Polyhedral Geometry and Linear Optimization

Summer Semester 2010

Andreas Paffenholz

Contents

Preface

These lecture notes grew out of a BMS class *Discrete Mathematics II* that I gave in the Winter Semester 2008/2009 at *Freie Universität Berlin*. I made several improvements during the summer semester 2010, when I gave a course on a similar topic, again at *Freie Universität Berlin*.

The course should give an introduction to the theory of discrete mathematical optimization from a discrete geometric view point, and present applications of this in geometry and graph theory. It covers convex polyhedral theory, the Simplex Method and Duality, integer polyhedra, unimodularity, TDI systems, cutting plane methods, and non-standard optimization methods coming from toric algebra.

These lecture notes originate from various sources. I got the idea for this course from a series of lectures by Rekha Thomas given at the PreDoc course at FU Berlin "Integer Points in Polyhedra" in Summer 2007 (unpublished). Most of the material can be found in the book of Schrijver [Sch86]. A not so dense treatment of polyhedral theory can be found in Ziegler's book [Zie95] and the book of Barvinok [Bar02]. The part about optimization in graphs is based on another book of Schrijver [Sch03]. For the parts on unimodularity and TDI systems I have also taken some material from the book of Korte and Vygen [KV08]. For the more practical chapters about duality and the simplex method I have mainly looked into Chvátal's book [Chv83]. An easy introduction is the book of Gärtner and Matoušek [MG07]. For more on linear programming, the excellent lecture notes of Grötschel and Möhring are a good place to look at. They are available online.

These lecture notes have benefitted from comments by participants in my courses and many other readers. In particular, Christian Haase, Silke Horn, Kaie Kubjas, Benjamin Lorenz, and Krishnan Narayanan gave valuable hints for a better exposition or informed me about typos or omissions in the text. There are certainly still many errors in this text. If you find one it would be nice if you write me an email. Any other feedback is also welcome.

Darmstadt, February 2013

Cones 1

This and the following two sections will introduce the basic objects in discrete geometry and optimization: cones, polyhedra, and linear programs. Each section also brings some new methods necessary to work with these objects. We will see that cones and polyhedra can both be defined in two different ways. We have an **exterior** description as the intersection of some half spaces, and we have an **interior** description as convex and conic combinations of a finite set of points. The main therem in this and the next section will be two versions of the MINKOWSKI-WEYL Theorem relating the two descriptions. This will directly lead to the well known linear programming duality, which we discuss in the third section. The basic tool for the proof of these duality theorems is the FARKAS Lemma. This is sometimes is also referred to as the **Fundamental Theorem of Linear Inequalities**. It is an example of an **alternative theorem**, i.e. it states that of two given options always exactly one is true. The FARKAS Lemma comes in many variants, and we will encounter several other versions and extensions in the next sections. Before we can actually start with discrete geometry we need to review some material from linear algebra and fix some terminology.

Fundamental Theorem of Linear Inequalities
alternative theorem

Definition 1.1. *Let* $\mathbf{x}_1, \ldots, \mathbf{x}_k \in \mathbb{R}^n$*, and* $\lambda_1, \ldots, \lambda_k \in \mathbb{R}$*. Then* $\sum_{i=1}^k \lambda_i \mathbf{x}_i$ *is called a* **linear combination** *of the vectors* $\mathbf{x}_1, \ldots, \mathbf{x}_k$*. It is further a*

(1) **conic combination***, if* $\lambda_i \geq 0$*,*
(2) **affine combination***, if* $\sum_{i=1}^k \lambda_i = 1$*, and a*
(3) **convex combination***, if it is conic and affine.*

linear combination
conic combination
affine combination
convex combination

The **linear** *(***conic***,* **affine***,* **convex***)* **hull** *of a set* $X \subseteq \mathbb{R}^n$ *is the set of all points that are a linear (conic, affine, convex) combination of some finite subset of* X*. It is denoted by* $\operatorname{lin}(X)$ *(or,* $\operatorname{cone}(X)$, $\operatorname{aff}(X)$, $\operatorname{conv}(X)$*, respectively).* X *is a* **linear space** *(***cone***,* **affine space***,* **convex set***) if* X *equals its linear hull (or conic hull, affine hull, convex hull, respectively). Figure 1 illustrates the affine, conic, and convex hull of the set* $X := \{(-1,5),(1,2)\}$*. The linear span of* X *is the whole plane.*

linear hull
conic hull
affine hull
convex hull
linear space
cone
affine space
convex set

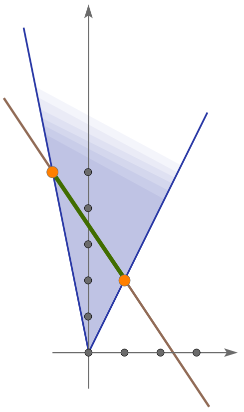


Figure 1.1

We denote the dual space of $\mathbb{R}^n$ by $(\mathbb{R}^n)^*$. This is the vector space of all **linear functionals** $\mathbf{a} : \mathbb{R}^n \to \mathbb{R}$. Given some basis $\mathbf{e}_1, \ldots, \mathbf{e}_n$ of $\mathbb{R}^n$ and the corresponding basis $\mathbf{e}_1^*, \ldots, \mathbf{e}_n^*$ of $(\mathbb{R}^n)^*$ (using the standard scalar product) we can write elements $\mathbf{x} \in \mathbb{R}^n$ and $\mathbf{a} \in (\mathbb{R}^n)^*$ as (column) vectors

$$\mathbf{x} := \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} \qquad \mathbf{a} := \begin{pmatrix} a_1 \\ \vdots \\ a_n \end{pmatrix}.$$

To save space we will often write vectors as row vectors in this text. A linear functional $\mathbf{a} : \mathbb{R}^n \to \mathbb{R}$ then has the form

$$\mathbf{a}(\mathbf{x}) = \mathbf{a}^t \mathbf{x} = a_1 x_1 + \cdots a_n x_n .$$

Using the standard scalar product we can (and most often will) identify $\mathbb{R}^n$ and its dual space $(\mathbb{R}^n)^*$, and view a linear functional as a (row) vector $\mathbf{a}^t \in \mathbb{R}^n$.

Let $B \in \mathbb{R}^{m \times n}$ be a matrix with column vectors $\mathbf{b}_1, \ldots, \mathbf{b}_n$. Then $\operatorname{cone}(B)$ denotes the conic hull $\operatorname{cone}(\{\mathbf{b}_1, \ldots, \mathbf{b}_n\})$ of these vectors. Similarly, if $A \in \mathbb{R}^{m \times n}$ is a matrix of functionals $\mathbf{a}_1^t, \ldots, \mathbf{a}_m^t$ (i.e. the rows of A), then $\operatorname{cone}(A) := \operatorname{cone}(\mathbf{a}_1^t, \ldots, \mathbf{a}_m^t) \subseteq (\mathbb{R}^n)^*$. Define $\operatorname{lin}(B)$, $\operatorname{aff}(B)$, $\operatorname{conv}(B)$, $\operatorname{lin}(A)$, $\operatorname{aff}(A)$, and $\operatorname{conv}(A)$ similarly.

linear half-space
affine half-space

Definition 1.2. *For any non-zero linear functional $\mathbf{a}^t \in \mathbb{R}^n$ and $\delta \in \mathbb{R}$ the set*

$$\begin{aligned} &\{\mathbf{x} \mid \mathbf{a}^t\mathbf{x} \le 0\} && \text{is a } \textbf{linear half-space}\text{, and} \\ &\{\mathbf{x} \mid \mathbf{a}^t\mathbf{x} \le \delta\} && \text{is an } \textbf{affine half-space.} \end{aligned}$$

linear hyperplane
affine hyperplane

Their boundaries $\{\mathbf{x} \mid \mathbf{a}^t\mathbf{x} = 0\}$ and $\{\mathbf{x} \mid \mathbf{a}^t\mathbf{x} = \delta\}$ are a **linear** *and* **affine hyperplane** *respectively.*

We can now define the basic objects that we will study in this course.

polyhedral cone
finitely constrained cone

Definition 1.3. *(1) A* **polyhedral cone** *(or a* **finitely constrained cone***) is a subset $C \subseteq \mathbb{R}^n$ of the form*

$$C := \{\mathbf{x} \in \mathbb{R}^n \mid A\mathbf{x} \le \mathbf{0}\}$$

for a matrix $A \in \mathbb{R}^{m \times n}$ of row vectors (linear functionals).

(2) For a finite number of vectors $\mathbf{b}_1, \ldots, \mathbf{b}_r \in \mathbb{R}^n$ the set

$$C = \operatorname{cone}(\mathbf{b}_1, \ldots, \mathbf{b}_r) := \left\{\sum_{i=1}^r \lambda_i \mathbf{b}_i \mid \lambda_i \ge 0\right\} = \{B\lambda \mid \lambda \ge 0\}.$$

finitely generated

is a **finitely generated cone** *C, where B is the matrix with columns $\mathbf{b}_1, \ldots, \mathbf{b}_r$.*

Definition 1.1 immediately implies that a finitely generated cone is a cone. Note that any A uniquely defines a cone, but the same cone can be defined by different constraint matrices. Let $\lambda, \mu \ge 0$ and $\mathbf{a}_1^t, \mathbf{a}_2^t$ two rows of A. Then

$$C := \{\mathbf{x} \mid A\mathbf{x} \le \mathbf{0}\} = \{\mathbf{x} \mid A\mathbf{x} \le \mathbf{0},\ (\lambda\mathbf{a}_1^t + \mu\mathbf{a}_2^t)\mathbf{x} \le 0\}$$

Similarly, a generating set uniquely defines a cone, but adding the sum of all generators as a new generator does not change the cone.

dimension

Definition 1.4. *The* **dimension** *of a cone C is $\dim(C) := \dim(\operatorname{lin}(C))$.*

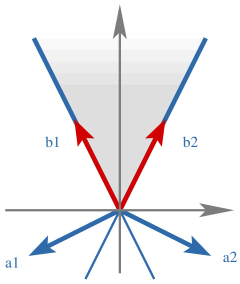


Figure 1.2

Example 1.5. Let $A := \begin{pmatrix} -2 & -1 \\ 2 & -1 \end{pmatrix}$, $\mathbf{b}_1 := \begin{pmatrix} -1 \\ 2 \end{pmatrix}$, and $\mathbf{b}_2 := \begin{pmatrix} 1 \\ 2 \end{pmatrix}$.

