^{\mathrm{T}}\boldsymbol{y}$ 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

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

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 minimum weight.

Idea of the algorithm

- If there exists an $M_{\boldsymbol{x}}$-augmenting path P in $(V, E_{\boldsymbol{y}})$, then use P to augment $M_{\boldsymbol{x}}$.
- Otherwise, use a frustrated $M_{\boldsymbol{x}}$-alternating tree in $(V, E_{\boldsymbol{y}})$ to update the dual solution $\boldsymbol{y}$ and enlarge $E_{\boldsymbol{y}}$.

Algorithm for minimum weight perfect matchings in a bipartite graph

Algorithm

Init: $M := \emptyset$ and $\boldsymbol{y} = \boldsymbol{0}$
while *G contains an M-exposed vertex r* **do**
  $A := \emptyset$ and $B = \{r\}$ `# Build an` M`-alternating tree from` r.
  while *r is M-exposed* **do**
    if *there exists $uv \in E_{\boldsymbol{y}}$ with $u \in B$ and $v \notin A \cup B$* **then**
      if *v is M-covered* **then**
        Use uv to extend T ①
      else
        Use uv to augment M ②
    else if *there exists $uv \in E$ with $u \in B$ and $v \notin A \cup B$* **then**
      $\epsilon = \min\{c_{uv} - \boldsymbol{y}_u - \boldsymbol{y}_v;\ u, v \in E, u \in B, v \notin A \cup B\}$
      $\boldsymbol{y}_u := \boldsymbol{y}_u + \epsilon$ for all $u \in B$
      $\boldsymbol{y}_v := \boldsymbol{y}_v - \epsilon$ for all $v \in A$ ③ ④
    else
      return There is no perfect matching in G. ⑤
return Minimum weight perfect matching M

1. Note that T uses only tight edges.
2. Invariants are satisfied since M is augmented by edges of T which are tight.
3. Observe that $\boldsymbol{y}$ remains a dual feasible solution. Furthermore, no edge is removed from the tight set $E_{\boldsymbol{y}}$ and at least one edge become tight. Therefore, all invariants remain satisfied.
4. In the next iteration, an edge uv minimizing ϵ is used to extend T or augment M.
5. T is a frustrated M-alternating tree in G. Also note that the dual problem is unbounded since ϵ is unbounded in this case.

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 ①

- $E(G)$ 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

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

Remark

There exists a graph G with odd circuit C such that $\operatorname{def}(G) < \operatorname{def}(G \times C)$.

1. Formally, $E(G \times C) = \{uv;\ uv \in E(G),\ u, v \in V(G) \setminus V(C)\} \cup \{uc';\ \exists v \in V(C) :\ uv \in E(G), u \in V(G) \setminus V(C)\}$.

Perfect matching in general graphs

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 $M' \setminus E(C)$
- 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.

Proof is based on Tutte's matching theorem

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

Perfect matchings algorithm in a non-weighted graph

Algorithm

```
Init: M := ∅
while G contains an M-exposed vertex r do
    M' = M, G' = G and T = ({r}, ∅)
    while there exists uv ∈ E(G') with u ∈ B and v ∉ A do
        if v ∈ B then
            Use uv to shrink and update M' and T
        else if v is M'-covered then
            Use uv to extend T
        else
            Use uv to augment M'
            Extend M' to a matching M of G
            break # Terminate the inner loop.
    if r is still M-exposed then
        return There is no perfect matching
return Perfect matching M
```

Algorithm

Init: $M' := \emptyset$, $G' = G$
while *G' contains an M'-exposed vertex r* **do**
  $T = (\{r\}, \emptyset)$
  while *r is M'-exposed* **do**
    if *there exists $uv \in E(G')$ with $u \in B$ and $v \notin A$* **then**
      if $v \in B$ **then**
        Use uv to shrink and update M' and T
      else if *v is M'-covered* **then**
        Use uv to extend T
      else
        Use uv to augment M'
    else if *there exists a pseudonode u in A* **then**
      Expand u into a circuit and update T, M' and G'
    else
      return There is no perfect matching
Expand all pseudonodes and obtain M from M'
return Perfect matching M

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

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

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.

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

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

Dual

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

Notation: Reduced cost

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

An edge e is tight if $\bar{\boldsymbol{c}}_e = 0$ and let E_y be the set of tight edges.

Complementary slackness

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

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

Init: $M' := \emptyset$, $G' = G$
while *G' contains an M'-exposed vertex r* **do**
    $T = (\{r\}, \emptyset)$
    while *r is M'-exposed* **do**
        if *there exists $uv \in E_=(G')$ with $u \in B$ and $v \notin A$* **then**
            if $v \in B$ **then**
                Use uv to shrink and update M' and T
            else if *v is M'-covered* **then**
                Use uv to extend T
            else
                Use uv to augment M'
        else if *there exists a pseudonode u in A with $y'_u = 0$* **then**
            Expand u into a circuit and update T, M' and G'
        else
            Determine ϵ and change $\boldsymbol{y}$
            if $\epsilon = \infty$ **then**
                return There is no perfect matching

Expand all pseudonodes and obtain M from M'
return Perfect matching M

Maximum-weight (general) matching

Reduction to perfect matching problem

Let G be a graph with weights $\boldsymbol{c} \in \mathbb{R}^E$.

- 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 $\mathrm{oc}(G \setminus A) \leq |A|$ for every $A \subseteq V$.