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 2016/17
Last change on May 11, 2017

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

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

Matrix notation of Linear programming problem

Formulation using linear programming

$$
\begin{array}{lrcrcrcl}
\text{Minimize} & 0.75x_1 & + & 0.5x_2 & + & 0.15x_3 & & \\
\text{subject to} & 35x_1 & + & 0.5x_2 & + & 0.5x_3 & \geq & 0.5 \\
 & 60x_1 & + & 300x_2 & + & 10x_3 & \geq & 15 \\
 & 30x_1 & + & 20x_2 & + & 10x_3 & \geq & 4 \\
 & & & & & x_1, x_2, x_3 & \geq & 0
\end{array}
$$

Matrix notation

- Minimize

$$
\begin{pmatrix} 15 \\ 10 \\ 3 \end{pmatrix}^T \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix}
$$

- Subject to

$$
\begin{pmatrix} 35 & 0.5 & 0.5 \\ 60 & 300 & 10 \\ 30 & 20 & 10 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} \geq \begin{pmatrix} 0.5 \\ 15 \\ 4 \end{pmatrix}
$$

- a $x_1, x_2, x_3 \geq 0$

Notation: Vector and matrix

Matrix

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

Vector

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

Scalar

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

Special vectors

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

Transpose

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

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

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

Mathematical optimization

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 $A\boldsymbol{x} = \boldsymbol{b}, \boldsymbol{x} \geq \boldsymbol{0}$

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 $A\boldsymbol{x} \leq \boldsymbol{b}, -A\boldsymbol{x} \leq -\boldsymbol{b}, -\boldsymbol{x} \leq \boldsymbol{0}$

Conversion from the canonical form to the equation form

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

Basic terminology

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

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

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

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

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

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 $\boldsymbol{x} \geq 0$

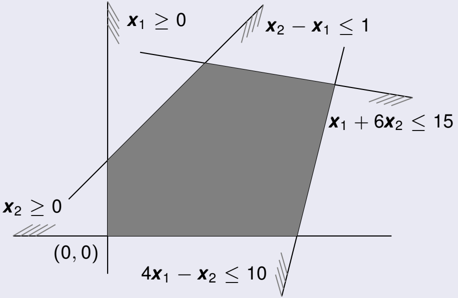
Graphical method: Set of feasible solutions

Example

Draw the set of all feasible solutions $(\boldsymbol{x}_1, \boldsymbol{x}_2)$ satisfying the following conditions.

$$\begin{array}{rcrcl} \boldsymbol{x}_1 & + & 6\boldsymbol{x}_2 & \leq & 15 \\ 4\boldsymbol{x}_1 & - & \boldsymbol{x}_2 & \leq & 10 \\ -\boldsymbol{x}_1 & + & \boldsymbol{x}_2 & \leq & 1 \\ & & \boldsymbol{x}_1, \boldsymbol{x}_2 & \geq & 0 \end{array}$$

Solution

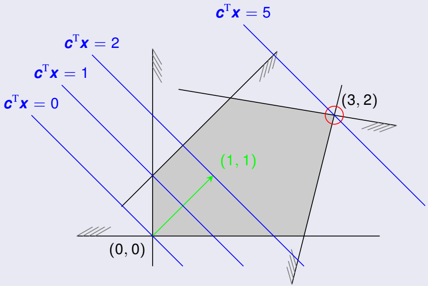


Example

Find the optimal solution of the following problem.

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

Solution



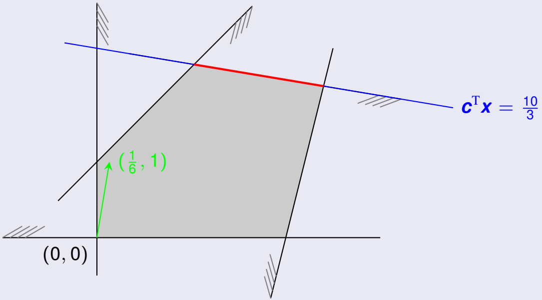
Graphical method: Multiple optimal solutions

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



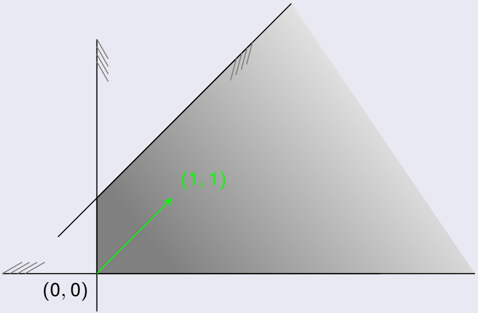
Graphical method: Unbounded problem

Example

Show that the following problem is unbounded.

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

Solution

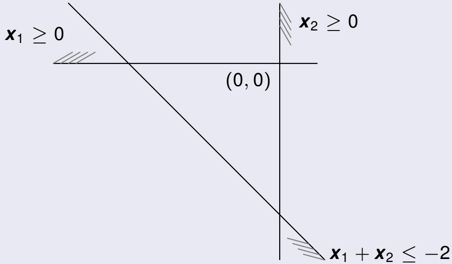


Example

Show that the following problem has no feasible solution.

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

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 the *ellipsoid* and the *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

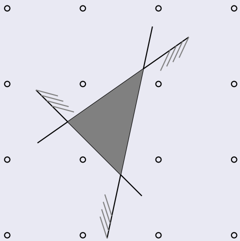
Variables: cover $\boldsymbol{x} \in \{0, 1\}^V$ (i.e. $\mathbf{0} \leq \boldsymbol{x} \leq \mathbf{1}$ and $\boldsymbol{x} \in \mathbb{Z}^V$)

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

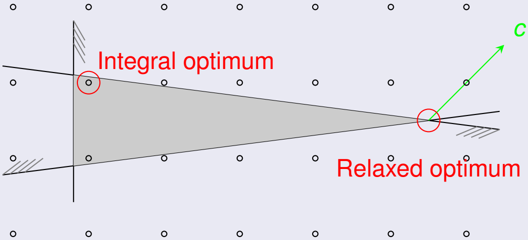
Objective function: Minimize $\mathbf{1}^{\mathrm{T}} \boldsymbol{x}$

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)
- Storage facilities for 1 ton of ice cream cost a per 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

- Variable $\boldsymbol{x}_i$ determines the amount of produced ice cream in month $i \in \{0, \ldots, n\}$
- Variable $\boldsymbol{s}_i$ determines the amount of stored ice cream from month $i - 1$ month i
- The stored quantity is computed by $\boldsymbol{s}_i = \boldsymbol{s}_{i-1} + \boldsymbol{x}_i - \boldsymbol{d}_i$ for every $i \in \{1, \ldots, n\}$
- Durability is ensured by $\boldsymbol{s}_i \leq \boldsymbol{d}_i$ for all $i \in \{1, \ldots, n\}$
- Non-negativity of the production and the storage $\boldsymbol{x}, \boldsymbol{s} \geq \mathbf{0}$
- 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
- Let $\boldsymbol{y}_i \geq 0$ and $\boldsymbol{z}_i \geq 0$ be the increment and the decrement of production, reps., and $\boldsymbol{x}_i - \boldsymbol{x}_{i-1} = \boldsymbol{y}_i - \boldsymbol{z}_i$
- Linear programming problem formulation

$$\begin{array}{lccll} \text{Minimize} & b \sum_{i=1}^{n} (\boldsymbol{y}_i + \boldsymbol{z}_i) + a \sum_{i=1}^{n} \boldsymbol{s}_i & & & \\ \text{subject to} & \boldsymbol{s}_{i-1} - \boldsymbol{s}_i + \boldsymbol{x}_i & = & \boldsymbol{d}_i & \text{for } i \in \{1, \ldots, n\} \\ & \boldsymbol{s}_i & \leq & \boldsymbol{d}_i & \text{for } i \in \{1, \ldots, n\} \\ & \boldsymbol{x}, \boldsymbol{s}, \boldsymbol{y}, \boldsymbol{z} & \geq & \mathbf{0} & \end{array}$$

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

Finding shortest paths from a vertex s in an oriented graph

Linear programming problem

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

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

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

Outline

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 $\alpha \in T$.

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

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 $\boldsymbol{x} = \boldsymbol{0}$ 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 $V = \{\boldsymbol{x};\ A\boldsymbol{x} = \boldsymbol{0}\}$. Hence, $V + \boldsymbol{a} = \{\boldsymbol{x} + \boldsymbol{a};\ A\boldsymbol{x} = \boldsymbol{0}\} = \{\boldsymbol{y};\ A\boldsymbol{y} - A\boldsymbol{a} = \boldsymbol{0}\} = \{\boldsymbol{y};\ A\boldsymbol{y} = \boldsymbol{b}\}$ 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, $L - \boldsymbol{y} = \{\boldsymbol{x} - \boldsymbol{y};\ A\boldsymbol{x} = \boldsymbol{b}\} = \{\boldsymbol{z};\ A\boldsymbol{y} + A\boldsymbol{z} = \boldsymbol{b}\} = \{\boldsymbol{z};\ A\boldsymbol{z} = \boldsymbol{0}\}$ is a linear space since $A\boldsymbol{y} = \boldsymbol{b}$.