Then $C := \{\mathbf{x} \mid A\mathbf{x} \le \mathbf{0}\}$ and $C' := \operatorname{cone}(\mathbf{b}_1, \mathbf{b}_2)$ define the same subset of $\mathbb{R}^2$. It is the shaded area in Figure 1.2. Note that we have drawn the functionals $\mathbf{a}_1^t, \mathbf{a}_2^t$ in the same picture, together with the lines $\mathbf{a}_i^t\mathbf{x} = 0$, $i = 1, 2$. ◊

We want to show that the two definitions of a polyhedral and a finitely generated cone actually describe the same objects, i.e. any finitely constrained cone has a finite generating set, and any finitely generated cone can equally be described as the intersection of finitely many linear half spaces.

Fourier-Motzkin *elimination*

The main tool we need in the proof is a method to solve systems of linear inequalities, known as FOURIER-MOTZKIN elimination. The analogous task for a system of linear equations is efficiently solved by the well known GAUSSian elimination. We will basically exploit the same idea for linear inequalities. However, in contrast to GAUSSian elimination, it will not be efficient (nevertheless it is still one of the best algorithms known for this problem). We start with an example to explain the idea, before we give a formal proof. We consider the following system of linear inequalities.

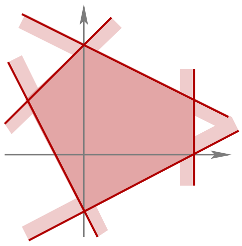


Figure 1.3

$$\begin{array}{rcrcl} -x & + & y & \le & 2 \\ x & + & 2y & \le & 4 \\ -2x & - & y & \le & 1 \\ x & - & 2y & \le & 2 \\ x & & & \le & 2 \end{array} \tag{1.1}$$

See Figure 1.3 for an illustration fo the solution set. For a given x we want to find conditions that guarantee the existence of a y such that (x, y) is a solution. We rewrite the inequalities, solving for y. The first two conditions impose an upper bound on y,

$$\begin{aligned} y &\leq 2 + \quad x \\ y &\leq 2 - \frac{1}{2}x \end{aligned}$$

the third and forth a lower bound,

$$\begin{aligned} -1 - \ 2x &\leq y \\ -1 + \frac{1}{2}x &\leq y \end{aligned}$$

while the last one does not affect y:

$$x \leq 2$$

Hence, for a given x, the system (1.1) has a solution for y if and only if

$$\max(-1-2x, -1+\frac{1}{2}x) \leq y \leq \min(2+x, 2-\frac{1}{2}x), \tag{1.2}$$

that is, if and only if the following system of inequalities holds:

$$\begin{array}{rcrcrcr} -1 & - & 2x & \leq & 2 & + & x \\ -1 & - & 2x & \leq & 2 & - & \frac{1}{2}x \\ -1 & + & \frac{1}{2}x & \leq & 2 & + & x \\ -1 & + & \frac{1}{2}x & \leq & 2 & - & \frac{1}{2}x \end{array}$$

Rewriting this in standard form, and adding back the constraint that does not involve y, we obtain

$$\begin{array}{rcl} -x & \leq & 1 \\ -x & \leq & 2 \\ -x & \leq & 6 \\ x & \leq & 3 \\ x & \leq & 2 \end{array} \tag{1.3}$$

This system (1.3) of inequalities has a solution if and only if the system (1.1) has a solution. However, it has one variable less. We can now iterate to obtain

$$\max(-6, -2, -1) \ \leq \ x \ \leq \ \min(2, 3). \tag{1.4}$$

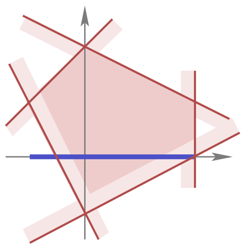


Figure 1.4

Now both the minimum and maximum do not involve a variable anymore, so we can compute them to obtain that (1.1) has a solution if and only if

$$-1 \leq x \leq 2$$

This is satisfiable, so (1.1) does have a solution. (1.4) tells us, that any x between -1 and 2 is good. If we have fixed some x, we can plug it into (1.2) to obtain a range of solutions for y. The range for x is just the projection of the original set onto the x-axis. See Figure 1.4.

In general, by iteratively applying this elimination procedure to a system of linear inequalities we obtain a method to check whether the system has a solution (and we can even compute one, by substituting solutions).

Theorem 1.6 **(FOURIER-MOTZKIN-Elimination).** *Let $A\mathbf{x} \leq \mathbf{b}$ be a system of linear inequalities with $n \geq 1$ variables and m inequalities.*

Then there is a system $A'\mathbf{x}' \leq \mathbf{b}'$ with $n-1$ variables and at most $\max(m, \frac{m^2}{4})$ inequalities, such that $\mathbf{s}'$ is a solution of $A'\mathbf{x}' \leq \mathbf{b}'$ if and only if there is $\mathbf{s}_0 \in \mathbb{R}$ such that $\mathbf{s} := (s_0, \mathbf{s}')$ is a solution of $A\mathbf{x} \leq \mathbf{b}$.

Proof. We classify the inequalities depending on the coefficient a_{i0} of x_0. Let U be the indices of inequalities with $a_{i0} > 0$, E those with $a_{i0} = 0$ and L the rest. We multiply all inequalities in L and U by $\frac{1}{|a_{i0}|}$.

Now we eliminate x_0 by adding any inequality in L to any inequality in U. That is, our new system consists of the inequalities

$$\begin{aligned} \mathbf{a}_j'^t\mathbf{x}' + \mathbf{a}_k'^t\mathbf{x}' &\le b_j + b_k && \text{for } j \in L, k \in U \\ \mathbf{a}_l'^t\mathbf{x}' &\le b_l && \text{for } l \in E. \end{aligned} \tag{1.5}$$

Any solution $\mathbf{x}$ of the original system yields a solution $\mathbf{x}'$ of the new system by just forgetting the first coordinate. We have at most $|L| \cdot |U| + |E|$ many inequalities, which proves the bound.

Now assume, our new system has a solution $\mathbf{x}'$. Then (1.5) implies

$$\mathbf{a}_j'^t\mathbf{x}' - b_j \le b_k - \mathbf{a}_k'^t\mathbf{x}' \qquad \text{for all } j \in L, k \in U$$

which implies

$$\max_{j\in L}(\mathbf{a}_j'^t\mathbf{x}' - b_j) \le \min_{k\in U}(b_k - \mathbf{a}_k'^t\mathbf{x}') . \tag{1.6}$$

I f we take for x_0 any value in between, then

$$\begin{aligned} x_0 + \mathbf{a}_k'^t\mathbf{x}' &\le b_k && \text{for all } k \in U \\ -x_0 + \mathbf{a}_j'^t\mathbf{x}' &\le b_j && \text{for all } j \in L \end{aligned}$$

The coefficient of x_0 is 0 for all inequalities in E. Hence, they are also satisfied by $\mathbf{x} = (x_0, \mathbf{x}')$ and we have found a solution of $A\mathbf{x} \le \mathbf{b}$. □

Remark 1.7. (1) If $A, \mathbf{b}$ are rational, then so are $A', \mathbf{b}'$.

(2) Any inequality in the new system is a conic combination of inequalities of the original system. This implies that there is a matrix $U \ge 0$ such that $A' = UA$ and $\mathbf{b}' = U\mathbf{b}$.

(3) U or L may be empty. In this case only the inequalities in the set E survive. More specifically, suppose that $L = \varnothing$. Then, given any solution $\mathbf{x}'$ of the projected system, we can choose

$$x_0 := \min_{i:a_{i0}>0} \frac{b_i - \sum_{k=1}^n a_{ik}x_k}{a_{i0}}$$

and obtain a solution to the original system. Further, if $a_{i0} > 0$ for all i, then the projected system is empty, hence any point $\mathbf{x}' \in \mathbb{R}^{n-1}$ lifts to a solution of $A\mathbf{x} \le \mathbf{b}$. ◊

Now that we have studied our new tool, let us get back to cones.

Weyl's *Theorem*

Theorem 1.8 **(WEYL's Theorem).** *A non-empty finitely generated cone C is polyhedral.*

Proof. Let $C = \{B\boldsymbol{\lambda} \mid \boldsymbol{\lambda} \ge 0\}$ be a finitely generated cone with a matrix of generators $B \in \mathbb{R}^{n\times r}$. Then

$$\begin{aligned} C &= \{\mathbf{x} \in \mathbb{R}^n \mid \exists\, \boldsymbol{\lambda} \in \mathbb{R}^r \,:\, \mathbf{x} = B\boldsymbol{\lambda}, \boldsymbol{\lambda} \ge \mathbf{0}\} \\ &= \{\mathbf{x} \in \mathbb{R}^n \mid \exists\, \boldsymbol{\lambda} \in \mathbb{R}^r \,:\, \mathbf{x} - B\boldsymbol{\lambda} \le \mathbf{0},\ -\mathbf{x} + B\boldsymbol{\lambda} \le \mathbf{0},\ -\boldsymbol{\lambda} \le \mathbf{0}\} \end{aligned}$$

This set is the projection onto the first n coordinates of the set

$$C' := \{(\mathbf{x}, \boldsymbol{\lambda}) \in \mathbb{R}^{n+r} \mid \mathbf{x} - B\boldsymbol{\lambda} \le \mathbf{0},\ -\mathbf{x} + B\boldsymbol{\lambda} \le \mathbf{0},\ -\boldsymbol{\lambda} \le \mathbf{0}\} . \tag{1.7}$$

Using FOURIER-MOTZKIN elimination to eliminate the variables $\lambda_1, \dots, \lambda_r$ from the system of linear inequalities defining C' we can write the cone C as

$$C = \{\mathbf{x} \in \mathbb{R}^n \ \mid Ax \le \mathbf{0}\}$$

for some matrix $A \in \mathbb{R}^{m\times n}$. Hence, C is polyhedral. □

Note that the proof of this theorem is constructive. Given any finitely generated cone C we can write it in the form (1.7) and apply FOURIER-MOTZKIN elimination to obtain a corresponding system of linear inequalities. Using WEYL's theorem we can now prove a first variant of the FARKAS Lemma.

