Linear Programs

Niklas Rieken

July 20, 2019

A *linear program* (*LP*) consists of *variables* $x = (x_1, \dots, x_n)$, an *objective function* $z(x) = c^T x$ with $c \in \mathbb{Q}^n$ that can either be maximized or minimized, and *constraints* $a^T x \sim b$ with $a \in \mathbb{Q}^n, b \in \mathbb{Q}, \sim \in \{\leq, =, \geq\}$. An LP is in *canonical form* if it is written as

$$\begin{aligned} \text{maximize } & c_1 x_1 + c_2 x_2 + \ldots + c_n x_n & \\ \text{subject to } & a_{11} x_1 + a_{12} x_2 + \ldots + a_{1n} x_n & \leq b_1 \\ & a_{21} x_1 + a_{22} x_2 + \ldots + a_{2n} x_n & \leq b_2 \\ & \vdots & \\ & a_{m1} x_1 + a_{m2} x_2 + \ldots + a_{mn} x_n & \leq b_m \\ & x_1, x_2, \ldots, x_n & \geq 0 \end{aligned}$$

or shorter

$$\begin{aligned} \text{maximize } & \sum_{j=1}^n c_j x_j & \\ \text{subject to } & \sum_{j=1}^n a_{ij} x_j \leq b_i & i \in \{1, \ldots, m\} \\ & x_j \geq 0 & j \in \{1, \ldots, n\} \end{aligned}$$

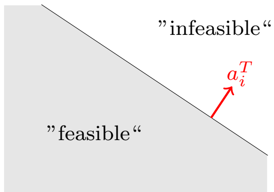
or even shorter

$$\begin{aligned} \text{maximize } & c^T x \\ \text{subject to } & Ax \leq b \\ & x \geq 0. \end{aligned}$$

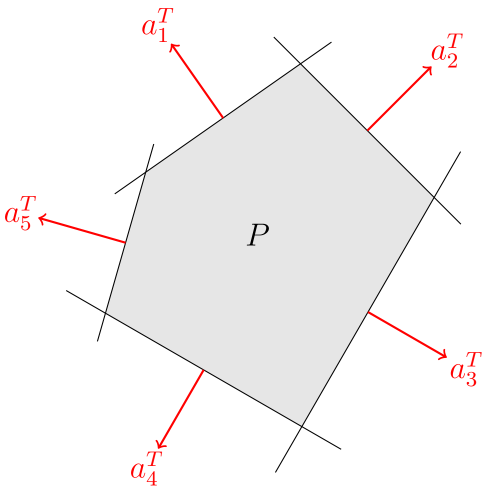
We call the x_j ($j \in \{1, \ldots, n\}$) *decision variables*, c^T the *objective function coefficients*, A the *coefficient matrix*, and b the *right hand side* of the LP. Moreover, we write a_i^T for the coefficients of the i-th constraint. Note that every LP has an "equivalent" LP in canonical form:

- $\min_x c^T x = \max_x -c^T x$,
- $a_i^T x \geq b_i$ iff $-a_i^T x \leq -b_i$,
- $a_i^T x = b_i$ iff $a_i^T x \leq b_i$ and $a_i^T x \geq b_i$,
- $x_j \leq 0$ iff $-x_j \geq 0$,
- x_j unbounded: Replace $x_j = x_j^+ - x_j^-$ with $x_j^+, x_j^- \geq 0$.

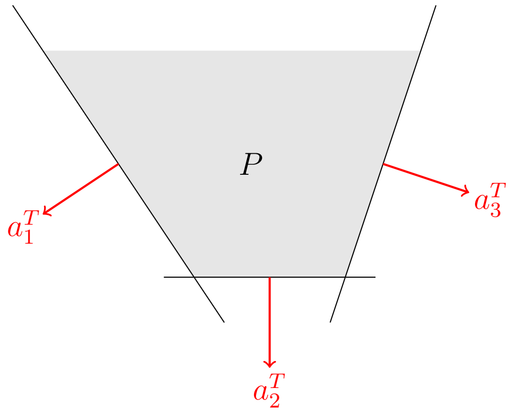
Every constraint (from an LP in canonical form) defines a *halfspace* in $\mathbb{R}^n$ and its separating hyperplane has a_i^T as its normal vector.



We call the set of points that satisfy all constraints *feasible region* $P := \{x \in \mathbb{R}^n \mid Ax \leq b, x \geq 0\}$. P is a *polyhedron*, i.e. the intersection of finitely many halfspaces. If P is bounded, it is called a *polytope*.