Convex set

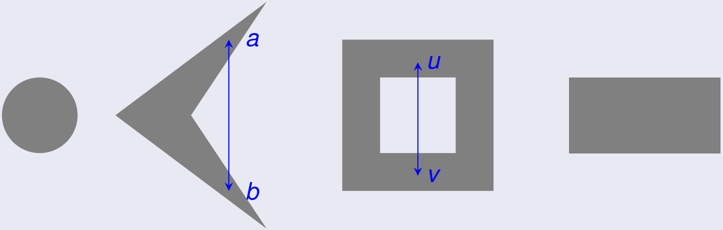
Observation

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



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 $\operatorname{conv}(S) = \bigcap_{M \supseteq S, M \text{ convex}}$ 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 $\sum_{i=1}^k \alpha_i \mathbf{v}_i$ be a linear combination of S and $c \in \mathbb{R}$ be a scalar. Then, $c \sum_{i=1}^k \alpha_i \mathbf{v}_i = \sum_{i=1}^k (c\alpha_i) \mathbf{v}_i$ 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 $\sum_{i=1}^k \alpha_i \mathbf{u}_i$ and $\sum_{j=1}^l \beta_j \mathbf{v}_j$ be two convex combinations of S. In order to prove that the set of all convex combinations of S contains the line segment between $\sum_{i=1}^k \alpha_i \mathbf{u}_i$ and $\sum_{j=1}^l \beta_j \mathbf{v}_j$, let us consider $\gamma_1, \gamma_2 \geq 0$ such that $\gamma_1 + \gamma_2 = 1$. Then, $\gamma_1 \sum_{i=1}^k \alpha_i \mathbf{u}_i + \gamma_2 \sum_{j=1}^l \beta_j \mathbf{v}_j = \sum_{i=1}^k (\gamma_1 \alpha_i) \mathbf{u}_i + \sum_{j=1}^l (\gamma_2 \beta_j) \mathbf{v}_j$ 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 $\sum_{i=1}^k \alpha_i \mathbf{v}_i$ be an affine combination of S. Since $S - \mathbf{v}_k$ is a linear space, the linear combination $\sum_{i=1}^k \alpha_i (\mathbf{v}_i - \mathbf{v}_k)$ of $S - \mathbf{v}_k$ belongs into $S - \mathbf{v}_k$. Hence, $\mathbf{v}_k + \sum_{i=1}^k \alpha_i (\mathbf{v}_i - \mathbf{v}_k) = \sum_{i=1}^k \alpha_i \mathbf{v}_i$ belongs to S.
6. We prove by induction on k that S contains every convex combination $\sum_{i=1}^k \alpha_i \mathbf{v}_i$ of S. The statement holds for $k \leq 2$ by the definition of a convex set. Let $\sum_{i=1}^k \alpha_i \mathbf{v}_i$ 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 $\sum_{i=1}^k \alpha_i \mathbf{v}_i = \mathbf{v}_k \in S$. Hence, $\sum_{i=1}^k \alpha_i \mathbf{v}_i = (1 - \alpha_k) \sum_{i=1}^k \frac{\alpha_i}{1-\alpha_k} \mathbf{v}_i + \alpha_k \mathbf{v}_k = (1 - \alpha_k) \mathbf{y} + \alpha_k \mathbf{v}_k$ 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 $\text{conv}(S) \subseteq T$. The definition states that $\text{conv}(S) = \bigcap_{M \supseteq S, M \text{ convex}} M$ and we proved that T is a convex set containing S, so T is included in this intersection which implies that $\text{conv}(S)$ is a subset of T.
   In order to prove $\text{conv}(S) \supseteq T$, we again consider the intersection $\text{conv}(S) = \bigcap_{M \supseteq S, M \text{ convex}} M$. 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 $\text{conv}(S) \supseteq T$.

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 (Homework)

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

Observation

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

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

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 $\boldsymbol{x} \in \text{conv}(S)$. Let $\boldsymbol{x} = \sum_{i=1}^k \alpha_i \boldsymbol{x}_i$ be a convex combination of points of S with the smallest k. If $\boldsymbol{x}_1, \dots, \boldsymbol{x}_k$ 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

Notation

Notation used in the Simplex method

- Equation form: Maximize $\mathbf{c}^{\mathrm{T}}\mathbf{x}$ such that $A\mathbf{x} = \mathbf{b}$ and $\mathbf{x} \geq \mathbf{0}$ where $A \in \mathbb{R}^{m \times n}$ and $\mathbf{b} \in \mathbb{R}^m$.
- 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, $\mathbf{x}_B$ denotes the coordinates of $\mathbf{x}$ 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

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

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

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

1. For a system $A\boldsymbol{x} = \boldsymbol{b}$ with n variables and n linearly independent conditions, there exists the inverse matrix A^{-1} and the only feasible solution of $A\boldsymbol{x} = \boldsymbol{b}$ is $\boldsymbol{x}^\star = A^{-1}\boldsymbol{b}$.
2. Consider a system $A\boldsymbol{x} \leq \boldsymbol{b}$ with $n = \operatorname{rank}(A)$ variables and $m \geq n$ conditions and select n linearly independent rows $A'\boldsymbol{x} \leq \boldsymbol{b}'$. Then, the system $A'\boldsymbol{x} = \boldsymbol{b}'$ has a solution $\boldsymbol{x}^\star = A'^{-1}\boldsymbol{b}'$.
   Moreover, if $A\boldsymbol{x}^\star \leq \boldsymbol{b}$, then $\boldsymbol{x}^\star$ is a vertex of the polyhedron $A\boldsymbol{x} \leq \boldsymbol{b}$. ①
3. Consider the equation form $A\boldsymbol{x} = \boldsymbol{b}$ and $\boldsymbol{x} \geq \boldsymbol{0}$ and let N be $n - m$ rows of $\boldsymbol{x} \geq \boldsymbol{0}$.
   If rows of the system $A\boldsymbol{x} = \boldsymbol{0}$ and $\boldsymbol{x}_N = \boldsymbol{0}$ are linearly independent, then $\boldsymbol{b} = A\boldsymbol{x} = A_B\boldsymbol{x}_B + \mathrm{A}_N\boldsymbol{x}_N = A_B\boldsymbol{x}_B$, so $\boldsymbol{x}^\star = (\boldsymbol{x}^\star_B, \boldsymbol{x}^\star_N) = (A_B^{-1}\boldsymbol{b}, \boldsymbol{0})$ where $B = \{1, \ldots, n\} \setminus N$.
   Moreover, if $\boldsymbol{x}^\star_B \geq \boldsymbol{0}$, then $\boldsymbol{x}^\star$ is a vertex of $A\boldsymbol{x} = \boldsymbol{b}$ and $\boldsymbol{x} \geq \boldsymbol{0}$.
4. Consider the equation form again. If we choose m linearly independent columns B of A, then conditions $A\boldsymbol{x} = \boldsymbol{b}$ and $\boldsymbol{x}_N = \boldsymbol{0}$ are linearly independent.

1. The solution $\mathbf{x}^\star = A'^{-1}\mathbf{b}'$ will be called a basis solution. Vertices of a polyhedron will be formally defined later, so we use a geometrical intuition now.

Basic feasible solutions

Definitions

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

- A set $B \subseteq \{1, \dots, n\}$ of linearly independent columns of A is called a basis. ①
- 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 a *basic feasible solution*.
- $\mathbf{x}_B$ are called basis variables and $\mathbf{x}_N$ are called non-basis variables. ②

Observation

A feasible solution $\mathbf{x}$ of a $A\mathbf{x} = \mathbf{b}$ and $\mathbf{x} \geq \mathbf{0}$ is basis if and only if columns of A_K are linearly independent where $K = \{j \in \{1, \dots, n\}\,;\ \mathbf{x}_j > 0\}$. ③ ④

Observation

Linear program in the equation form has at most $\binom{n}{m}$ basis solutions. ⑤

1. Observe that $B \subseteq \{1, \ldots, n\}$ is a basis if and only if A_B is a regular matrix.
2. Remember that non-basis variables are always equal to zero.
3. 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 $\boldsymbol{x}$.
4. Note that basis variables can also be zero. In this case, the basis B corresponding to a basis solution $\boldsymbol{x}$ may not be unique since there may be many ways to extend K into a basis B. This is called degeneracy.
5. There are $\binom{n}{m}$ subsets $B \subseteq \{1, \ldots, n\}$ and for some of these subsets A_B may not be regular.

Optimal basis feasible solutions

Theorem

If the linear program $\max \mathbf{c}^{\mathrm{T}}\mathbf{x}$ subject to $A\mathbf{x} = \mathbf{b}$ and $\mathbf{x} \geq \mathbf{0}$ has a feasible solution and the objective function is bounded from above of the set of all feasible solutions, then there exists an optimal solution.
Moreover, if an optimal solution exists then there is a basis feasible solution which is optimal. ①

Lemma

If the objective function of a linear program in the equation form is bounded above, then for every feasible solution $\mathbf{x}'$ there exists a basis feasible solution $\mathbf{x}^\star$ with the same or larger value of the objective function, i.e. $\mathbf{c}^{\mathrm{T}}\mathbf{x}^\star \geq \mathbf{c}^{\mathrm{T}}\mathbf{x}'$. ②

1. If the problem is bounded, one may try to find the optimal solution by finding all basis feasible solutions. However, this is not an efficient algorithm since the number of basis grows exponentially.

