Combinatorial Optimization in Computer Vision

Chapter 3: Linear Programming

WS 2011/12

Ulrich Schlickewei
Computer Vision and Pattern Recognition Group
Technische Universität München

Plan for Today

1. Standard Form of Linear Programs
2. Convex Polyhedra and Their Vertices
3. The Simplex Algorithm
4. Duality

Definition Linear Program

Definition: A linear program is an optimization problem of the form

$$\begin{aligned} \min_{x \in \mathbb{R}^n} \quad & f(x) \\ \text{s.t.} \quad & g_i(x) = b_i, \\ & h_j(x) \leq c_j \end{aligned}$$

($1 \leq i \leq m$, $1 \leq j \leq l$) where all functions f, g_i and h_j are **linear**.

Example:

$$\begin{aligned} \min_{x \in \mathbb{R}^3} \quad & 2x_1 + 3x_2 \\ \text{s.t.} \quad & x_1 + x_2 = 0, \\ & x_1 - 5x_2 \leq 3. \end{aligned}$$

Towards the Standard Form I

Linear Optimization Function

$$f(x) = p^\top \cdot x \ \text{ for } p \in \mathbb{R}^n$$

Linear Equality Constraints

$$\begin{pmatrix} g_1(x) \\ \vdots \\ g_m(x) \end{pmatrix} = A \cdot x \ \text{ for } A \in \mathbb{R}^{mn} \quad \rightsquigarrow Ax = b.$$

Linear Inequality Constraints

$$\begin{pmatrix} h_1(x) \\ \vdots \\ h_l(x) \end{pmatrix} = B \cdot x \ \text{ for } B \in \mathbb{R}^{ln} \quad \rightsquigarrow Bx \leq c$$

Towards the Standard Form II

Aim: Introduction of slack variables in order to get constraints of the standard form

$$
\begin{aligned}
Ax &= b \\
x &\geq \mathbf{0}.
\end{aligned}
$$

Towards the Standard Form III

Introduce slack variables $y_1 \geq 0, \ldots, y_l \geq 0$ to reformulate the inequality constraint $Bx \leq c$ as

$$\begin{pmatrix} B & \mathbb{1} \end{pmatrix} \cdot \begin{pmatrix} x \\ y \end{pmatrix} = c.$$

$$\boxed{\begin{aligned} \min_{(x,y)\in\mathbb{R}^{n+l}} \quad & \begin{pmatrix} p \\ \mathbf{0} \end{pmatrix}^\top \cdot \begin{pmatrix} x \\ y \end{pmatrix} \\ \text{s.t.} \quad & \begin{pmatrix} A & \mathbf{0} \\ B & \mathbb{1} \end{pmatrix} \cdot \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} b \\ c \end{pmatrix}, \\ & y \geq \mathbf{0}. \end{aligned}}$$

Towards the Standard Form IV

Decompose x in $x^+ + x^-$ with

$$x_i^+ = \max(0, x_i) \geq 0$$
$$x_i^- = \max(0, -x_i) \geq 0.$$

$$\boxed{\begin{aligned} \min_{(x,y)\in\mathbb{R}^{n+l}} \quad & \begin{pmatrix} p \\ -p \\ \mathbf{0} \end{pmatrix}^\top \cdot \begin{pmatrix} x^+ \\ x^- \\ y \end{pmatrix} \\ \text{s.t.} \quad & \begin{pmatrix} A & -A & \mathbf{0} \\ B & -B & \mathbb{1} \end{pmatrix} \cdot \begin{pmatrix} x^+ \\ x^- \\ y \end{pmatrix} = \begin{pmatrix} b \\ c \end{pmatrix}, \\ & \begin{pmatrix} x^+ \\ x^- \\ y \end{pmatrix} \geq \mathbf{0}. \end{aligned}}$$

Standard Form of LP

Proposition: Any linear program is equivalent to a linear program in standard form

$$\begin{aligned} \min_{x \in \mathbb{R}^n} \quad & p^\top \cdot x \\ \text{s.t.} \quad & Ax = b \\ & x \geq \mathbf{0}. \end{aligned}$$

Remark:

- standard form useful for studying LPs
- In practice slack variables can be expensive and are often avoided.