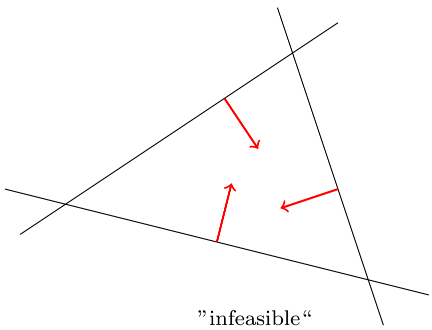


”bounded“

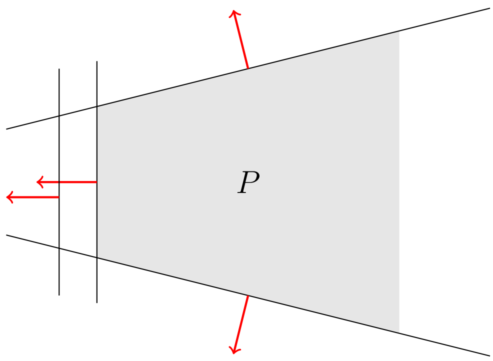


”unbounded“

If the constraints contradict themselves (for instance, $x + y \leq 1$, $x \geq 2$, and $x, y \geq 0$), then $P = \varnothing$, a corresponsing LP is called *infeasible*. Constraints might also be *redundant*, i.e. a constraint can be omitted without increasing the polyhedron (for example, $x+y \geq 0$ and $x, y \geq 0$).



”infeasible“



”redundant“

Polyhedra are *convex*, i.e. they satisfy Jensen's inequality: For all $x, y \in P$ and all $\vartheta \in [0, 1]$, $(1-\vartheta)x + \vartheta y \in P$.

Proof. Let $x, y \in P$ and $\vartheta \in [0, 1]$, then

$$\underbrace{(1-\vartheta)}_{\geq 0}\underbrace{x}_{\geq 0} + \underbrace{\vartheta}_{\geq 0}\underbrace{y}_{\geq 0} \geq 0$$

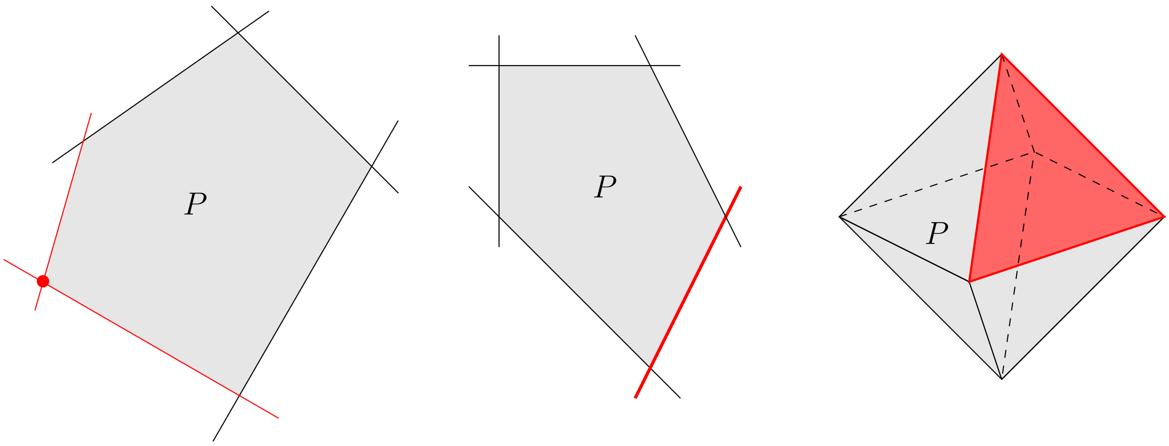
and moreover,

$$\begin{aligned} A((1-\vartheta)x+\vartheta y) &= (1-\vartheta)\underbrace{Ax}_{\leq b}+\vartheta\underbrace{Ay}_{\leq b} \\ &\leq (1-\vartheta)b+\vartheta b \\ &= b. \end{aligned}$$

□

The *dimension* $\dim P$ is the dimension of the smallest affine subspace containg P. If $P \subseteq \mathbb{R}^n$ with $\dim P = n$, then P is called *full dimensional*. For a given $x \in P$, a constraint $a_i^T x \leq b_i$ is called *active* (or *binding*) if $a_i^T x = b_i$. A *face* with respect to $H \subseteq \{1, \ldots, m\}$ is

$$F \coloneqq \{x \in P \mid a_i^T x \leq b_i \text{ active in } x, i \in H\}.$$



”vertex (0-dim)“ ”edge (1-dim)“ ”facet ($(n-1)$-dim)“

Every face itself is again a polyhedron. A point $x \in P$ is called *feasible solution*, $x^* \in P$ such that $c^T x^* \geq c^T x$ for all $x \in P$ is called *optimal solution* (provided that the objective function should be maximized), $c^T x^*$ is called *optimum*.