2.
   - Let $\boldsymbol{x}^\star$ be a feasible solution with $\boldsymbol{c}^{\mathrm{T}}\boldsymbol{x}^\star \geq \boldsymbol{c}^{\mathrm{T}}\boldsymbol{x}'$ and the smallest possible size of the set $K = \left\{j \in \{1, \dots, n\} ;\ \boldsymbol{x}^\star_j > 0\right\}$. Let $N = \{1, \dots, n\} \setminus K$.
   - If columns of A_K are linearly independent, then $\boldsymbol{x}^\star$ is a basis solution.
   - There exists a non-zero vector $\boldsymbol{v}_K$ such that $A_K \boldsymbol{v}_K = \mathbf{0}$. Let $\boldsymbol{v}_N = \mathbf{0}$.
   - WLOG: $\boldsymbol{c}^{\mathrm{T}}\boldsymbol{v} \geq \mathbf{0}$ since we can replace $\boldsymbol{v}$ by $-\boldsymbol{v}$.
   - Consider the line $x(t) = \boldsymbol{x}^\star + t\boldsymbol{v}$ for $t \in \mathbb{R}$.
   - For every $t \in \mathbb{R}$: $Ax(t) = \boldsymbol{b}$ and $(x(t))_N = \mathbf{0}$.
   - For every $t \geq 0$: $\boldsymbol{c}^{\mathrm{T}}x(t) \geq \boldsymbol{c}^{\mathrm{T}}\boldsymbol{x}$.
   - If $\boldsymbol{c}^{\mathrm{T}}\boldsymbol{v} > 0$ and $\boldsymbol{v} \geq \mathbf{0}$, then points $x(t)$ are feasible for every $t \geq 0$ and the objective function $\boldsymbol{c}^{\mathrm{T}}\boldsymbol{x}(t) = \boldsymbol{c}^{\mathrm{T}}\boldsymbol{x}^\star + t\boldsymbol{c}^{\mathrm{T}}\boldsymbol{v}$ converges to infinity which contradicts assuptions.
   - If $\boldsymbol{v}_j < 0$ for some $j \in K$, then consider $j \in K$ with $\boldsymbol{v}_j < 0$ and minimal $\frac{\boldsymbol{x}^\star_j}{-\boldsymbol{v}_j}$. Let $\bar{t} = \frac{\boldsymbol{x}^\star_j}{-\boldsymbol{v}_j}$. Since $x(\bar{t}) \geq 0$ and $(x(\bar{t}))_j = 0$, the solution $x(\bar{t})$ is feasible with smaller number of positive components than $\boldsymbol{x}^\star$ which is a contradiction.
   - The remaining case is $\boldsymbol{c}^{\mathrm{T}}\boldsymbol{v} = 0$ and $\boldsymbol{v}_j \geq \mathbf{0}$. Since $\boldsymbol{v}_K$ is a non-trivial combination, there exists $j \in K$ with $\boldsymbol{v}_j > \mathbf{0}$. Replace $\boldsymbol{v}$ by $-\boldsymbol{v}$ and apply the previous case.

Convex polyhedrons

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

Examples

- n-dimensional hypercube: $\{\boldsymbol{x} \in \mathbb{R}^n;\ \mathbf{0} \leq \boldsymbol{x} \leq \mathbf{1}\}$
- n-dimensional crosspolytope: $\{\boldsymbol{x} \in \mathbb{R}^n;\ \sum_{i=1}^n |\boldsymbol{x}_i| \leq 1\}$ ①
- n-dimensional simplex: $\{\boldsymbol{x} \in \mathbb{R}^{n+1};\ \boldsymbol{x} \geq \mathbf{0},\ \mathbf{1}\boldsymbol{x} = 1\}$ ②

1. Formally, $\sum_{i=1}^{n} |\boldsymbol{x}_i| \leq 1$ is not a linear inequality. However, it can be replaced by 2^n linear inequalities $\boldsymbol{dx} \leq 1$ for all $\boldsymbol{d} \in \{-1, 1\}^n$.
2. n-dimensional simplex is a convex hull of $n + 1$ affinely independent points.

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.

Definition

If $\boldsymbol{\alpha}^{\mathrm{T}}\boldsymbol{x} \leq \beta$ is a supporting hyperplane of a polyhedron P, then $P \cap \left\{\boldsymbol{x};\ \boldsymbol{\alpha}^{\mathrm{T}}\boldsymbol{x} = \beta\right\}$ is called a *face* of P.
By convention, the empty set and P are also called 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.

Vertices

Observations

- The set of all optimal solutions of a linear program $\max \mathbf{c}^{\mathrm{T}}\mathbf{x}$ over a polyhedron P is a face of P. ①
- Every proper face of P is a set of all optimal solutions of a linear program $\max \mathbf{c}\mathbf{x}$ over a polyhedron P for some $\mathbf{c} \in \mathbb{R}^n$. ② ③
- Vertices are unique solutions of linear programs $\max \mathbf{c}^{\mathrm{T}}\mathbf{x}$ over P for some $\mathbf{c}$.

Theorem

Let P be the set of all solutions of a linear program in the equation form and $\mathbf{v} \in P$. Then the following statements are equivalent.

1. v is a vertex of a polyhedron P.
2. v is a basis feasible solution of the linear program. ④

Theorem

If the linear program $\max \mathbf{c}^{\mathrm{T}}\mathbf{x}$ subject to $A\mathbf{x} = \mathbf{b}$ and $\mathbf{x} \geq \mathbf{0}$ has a feasible solution and the objective function is bounded from above of the set of all feasible solutions, then there exists an optimal solution.
Moreover, if an optimal solution exists then there is a basis feasible solution which is optimal.

1. Let F be the set of all optimal solutions. If $F = \emptyset$ or $F = P$, then F is a face of P by definition. Otherwise, $d = \max\left\{\boldsymbol{c}^{\mathrm{T}}\boldsymbol{x};\ \boldsymbol{x} \in P\right\}$ exists. Since $\boldsymbol{c}^{\mathrm{T}}\boldsymbol{x} = d$ is a supporting hyperplane of P and $F = P \cap \left\{\boldsymbol{x};\ \boldsymbol{c}^{\mathrm{T}}\boldsymbol{x} = d\right\}$, it follows that F is a face of P.
2. A proper face F of P is defined as the intersection of P and a supporting hyperplane $\boldsymbol{c}^{\mathrm{T}}\boldsymbol{x} = d$, so F is the set of all optimal solutions of the linear program $\max \boldsymbol{c}^{\mathrm{T}}\boldsymbol{x}$ over P.
3. Note that P is also the set of all optimal solutions of a linear program for $\boldsymbol{c} = \mathbf{0}$. On the other hand, if P is non-empty and bounded, then the empty set cannot be express as a set of all optimal solutions for any $\boldsymbol{c}$.
4.
   - $\Rightarrow$ Follows from the following theorem.
   - $\Leftarrow$ Let B be the basis defining v and let $\boldsymbol{c}_B = \mathbf{0}$ and $\boldsymbol{c}_N = \mathbf{-1}$. Then $\boldsymbol{c}^{\mathrm{T}}\boldsymbol{v} = \boldsymbol{c}_B^{\mathrm{T}}\boldsymbol{v}_B + \boldsymbol{c}_N^{\mathrm{T}}\boldsymbol{v}_N = 0$ and for every feasible $\boldsymbol{x}$ it holds holds that $\boldsymbol{x} \geq \mathbf{0}$, so $\boldsymbol{c}^{\mathrm{T}}\boldsymbol{x} \leq 0$. Hence, $\boldsymbol{v}$ is a optimal solution of the linear program with the objective function $\max \boldsymbol{c}^{\mathrm{T}}\boldsymbol{x}$. Furthermore, $\boldsymbol{v}$ is the only optimal solution since every optimal solution $\boldsymbol{x}$ must satisfy $\boldsymbol{x}_N = 0$. In this case, $\boldsymbol{x}_B = A_B^{-1}\boldsymbol{b}$ is unique.

Example: Initial simplex tableau

Canonical form

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

Equation form

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

Simplex tableau

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

Example: Initial simplex tableau

Simplex tableau

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

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

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

Basis

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

New simplex tableau

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

Example: Next step

Simplex tableau

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

Next pivot

- Basis $B = \{2, 4, 5\}$ with a basis feasible solution $(0, 1, 0, 3, 1)$.
- This vertex has two incident edges but only one increases the objective function.
- The edge with 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

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

Example: Last step

Simplex tableau

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

Next pivot

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

New simplex tableau

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

Example: Optimal solution

Simplex tableau

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

No other pivot

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

Why this is an optimal solution?