***Theorem 1.9* (FARKAS Lemma, Geometric Version).** *For any matrix $B \in \mathbb{R}^{m \times n}$ and vector $\mathbf{b} \in \mathbb{R}^m$ exactly one of the following holds:*

Farkas *Lemma*

- *(1) there is $\boldsymbol{\lambda} \in \mathbb{R}^m$ such that $B\boldsymbol{\lambda} = \mathbf{b}$, $\boldsymbol{\lambda} \geq \mathbf{0}$, or*
- *(2) there is $\mathbf{a} \in \mathbb{R}^n$ such that $\mathbf{a}^t B \leq \mathbf{0}$ and $\mathbf{a}^t \mathbf{b} > 0$.*

Geometrically, this theorem means the following. Given a cone C generated by the columns of B and some vector $\mathbf{b}$,

- (1) either $\mathbf{b} \in C$, in which case there are non-negative coefficients that give a representation of $\mathbf{b}$ using the columns of B,
- (2) or $\mathbf{b} \notin C$, in which case we can find a hyperplane $H_{\mathbf{a}}$ given by its normal $\mathbf{a}$, such that $\mathbf{b}$ and C are on different sides of $H_{\mathbf{a}}$.

Proof (FARKAS Lemma). The two statements cannot hold simultaneously: Assume there is $\boldsymbol{\lambda} \geq \mathbf{0}$ such that $B\boldsymbol{\lambda} = \mathbf{b}$ and $\mathbf{a} \in \mathbb{R}^m$ such that $\mathbf{a}^t B \leq \mathbf{0}$ and $\mathbf{a}^t \mathbf{b} > 0$. Then

$$0 < \mathbf{a}^t \mathbf{b} = \mathbf{a}^t (B\boldsymbol{\lambda}) = \mathbf{a}^t (B\boldsymbol{\lambda}) \leq 0\,,$$

a contradiction. Let $C := \{B\boldsymbol{\mu} \mid \boldsymbol{\mu} \geq 0\}$. Then there is $\boldsymbol{\lambda} \geq \mathbf{0}$ such that $B\boldsymbol{\lambda} = \mathbf{b}$ if and only if $\mathbf{b} \in C$. By WEYL's Theorem the cone C is polyhedral and there is a matrix A such that

$$C = \{\mathbf{x} \mid A\mathbf{x} \leq \mathbf{0}\}\,. \tag{1.8}$$

Hence, $\mathbf{b} \notin C$ if and only if there is a functional $\mathbf{a}^t$ among the rows of A such that $\mathbf{a}^t \mathbf{b} > 0$. Clearly, $\mathbf{b}_j \in C$ for each column of B, hence $\mathbf{a}^t B \leq \mathbf{0}$. So $\mathbf{a}$ is as desired. □

Definition 1.10. *The* **polar (dual)** *of a cone $C \subset \mathbb{R}^n$ is the set*

polar dual of a cone

$$C^* := \{\mathbf{a} \in \mathbb{R}^n \mid \mathbf{a}^t \mathbf{x} \leq \mathbf{0} \ \textit{ for all } \ \mathbf{x} \in C\}.$$

See Figure 1.5 for and example of a cone and its dual.

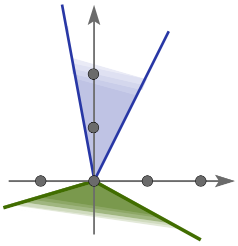


Figure 1.5

Proposition 1.11. *Let $C, D \subseteq \mathbb{R}^n$ be cones. Then the following holds:*

- *(1) $C \subseteq D$ implies $D^* \subseteq C^*$.*
- *(2) $C \subseteq C^{**}$.*
- *(3) $C^* = C^{***}$.*

Proof. (1) $\mathbf{a}^t \in D^*$ means $\mathbf{a}^t \mathbf{x} \leq 0$ for all $\mathbf{x} \in D \supseteq C$.

- (2) $\mathbf{x} \in C$ means $\mathbf{a}^t \mathbf{x} \leq 0$ for all $\mathbf{a} \in C^*$.
- (3) $\mathbf{a} \in C^{***}$ means $\mathbf{a}^t \mathbf{x} \leq 0$ for all $x \in C^{**}$, which is equivalent to $\mathbf{a} \in C^*$. □

Lemma 1.12. *Let C be a cone.*

- *(1) If $C = \{B\boldsymbol{\lambda} \mid \boldsymbol{\lambda} \geq 0\}$ then $C^* = \{\mathbf{a}^t \mid \mathbf{a}^t B \leq \mathbf{0}^t\}$.*
- *(2) if $C = \{\mathbf{x} \mid A\mathbf{x} \leq \mathbf{0}\}$, then $C = C^{**}$.*
- *(3) If C is finitely generated then $C^{**} = C$.*
- *(4) If $C = \{\mathbf{x} \mid A\mathbf{x} \leq \mathbf{0}\}$ is polyhedral for some $A \in \mathbb{R}^{m \times n}$, then $C^* = \{\boldsymbol{\lambda}^t A \mid \boldsymbol{\lambda} \geq \mathbf{0}\}$ is finitely generated.*

Proof. (1) Clearly, the inequalities are necessary. Let $\mathbf{a}$ satisfy $\mathbf{a}^t B \leq \mathbf{0}^t$. Then $\mathbf{x} = B\boldsymbol{\lambda} \in C$ implies $\mathbf{a}^t \mathbf{x} = \mathbf{a}^t B\boldsymbol{\lambda} \leq 0$, as $\boldsymbol{\lambda} \geq \mathbf{0}$.

- (2) Let $D := \{\boldsymbol{\lambda}^t A \mid \boldsymbol{\lambda} \geq 0\}$. (1) implies $C = D^*$ and Proposition 1.11(3) then shows $C^{**} = D^{***} = D^* = C$.

(3) Let $C = \{B\boldsymbol{\lambda} \mid \boldsymbol{\lambda} \geq 0\}$. Then (1) tells us that $C^* = \{\mathbf{a} \mid \mathbf{a}^t B \leq \mathbf{0}\}$.
By Proposition 1.11(2) we need only prove $C^{**} \subseteq C$. But this follows from the observation that, if $\mathbf{b} \notin C$, then, by the FARKAS Lemma, there is $\mathbf{a}^t$ such that $\mathbf{a}^t B \leq \mathbf{0}$, $\mathbf{a}^t\mathbf{b} > 0$. The first inequality implies $\mathbf{a}^t \in C^*$ and the second $\mathbf{b} \notin C^{**}$.

(4) C is the dual of $D := \{\boldsymbol{\lambda}^t A \mid \boldsymbol{\lambda} \geq 0\}$, i.e. $C = D^*$. By dualizing again and using (1) we have $C^* = D^{**} = D$. ☐

Minkowski's *Theorem*

***Theorem 1.13* (MINKOWSKI's Theorem).** *A polyhedral cone is non-empty and finitely generated.*

Proof. Let $C = \{\mathbf{x} \mid A\mathbf{x} \leq \mathbf{0}\}$. Then $\mathbf{0} \in C$ and C is not empty. Let $D := \{\boldsymbol{\lambda}^t A \mid \boldsymbol{\lambda} \geq 0\}$. By WEYL's Theorem, D is polyhedral. Hence D^* is finitely generated. But $D^* = C$, and so C is finitely generated. ☐

Weyl-Minkowski-*Duality*

Combining this with WEYL's Theorem we get the **WEYL-MINKOWSKI-Duality** for cones.

Theorem 1.14. *A cone is polyhedral if and only if it is finitely generated.* ☐

This finally proves the claimed equivalence of the two definitions of a cone in Definition 1.3. The FOURIER-MOTZKIN elimination also gives us a method to convert between the two representations (apply the method to the dual if you want to convert from a polyhedral representation to the generators). In the next chapters we will see that, although mathematically equivalent, there are properties of cones that are trivial to compute in one of the representations, but hard in the other. This affects also algorithmic questions. Efficient conversion between the two representations is the fundamental algorithmic problem in polyhedral theory.

Before we consider polyhedra and their representations in the next chapter we want to list some useful variants of the FARKAS lemma that follow directly from the above geometric version.

Proposition 1.15. *Let $B \in \mathbb{R}^{m\times n}$ and $\mathbf{b} \in \mathbb{R}^m$.*

(1) Either $B\boldsymbol{\lambda} = \mathbf{b}$ has a solution or there is $\mathbf{a}$ such that $\mathbf{a}^t B = \mathbf{0}$ and $\mathbf{a}^t b > 0$, but not both.

(2) Either $B\boldsymbol{\lambda} \leq \mathbf{b}$ has a solution, or there is $\mathbf{a} \leq \mathbf{0}$ such that $\mathbf{a}^t B = 0$ and $\mathbf{a}^t\mathbf{b} > 0$, but not both.

(3) Either $B\boldsymbol{\lambda} \leq \mathbf{b}$, $\boldsymbol{\lambda} \geq \mathbf{0}$ has a solution or there is $\mathbf{a} \leq \mathbf{0}$ such that $\mathbf{a}^t B \leq 0$ and $\mathbf{a}^t\mathbf{b} > 0$, but not both. ☐

We finish this chapter with an application of FOURIER-MOTZKIN elimination to linear programming. FOURIER-MOTZKIN elimination allows us to

(1) decide whether a linear program is feasible, and
(2) determine an optimal solution.

Let a linear program

$$\text{maximize} \quad \mathbf{c}^t\mathbf{x} \qquad\qquad \text{subject to} \quad A\mathbf{x} \leq \mathbf{b}$$

be given, with $A \in \mathbb{R}^{m\times n}$, $\mathbf{b} \in \mathbb{R}^m$, and $\mathbf{c} \in \mathbb{R}^n$. If we apply FOURIER-MOTZKIN elimination n-times to the system $A\mathbf{x} \leq \mathbf{b}$, then no variable is left and we have inequalities of the form

$$0 \leq \alpha_j \tag{1.9}$$

for right hand sides $\alpha_1, \ldots, \alpha_k$. By the FOURIER-MOTZKIN theorem, the system has a solution if and only if all inequalities in (1.9) have a solution, i.e. if all α_j are non-negative.

To obtain the optimal value, we add an additional variable x_{n+1} and extend our system as follows

$$B := \begin{pmatrix} A & 0 \\ -\mathbf{c}^t & 1 \end{pmatrix} \qquad \mathbf{d} := \begin{pmatrix} \mathbf{b} \\ 0 \end{pmatrix}.$$

In the system $B\binom{\mathbf{x}}{x_{n+1}} \leq \mathbf{d}$ we eliminate the first n variables, and we obtain upper and lower bounds on x_{n+1} in the form

$$\alpha_j \leq x_{n+1} \leq \beta_j$$

The minimum over the β_j is our optimal value, as this is the largest possible value such that there is $\mathbf{x} \in \mathbb{R}^n$ with

$$-\mathbf{c}^t\mathbf{x} + x_{n+1} \leq 0 \qquad \Longleftrightarrow \qquad x_{n+1} \leq \mathbf{c}^t\mathbf{x}\,.$$

If we need the minimum over $\mathbf{c}^t\mathbf{x}$ then we add the row $(\mathbf{c}^t, -1)$ instead. Observe however, that this procedure is far from practical, the number of inequalities may grow exponentially in the number of eliminated variables (can you find an example for this?).