Plan for Today

1. Standard Form of Linear Programs
2. Convex Polyhedra and Their Vertices
3. The Simplex Algorithm
4. Duality

Feasible Set of LPs

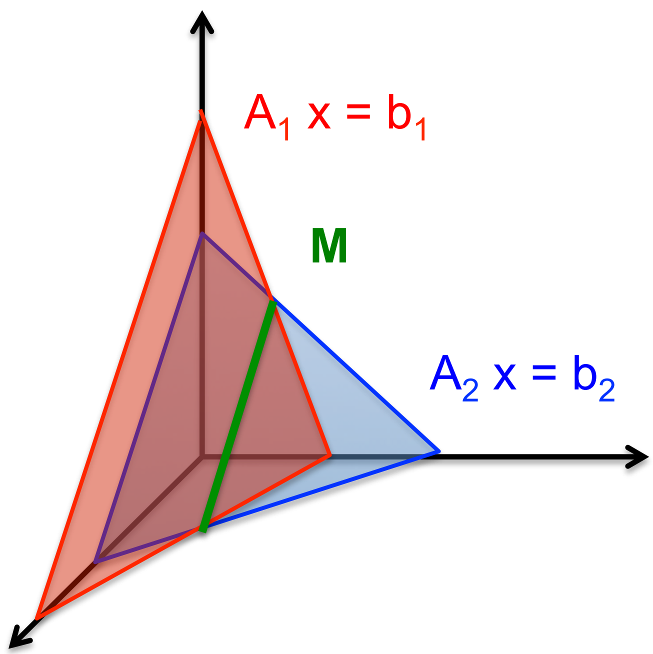
Plan

- feasible set of LP = polyhedron
- vertices of polyhedra
- If an LP has a minimum, then it is attained at a vertex.

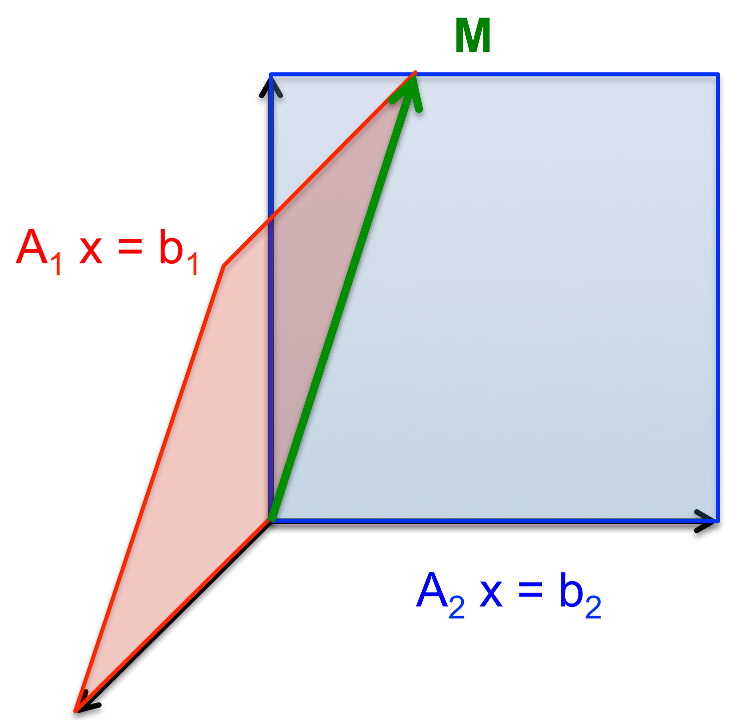
Polyhedra

Feasible set of LP is a polyhedron:

$$M = \{x \in \mathbb{R}^n \mid Ax = b,\ x \geq \mathbf{0}\}.$$



bounded polyhedron

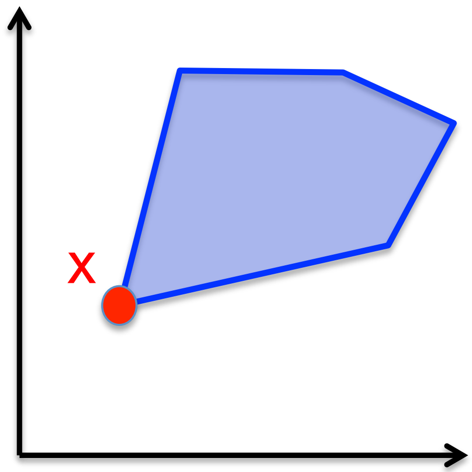


unbounded polyhedron

Vertices

Definition: Let M $\subset$ **R**^{n} be a convex set. A point x $\in$ M is a vertex if it does not lie on a line segment in M, that is

$$\Big(x = \lambda y + (1-\lambda) z \text{ for } y, z \in M, y \neq z \text{ and } \lambda \in [0,1]\Big) \implies \Big(\lambda = 0 \text{ or } \lambda = 1\Big).$$