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

Example: Unboundedness

Canonical form

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

Equation form

$$
\begin{array}{lrcrcrcrcl}
\text{Maximize} & x_1 & & & & & & & & \\
\text{subject to} & x_1 & - & x_2 & + & x_3 & & & = & 1 \\
 & -x_1 & + & x_2 & & & + & x_4 & = & 2 \\
 & & & & & & & x_1, x_2, x_3, x_4 & \geq & 0
\end{array}
$$

Initial simplex tableau

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

Example: Unboundedness

Simplex tableau

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

First pivot

- Basis $B = \{3, 4\}$ with a basis feasible solution $(0, 0, 1, 2)$.
- This vertex has two incident edges but only one increases the objective function.
- The edge with increasing objective function is $(t, 0, 1 - t, 2 + t)$.
- Feasible solutions for $x_3 = 1 - t \geq 0$ and $x_4 = 2 + t \geq 0$.
- Therefore, x_1 enters the basis and x_3 leaves the basis.

Simplex tableau

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

Example: Unboundedness

Simplex tableau

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

Unboundedness

- Basis $B = \{1, 4\}$ with a basis feasible solution $(1, 0, 0, 3)$.
- This vertex has two incident edges but only one increases the objective function.
- The edge with increasing objective function is $(1 + t, t, 0, 3)$.
- Every point $(1 + t, t, 0, 3)$ for $t \geq 0$ is feasible.
- The value of the objective function is $1 + t$.
- Therefore, this problem is unbounded.

Example: Degeneracy

Canonical form

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

Convert to the equation form by adding slack variables x_3 and x_4.

Initial simplex tableau

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

Basis feasible solution $(0, 0, 0, 2)$ with the basis $\{3, 4\}$.

Same solution with different basis

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

Basis feasible solution $(0, 0, 0, 2)$ with the basis $\{2, 4\}$.

Optimal simplex tableau

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

Optimal solution $(2, 2, 0, 0)$ with the basis $\{1, 2\}$.

Definition

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

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

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

Example

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

$$Q = \begin{pmatrix} 1 & -1 \\ -1 & 0 \end{pmatrix}$$

$$\boldsymbol{x}_B = \begin{pmatrix} \boldsymbol{x}_3 \\ \boldsymbol{x}_4 \end{pmatrix}, \boldsymbol{x}_N = \begin{pmatrix} \boldsymbol{x}_1 \\ \boldsymbol{x}_2 \end{pmatrix}, \boldsymbol{p} = \begin{pmatrix} 5 \\ 2 \end{pmatrix}$$

$$z_0 = 3,\ \boldsymbol{r}^{\mathrm{T}} = (1, 2)$$

Simplex tableau in general

Definition

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

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

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

Observation

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

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

1
- Since $A_B \boldsymbol{x}_B + A_N \boldsymbol{x}_N = \boldsymbol{b}$ and A_B is a regular matrix,
- it follows that $\boldsymbol{x}_B = A_B^{-1}\boldsymbol{b} - A_B^{-1} A_N \boldsymbol{x}_N$
- where $A_B^{-1}\boldsymbol{b} = \boldsymbol{p}$ and $A_B^{-1} A_N = Q$.
- The objective function is $\boldsymbol{c}_B^{\mathrm{T}}\boldsymbol{x}_B + \boldsymbol{c}_N^{\mathrm{T}}\boldsymbol{x}_N = \boldsymbol{c}_B^{\mathrm{T}}\mathrm{A}_B^{-1}\boldsymbol{b} - (\boldsymbol{c}_B^{\mathrm{T}} A_B^{-1} A_N + \boldsymbol{c}_N^{\mathrm{T}})\boldsymbol{x}_N$,
- where $\boldsymbol{c}_B^{\mathrm{T}}\mathrm{A}_B^{-1}\boldsymbol{b} = z_0$ and $\boldsymbol{c}_B^{\mathrm{T}} A_B^{-1} A_N + \boldsymbol{c}_N^{\mathrm{T}} = r^{\mathrm{T}}$.

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

If $\boldsymbol{r} \leq 0$, then the solution corresponding to a basis B is optimal.

Idea of the pivot step

Choose $v \in N$. Which is the last feasible point of the half-line $\boldsymbol{x}(t)$ for $t \geq 0$ where

- $\boldsymbol{x}_v(t) = t$
- $\boldsymbol{x}_{N\setminus\{v\}}(t) = \mathbf{0}$
- $\boldsymbol{x}_B(t) = \boldsymbol{p} + Q_{\star,v}t$?

Observation

If there exists a non-basis variable $\boldsymbol{x}_v$ such that $r_v > 0$ and $Q_{\star,v} \geq 0$, then the problem is unbounded.

Simplex tableau in general

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

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 problem is also unbounded.
- 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.

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, e.i. $\boldsymbol{c}^{\mathrm{T}}(\boldsymbol{x}_{new} - \boldsymbol{x}_{old})$ is maximal.

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.

Pivot step

Simplex tableau in general

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

Gaussian elimination

- New basis variables are $(B \setminus \{u\}) \cup \{v\}$ and new non-basis variables are $(N \setminus \{v\}) \cup \{u\}$
- Row $\boldsymbol{x}_u = \boldsymbol{p}_u + Q_{u,v}\boldsymbol{x}_v + \sum_{j \in N \setminus \{v\}} Q_{u,j}\boldsymbol{x}_j$ is replaced by
- row $\boldsymbol{x}_v = \frac{\boldsymbol{p}_u}{-Q_{u,v}} + \frac{1}{Q_{u,v}}\boldsymbol{x}_u + \sum_{j \in N \setminus \{v\}} \frac{Q_{u,j}}{-Q_{u,v}}\boldsymbol{x}_j$.
- Rows $\boldsymbol{x}_i = \boldsymbol{p}_i + Q_{i,v}\boldsymbol{x}_v + \sum_{j \in N \setminus \{v\}} Q_{i,j}\boldsymbol{x}_j$ for $i \in B \setminus \{u\}$ are replaced by
- rows $\boldsymbol{x}_i = (\boldsymbol{p}_i + \frac{Q_{i,v}}{-Q_{u,v}}\boldsymbol{p}_u) + \frac{Q_{i,v}}{Q_{u,v}}\boldsymbol{x}_u + \sum_{j \in N \setminus \{v\}} (Q_{i,j} + \frac{Q_{u,j}Q_{i,v}}{-Q_{u,v}})\boldsymbol{x}_j$.
- Objective function $z = z_0 + \boldsymbol{r}_v\boldsymbol{x}_v + \sum_{j \in N \setminus \{v\}} \boldsymbol{r}_j\boldsymbol{x}_j$ is replaced by
- objective function $z = (z_0 + \frac{\boldsymbol{p}_u}{-Q_{u,v}}) + \frac{\boldsymbol{r}_v}{-Q_{u,v}}\boldsymbol{x}_u + \sum_{j \in N \setminus \{v\}} (\boldsymbol{r}_j + \frac{\boldsymbol{r}_v Q_{i,v}}{-Q_{u,v}})\boldsymbol{x}_j$.

Observation

Pivot step does not change the set of all feasible solutions.

Bland's rule

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

Let B is a basis with the corresponding solution $\boldsymbol{x}'$ and let $\bar{B}$ a new basis with the corresponding solution $\bar{\boldsymbol{x}}$ after a single pivot step. Then, $\boldsymbol{x}' = \bar{\boldsymbol{x}}$ or $\boldsymbol{c}^{\mathrm{T}}\boldsymbol{x}' < \boldsymbol{c}^{\mathrm{T}}\bar{\boldsymbol{x}}$. ①

Observation

If the simplex method loops endlessly, then basis occuring in the loop correspond to the same vertex. ②

Theorem

The simplex method with Bland's pivot rule is always finite. ③

1. Consider the half-line $\boldsymbol{x}(t)$ providing the pivot step and let $\bar{t} = \max\{t \geq 0;\ \boldsymbol{x}(t) \geq 0\}$. Clearly, $\boldsymbol{c}^{\mathrm{T}}\bar{\boldsymbol{x}} = \boldsymbol{x}(\bar{t})$. If $\bar{t} = 0$, then $\bar{\boldsymbol{x}} = \boldsymbol{x}(0) = \boldsymbol{x}'$. If $\bar{t} > 0$, then $\boldsymbol{c}^{\mathrm{T}}\bar{\boldsymbol{x}} = \boldsymbol{c}^{\mathrm{T}}\boldsymbol{x}(\bar{t}) = z_0 + \boldsymbol{r}_v\bar{t} > z_0 = \boldsymbol{c}^{\mathrm{T}}\boldsymbol{x}'$ since $\boldsymbol{r}_t > 0$.
2. Consider that the simplex method iteraters over basis $B^(1), \dots, B^{(k)}, B^{(k+1)} = B^{(1)}$ with the corresponding solutions $\boldsymbol{x}^{(1)}, \dots, \boldsymbol{x}^{(k)}, \boldsymbol{x}^{(k+1)} = \boldsymbol{x}^{(1)}$. By the previous observation holds that $\boldsymbol{c}^{\mathrm{T}}\boldsymbol{x}^{(1)} \leq \boldsymbol{c}^{\mathrm{T}}\boldsymbol{x}^{(2)} \leq \dots \leq \boldsymbol{c}^{\mathrm{T}}\boldsymbol{x}^{(k)} \leq \boldsymbol{c}^{\mathrm{T}}\boldsymbol{x}^{(k+1)} = \boldsymbol{c}^{\mathrm{T}}\boldsymbol{x}^{(1)}$. Hence, $\boldsymbol{c}^{\mathrm{T}}\boldsymbol{x}^{(1)} = \dots = \boldsymbol{c}^{\mathrm{T}}\boldsymbol{x}^{(k+1)}$ and the previous observation implies that $\boldsymbol{x}^{(1)} = \dots = \boldsymbol{x}^{(k+1)}$.
3. For the sake of contradiction, we assume that the simplex method with Bland's pivot rule loops endlessly. Consider all basis in the loop. Let F be the set of all entering variables and let $\boldsymbol{x}_v \in F$ be the variable with largest index. Let B be a basis in the loop just before $\boldsymbol{x}_v$ enters. Note that variables of $B \setminus F$ and $N \setminus B$ are always basis and non-basis variables during the loop, respectively. Consider the following auxiliary problem.