Polyhedra 2

Now we generalize the results of the previous section to polyhedra and collect basic geometric and combinatorial properties. A polyhedron is a quite natural generalization of a cone. We relax the type of half spaces we use in the definition and allow affine instead of linear boundary hyperplanes. We will see that we can use the theory developed in the previous section to state an affine version of our duality theorem. We will continue this in Section 4 with the study of faces of polyhedra after we have discussed Linear Programming and Duality in the next section.

Definition 2.1. *(1) A* **polyhedron** *is a subset $P \subseteq \mathbb{R}^n$ of the form*

$$P = P(A, \mathbf{b}) = \{\mathbf{x} \in \mathbb{R}^n \mid Ax \leq \mathbf{b}\}$$

for a matrix $A \in \mathbb{R}^{m \times n}$ of row vectors and a vector $\mathbf{b} \in \mathbb{R}^m$.

(2) A **polytope** *is the convex hull of a finite set of points in $\mathbb{R}^n$.*

polyhedron

polytope

Definition 2.2. *The* **dimension** *of a non-empty polyhedron P is $\dim(P) := \dim(\operatorname{aff}(P))$. The dimension of $\varnothing$ is -1.*

dimension

Example 2.3. (1) Polyhedral cones are polyhedra.

(2) The matrix $A := \begin{pmatrix} -1 & 1 \\ 1 & 0 \\ 0 & 1 \end{pmatrix}$ and the vector $\mathbf{b} := \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix}$ define the polyhedron shown in Figure 2.1.

(3) The system of linear inequalities and equations

$$\begin{aligned} B\mathbf{x} + C\mathbf{y} &\leq \mathbf{c} \\ D\mathbf{x} + E\mathbf{y} &= \mathbf{d} \\ \mathbf{x} \quad &\geq \mathbf{0} \end{aligned} \tag{*}$$

for compatible matrices B, C, D and E and vectors $\mathbf{c}, \mathbf{d}$ is a polyhedron:

Put $\quad A := \begin{pmatrix} B & C \\ D & E \\ -D & -E \\ -I & 0 \end{pmatrix} \qquad b := \begin{pmatrix} \mathbf{c} \\ \mathbf{d} \\ -\mathbf{d} \\ \mathbf{0} \end{pmatrix}$

(4) $\mathbb{R}^n = \{\mathbf{x} \mid \mathbf{0}^t\mathbf{x} \leq 0\}$, $\varnothing = \{\mathbf{x} \mid \mathbf{0}^t\mathbf{x} \leq -1\}$ and affine spaces are polyhedra.

(5) cubes $\operatorname{conv}(\{0,1\}^n) \subset \mathbb{R}^n$ are polytopes. ◊

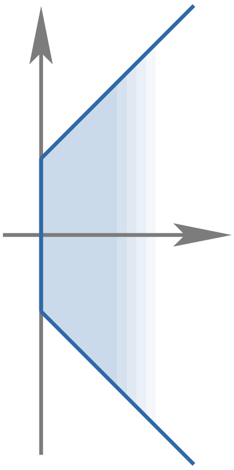


Figure 2.1

Proposition 2.4. *Arbitrary intersections of polyhedra with a affine spaces or polyhedra are polyhedra.*

Proof. $P(A, \mathbf{b}) \cap P(A', \mathbf{b}') = \{\mathbf{x} \mid A\mathbf{x} \leq \mathbf{b}, A'\mathbf{x} \leq \mathbf{b}'\}$. ☐

A polyhedron defined by linear inequalities is a finitely generated cone by our considerations in the previous section, and we will see in the next theorem that any bounded polyhedron is a polytope. The general case interpolates between these two by taking a certain sum of a cone and a polytope.

Definition 2.5. *For two subsets $P, Q \subseteq \mathbb{R}^n$, the set*

$$P + Q := \{p + q \mid p \in P,\ q \in Q\}$$

Minkowski *sum*

is the **MINKOWSKI sum** *of P and Q.*

We will prove later that Minkowski sums of polyhedra are again polyhedra. The next theorem gives the analogue of WEYL's Theorem for polyhedra. We will prove it by reducing a general polyhedron P to a cone, the **homogenization** homog(P) of P.

homogenization

Affine Weyl *Theorem*

Theorem 2.6 **(Affine WEYL Theorem).** *Let $B \in \mathbb{R}^{n\times p}$, $C \in \mathbb{R}^{n\times q}$, and*

$$P := \operatorname{conv}(B) + \operatorname{cone}(C) = \{B\boldsymbol{\lambda} + C\boldsymbol{\mu} \mid \boldsymbol{\lambda}, \boldsymbol{\mu} \geq 0, \sum \lambda_i = 1\}\,.$$

Then there is $A \in \mathbb{R}^{m\times n}$ and $\mathbf{b} \in \mathbb{R}^m$ such that $P = \{\mathbf{x} \in \mathbb{R}^n \mid A\mathbf{x} \leq \mathbf{b}\}$.

Proof. For $P = \emptyset$ we can take $A = 0$ and $\mathbf{b} = -1$. If $P \neq \emptyset$, but $B = 0$, then the theorem reduces to WEYL's Theorem for cones. So assume that $P \neq \emptyset$ and $p > 0$. We define the following finitely generated cone:

$$\begin{aligned} Q &:= \left\{ \begin{pmatrix} \mathbf{1}^t & \mathbf{0}^t \\ B & C \end{pmatrix} \begin{pmatrix} \boldsymbol{\lambda} \\ \boldsymbol{\mu} \end{pmatrix} \;\middle|\; \boldsymbol{\lambda}, \boldsymbol{\mu} \geq 0 \right\} \\ &= \left\{ \begin{pmatrix} \sum \lambda_i \\ B\boldsymbol{\lambda} + C\boldsymbol{\mu} \end{pmatrix} \;\middle|\; \boldsymbol{\lambda}, \boldsymbol{\mu} \geq 0 \right\}. \end{aligned}$$

Q is the **homogenization** of P. We obtain the following correspondence between points in P and in Q:

$$\mathbf{x} \in P \quad \Longleftrightarrow \quad \begin{pmatrix} 1 \\ \mathbf{x} \end{pmatrix} \in Q\,. \tag{2.1}$$

Q is a cone, so by WEYL's Theorem 1.8 there is a matrix A' such that

$$Q = \{ \left(\begin{smallmatrix} x_0 \\ \mathbf{x} \end{smallmatrix}\right) \mid A' \left(\begin{smallmatrix} x_0 \\ \mathbf{x} \end{smallmatrix}\right) \leq \mathbf{0} \}\,.$$

Write A' in the form $A' = (-\mathbf{b}|A)$ by separating the first column. Using the correspondence (2.1) between points in P and in Q we obtain $P = \{\mathbf{x} \in \mathbb{R}^n \mid A\mathbf{x} \leq \mathbf{b}\}$. ☐

In particular, we obtain from this Theorem, that the Minkowski sum of a polytope and a cone is a polyhedron. Using the same trick of homogenizing a polyhedron we can also prove the analogue of MINKOWSKI's Theorem.

Affine Minkowski *Theorem*

Theorem 2.7 **(Affine MINKOWSKI Theorem).** *Let $P = \{\mathbf{x} \in \mathbb{R}^n \mid A\mathbf{x} \leq \mathbf{b}\}$ for some $A \in \mathbb{R}^{m\times n}$ and $\mathbf{b} \in \mathbb{R}^m$. Then there is $B \in \mathbb{R}^{n\times p}$ and $C \in \mathbb{R}^{n\times q}$ such that*

$$P = \operatorname{conv}(B) + \operatorname{cone}(C)\,.$$

Proof. If $P = \emptyset$, then we let B and C be empty matrices (i.e. we take $p = q = 0$). Otherwise, we define a polyhedral cone $Q \subseteq \mathbb{R}^{n+1}$ as

$$Q := \left\{ \begin{pmatrix} x_0 \\ \mathbf{x} \end{pmatrix} \;\middle|\; \begin{pmatrix} -1 & \mathbf{0}^t \\ -\mathbf{b} & A \end{pmatrix} \begin{pmatrix} x_0 \\ \mathbf{x} \end{pmatrix} \leq \mathbf{0} \right\}$$

Again, in the same way as in the previous proof, we obtain the correspondence $\mathbf{x} \in P$ if and only if $(1, \mathbf{x})^t \in Q$ between points of P and Q. Q is the **homogenization** of the polyhedron P as a polyhedral cone. The defining inequalities of Q imply $x_0 \geq 0$ for all $(x_0, \mathbf{x})$ in Q. By MINKOWSKI's Theorem for cones there is $M \in \mathbb{R}^{r\times(n+1)}$ such that

$$Q = \{M\boldsymbol{\eta} \mid \boldsymbol{\eta} \geq 0\}.$$

The columns of M are the generators of Q. We can reorder and scale these generators with a positive scalar without changing the cone Q. Hence, we can write M in the form

$$M = \begin{pmatrix} \mathbf{1}^t & \mathbf{0}^t \\ B & C \end{pmatrix}$$

for some $B \in \mathbb{R}^{n\times p}$, $C \in \mathbb{R}^{n\times q}$ with $p+q=r$. Split $\boldsymbol{\eta} = (\boldsymbol{\lambda}, \boldsymbol{\mu}) \in \mathbb{R}^p \times \mathbb{R}^q$ accordingly. Then

$$Q = \left\{ \begin{pmatrix} \mathbf{1}^t\boldsymbol{\lambda} \\ B\boldsymbol{\lambda} + C\boldsymbol{\mu} \end{pmatrix} \,\middle|\, \boldsymbol{\lambda}, \boldsymbol{\mu} \geq \mathbf{0} \right\} .$$

P is the subset of all points with first coordinate $x_0 = 1$, so $P = \operatorname{conv}(B) + \operatorname{cone}(C)$. □

Combining the affine versions of WEYL's and MINKOWSKI's Theorem we obtain a duality between the definition of polyhedra as solution sets of systems of linear inequalities and Minkowski sums of a convex and conic hull of two finite point sets.