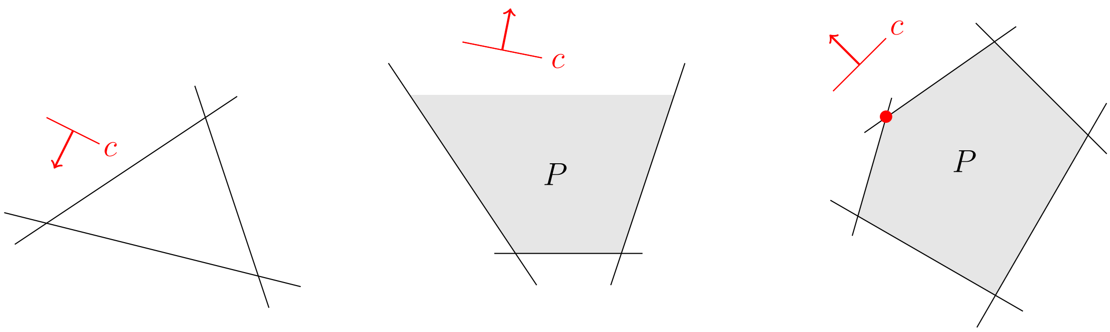
Theorem 1. *For every LP exactly one of the following hold:*

(i) The LP is infeasible, i.e. $P = \varnothing$.

(ii) The optimum is unbounded, i.e. for all $M > 0$, there exists $x \in P$ with $c^T x \geq M$.

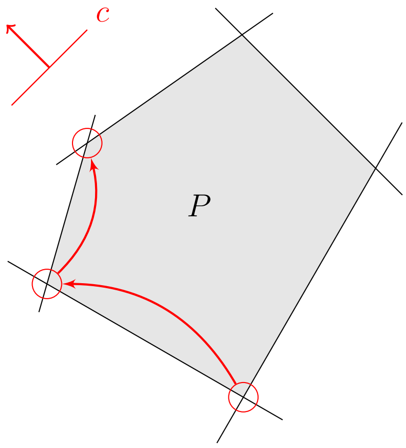
(iii) There exists a finite optimal solution.

If the last case is true, then the optimal sollution is assumed at a vertex of P.



Lemma 2. *An n-dimensional polyhedron given by m constraints has at most $\binom{m}{n}$ vertices, i.e. finitely many.*

Thus, of all (usually infinitely many) feasible points, only finitely many are relevant. However, enumerating all those points and choosing the best solution (*brute force*) is still very inefficient.
Idea for the Simplex algorithm: Move from vertex to vertex such that the objective value only increases (decreases, respectively, if objective is to minimize).



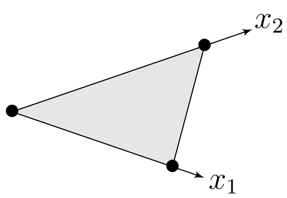
We introduce *slack variables* to fill gaps between constraints and corresponding right hand side.

$$y \coloneqq b - Ax, \quad y \geq 0.$$

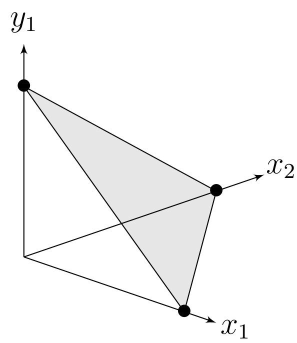
This yields to *standard form* for LPs:

$$\begin{aligned} &\text{minimize } c^T x \\ &\text{subject to } Ax + y = b \\ &\qquad x, y \geq 0 \end{aligned} \qquad\qquad \begin{aligned} &\text{minimize } c^T x \\ &\text{subject to } Ax = b \\ &\qquad x \geq 0 \end{aligned}$$

where the LP on the right is obtained by setting $x_{n+i} = y_i$ for $i \in \{1, \dots, m\}$ and we extend c and A in the obvious way.
The transformation from canonical form to standard form preserves dimension and vertices of the polyhedron.



$P = \{x \in \mathbb{R}^2_+ \mid x_1 + x_2 \leq 1\}$



$P' = \{(x, y) \in \mathbb{R}^3_+ \mid x_1 + x_2 + y_1 = 1\}$

Note that the slack variable $y_i = 0$ iff the i-th constraint is active. We can also interpret x as slack variable, just note that $x_j = 0$ iff $x_j \geq 0$ active.
Let $A_1, A_2, \dots, A_n, A_{n+1}, \dots, A_{n+m}$ be the columns of A (the latter m columns are unit vectors for slack varaiables). For $J \subseteq (1, \dots, n+m)$, let A_J denote the matrix consisting of columns A_j with $j \in J$, e.g. for

$$A = \begin{pmatrix} 3 & 7 & 0 & -1 & 1 & 0 \\ -1 & -1 & -2 & 2 & 0 & 1 \end{pmatrix}, \quad J = (5, 2) \implies A_J = \begin{pmatrix} 1 & 7 \\ 0 & -1 \end{pmatrix}.$$