$$\begin{array}{llcl}
\text{Maximize} & \boldsymbol{c}^{\mathrm{T}}\boldsymbol{x} & & \\
\text{subject to} & A\boldsymbol{x} & = & \boldsymbol{b} \\
 & \boldsymbol{x}_{F\setminus\{v\}} & \geq & \boldsymbol{0} \\
 & \boldsymbol{x}_v & \leq & 0 \\
 & \boldsymbol{x}_{N\setminus F} & = & \boldsymbol{0} \\
 & \boldsymbol{x}_{B\setminus F} & \in & \mathbb{R}^{|B\setminus F|}
\end{array} \qquad (\star)$$

We prove that $(\star)$ has an optimal solution and it is also unbounded which is a contradiction.

- $\boldsymbol{r}_v > 0$ since $\boldsymbol{x}_v$ is the entering variable
- $\boldsymbol{r}_i \le 0$ for every $i \in (F \cap N) \setminus \{v\}$ since $\boldsymbol{x}_v$ is the improving variable with the smallest index (Bland's rule)
- For every solution $\boldsymbol{x}$ satisfying $(\star)$ holds that $\boldsymbol{c}^{\mathrm{T}}\boldsymbol{x} = z_0 + \boldsymbol{r}^{\mathrm{T}}\boldsymbol{x}_N = z_0 + \boldsymbol{r}_v^{\mathrm{T}}\boldsymbol{x}_v + \boldsymbol{r}^{\mathrm{T}}_{(F\cap N)\setminus\{v\}}\boldsymbol{x}_{(F\cap N)\setminus\{v\}} + \boldsymbol{r}^{\mathrm{T}}_{N\setminus F}\boldsymbol{x}_{N\setminus F} \le z_0$.
- Hence, the solution corresponding to the basis B is an optimal solution to $(\star)$.

Now, we prove that $(\star)$ is unbounded.

- Let B be a basis in the loop just before $\boldsymbol{x}_v$ leaves and let Q', $\boldsymbol{p}'$ and $\boldsymbol{r}'$ be the parameter of the simplex tableau corresponding to B'.
- Let $\boldsymbol{x}_u$ be the entering variable. Hence, $\boldsymbol{r}'_u > 0$.
- $Q'_{v,u} < 0$ since v is the leaving variable.
- From Bland's rule it follows that $Q'_{i,u} \ge 0$ for every $i \in (F \cap B') \setminus \{v\}$
- $\boldsymbol{p}'_{F\cap B'} = \boldsymbol{0}$ since degenerated basis variables are zero
- Consider the half-line $\boldsymbol{x}(t)$ for $t \ge 0$ where $\boldsymbol{x}_u(t) = t$ and $\boldsymbol{x}_{N'\setminus\{v\}}(t) = \boldsymbol{0}$ and $\boldsymbol{x}_{B'}(t) = \boldsymbol{p}' + Q'_{\star,v}t$.
- $\boldsymbol{x}_{(F\cap N')\setminus\{u\}}(t) = \boldsymbol{0}$ since non-basis variables remains zero
- $\boldsymbol{x}_i(t) = \boldsymbol{p}'_i + Q'_{i,u}t \ge 0$ for every $i \in (F \cap B') \setminus \{v\}$
- Hence, $\boldsymbol{x}_{F\setminus\{v\}}(t) \ge \boldsymbol{0}$
- $\boldsymbol{x}_v(t) = \boldsymbol{p}'_v + Q'_{v,u}t \le 0$
- $\boldsymbol{x}_{N'\setminus F}(t) = \boldsymbol{0}$ since non-basis variables remains zero
- Hence, $\boldsymbol{x}(t)$ satisfies $(\star)$ for every $t \ge 0$
- $\boldsymbol{r}'_u > 0$ since $\boldsymbol{x}_u$ is the entering variable
- $\boldsymbol{c}^{\mathrm{T}}\boldsymbol{x}(t) = z'_0 + \boldsymbol{r}'_u t \to \infty$ for $t \to \infty$
- Hence, $(\star)$ is unbounded.

Initial feasible basis

Linear programming problem in the equation form

- Maximize $\boldsymbol{c}^{\mathrm{T}}\boldsymbol{x}$ subject to $A\boldsymbol{x} = \boldsymbol{b}$ and $\boldsymbol{x} \geq 0$.
- Assume that $\boldsymbol{b} \geq \mathbf{0}$ ①

Auxiliary problem

We add auxiliary variables $\boldsymbol{y} \in \mathbb{R}^m$ to obtain the auxiliary problem maximize $-\boldsymbol{y}_1 - \cdots - \boldsymbol{y}_m$ subject to $A\boldsymbol{x} + I\boldsymbol{y} = \boldsymbol{b}$ a $\boldsymbol{x}, \boldsymbol{y} \geq \mathbf{0}$.

Observation

Initial feasible basis for the auxiliary problem is $B = \{\boldsymbol{y}_1, \ldots, \boldsymbol{y}_m\}$ with the initial tableau

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

Observation

The following statements are equivalent

1. The original problem $\max \left\{\boldsymbol{c}^{\mathrm{T}}\boldsymbol{x};\ A\boldsymbol{x} = \boldsymbol{b},\ \boldsymbol{x} \geq 0\right\}$ has a feasible solution.
2. Optimal value of the objective function of the auxiliary problem is 0.
3. Auxiliary problem has a feasible solution satisfying $\boldsymbol{y} = \mathbf{0}$.

1. We multiply every equation with negative right hand side by -1.

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

$$
\begin{array}{lrcrcrl}
\text{Maximize} & 2x_1 & + & 3x_2 & & & \\
\text{subject to} & 4x_1 & + & 8x_2 & \leq & 12 & / \cdot y_1 \\
 & 2x_1 & + & x_2 & \leq & 3 & / \cdot y_2 \\
 & 3x_1 & + & 2x_2 & \leq & 4 & / \cdot y_3 \\
 & & & x_1, x_2 & \geq & 0 &
\end{array}
$$

Observations

- Every feasible solution $\boldsymbol{x}$ and non-negative combination $\boldsymbol{y}$ satisfies $(4y_1 + 2y_2 + 3y_3)x_1 + (8y_1 + y_2 + 2y_3)x_2 \leq 12y_1 + 3y_2 + 4y_3$.
- If $4y_1 + 2y_2 + 3y_3 \geq 2$ and $8y_1 + y_2 + 2y_3 \geq 3$, then $12y_1 + 2y_2 + 4y_3$ is an upper for the objective function.

Dual program ①

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

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

Duality of linear programming: General

Primal linear program

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

Dual linear program

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

Weak duality theorem

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

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 $\mathbf{x}$ and $\mathbf{y}$ such that $\mathbf{c}^{\mathrm{T}}\mathbf{x} = \mathbf{b}^{\mathrm{T}}\mathbf{y}$

Dualization

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

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

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

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

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

Feasibility versus optimality

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

Using duality

The optimal solutions of linear programs

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

are exactly feasible solutions satisfying

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

Complementary slackness

Theorem

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

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

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

Proof of duality using simplex method with Bland's rule

Notation

- Primal: Maximize $\mathbf{c}^{\mathrm{T}}\mathbf{x}$ subject to $A\mathbf{x} \leq \mathbf{b}$ and $\mathbf{x} \geq \mathbf{0}$
- Primal with slack variables: Maximize $\bar{\mathbf{c}}^{\mathrm{T}}\bar{\mathbf{x}}$ subject to $\bar{A}\bar{\mathbf{x}} = \mathbf{b}$ and $\bar{\mathbf{x}} \geq \mathbf{0}$ ①
- Dual: Minimize $\mathbf{b}^{\mathrm{T}}\mathbf{y}$ subject to $A^{\mathrm{T}}\mathbf{y} \geq \mathbf{c}$ and $\mathbf{y} \geq \mathbf{0}$