Characterization of Vertices I

Let x $\in$ M. Let I_x be the set of indices of non-zero entries of x

$$I_x := \{i \in \{1, \dots, n\} \mid x_i \neq 0\}.$$

Example:

$$x = (0, \boxed{4,}\boxed{3,} 0, \boxed{4,}\boxed{6,} 0) \in \mathbb{R}^7$$

$$\Rightarrow \quad I_x = \{2, 3, 5, 6\} \subset \{1, \dots, 7\}$$

Characterization of Vertices II

Denote by A^i the i-th column of A. Then

$$A \cdot x = \begin{pmatrix} A^1 & \cdots & \boxed{A^{j_1}} & \cdots & \boxed{A^{j_l}} & \cdots & A^n \end{pmatrix} \cdot \begin{pmatrix} 0 \\ \vdots \\ \boxed{x_{j_1}} \\ \vdots \\ \boxed{x_{j_l}} \\ \vdots \\ 0 \end{pmatrix} = \sum_{j \in I_x} x_j A^j = b.$$

Characterization of Vertices III

Notation: Let $x \in M$, let $l = |I_x|$.

- For a $(m \times n)$-matrix A, we denote by A_x the $(m \times l)$-matrix consisting of the columns with indices in $I_\mathbf{x}$

$$A_x = (A^j)_{j \in I_x} = \begin{pmatrix} A^{j_1} & \cdots & A^{j_l} \end{pmatrix} \in \mathbb{R}^{ml}.$$

- For a vector $y \in R^\mathbf{n}$ we denote by $y_\mathbf{x} \in R^\mathbf{l}$ the vector consisting of the entries with indices in $I_\mathbf{x}$

$$y_x = (y_j)_{j \in I_x} = \begin{pmatrix} y_{j_1} \\ \vdots \\ y_{j_l} \end{pmatrix}$$

Characterization of Vertices IV

Proposition: The following are equivalent

a) x is a vertex of M

b) $A^{j_1}, \ldots, A^{j_l}$ are linearly independent

c) A_x has full rank: $\mathrm{rk}(A_x) = l.$

Proof I

- Items b) and c) are equivalent by definition.
- To prove b) $\Rightarrow$ a) we prove ($\neg$ a)) $\Rightarrow$ ($\neg$ b)).

Assume that x is not a vertex of M. Then there exist $y \neq z \in M$ and $\lambda \in (0,1)$ with $x = \lambda\, y + (1-\lambda)\, z$.

Then $I_{y-z} \subset I_x$, because $x_j = 0 \Rightarrow y_j = 0$ and $z_j = 0$ (use that $y, z \geq 0$). Since $(y-x) \neq 0$, we get $(y-z)_x \neq 0$.

But then $0 = b-b = A(y-z) = A_x\,(y-z)_x$. With $(y-z)_x \neq 0$, this implies that the columns of A_x are linearly dependent.

Proof II

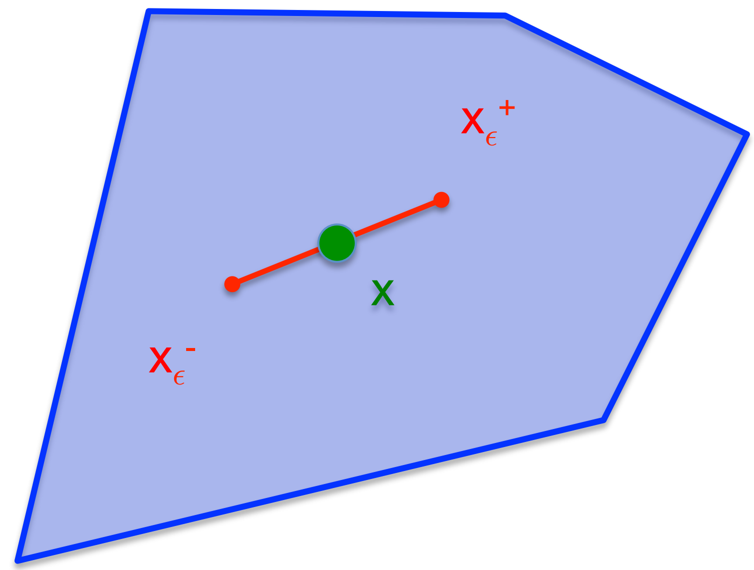
To prove a) $\Rightarrow$ b) we prove ($\neg$ b)) $\Rightarrow$ ($\neg$ a)).