Affine Minkowski-Weyl-*Duality*

Theorem 2.8 **(Affine MINKOWSKI-WEYL-Duality).** *A subset $P \subset \mathbb{R}^n$ is a polyhedron if and only if it is the* MINKOWSKI *sum of a polytope and a finitely generated cone.* □

Using this duality we can characterize all polyhedra that are polytopes.

Corollary 2.9. *$P \subseteq \mathbb{R}^n$ is a polytope if and only if P is a bounded polyhedron.*

Proof. Any polytope $P = \{\lambda^t A \mid \lambda \geq 0, \sum \lambda_i = 1\}$ is a polyhedron by the affine WEYL Theorem, and it is contained in the ball with radius $\sum \|a_i\|$ around any point in P. Conversely, any polyhedron can be written in the form $P = \{B\boldsymbol{\lambda} + C\boldsymbol{\mu} \mid \boldsymbol{\lambda}, \boldsymbol{\mu} \geq \mathbf{0},\ \sum \lambda_i = 1\}$. If it is bounded then $C = 0$, as otherwise $rC\boldsymbol{\lambda} \in P$ for any $r \geq 0$. □

Here are some examples to illustrate this theorem.

standard simplex

Example 2.10. (1) The 3-dimensional **standard simplex** is the polytope defined as the convex hull of the three unit vectors,

$$\Delta_3 := \operatorname{conv}\left(\begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} \right) = \left\{ \begin{pmatrix} x \\ y \\ z \end{pmatrix} \,\middle|\, x, y, z \geq 0,\ x + y + z = 1 \right\} .$$

(2) Here is an example of an unbounded polyhedron that is not a cone.

$$\begin{aligned} P_2 &= \left\{ \begin{pmatrix} x \\ y \end{pmatrix} \,\middle|\, \begin{pmatrix} -2 & -1 \\ -1 & -2 \\ 2 & -1 \\ -1 & 2 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} \leq \begin{pmatrix} -3 \\ -3 \\ -3 \\ -3 \end{pmatrix} \right\} \\ &= \operatorname{conv}\left(\begin{pmatrix} 3 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 3 \end{pmatrix}, \begin{pmatrix} 1 \\ 1 \end{pmatrix} \right) + \operatorname{cone}\left(\begin{pmatrix} 1 \\ 2 \end{pmatrix}, \begin{pmatrix} 2 \\ 1 \end{pmatrix} \right) . \end{aligned}$$

(3) We have already seen that the representation of a polyhedron as the set of solutions of a linear system of inequalities is not unique. The same is true for its representation as a Minkowski sum of a polytope and a cone.

$$\begin{aligned} P_3 &= \left\{ \begin{pmatrix} x \\ y \\ z \end{pmatrix} \,\middle|\, \begin{pmatrix} 1 & 1 & -1 \\ -1 & -1 & -1 \\ 0 & 0 & -1 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix} \leq \begin{pmatrix} 1 \\ 1 \\ -1 \end{pmatrix} \right\} \\ &= \operatorname{conv}\left(\begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}, \begin{pmatrix} -1 \\ -1 \\ 1 \end{pmatrix} \right) + \operatorname{cone}\left(\begin{pmatrix} 1 \\ -1 \\ 0 \end{pmatrix} \begin{pmatrix} -1 \\ 1 \\ 0 \end{pmatrix} \begin{pmatrix} 1 \\ 1 \\ 2 \end{pmatrix} \begin{pmatrix} -1 \\ -1 \\ 2 \end{pmatrix} \right) \\ &= \operatorname{conv}\left(\begin{pmatrix} 2 \\ 0 \\ 1 \end{pmatrix}, \begin{pmatrix} -2 \\ 0 \\ 1 \end{pmatrix} \right) + \operatorname{cone}\left(\begin{pmatrix} 1 \\ -1 \\ 0 \end{pmatrix} \begin{pmatrix} -1 \\ 1 \\ 0 \end{pmatrix} \begin{pmatrix} 1 \\ 1 \\ 2 \end{pmatrix} \begin{pmatrix} -1 \\ -1 \\ 2 \end{pmatrix} \right) \end{aligned}$$

(4) FOURIER-MOTZKIN elimination allows us to explicitely compute one representation of a polyhedron from the other. We give an explicit example. Let Id_3 be the (3×3)-identity matrix and

$$P := \operatorname{conv}(\mathrm{Id}_3) = \{\mathrm{Id}_3\boldsymbol{\lambda} \mid \boldsymbol{\lambda} \geq 0,\ \sum \lambda_{\mathrm{i}} = 1\} .$$

We want to compute a polyhedral description. As in the proof of the Affine WEYL Theorem, we extend the system to

$$\overline{P} := \left\{ \begin{pmatrix} \mathbf{1}^t \\ \mathrm{Id}_3 \end{pmatrix} \boldsymbol{\lambda} \mid \boldsymbol{\lambda} \geq 0 \right\} .$$

We apply FOURIER-MOTZKIN-elimination to the system

$$\begin{array}{rrr}
x_0 - \lambda_1 - \lambda_2 - \lambda_3 \leq 0 & -x_0 + \lambda_1 + \lambda_2 + \lambda_3 \leq 0 & \\
x_1 - \lambda_1 \leq 0 & -x_1 + \lambda_1 \leq 0 & -\lambda_1 \leq 0 \\
x_2 - \lambda_2 \leq 0 & -x_2 + \lambda_2 \leq 0 & -\lambda_2 \leq 0 \\
x_3 - \lambda_3 \leq 0 & -x_3 + \lambda_3 \leq 0 & -\lambda_2 \leq 0
\end{array}$$

Eliminating all λ_j from this we obtain

$$\begin{array}{rr}
x_0 - x_1 - x_2 - x_3 \leq 0 & -x_1 \leq 0 \\
-x_0 + x_1 + x_2 + x_3 \leq 0 & -x_2 \leq 0 \\
-x_0 \leq 0 & -x_3 \leq 0
\end{array}$$

so that $\overline{P} = \{\overline{\mathbf{x}} = \binom{x_0}{\mathbf{x}} \mid \overline{A}\overline{\mathbf{x}} \leq 0\}$ for

$$\overline{A} := \begin{pmatrix} 1 & -1 & -1 & -1 \\ -1 & 1 & 1 & 1 \\ -1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix}$$

Separating the first column we obtain $P = \{\mathbf{x} \mid x_i \geq 0, x_1 + x_2 + x_3 = 1\}$.

unit cube

(5) The 3-dimensional **unit cube** is

$$\begin{aligned} C_3 &:= \operatorname{conv}\left(\begin{pmatrix} 0 & 1 & 0 & 0 & 1 & 1 & 0 & 1 \\ 0 & 0 & 1 & 0 & 1 & 0 & 1 & 1 \\ 0 & 0 & 0 & 1 & 0 & 1 & 1 & 1 \end{pmatrix} \right) \\ &= \{(x_1, x_2, x_3)^t \mid 0 \leq x_i \leq 1,\ 1 \leq i \leq 3\} . \end{aligned}$$

◊

Summarizing our results we have two different descriptions of polyhedra,

(1) either as the solution set of a system of linear inequalities
(2) or as the MINKOWSKI sum of a polytope and a finitely generated cone.

exterior description
H-description
interior description
V-description

The first description is called the **exterior** or **H-description** of a polyhedron, the second is the **interior** or **V-description**.

Both are important for various problems in polyhedral geometry. When working with polyhedra it is often important to choose the right representation. There are many properties of a polyhedron that are almost trivial to compute in one representation, but very hard in the other. Here are two examples.

- ► Using the inequality description it is immediate that intersections of polyhedra are polyhedra, and
- ► using the interior description, it is straight forward (and we will do it in the next proposition) to show that affine images and MINKOWSKI sums of polyhedra are polyhedra.

affine map
affinely equivalent

A map $f : \mathbb{R}^n \longrightarrow \mathbb{R}^d$ is an **affine map** if there is a matrix $M \in \mathbb{R}^{d \times n}$ and a vector $\mathbf{b} \in \mathbb{R}^d$ such that $f(\mathbf{x}) = M\mathbf{x} + \mathbf{b}$. Two polytopes $P \in \mathbb{R}^n$ and $Q \in \mathbb{R}^d$ are **affinely equivalent** if there is an affine map $f : \mathbb{R}^n \to \mathbb{R}^d$ such that $f(P) = Q$. For any polyhedron $P \in \mathbb{R}^n$ with $\dim(P) = d$ there is an affinely equivalent polyhedron $Q \in \mathbb{R}^d$. The first part of the next proposition supports this definition.

Proposition 2.11. *(1) Affine images of polyhedra are polyhedra.*
(2) MINKOWSKI *sums of polyhedra are polyhedra.*

Proof. (1) Let $P := \{B\boldsymbol{\lambda} + C\boldsymbol{\mu} \mid \boldsymbol{\lambda}, \boldsymbol{\mu} \geq 0, \sum \lambda_i = 1\}$ for $B \in \mathbb{R}^{n\times p}$, $C \in \mathbb{R}^{n\times q}$ be a polyhedron and $f : \mathbf{x} \longmapsto M\mathbf{x} + \mathbf{t}$ with $M \in \mathbb{R}^{d\times n}$, $\mathbf{t} \in \mathbb{R}^d$ an affine map. Let T be the $(n \times r)$-matrix whose columns are copies of $\mathbf{t}$. Let

$$\overline{B} := MB + T \qquad \text{and} \qquad \overline{C} := MC\,.$$

Then

$$\begin{aligned} f(P) &= \left\{M(B\boldsymbol{\lambda} + C\boldsymbol{\mu}) + \mathbf{t} \mid \boldsymbol{\lambda}, \boldsymbol{\mu} \geq 0, \textstyle\sum \lambda_i = 1\right\} \\ &= \left\{\overline{B}\boldsymbol{\lambda} + \overline{C}\boldsymbol{\mu} \mid \boldsymbol{\lambda}, \boldsymbol{\mu} \geq 0, \textstyle\sum \lambda_i = 1\right\} \end{aligned}$$

which is again a polyhedron.