Simplex tableau

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

Simplex tableau is unique for every basis B

- $Q = -\bar{A}_B^{-1}\bar{A}_N$
- $\mathbf{p} = \bar{A}_B^{-1}\mathbf{b}$
- $z_0 = \bar{\mathbf{c}}^{\mathrm{T}}{}_B\bar{A}_B^{-1}\mathbf{b}$
- $\mathbf{r} = \bar{\mathbf{c}}_N - (\bar{\mathbf{c}}^{\mathrm{T}}{}_B\bar{A}_B^{-1}\bar{A}_N)^{\mathrm{T}}$

Lemma

If B is a basis with an optimal solution $\bar{\mathbf{x}}^\star$ of the primal problem, then $\mathbf{y}^\star = (\bar{\mathbf{c}}^{\mathrm{T}}\bar{A}_B^{-1})^{\mathrm{T}}$ is an optimal solution of the dual problem and $\mathbf{c}^{\mathrm{T}}\mathbf{x}^\star = \mathbf{b}^{\mathrm{T}}\mathbf{y}^\star$. ②

1. $\bar{\boldsymbol{x}}$ is obtained from $\boldsymbol{x}$ by adding slack variables. So, $\bar{A} = (A|I)$ and $\bar{\boldsymbol{c}}^{\mathrm{T}} = (\boldsymbol{c}^{\mathrm{T}}, \mathbf{0})$.
2.
   - The primal optimal solution is $\bar{\boldsymbol{x}}^{\star}_B = \bar{A}_B^{-1}\boldsymbol{b}$ and $\bar{\boldsymbol{x}}_N = \mathbf{0}$
   -

Proof of duality using simplex method with Bland's rule

Lemma

If B is a basis with an optimal solution $\bar{\boldsymbol{x}}^\star$ of the primal problem, then $\boldsymbol{y}^\star = (\bar{\boldsymbol{c}}^{\mathrm{T}} \bar{A}_B^{-1})^{\mathrm{T}}$ is an optimal solution of the dual problem and $\boldsymbol{c}^{\mathrm{T}} \boldsymbol{x}^\star = \boldsymbol{b}^{\mathrm{T}} \boldsymbol{y}^\star$.

Duality theorem (shorted version)

If the primal problem is feasible and bounded, the dual problem has an optimal solution with the same optimum value as the primal.

Corollary of the weak duality theorem

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

Duality theorem (longer version)

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 $\boldsymbol{c}^{\mathrm{T}} \boldsymbol{x} = \boldsymbol{b}^{\mathrm{T}} \boldsymbol{y}$

Fourier–Motzkin elimination: Example

Goal: Find a feasible solution

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

Express the variable x in each condition

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

Eliminate the variable x

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

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

Fourier–Motzkin elimination: Example

Rewrite into a system of inequalities

Real numbers y and z satisfy
$\max\left\{7+5y-2z, -4+\frac{2}{3}y+2z\right\} \leq \min\left\{5+\frac{5}{2}y-2z, 3+2y-z, 3-2y+\frac{1}{5}z\right\}$ if and only they satisfy

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

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

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

Proof

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

$$\begin{aligned} j \in C, k \in F: \quad (A_{j,\star} + A_{k,\star})x &\leq b_j + b_k \quad (1) \\ l \in L: \quad A_{l,\star}x &\leq b_l \quad (2) \end{aligned}$$

4. Assuming $A'x' \leq b'$ has a solution x', we find a solution x of $Ax \leq b$:
   - (1) is equivalent to $A'_{k,\star}x' - b_k \leq b_j - A'_{j,\star}x'$ for all $j \in C, k \in F$,
   - which is equivalent to $\max_{k \in F}\left\{A'_{k,\star}x' - b_k\right\} \leq \min_{j \in C}\left\{b_j - A'_{j,\star}x'\right\}$
   - Choose x_1 between these bounds and $x = (x_1, x')$ satisfies $Ax \leq b$

Farkas lemma

Proposition (Farkas lemma, 3rd version)

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

Proof (overview)

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

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

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

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

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

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

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

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 $\boldsymbol{y}^\mathrm{T}\boldsymbol{b} \geq 0$.

Proof of the equivalence of variants of Farkas lemma

Exercise :)

Farkas lemma

Definition

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

Proposition (Farkas lemma geometrically)

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

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

Proposition (Farkas lemma)

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

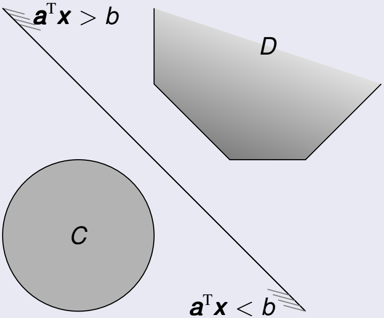
1. There exists a vector $\boldsymbol{x} \in \mathbb{R}^n$ satisfying $A\boldsymbol{x} = \boldsymbol{b}$ and $\boldsymbol{x} \geq \mathbf{0}$.
2. There exists a vector $\boldsymbol{y} \in \mathbb{R}^m$ satisfying $\boldsymbol{y}^{\mathrm{T}}A \geq \mathbf{0}$ and $\boldsymbol{y}^{\mathrm{T}}b < \mathbf{0}$.

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



Mathematical analysis

Definition

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

Theorem

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

Theorem

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

Infimum and supremum

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

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 $C \subseteq \{\boldsymbol{x}; \boldsymbol{a}^{\mathrm{T}}\boldsymbol{x} < b\}$ and $D \subseteq \{\boldsymbol{x}; \boldsymbol{a}^{\mathrm{T}}\boldsymbol{x} > b\}$.

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 $\{\boldsymbol{c}_{k_n}\}_{n=1}^{\infty}$ converging to $\boldsymbol{c} \in C$.
    4. There exists $z \in \mathbb{R}$ such that for every $n \in \mathbb{N}$ the distance $||\boldsymbol{d}_n - \boldsymbol{c}||$ is at most z. ①
    5. Since the set $D \cap \{\boldsymbol{x} \in \mathbb{R}^n;\ ||\boldsymbol{x} - \boldsymbol{c}|| \leq z\}$ is compact, the sequence $\{\boldsymbol{d}_{k_n}\}_{n=1}^{\infty}$ has a subsequence $\{\boldsymbol{d}_{l_n}\}_{n=1}^{\infty}$ converging to $\boldsymbol{d} \in D$.
    6. 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 an elementary operation, 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. $$\begin{aligned} ||\boldsymbol{d} - y||^2 &\geq ||\boldsymbol{d} - \boldsymbol{c}||^2 \\ (\boldsymbol{d} - \boldsymbol{c} - \alpha(\boldsymbol{c}' - \boldsymbol{c}))^{\mathrm{T}}(\boldsymbol{d} - \boldsymbol{c} - \alpha(\boldsymbol{c}' - \boldsymbol{c})) &\geq (\boldsymbol{d} - \boldsymbol{c})^{\mathrm{T}}(\boldsymbol{d} - \boldsymbol{c}) \\ \alpha^2(\boldsymbol{c}' - \boldsymbol{c})^{\mathrm{T}}(\boldsymbol{c}' - \boldsymbol{c}) - 2\alpha(\boldsymbol{d} - \boldsymbol{c})^{\mathrm{T}}(\boldsymbol{c}' - \boldsymbol{c}) &\geq 0 \\ \frac{\alpha}{2}||\boldsymbol{c}' - \boldsymbol{c}||^2 + \boldsymbol{a}^{\mathrm{T}}\boldsymbol{c} &\geq \boldsymbol{a}^{\mathrm{T}}\boldsymbol{c}' \end{aligned}$$

Relations between Farkas lemma, duality and linear programming

Farkas lemma

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 $\boldsymbol{y}^{\mathrm{T}} A = \boldsymbol{0}^{\mathrm{T}}$ satisfies $\boldsymbol{y}^{\mathrm{T}}\boldsymbol{b} \ge 0$.

Feasibility of a linear programming problem

Problem $\max\left\{\boldsymbol{c}^{\mathrm{T}}\boldsymbol{x};\ A\boldsymbol{x} \le \boldsymbol{b}\right\}$ is infeasible if and only if there exists a non-negative combination $\boldsymbol{y}$ of inequalities $A\boldsymbol{x} \le \boldsymbol{b}$ such that $\boldsymbol{y}^{\mathrm{T}}A = \boldsymbol{0}$ and $\boldsymbol{y}^{\mathrm{T}}\boldsymbol{b} < 0$.

Boundedness of a linear programming problem

- If the problem $\max\left\{\boldsymbol{c}^{\mathrm{T}}\boldsymbol{x};\ A\boldsymbol{x} \le \boldsymbol{b}\right\}$ is bounded and feasible, then $\boldsymbol{c}$ is a non-negative combination $\boldsymbol{y}$ of rows of A, i.e. $\boldsymbol{c}^{\mathrm{T}} = \boldsymbol{y}^{\mathrm{T}}A$.
- If $\boldsymbol{c}$ is a non-negative combination $\boldsymbol{y}$ of rows of A, then the problem $\max\left\{\boldsymbol{c}^{\mathrm{T}}\boldsymbol{x};\ A\boldsymbol{x} \le \boldsymbol{b}\right\}$ is bounded.