Assume that the columns of A_x are linearly dependent. Let $y' \in \mathbf{R}^l$ non-zero with $A_x\, y' = 0$.

Let $y \in \mathbf{R}^n$ with $I_y \subset I_x$ and $y_x = y'$. Then $Ay = 0$. Let $\epsilon > 0$ be small. Let $x_\epsilon^+ = x + \epsilon\, y$ and $x_\epsilon^- = x - \epsilon\, y$.

Then $Ax_\epsilon^+ = b$ and $x_\epsilon^+ \geq 0$ for ϵ sufficiently small. $\Rightarrow x_\epsilon^+ \in M$.
Similarly $x_\epsilon^- \in M$.

But $x = ½\, x_\epsilon^+ + ½\, x_\epsilon^-$, which implies that x is not a vertex.



Consequences I

Corollary 1: The polyhedron M has at most finitely many vertices.

Proof: Let x be a vertex of M. Then A_x has maximal rank and consequently x_x is the uniquely determined solution of $A_x\, x_x = b$. Thus, I_x determines x uniquely.

The corollary follows from the fact that there are only finitely many subsets $I \subset \{1,...,n\}$.

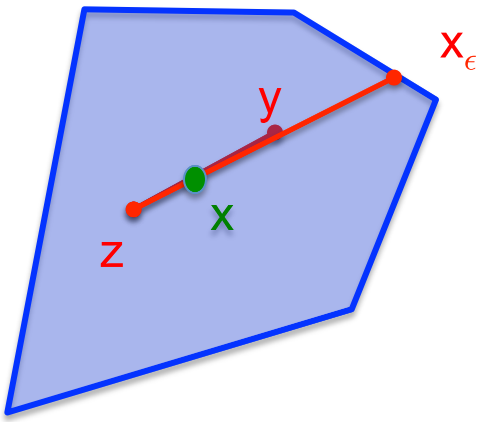
Consequences II

Corollary 2: If M is non-empty, then M has a vertex.

Proof: Let $x \in M$ with $|I_x|$ minimal. We will show that x is a vertex.
If $x = \lambda\, y + (1-\lambda)\, z$ for $y \neq z \in M$, then $I_y, I_z \subset I_x$.

W.l.o.g. assume that $(y-z)_x$ has a positive entry.
Define $x_\epsilon = x + \epsilon\,(y - z) \in M$.

Choose ϵ in such a way that one entry of $(x_\epsilon)_x$ is 0.



Then $x_\epsilon \in M$ and $|I_{x_\epsilon}| < |I_x|$. Contradiction.

Consequences III

Corollary 3: If $f(x) = p^\top x$ attains a minimum on M, then it attains the minimum at a vertex of M.

Proof: Let $m = \min_{x \in M} p^\top x$. Choose $u \in M$ with $p^\top u = m$ such that $|I_u|$ is minimal, i.e.

$$|I_u| = \min\{ |I_v| : v \in M, p^\top v = m\}.$$

If u was not a vertex of M, we could construct as above $v \neq w \in M$ and $\lambda \in (0,1)$ with $u = \lambda v + (1-\lambda) w$ and $|I_v| < |I_u|$

But then $f(u) = \lambda f(v) + (1-\lambda) f(w)$. By minimality of f(u), we get $f(u) = f(v) = f(w)$.

Contradiction to the minimality of $|I_u|$.

Summary on Polyhedra

Alltogether, we have proved

Theorem:

- If the polyhedron M is non-empty, then it has at least one vertex. There are only finitely many vertices.
- If the linear program admits a minimum over M, then this minimum is attained at a vertex of M.

$\Rightarrow$ Solution strategy: Search the vertices of M

Polyhedra as Simplices

One can prove even more.

Theorem: Let $M \subset \mathbf{R}^n$ be a polyhedron. Assume that M is bounded. Let $v_1,...,v_k$ be the vertices of M. Then M is the convex hull of ist vertices

$$M = \mathrm{conv}(v_1,...,v_k).$$

In this case the linear program takes a minimum on M.