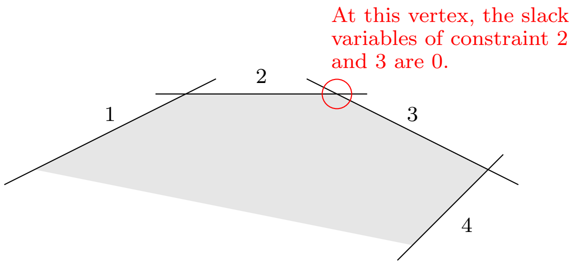
A *basis* $B = (B_1, \ldots, B_m) \subseteq (1, \ldots, n+m)$ is a subset of m column indices such that the corresponding columns are linearly independent. $N = (1, \ldots, n+m) \setminus B$ is called *non basis*. Variables x_j with $j \in B$ are called *basic variables*, and *non-basic variables* if $i \in N$. A vector $x \in \mathbb{Q}^{n+m}$ is a *basic solution* to $Ax = b$, $x \geq 0$ if there is a basis B such that

- $A_B x_B = b$ (*uniqueness*),
- $x_N = 0$ (*at boundary, vertex*).

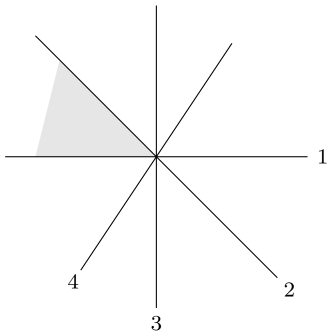
If additionally $x_B \geq 0$ holds x is called *feasible basic solution.*

Theorem 3. *Every feasible basic solution corresponds to exactly one vertex of P.*

Basic solution are also called *extreme point solutions.*
We need $\dim P$ constraints to describe a vertex. The non-basic variables correspond to the active constraints.



Note that the basis in one vertex is not necessarily unique. We call those vertices *degnerate.*



Possible bases:
$B = (1,2), (1,3), (1,4), (2,3), \ldots$

Given a standard form LP with $Ax = b$, $x \geq 0$ and basis B:

$$\begin{aligned} & A_B x_B + A_N x_N && = b \\ \iff & x_B && = \underbrace{A_B^{-1} b}_{=:\bar{b}} - \underbrace{A_B^{-1} A_N}_{=:\bar{A}_N} \underbrace{x_N}_{=0}. \end{aligned}$$

Using the objective function $z(x) = c^T x$:

$$\begin{aligned} z(x) &= c^T x \\ &= c_B^T x_B + c_N^T x_N \\ &= c_B^T (A_B^{-1} b - A_B^{-1} A_N x_N) + c_N^T x_N \\ &= c_B^T A_B^{-1} - c_B^T A_B^{-1} A_N x_N + c_N^T x_N \\ &= \underbrace{c_B^T A_B^{-1} b}_{=:\bar{z}} + \underbrace{(c_N^T - c_B^T A_B^{-1} A_N)}_{=:\bar{c}_N^T \text{ ”reduced cost“}} \underbrace{x_N}_{=0}. \end{aligned}$$

Optimality condition: Basis B (and the corresponding vertex) optimal if reduced cost $\bar{c}_N^T \leq 0$. Intuitively, no non-basic variable can be increased without decreasing the value of z.
Otherwise we can find a non-basic variable that can ”improve“ the objective value. This means we deactivate a constraint (increasing its slack) and move to another vertex:

- Initially: Vertex given by B, N.
- If there is a non-basic variable x_s, $s \in N$, with $\bar{c}_s > 0$, it is beneficial to increase x_s (currently $x_s = 0$).
- Since $x_B = A_B^{-1}b - A_B^{-1}A_N x_N$, the values of basic-variables decrease if $A_B^{-1}A_s > 0$.
- The maximum value for x_s is determined by "the first" basic variable which becomes 0.
- If this never happens, i.e. $A_B^{-1}A_s \leq 0$, then x_s can be arbitrarily increased; in this case the objective value is unbounded.

The first basis: If $b \geq 0$, all slack variables are a feasible basis, i.e. $B = (n+1, \ldots, n+m)$, and hence, $A_B = A_B^{-1} = \mathbb{1}_m$ and $\bar{A}_N = A_N$, $\bar{b} = b$. The first basic solution is then $x_{n+i} = b_i$ for $i \in \{1, \ldots, m\}$ and $x_1 = \ldots = x_n = 0$ (i.e. the origin). Since $c_B^T = 0$ (all slack), we have $\bar{c}_N^T = C_N^T$ and $\bar{z} = 0$. For calculation by hand, we can store all coefficients in a *dictionary* (or *tableau*):

	$\bar{c}_N^T$	0	$\bar{z}$
x_B	A_N	$\mathbb{1}_m$	b
	x_N	x_B	