Farkas lemma also follows from duality

$$\max\left\{\boldsymbol{0}^{\mathrm{T}}\boldsymbol{x};\ A\boldsymbol{x} \le \boldsymbol{b}\right\} = \min\left\{\boldsymbol{b}^{\mathrm{T}}\boldsymbol{y};\ A^{\mathrm{T}}\boldsymbol{y} = \boldsymbol{0},\ \boldsymbol{y} \ge \boldsymbol{0}\right\}$$

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

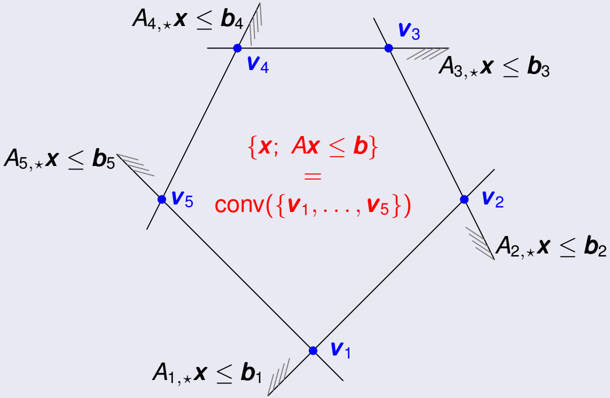
Let $P = \{\boldsymbol{x} \in \mathbb{R}^n;\ A'\boldsymbol{x} = \boldsymbol{b}',\ A''\boldsymbol{x} \le \boldsymbol{b}''\}$ be a *minimal defining system* of a polyherdon P. Let $P' = \{\boldsymbol{x} \in P;\ A''_{i,\star}\boldsymbol{x} = \boldsymbol{b}''_i\}$ for some row i of $A''\boldsymbol{x} \le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

1. There exists $x \in P \setminus P'$. Observe that x is not an affine combination of P'. Hence, $\dim(P') + 1 = \dim(P' \cup \{x\}) \leq \dim(P)$.

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 S_i and S_j, resp.
     Since $\boldsymbol{x}$ is a also a convex combination of $\boldsymbol{u}$ and $\boldsymbol{v}$, we have $\boldsymbol{x} \in \text{conv}(S)$.

Theorem (Minkowski-Weyl)

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

Lemma

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

Corollary

$$\left\{\binom{\boldsymbol{\alpha}}{\beta};\ \boldsymbol{\alpha}^{\mathrm{T}}\boldsymbol{v} \leq \beta\ \forall \boldsymbol{v} \in V\right\} = \left\{\binom{\boldsymbol{\alpha}}{\beta};\ \boldsymbol{\alpha}^{\mathrm{T}}\boldsymbol{v} \leq \beta\ \forall \boldsymbol{v} \in \text{conv}(V)\right\}$$

Lemma

Let $C \subseteq \mathbb{R}^n$ be a closed and convex set and let Q_1 be the set of all $\binom{\boldsymbol{\alpha}}{\beta}$ such that the condition $\boldsymbol{\alpha}^{\mathrm{T}}\boldsymbol{v} \leq \beta$ is satisfied by all points $\boldsymbol{v} \in C$. Let $\boldsymbol{x} \in \mathbb{R}^n$. Then, $\boldsymbol{x} \in C$ if and only $\boldsymbol{\alpha}^{\mathrm{T}}\boldsymbol{x} \leq \beta$ for every $\binom{\boldsymbol{\alpha}}{\beta} \in Q_1$. (1)

1. $\Rightarrow$: Trivial

   $\Leftarrow$: If $\{\boldsymbol{x}\} \cap C = \emptyset$, then by hyperplane separation theorem there exists a hyperplane separating $\{\boldsymbol{x}\}$ and C: $\boldsymbol{\alpha}^{\mathrm{T}}\boldsymbol{x} > \beta$ and $\boldsymbol{\alpha}^{\mathrm{T}}\boldsymbol{v} < \beta$ for every $\boldsymbol{v} \in C$. Hence, $\binom{\boldsymbol{\alpha}}{\beta} \in Q_1$ but $\boldsymbol{\alpha}^{\mathrm{T}}\boldsymbol{x} \leq \beta$ fails.

Minkowski-Weyl

Theorem (Minkowski-Weyl)

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

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

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

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

(1) $\Leftrightarrow$ (2) Lemma with $C = \text{conv}(V)$.

(2) $\Leftrightarrow$ (3) By Corollary, $Q_1 = Q_2$.

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

(4) $\Leftrightarrow$ (5) Lemma.

Convex hull of vertices of a polytope

Theorem

Let P be a polytope 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, $P = \operatorname{conv}(V)$.

Proof

- Let V_0 be an 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 $V \subseteq V_e \subseteq V_0 \subseteq V$.
- $V \subseteq V_e$: Let $\boldsymbol{z} \in V$ be a vertex.
  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 $\boldsymbol{z} \in V_0$.

Convex hull of vertices of a polytope

Theorem

Let P be a polytope 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, $P = \operatorname{conv}(V)$.

Proof

Let V_0 be an inclusion minimal set such that $P = \operatorname{conv}(V_0)$. We prove that $V_0 \subseteq V$.

1. Let $\boldsymbol{z} \in V_0$ and $D = \operatorname{conv}(V_0 \setminus \{\boldsymbol{z}\})$.
2. Minkovsky-Weil's theorem $\Rightarrow$ V_0 is finite $\Rightarrow$ D is compact.
3. By the separation theorem we separate $\{\boldsymbol{z}\}$ and D: $\boldsymbol{c}^{\mathrm{T}}\boldsymbol{x} < r < \boldsymbol{c}^{\mathrm{T}}\boldsymbol{z}$ for all $\boldsymbol{x} \in D$.
4. Let $t = \boldsymbol{c}^{\mathrm{T}}\boldsymbol{z}$. We prove that $A = \left\{\boldsymbol{x};\ \boldsymbol{c}^{\mathrm{T}}\boldsymbol{x} = t\right\}$ is a supporting hyperplane of P.
5. Clearly, $\boldsymbol{c}^{\mathrm{T}}\boldsymbol{x} \le r$ for every $x \in P$ and $\boldsymbol{z} \in A \cap P$.
6. For a sake of contradiction, let $\boldsymbol{z}' \in A \cap P$ and $\boldsymbol{z} \neq \boldsymbol{z}'$.
7. Let $\boldsymbol{z}' = \alpha_0\boldsymbol{z} + \alpha_1\boldsymbol{x}_1 + \cdots + \alpha_k\boldsymbol{x}_k$ be a convex combination of V_0.
8. From $\boldsymbol{z} \neq \boldsymbol{z}'$ it follows that $\alpha_0 < 1$ and WLOG $\alpha_1 > 0$.
9. It holds that $\alpha_0\boldsymbol{c}^{\mathrm{T}}\boldsymbol{z} = \alpha_0 t$ and $\alpha_1\boldsymbol{c}^{\mathrm{T}}\boldsymbol{x}_1 < \alpha_1 t$ and $\alpha_i\boldsymbol{c}^{\mathrm{T}}\boldsymbol{x}_i \le \alpha_i t$ for all $i = 1, \ldots, k$.
10. Hence, $\boldsymbol{c}^{\mathrm{T}}\boldsymbol{z}' = \alpha_0\boldsymbol{c}^{\mathrm{T}}\boldsymbol{z} + \alpha_1\boldsymbol{c}^{\mathrm{T}}\boldsymbol{x}_1 + \sum_{i=2}^{k}\alpha_i\boldsymbol{c}^{\mathrm{T}}\boldsymbol{x}_i < \alpha_0 t + \alpha_1 t + \sum_{i=2}^{k}\alpha_i t = t$.
11. This contradicts the assumption that $\boldsymbol{z}' \in A$.

Example: Maximal weighted perfect matching in a graph (V, E, w)

Integer linear program

max $\boldsymbol{w}^{\mathrm{T}}\boldsymbol{x}$ subject to $A\boldsymbol{x} = \boldsymbol{1}$ and $\boldsymbol{x} \in \{0, 1\}$ where A is the incidence matrix
Relaxed program: replace $\boldsymbol{x} \in \{0, 1\}$ by $\boldsymbol{0} \leq \boldsymbol{x} \leq \boldsymbol{1}$
Matching polytope $P = \left\{\boldsymbol{x} \in \mathbb{R}^{E};\ A\boldsymbol{x} = \boldsymbol{1}, \boldsymbol{0} \leq \boldsymbol{x} \leq \boldsymbol{1}\right\}$

Bipartite graphs

If the graph is bipartite, then every vertex of P is a perfect matching.

Corollary

If the graph is bipartite, every optimal basis solution is a perfect matching.

Non-bipartite graph (example)

For the triangle, $(\frac{1}{2}, \frac{1}{2}, \frac{1}{2})$ is a vertex of P (and the only point of P).