(2) Let $P = \operatorname{conv}(\mathbf{b}_1, \ldots, \mathbf{b}_r) + \operatorname{cone}(\mathbf{y}_1, \ldots, \mathbf{y}_s)$, $P' = \operatorname{conv}(\mathbf{b}'_1, \ldots, \mathbf{b}'_{r'}) + \operatorname{cone}(\mathbf{y}'_1, \ldots, \mathbf{y}'_{s'})$ be two polyhedra. We claim that their Minkowski sum is

$$P + P' = \operatorname{conv}(\mathbf{b}_i + \mathbf{b}'_j \mid 1 \leq i \leq r, 1 \leq j \leq r') + \operatorname{cone}(\mathbf{y}_i, \mathbf{y}'_j \mid 1 \leq i \leq s, 1 \leq j \leq s')\,.$$

We prove both inclusions for this equality. Let

$$\mathbf{p} := \left(\textstyle\sum_{i=1}^r \lambda_i \mathbf{b}_i + \sum_{j=1}^s \mu_j \mathbf{y}_j\right) + \left(\sum_{i=1}^{r'} \lambda'_i \mathbf{b}'_i + \sum_{j=1}^{s'} \mu'_j \mathbf{y}'_j\right) \in P + P'$$

with $\sum_{i=1}^r \lambda_i = \sum_{i=1}^{r'} \lambda'_i = 1$. Then

$$\mathbf{y} := \textstyle\sum_{j=1}^s \mu_j \mathbf{y}_j + \sum_{j=1}^{s'} \mu'_j \mathbf{y}'_j \in \operatorname{cone}(\mathbf{y}_i, \mathbf{y}'_j \mid 1 \leq i \leq s, 1 \leq j \leq s')\,,$$

so we have to show that

$$\mathbf{x} := \textstyle\sum_{i=1}^r \lambda_i \mathbf{b}_i + \sum_{j=1}^{s'} \mu'_j \mathbf{y}'_j \in \operatorname{conv}(\mathbf{b}_i + \mathbf{b}'_j \mid 1 \leq i \leq r, 1 \leq j \leq r') \tag{2.2}$$

We successively reduce this to a convex combination of $\mathbf{b}_i + \mathbf{b}'_j$ in the following way. We can assume that all coefficients in this sum are strictly positive (remove all other summands from the sum). Choose the smallest coefficient among $\lambda_1, \ldots, \lambda_r, \lambda'_1, \ldots, \lambda'_{r'}$. Without loss of generality let this be λ_1. Define $\lambda''_1 := \lambda'_1 - \lambda_1 \geq 0$, $\lambda''_i := \lambda'_i$, $2 \leq i \leq r'$, and $\eta_{11} := \lambda_1$. Then

$$\mathbf{x} - \eta_{11}(\mathbf{b}_1 + \mathbf{b}'_1) = \textstyle\sum_{i=2}^r \lambda_i \mathbf{b}_i + \sum_{i=1}^{r'} \lambda''_i \mathbf{b}'_i\,.$$

and $\sum_{i=2}^r \lambda_i = \sum_{i=1}^{r'} \lambda''_i$. The sum on the right hand side contains at least one summand less than the sum in (2.2), so repeating this procedure a finite number of times gives a representation of $\mathbf{x}$ as

$$\mathbf{x} = \sum_{\substack{1\leq i\leq r \\ 1\leq j\leq r'}} \eta_{ij}(\mathbf{b}_i + \mathbf{b}'_j)$$

for non-negative coefficients η_{ij} summing to 1. The other inclusion is obvious. □

Remark 2.12. More complicated examples require the use of a computer to assist the computations. One suitable option for this is the software tool `polymake`. It can (among other things) transform between V- and H-representations of polytopes (using some other software, either `cdd` or `lrs`). It uses the homogenization of a polyhedron for representation, so that a matrix of generators B should be given to `polymake` as $\mathbf{1}|B^t$, and inequalities $A\mathbf{x} \leq \mathbf{b}$ are written as $\mathbf{b}|-A$. `polymake` can also compute optimal solutions of linear programs. See `http://www.polymake.de` for more information. ◊

We will see in the next section that in Linear Programming we are given the exterior description of a polyhedron, while finding the optimum requires to find (at least some of) the extreme points. If we are also given the V-description then linear programming is a linear time algorithm (in the input size, which can be very large). So the complexity of linear programming is to some extend related to the complexity of finding an interior from the exterior description.

Given an exterior description, there is no polynomial bound on the number of points in an interior description, unless the dimension is fixed. A simple example for this behaviour is the cube

$$\mathrm{C}_n := \{\mathbf{x} \in \mathbb{R}^n \mid -1 \le x_i \le 1,\ 1 \le i \le n\}\,.$$

C_n needs $2n$ linear inequalities for its exterior description. For the interior description, it is the convex hull of the 2^n points with coordinates $+1$ and -1. Conversely, the cross polytope

$$\mathrm{Cr}_n := \{\mathbf{x} \in \mathbb{R}^n \mid \mathbf{e}^t\mathbf{x} \le 1, \mathbf{e} \in \{+1,-1\}^n\}\,.$$

needs 2^n inequalities, but it is the convex hull of the $2n$ points $\pm\mathbf{e}_i$, $1 \le i \le n$. We study this problem in more detail in Sections 4 and 5

characteristic cone
recession cone

Definition 2.13. *The* **characteristic** *or* **recession cone** *of a convex set $P \subseteq \mathbb{R}^n$ is the cone*

$$\mathrm{rec}(P) := \{\mathbf{y} \in \mathbb{R}^n \mid \mathbf{x} + \lambda\mathbf{y} \in P \text{ for all } \mathbf{x} \in P, \lambda \ge 0\}.$$

lineality space

The **lineality space** *of a polyhedron is the linear subspace*

$$\begin{aligned}\mathrm{lineal}(P) &:= \mathrm{rec}(P) \cap (-\,\mathrm{rec}(P)) \\ &= \{\mathbf{y} \in \mathbb{R}^n \mid \mathbf{x} + \lambda\mathbf{y} \in P \text{ for all } x \in P, \lambda \in \mathbb{R}\}.\end{aligned}$$

pointed polyhedron

A polyhedron P is **pointed** *if* $\mathrm{lineal}(P) = \{0\}$.

$\mathrm{lineal}(P)$ is a linear subspace of $\mathbb{R}^n$. Let W be a complementary subspace of $\mathrm{lineal}(P)$ in $\mathbb{R}^n$. Then

$$P = \mathrm{lineal}(P) + (P \cap W)\,. \tag{2.3}$$

as a Minkowski sum of a linear space and a convex set $P \cap W$ whose lineality space is $\mathrm{lineal}(P \cap W) = \{0\}$. So any polyhedron P is the Minkowski sum of a linear space and a pointed polyhedron. Note that this decomposition is different from the one used in the `Minkowski-Weyl` Duality.

Example 2.10 continued. Consider again the polyhedron P_3. We compute the lineality space L:

$$L := \mathrm{lineal}(P_3) = \mathrm{lin}\left(\begin{pmatrix} 1 \\ -1 \\ 0 \end{pmatrix}\right)$$

if we choose a transversal subspace

$$W := \left\{\begin{pmatrix} x \\ y \\ z \end{pmatrix} \,\middle|\, x = y\right\}$$

then

$$Q := P \cap W := \mathrm{conv}\left(\begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}, \begin{pmatrix} -1 \\ -1 \\ 1 \end{pmatrix}\right) + \mathrm{cone}\left(\begin{pmatrix} 1 \\ 1 \\ 10 \end{pmatrix}, \begin{pmatrix} -1 \\ 1 \\ 10 \end{pmatrix}\right)$$

and P splits as $P = L + Q$. ◊

Proposition 2.15. *Let $P = \{\mathbf{x} \mid A\mathbf{x} \le \mathbf{b}\} = \mathrm{conv}(V) + \mathrm{cone}(Y)$ be a polyhedron.*

(1) $\mathrm{rec}(P) = \{\mathbf{y} \mid A\mathbf{y} \le \mathbf{0}\} = \mathrm{cone}(Y)$.
(2) $\mathrm{lineal}(P) = \{\mathbf{y} \mid A\mathbf{y} = \mathbf{0}\}$.
(3) $P + \mathrm{rec}(P) = P$.
(4) *P polytope if and only if* $\mathrm{rec}(P) = \{\mathbf{0}\}$.

Proof. (1) Let $C := \{\mathbf{x} \mid A\mathbf{x} \le \mathbf{0}\}$. If $\mathbf{y} \in \mathbb{R}^n$ satisfies $A\mathbf{y} \le \mathbf{0}$ and $\mathbf{x} \in P$, then $A(\mathbf{x} + \lambda\mathbf{y}) = A\mathbf{x} + A\mathbf{y} \le \mathbf{b}$, so $C \subseteq \operatorname{rec}(P)$.

Conversely, assume $\mathbf{y} \in \operatorname{rec}(P)$ and there is a row $\mathbf{a}_k^t$ of A such that $s := \mathbf{a}_k^t\mathbf{y} > 0$. Let $r := \inf_{\mathbf{x}\in P}(b_k - \mathbf{a}_k^t\mathbf{x})$. Then $b_k - \mathbf{a}_k^t(\mathbf{x}+\mathbf{y}) = r - s < r$. This is a contradiction to the choice of r.

Now consider the second equality. Let $Q := \operatorname{conv}(V)$ and $D = \operatorname{cone}(Y)$. Clearly $D \subseteq \operatorname{rec}(P)$. Assume $\mathbf{y} \in \operatorname{rec}(P)$, $\mathbf{y} \notin D$. By the FARKAS Lemma, there is a functional $\mathbf{c}$ such that $\mathbf{c}^t Y \le 0$, but $\mathbf{c}^t\mathbf{y} > 0$. Choose any $\mathbf{p} \in Q$. By assumption, $\mathbf{p} + n\mathbf{y} \in P$ for $n \in \mathbb{N}$, so there are $\mathbf{p}_n \in Q$, $\mathbf{q}_n \in D$ such that $\mathbf{p} + n\mathbf{y} = \mathbf{p}_n + \mathbf{q}_n$. Q is bounded, so there is a constant $M > 0$ such that $\mathbf{c}^t\mathbf{x} \le M$ for all $\mathbf{x} \in Q$. Apply $\mathbf{c}^t$ to the sequence $\mathbf{p} + n\mathbf{y}$ to obtain

$$\mathbf{c}^t\mathbf{p} + n\mathbf{c}^t\mathbf{y} = \mathbf{c}^t\mathbf{p}_n + \mathbf{c}^t\mathbf{q}_n.$$

By construction, the right side of this equation is bounded above for all n, while the left side tends to infinity for $n \to \infty$. This is a contradiction, so $\operatorname{rec}(P) \subseteq D$.

(2) Follows immediately from (1).

(3) "$\subseteq$": $\mathbf{x} \in P, \mathbf{y} \in \operatorname{rec}(P)$, then $\mathbf{x} + \mathbf{y} \in P$ by definition.
"$\supseteq$": $\mathbf{0} \in \operatorname{rec}(P)$.