Plan for Today

1. Standard Form of Linear Programs
2. Convex Polyhedra and Their Vertices
3. The Simplex Algorithm
4. Duality

Basic Idea

Input: LP in standard form

$$\begin{aligned} \min_{x \in \mathbb{R}^n} \quad & p^\top \cdot x \\ \text{s.t.} \quad & Ax = b \\ & x \geq \mathbf{0}. \end{aligned}$$

Idea: walk through the vertices of the polyhedron $M = \{x \in \mathbf{R}^n \mid Ax = b, x \geq 0\}$. Decrease the value of the objective function in each step.

Phase I: Find a starting vertex x of M.

Phase II: Determine an edge e connecting x with a neighbor vertex such that the objective function decreases along e. Repeat this step until convergence.

Phase I

To find a starting vertex, we solve an auxiliary LP.

If $b \geq 0$, we use

$$\begin{aligned} \min_{(x,z)^\top \in \mathbb{R}^{n+m}} \quad & \sum_i z_i \\ \text{s.t.} \quad & Ax + z = b \\ & \begin{pmatrix} x \\ z \end{pmatrix} \geq \mathbf{0}. \end{aligned}$$

If some $b_j \leq 0$, we multiply the corresponding line in A with -1 and change b_j into $-b_j$.

Phase I (continued)

- A starting vertex for the auxiliary LP (ALP)

$$
\begin{aligned}
\min_{(x,z)^\top \in \mathbb{R}^{n+m}} \quad & \sum_i z_i \\
\text{s.t.} \quad & Ax + z = b \\
& \begin{pmatrix} x \\ z \end{pmatrix} \geq \mathbf{0}.
\end{aligned}
$$

  is $(x_0,z_0) = (0,b)$.
- We solve (ALP) by stepping to Phase II with this starting vertex.
- A basic solution (i.e. a solution which is a vertex) of (ALP) is a starting vertex for the original LP.
- If (ALP) is not solvable, then the orginal LP is not solvable.

Phase II

In Phase II we have to walk along edges through the vertices of $M = \{x \in \mathbf{R}^n \mid Ax = b, x \geq 0\}$ and to successively decrease the value of the objective function.

To do so, we use the characterization of vertices x by their index set I_x.

Recall that the columns of A_x are linearly independent if x is a vertex.

Phase II (continued)

- For each vertex x, choose a maximal set $I \subset \{1,...n\}$ such that
  - the columns of A_I are linearly independent and
  - $I_x \subset I$.

  (Note that $|I| = rk(A)$.)

- Conversely, given $I \subset \{1,...,n\}$ with $|I| = rk(A)$ such that the columns of A_I are linearly independent, consider the linear equation $A_I\, y = b$. If there is a solution y‘ of this system with $y' \geq 0$, then y‘ can be extended by zeros to a vertex y of M with $I_y \subset I$.

- If a vertex x satisfies |I_x| = rk(A), then the basis set I is uniquely defined. Such a vertex is called non-degenerate.
  If |I_x| < rk(A), there could be several basis sets I. Such vertices are called degenerate. In the simplex algorithm one has to take care to avoid cycling caused by degenerate vertices.

Phase II (continued)

Assume we are at vertex x and we have chosen a basis set I. To pass to a neighboring vertex y two choices have to be made:

- incoming index: choose $j \in \{1,...,n\}$ with $j \notin I$.
- outgoing index: choose $k \in I$.

Then replace k by j: $I' = (I \setminus \{k\}) \cup \{j\}$.

Phase II (continued)

- Algorithmically, this comes down to solving the equation $A_J y = b$. This can be done with a warm start, using that the equation $A_I x = b$ has been solved before.
- In practice, the linear program is written down in a so-called tableau. Solving the equation $A_J y = b$ is then achieved by performing elementary Gauss elimination steps.
- The choice of j and k is guided by two principles
  - decrease the objective function
  - ensure that y is feasible, i.e. that $y \geq 0$

Some Facts

- If an appropriate strategy for the choice of j and k is chosen, the simplex algorithm terminates. This means that it either finds a solution of the LP after finitely many steps or it detects that the LP is unfeasible or unbounded.
- The worst case of the algorithm is exponential in the problem size.
- In practice, the simplex algorithm is very performant and it is frequently used.
- The ellipsoid method, proposed in 1976/77, was the first polynomial time algorithm for Linear Programming.