Outline

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 $A\boldsymbol{x} \le \boldsymbol{b}$ where $A \in \mathbb{R}^{m \times n}$ and $\boldsymbol{b} \in \mathbb{R}^m$.

Mix integer linear programming

Some variables are integer and others are real.

Relaxed problem and solution

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

Observation

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

Rational and integral polyhedrons

Definition: Rational polyhedron

A polyhedron P is called rational if it is defined by a rational linear system $P = \{\boldsymbol{x};\ A\boldsymbol{x} \le \boldsymbol{b}\}$ where $A \in \mathbb{Q}^{m \times n}$ and $\boldsymbol{b} \in \mathbb{Q}^m$. ①

Observation

Every vertex of a rational polyhedron in the canonical form $P = \{\boldsymbol{x};\ A\boldsymbol{x} = \boldsymbol{b}, \boldsymbol{x} \ge \boldsymbol{0}\}$ is rational. ②

Definition: Integral polyhedron

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

Observation

Let P be a rational polyhedron which has a vertex. Then, P is integral if and only if every vertex of P is integral. ③

Theorem

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

1. If P is a rational polyherdon, then there exists an integral linear system $P = \{\boldsymbol{x};\ A'\boldsymbol{x} \leq \boldsymbol{b}'\}$ where $A' \in \mathbb{Z}^{m \times n}$ and $\boldsymbol{b} \in \mathbb{Z}^m$ since we can multiply every row of $A\boldsymbol{x} \leq \boldsymbol{b}$ so that the resulting system consists of integers.
2. Every vertex of P is a basis feasible solution with a basis B and coordinates $\boldsymbol{x}_B = A_B^{-1}\boldsymbol{b}$ and $\boldsymbol{x} = \boldsymbol{0}$. Since A_B is regular and rational, the inverse matrix A_B^{-1} is also rational, so $\boldsymbol{x}_B = A_B^{-1}\boldsymbol{b}$ is rational.
3. Since a vertex is an non-empty face, every vertex of an integral polyhedron must be integral. Since P has a vertex, every face contains a vertex and this vertex must be integral.

Theorem

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

Proof

- $\Rightarrow$ Every vertex of P is integral, so optimal values are integrals. ①
- $\Leftarrow$ Let $\boldsymbol{v}$ be a vertex of P. We prove that $\boldsymbol{v}_1$ is an integer.
  1. Let $\boldsymbol{c}$ be an integer vector such that $\boldsymbol{v}$ is the only optimal solution. ②
  2. We can scale the vector $\boldsymbol{c}$ by a sufficiently large integer k so that $\boldsymbol{v}$ is also the optimal vertex for objective vector $(k\boldsymbol{c} + e_1)$ where $e_1 = (1, 0, \dots, 0)^{\mathrm{T}}$. ③
  3. Hence, $\boldsymbol{c}^{\mathrm{T}}\boldsymbol{v}$, $(k\boldsymbol{c} + e_i)^{\mathrm{T}}\boldsymbol{v}$ and $\boldsymbol{v}_1 = (k\boldsymbol{c} + e_i)^{\mathrm{T}}\boldsymbol{v} - k\boldsymbol{c}^{\mathrm{T}}\boldsymbol{v}$ are integers.

1. If a polytope is integral, then the face of all optimal solution contains an integral point $\boldsymbol{x}^\star$, so the dot product of $\boldsymbol{x}^\star$ and an integral vector $\boldsymbol{c}$ is an integer.
2. Assume that $P = \{\boldsymbol{x};\ A\boldsymbol{x} \le \boldsymbol{b}\}$ where A and $\boldsymbol{b}$ are integral. Let $A'\boldsymbol{x} = \boldsymbol{b}'$ be the subsystem of $A\boldsymbol{x} \le \boldsymbol{b}$ which $\boldsymbol{v}$ satisfies all inequalities in equations. We sum up all equations $A'\boldsymbol{x} = \boldsymbol{b}'$ into $\boldsymbol{c}\boldsymbol{x} = d$. We know that $\boldsymbol{c}\boldsymbol{x} = d$ is a supporting hyperplane for $\boldsymbol{v}$.
3. Choose a positive integer k to be at least $\max\left\{\frac{\boldsymbol{u}_1 - \boldsymbol{v}_1}{\boldsymbol{c}^{\mathrm{T}}\boldsymbol{v} - \boldsymbol{c}^{\mathrm{T}}\boldsymbol{u}};\ u \text{ vertex of } P\right\}$.

Total unimodularity

Questions

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

Proposition

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

Proof

$\Leftarrow$
- Cramer's rule: $A^{-1}_{j,i} = \frac{\det B}{\det A}$ where B is a matrix obtained from A by replacing the i-th column by e_j.
- Hence, A^{-1} is integral, so $A^{-1}b$ is integral for every integral $\boldsymbol{b}$

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

Definition

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

Theorem

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

Proof

$\Leftarrow$

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

$\Rightarrow$

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

Totally unimodular matrix

Definition

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

Exercise

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

Exercise

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

Theorem: Hoffman-Kruskal

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

Proof

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

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

Totally unimodular matrix: Application

Observation

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

Proof

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

$k = 1$ Trivial

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

Corollary

The incidence matrix of an oriented graph is totally unimodular.

Observation: Other totally unimodular (TU) matrices

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

Definition: Network flow

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

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

Capacity: $0 \leq f \leq c$

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

Theorem

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

Proof

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

Duality of the network flow problem

Primal: Network flow

Maximize $\boldsymbol{f}_{ts}$ subject to $A\boldsymbol{f} = \mathbf{0}$, $\boldsymbol{f} \leq \boldsymbol{c}$ and $\boldsymbol{f} \geq \mathbf{0}$.

Primal dual

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

Observation

Dual problem has an integral optimal solution.

Complementary slackness

- $\boldsymbol{f}_{uv} = \boldsymbol{c}_{uv}$ or $\boldsymbol{z}_{uv} = 0$ for every edge uv ②
- $\boldsymbol{f}_{uv} = 0$ or $-\boldsymbol{y}_u + \boldsymbol{y}_v + \boldsymbol{z}_{uv} = 0$ for every edge $uv \neq ts$
- $\boldsymbol{f}_{ts} = 0$ or $-\boldsymbol{y}_t + \boldsymbol{y}_s + \boldsymbol{z}_{ts} = 1$ ③

Observation

Every feasible solution defines a cut where $Z = \{uv \in E;\ z_{uv} > 0\}$ are cut edges and $U = \{u \in V;\ \boldsymbol{y}_u > \boldsymbol{y}_t\}$ is partition of vertices. Moreover, the minimal cut equals the maximal flow. ④

1. Observe that if $(\boldsymbol{y}, \boldsymbol{z})$ is a feasible solution to the dual problem, then $(\boldsymbol{y} + \alpha, \boldsymbol{z})$ is a feasible solution for every $\alpha \in \mathbb{R}$, so we can assume that $\boldsymbol{y}_t = 1$.
2. If $\boldsymbol{c}_{ts}$ is sufficiently large, then $\boldsymbol{f}_{ts} < \boldsymbol{c}_{ts}$ in every feasible solution, so $\boldsymbol{z}_{ts} = 0$.
3. Since $\boldsymbol{z}_{ts} = 0$, we have $\boldsymbol{y}_s \geq \boldsymbol{y}_t + 1$. If the graph has a non-trivial flow, then $vf_{ts} > 0$, so $\boldsymbol{y}_s = \boldsymbol{y}_t + 1 = 1$.
4. For every edge uv with $u \notin U$ and $v \in U$, we have $\boldsymbol{z}_{uv} \geq \boldsymbol{y}_u - \boldsymbol{y}_v > 0$, so $uv \in Z$. Furthermore, if $\boldsymbol{f}$ and $(\boldsymbol{y}, \boldsymbol{z})$ are optimal solutions, then the complementarity slackness implies that for every $uv \in Z$ it holds that $\boldsymbol{f}_{uv} = \boldsymbol{c}_{uv}$ and for every edge uv with $u \in U$ and $v \notin U$ it holds that $-\boldsymbol{y}_u + \boldsymbol{y}_v + \boldsymbol{z}_{uv} > -\boldsymbol{y}_t + \boldsymbol{y}_t + 0 > 0$, so the complementary slackness implies that $\boldsymbol{f}_{uv} = 0$.

Gomory-Chvátal cutting plane: Example

Interger linear programming problem

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

Relaxed problem

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

Cutting plane 1

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

Cutting plane 2

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

Gomory-Chvátal cutting plane proof

System of inequalities

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

Definition: Gomory-Chvátal cutting plane

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

Definition: Gomory-Chvátal cutting plane proof

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

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

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

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

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

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

Branch and bound

Branch

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

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

Bound

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

Observation

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

Outline

Definition

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

Integer linear programming formulation

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

Relaxed problem

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

Algorithm

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

Observation

S_{LP} is a vertex cover.

Observation

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

Proof

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

Maximum independent set problem

Definition

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

Integer linear programming formulation

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

Relaxed problem

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

Relaxed solution

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

Optimal integer solution

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

Conclusion

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

Inapproximability of the minimmum independent set problem

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

Outline