(4) P bounded if and only if $\operatorname{rec}(P) = \{\mathbf{0}\}$. ☐

Linear Programming and Duality 3

This section introduces linear programming as an optimization problem of a linear functional over a polyhedron. We explain standard terminology and conversions between different representations of a linear program, before we define dual linear programs and prove the Duality Theorem. This theorem will be an important tool for the study of faces of polyhedra in the next section.

linear programming
objective function

Linear programming is a technique for minimizing or maximizing a linear **objective function** over a set P defined by linear inequalities and equalities. More explicitly, let $A\in\mathbb{R}^{p\times n}$, $E\in\mathbb{R}^{q\times n}$, $\mathbf{b}\in\mathbb{R}^p$, $\mathbf{f}\in\mathbb{R}^q$, and $\mathbf{c}\in(\mathbb{R}^n)^*$. Then we want to find the maximum of $\mathbf{c}^t\mathbf{x}$ subject to the constraints

$$\begin{aligned} A\mathbf{x} &\le \mathbf{b}\\ E\mathbf{x} &= \mathbf{f}. \end{aligned}$$

This system defines a polyhedron $P\subseteq\mathbb{R}^n$. Techniques to explicitly and efficiently compute the maximal value $\mathbf{c}^t\mathbf{x}$ use algebraic manipulations on the representation of the polyhedron P to obtain a representation in which the maximal value can be directly read of from the system. We introduce two representations that we will use quite often and explain how to convert between them. We will first do the algebraic manipulations and later see what this means in geometrical terms.

standard form

***Definition 3.1**. A linear program in* **standard form** *is given by a matrix $A\in\mathbb{R}^{m\times n}$, a vector $\mathbf{b}\in\mathbb{R}^m$, and a cost vector $\mathbf{c}^t\in\mathbb{R}^n$ by*

$$\begin{array}{lll} \textit{maximize} & \mathbf{c}^t\mathbf{x} & \\ \textit{subject to} & A\mathbf{x}=\mathbf{b} & \\ & \mathbf{x}\ge\mathbf{0}. & \text{(non-negativity constraints)} \end{array}$$

canonical form

A program in **canonical form** *is given as*

$$\begin{array}{lll} \textit{maximize} & \mathbf{c}^t\mathbf{x} & \\ \textit{subject to} & A\mathbf{x}\le\mathbf{b} & \\ & \mathbf{x}\ge\mathbf{0}. & \text{(non-negativity constraints)} \end{array}$$

Note that the definition of a standard form is far from unique in the literature. To justify this definition we show that any linear program can be transformed into standard form (or into any other) without changing the solution set. Here is an example that should explain all necessary techniques. Consider the linear program

$$\text{maximize}\quad 2x_1+3x_2 \qquad \text{subject to}\quad \begin{aligned} 3x_1-x_2 &\ge 2\\ 4x_1+x_2 &\le 5\\ x_1 &\le 0 \end{aligned}$$

we can reverse the inequality signs and pass to equations using additional auxiliary variables x_3, x_4

$$\text{maximize}\quad 2x_1+3x_2 \qquad \text{subject to}\quad \begin{aligned} -3x_1+x_2+x_3 \qquad &= -2\\ 4x_1+x_2 \qquad +x_4 &= \ \ 5\\ x_1 &\le \ \ 0\\ x_3,x_4 &\ge \ \ 0 \end{aligned}$$

correct the variable constraints by substituting $y_1 := -x_1$

$$\begin{array}{lll} \text{maximize} & -2y_1 + 3x_2 & \text{subject to} \quad \begin{aligned} 3y_1 + x_2 + x_3 \quad\quad &= -2 \\ -4y_1 + x_2 \quad\quad + x_4 &= \ \ 5 \\ x_3, x_4, y_1 &\geq \ \ 0 \end{aligned} \end{array}$$

add constraints for x_2 by substituting $x_2 = y_2 - y_3$

$$\begin{array}{lll} \text{maximize} & -2y_1 + 3y_2 - 3y_3 & \text{subject to} \quad \begin{aligned} 3y_1 + y_2 - y_3 + x_3 \quad\quad &= -2 \\ -4y_1 + y_2 - y_3 \quad\quad + x_4 &= \ \ 5 \\ x_3, x_4, y_1, y_2, y_3 &\geq \ \ 0 \end{aligned} \end{array}$$

normalize by renaming $y_1 \to x_1$, $y_2 \to x_2$, $y_3 \to x_3$, $x_3 \to x_4$, $x_4 \to x_5$,

$$\begin{array}{lll} \text{maximize} & -2x_1 + 3x_2 - 3x_3 & \text{subject to} \quad \begin{aligned} 3x_1 + x_2 - x_3 + x_4 \quad\quad &= -2 \\ -4x_1 + x_2 - x_3 \quad\quad + x_5 &= \ \ 5 \\ x_1, x_1, x_3, x_4, x_5 &\geq \ \ 0 \end{aligned} \end{array}$$

If we want to write this in matrix form, A, b and c are given by

$$A = \begin{pmatrix} 3 & 1 & -1 & 1 & 0 \\ -4 & 1 & -1 & 0 & 1 \end{pmatrix} \qquad \mathbf{b} = \begin{pmatrix} -2 \\ 5 \end{pmatrix} \qquad \mathbf{c}^t = \begin{pmatrix} -2 & 3 & -3 & 0 & 0 \end{pmatrix}.$$

An optimal solution of this transformed system is given by $(-1, 1, 0, 0, 0)$, which we can transform back to an optimal solution $(1, 1)$ of the original system. Here is the general recipe:

$$\begin{array}{rcl} \text{minimize } \mathbf{c}^t\mathbf{x} & \longleftrightarrow & \text{maximize } -\mathbf{c}^t\mathbf{x}. \\ \mathbf{a}_i^t\mathbf{x} \geq b_i & \longleftrightarrow & -\mathbf{a}_i^t\mathbf{x} \leq -b_i \\ \mathbf{a}_i x = b_i & \longleftrightarrow & \mathbf{a}_i^t\mathbf{x} \leq b_i \text{ and } \mathbf{a}_i\mathbf{x} \geq b_i. \\ \mathbf{a}_i\mathbf{x} \leq b_i & \longleftrightarrow & \mathbf{a}_i^t\mathbf{x} + s_i = b_i \text{ and } s_i \geq 0. \\ x_i \in \mathbb{R} & \longleftrightarrow & x_i = x_i^+ - x_i^-,\ x_i^+, x_i^- \geq 0. \end{array}$$

slack variables

Definition 3.2. *The variables s_i introduced in rule* (4) *are called* **slack variables**.

feasible
feasible solution
infeasible
unbounded
optimal solution
optimal value

Definition 3.3. *A linear maximization program is*

1. **feasible**, *if there is $\mathbf{x} \in \mathbb{R}^n$ satisfying all constraints. $\mathbf{x}$ is then called a* **feasible solution**.
2. **infeasible** *if it is not feasible.*
3. **unbounded**, *if there is no $M \in \mathbb{R}$ such that $\mathbf{c}^t\mathbf{x} \leq M$ for all feasible $\mathbf{x} \in \mathbb{R}^n$.*

An **optimal solution** *is a feasible solution $\overline{\mathbf{x}}$ such that $\mathbf{c}^t\mathbf{x} \leq \mathbf{c}^t\overline{\mathbf{x}}$ for all feasible $\mathbf{x}$. Its value $\mathbf{c}^t\mathbf{x}$ is the* **optimal value** *of the program.*

Note that an optimal solution need not be unique. We will learn how to compute the maximal value of a linear program in Section 6. We can completely characterize the set of linear functionals that lead to a bounded linear program.

Proposition 3.4. *Let $A \in \mathbb{R}^{m \times n}$, $\mathbf{b} \in \mathbb{R}^m$, and $P := \{\mathbf{x} \in \mathbb{R}^n \mid A\mathbf{x} \leq \mathbf{b}\}$.*

1. *The linear program $\max(\mathbf{c}^t\mathbf{x} \mid \mathbf{x} \in P)$ is unbounded if and only if there is a $\mathbf{y} \in \operatorname{rec}(P)$ such that $\mathbf{c}^t\mathbf{y} > 0$.*
2. *Assume $P \neq \emptyset$. Then $\max(\mathbf{c}^t\mathbf{x} \mid \mathbf{x} \in P)$ is feasible bounded if and only if $\mathbf{c} \in \operatorname{rec}(P)^*$.*

Proof. (1) If $\max(\mathbf{c}^t\mathbf{x} \mid A\mathbf{x} \leq \mathbf{b})$ is unbounded, then $\min(\mathbf{z}^t\mathbf{b} \mid \mathbf{z}^tA = \mathbf{c}^t, \mathbf{z} \geq 0)$ is infeasible. Hence, there is no $\mathbf{z} \geq 0$ such that $\mathbf{z}^tA = \mathbf{c}^t$. The FARKAS Lemma (in the version of Proposition 1.15) gives us a vector $\mathbf{y} \leq 0$ such that $A\mathbf{y} \leq \mathbf{0}$ and $\mathbf{y}^t\mathbf{b} > 0$.

Now assume that there is such a vector $\mathbf{y}$. Let $\mathbf{x} \in P$, then $\mathbf{x} + \lambda\mathbf{y} \in P$ for all $\lambda \geq 0$. Hence, $\mathbf{c}^t(\mathbf{x} + \lambda\mathbf{y}) = \mathbf{c}^t\mathbf{x} + \lambda\mathbf{c}^t\mathbf{y}$ is unbounded.

(2) As $P \neq \emptyset$, the program is feasible. There is no $\mathbf{y} \in \operatorname{rec}(P)$ with $\mathbf{c}^t\mathbf{y} > 0$ if and only if $\mathbf{c} \in \operatorname{rec}(P)^*$ by the definition of the dual cone. □

Now we want to show that we can associate to any linear program another linear program, called the **dual linear program** whose feasible solutions provide some information about the original linear program. We start with some general observations and an example.