Plan for Today

1. Standard Form of Linear Programs
2. Convex Polyhedra and Their Vertices
3. The Simplex Algorithm
4. Duality

Duality

- For every LP there is an associated dual LP.
- Solving the LP is equivalent to solving the dual LP.
- We will first see duality for LPs in standard inequality form.
- Then in a second step we derive duality for LPs in standard equality form (i.e. in standard form as we have seen in 1.).

Symmetric Duality for LPs in Standard Inequality Form

Consider the following LP in standard inequality form:

$$\begin{aligned} \min_{x \in \mathbb{R}^n} \quad & p^\top x \\ \text{s.t.} \quad & Ax \geq b \\ & x \geq 0. \end{aligned}$$

The associated dual LP is

$$\begin{aligned} \max_{y \in \mathbb{R}^m} \quad & b^\top y \\ \text{s.t.} \quad & A^\top y \leq p \\ & y \geq 0. \end{aligned}$$

Weak Duality

Theorem: Let $x \in \mathbf{R}^n$ be feasible for the primal LP, let $y \in \mathbf{R}^m$ be feasible for the dual LP. Then

$$b^\top y \leq p^\top x.$$

Proof: By primal feasibility of x, we have $b \leq Ax$, by dual feasibility of y, we have $p \geq A^\top y$. Using $x \geq 0$ and $y \geq 0$, this implies

$$b^\top y \leq x^\top A^\top y \leq x^\top p.$$

Strong Duality

Theorem: If a standard inequality LP is bounded and feasible, then so is its dual. Denote by x^* and y^* the solutions of the primal resp. the dual problem. Then

$$b^\top y^* = p^\top x^*.$$

Duality for LPs in Standard Equality Form I

Consider now a LP in standard equality form

$$\begin{aligned} \min_{x \in \mathbb{R}^n} \quad & p^\top \cdot x \\ \text{s.t.} \quad & Ax = b \\ & x \geq \mathbf{0}. \end{aligned}$$

This can easily be transformed to a LP in standard inequality form

$$\begin{aligned} \min_{x \in \mathbb{R}^n} \quad & p^\top \cdot x \\ \text{s.t.} \quad & \begin{pmatrix} A \\ -A \end{pmatrix} x \geq \begin{pmatrix} b \\ -b \end{pmatrix} \\ & x \geq \mathbf{0}. \end{aligned}$$

Duality for LPs in Standard Equality Form II

The dual LP of

$$
\begin{aligned}
\min_{x \in \mathbb{R}^n} \quad & p^\top \cdot x \\
\text{s.t.} \quad & \begin{pmatrix} A \\ -A \end{pmatrix} x \geq \begin{pmatrix} b \\ -b \end{pmatrix} \\
& x \geq \mathbf{0}.
\end{aligned}
$$

is given by

$$
\begin{aligned}
\max_{(y_1, y_2)^\top \in \mathbb{R}^{2m}} \quad & b^\top \cdot y_1 - b^\top \cdot y_2 \\
\text{s.t.} \quad & A^\top y_1 - A^\top y_2 \leq p \\
& y_1 \geq \mathbf{0},\ y_2 \geq \mathbf{0}.
\end{aligned}
$$

Duality for LPs in Standard Equality Form III

This last problem

$$\begin{aligned}
\max_{(y_1, y_2)^\top \in \mathbb{R}^{2m}} \quad & b^\top \cdot y_1 - b^\top \cdot y_2 \\
\text{s.t.} \quad & A^\top y_1 - A^\top y_2 \leq p \\
& y_1 \geq \mathbf{0},\ y_2 \geq \mathbf{0}.
\end{aligned}$$

can be simplified by introducing $z = y_1 - y_2$. The resulting LP is

$$\begin{aligned}
\max_{z \in \mathbb{R}^m} \quad & b^\top z \\
\text{s.t.} \quad & A^\top z \leq p.
\end{aligned}$$

Summary

- We have seen that any LP can be written in standard form.
- The feasible set of an LP is a convex polyhedron. Such a polyhedron can be empty, bounded or unbounded.
- If the LP has a solution, the it has a solution which lies on a vertex of the polyhedron.
- The simplex algorithm searches the vertices of a polyhedron for the optimal solution.
- The dual LP associated with an LP is equivalent by the Strong Duality Theorem.