Let $A \in \mathbb{R}^{m\times n}$, $\mathbf{b} \in \mathbb{R}^m$, and $\mathbf{c} \in \mathbb{R}^n$. Consider the two linear programs in canonical form:

$$\max(\mathbf{c}^t\mathbf{x} \mid A\mathbf{x} \leq \mathbf{b},\, \mathbf{x} \geq \mathbf{0}) \tag{P}$$
$$\min(\mathbf{y}^t\mathbf{b} \mid \mathbf{y}^t A \geq \mathbf{c}^t,\, \mathbf{y} \geq \mathbf{0}) \tag{D}$$

Assume that the first linear program has a feasible solution $\mathbf{x}$, and the second linear program a feasible solution $\mathbf{y}$. Then

$$\mathbf{c}^t\mathbf{x} \leq \mathbf{y}^t A\mathbf{x} \leq \mathbf{y}^t\mathbf{b}\,,$$

where the first inequality holds, as $\mathbf{x} \geq 0$ and the second, as $\mathbf{y} \geq 0$. Thus, any feasible solution of (D) provides an upper bound for the value of (P). The best possible upper bound that we can construct this way is assumed by $\min(\mathbf{y}^t\mathbf{b} \mid \mathbf{y}^t A \geq \mathbf{c}^t, \mathbf{y} \geq \mathbf{0})$.

Example 3.5. Let

$$A := \begin{pmatrix} 8 & 6 \\ 2 & 6 \\ 3 & 5 \end{pmatrix} \qquad \mathbf{b} := \begin{pmatrix} 22 \\ 10 \\ 12 \end{pmatrix} \qquad \mathbf{c}^t := \begin{pmatrix} 2 & 3 \end{pmatrix}$$

and consider the linear program (P) and (D) as above. Then $\mathbf{y} = (1, 0, 0)$ is feasible for (D) and we obtain

$$2x_1 + 3x_2 \;\leq\; 8x_1 + 6x_2 \;\leq\; 22$$

so the optimum of (P) is at most 22. From the computation you can see that we overestimated the coefficients by a factor of at least 2, so a much better choice would be to take $\mathbf{y} = (\frac{1}{2}, 0, 0)$, which leads to

$$2x_1 + 3x_2 \leq 4x_1 + 3x_2 \;\leq\; 11\,.$$

So far, we have only used one of the inequalities. As long as we take non-negative scalars, we can also combine them. Choosing $\mathbf{y} = (\frac{1}{6}, \frac{1}{3}, 0)$ gives

$$2x_1 + 3x_2 = \frac{1}{6}(8x_1 + 6x_2 + 4x_1 + 12x_2) \leq \frac{1}{6}(22 + 20) = \frac{42}{6} = 7\,.$$

Hence, the optimum is at most 7. It is exactly 7, as $x_1 = 2, x_2 = 1$ is a feasible solution of (P). ◊

The program (D) is called the **dual linear program** for the linear program (P), which is then called the **primal linear program**. We have already proven the following proposition.

dual linear program
primal linear program

***Proposition 3.6* (Weak Duality Theorem).** *For each feasible solution* $\mathbf{y}$ *of* (D) *the value* $\mathbf{y}^t\mathbf{b}$ *provides an upper bound on the value of* (P)*, i.e. for each feasible solution* $\mathbf{x}$ *of* (P) *we have*

weak duality theorem

$$\mathbf{c}^t\mathbf{x} \leq \mathbf{y}^t\mathbf{b}.$$

In particular, if either of the programs is unbounded, then the other is infeasible. □

However, we can prove a much stronger relation between solutions of the primal and dual program.

duality theorem

Theorem 3.7 **(Duality Theorem).** *For the linear programs (P) and (D) exactly one of the following possibilities is true:*

(1) Both are feasible and their optimal values coincide.
(2) One is unbounded and the other is infeasible.
(3) Neither (P) nor (D) has a feasible solution.

A linear program can either be

(1) feasible and bounded (fb),
(2) infeasible (i), or
(3) feasible and unbounded (fu).

Hence, for the relation of (P) and (D) we a priori have 9 possibilities. Three of them are excluded by the weak duality theorem (wd), and another two are excluded by the duality theorem (d).

(D) (P)	(fb)	(i)	(fu)
(fb)	yes	(d)	(wd)
(i)	(d)	yes	yes
(fu)	(wd)	yes	(wd)

The four remaining cases can indeed occur.

(1) $\max(x \mid x \leq 1,\, x \geq 0)$ and $\min(y \mid y \geq 1,\, y \geq 0)$ are both bounded and feasible.
(2) $\max(x_1 + x_2 \mid -x_1 - 2x_2 \leq 1,\, x_1, x_2 \geq 0)$ has the solution $x = \mathbb{R}\mathbf{1}$, and the dual program $\min(y \mid -y \geq 1, -2y \geq 1,\, y \geq 0)$ is infeasible. These are the programs corresponding to

$$A = \begin{pmatrix} -1 & 2 \end{pmatrix} \qquad b = 1 \qquad \mathbf{c}^t = \begin{pmatrix} 1 & 1 \end{pmatrix}$$

(3) Consider the programs (P) and (D) for the input data

$$A := \begin{pmatrix} 1 & -1 \\ -1 & 1 \end{pmatrix} \qquad \mathbf{b} := \begin{pmatrix} 0 \\ -1 \end{pmatrix} \qquad \mathbf{c}^t := \begin{pmatrix} 1 & 1 \end{pmatrix}.$$

We can use the FARKAS Lemma to show that both programs are infeasible. Choose the following two functionals $\mathbf{u} = \mathbf{v} = \begin{pmatrix} 1 \\ 1 \end{pmatrix}$. Then

$$\begin{aligned} &\mathbf{u}^t A = 0 \qquad &&\mathbf{u}^t \mathbf{b} < 0 \qquad &&\mathbf{u} \geq 0 \\ &\mathbf{v}^t A \leq 0 \qquad &&\mathbf{v}^t \mathbf{c} > 0 \qquad &&\mathbf{v} \geq 0 \end{aligned}$$

Dual programs also exist for linear programs not in canonical form, and it is easy to generate it using the transformation rules between linear programs.

Proposition 3.8. *Let $A, B, C,$ and D be compatible matrices and $\mathbf{a}, \mathbf{b}, \mathbf{c}, \mathbf{d}$ corresponding vectors. Let a linear program*

$$\begin{aligned} \textit{maximize} \qquad & \mathbf{c}^t\mathbf{x} + \mathbf{d}^t\mathbf{y} \\ \textit{subject to} \qquad & A\mathbf{x} + B\mathbf{y} \leq \mathbf{a} \\ & C\mathbf{x} + D\mathbf{y} = \mathbf{b} \\ & \mathbf{x} \geq \mathbf{0} \end{aligned}$$

be given. Then its dual program is

$$\begin{aligned} \textit{minimize} \qquad & \mathbf{u}^t\mathbf{y} + \mathbf{v}^t\mathbf{b} \\ \textit{subject to} \qquad & \mathbf{u}^t A + \mathbf{v}^t C \geq \mathbf{c}^t \\ & \mathbf{u}^t B + \mathbf{v}^t D = \mathbf{d}^t \\ & \mathbf{u} \geq \mathbf{0} \end{aligned}$$

Proof. By our transformation rules we can write the primal as

$$\begin{array}{ll} \text{maximize} & \mathbf{c}^t\mathbf{x}+\mathbf{d}^t\mathbf{y}_1-\mathbf{d}^t\mathbf{v}_2 \\ \text{subject to} & A\mathbf{x}+B\mathbf{y}_1-B\mathbf{y}_2 \le \mathbf{a} \\ & C\mathbf{x}+D\mathbf{y}_1-D\mathbf{y}_2 \le \mathbf{b} \\ & -C\mathbf{x}-D\mathbf{y}_1+D\mathbf{y}_2 \le -\mathbf{b} \\ & \mathbf{x},\mathbf{y}_1,\mathbf{y}_2 \ge \mathbf{0} \end{array}$$

which translates to

$$\begin{array}{ll} \text{minimize} & \mathbf{u}^t\mathbf{a}+\mathbf{v}_1^t\mathbf{b}-\mathbf{v}_2^t\mathbf{b} \\ \text{subject to} & \mathbf{u}^tA+\mathbf{v}_1^tC-\mathbf{v}_2^tC \ge \mathbf{c}^t \\ & \mathbf{u}^tB+\mathbf{v}_1^tD-\mathbf{v}_2^tD \ge \mathbf{d}^t \\ & -\mathbf{u}^tB-\mathbf{v}_1^tD+\mathbf{v}_2^tD \ge \mathbf{d}^t \\ & \mathbf{u},\mathbf{v}_1,\mathbf{v}_2 \ge \mathbf{0} \end{array}$$

Set $\mathbf{v} := \mathbf{v}_1-\mathbf{v}_2$. and combine the second and third inequality to an equality. □

From this proposition we can derive a rule set that is quite convenient for quickly writing down the dual program. Let $A \in \mathbb{R}^{m\times n}$, $\mathbf{b} \in \mathbb{R}^m$, $\mathbf{c} \in \mathbb{R}^n$.

	primal		dual	
variables	$\mathbf{x} = (x_1,\dots,x_n)$		$\mathbf{y} = (y_1,\dots,y_m)$	
matrix	A		A^t	
right hand side	$\mathbf{b}$		$\mathbf{c}$	
objective function	max $\mathbf{c}^t\mathbf{x}$		min $\mathbf{y}^t\mathbf{b}$	
constraints	i-th constraint has	$\le$ $\ge$ $=$	$y_i \ge 0$ $y_i \le 0$ $y_i \in \mathbb{R}$	
	$x_j \ge 0$ $x_j \le 0$ $x_j \in \mathbb{R}$		j-th constraint has	$\ge$ $\le$ $=$

Observe that we have one-to-one correspondences

$$\begin{array}{rcl} \text{primal variables} & \Longleftrightarrow & \text{dual constraints} \\ \text{dual variables} & \Longleftrightarrow & \text{primal constraints} \end{array}$$

This fact will be used in the complementary slackness theorem at the end of this section. Now we finally prove the duality theorem.

Proof (Duality Theorem). By the considerations that we did after the statement of the theorem and the fact that **primal** and **dual** are interchangeable, it suffices to prove the following:

> If the linear program (P) is feasible and bounded, then also the linear program (D) is feasible and bounded with the same optimal value.

Assume that (P) has an optimal solution $\overline{\mathbf{x}}$. Let $\alpha := \mathbf{c}^t\overline{\mathbf{x}}$. Then the system

$$A\mathbf{x} \le \mathbf{b} \qquad \mathbf{c}^t\mathbf{x} \ge \alpha \qquad \mathbf{x} \ge \mathbf{0} \tag{*}$$

has a solution, but for any $\varepsilon > 0$, the system

$$A\mathbf{x} \leq \mathbf{b} \qquad \mathbf{c}^t\mathbf{x} \geq \alpha + \varepsilon \qquad \mathbf{x} \geq \mathbf{0} \tag{**}$$

has none. Consider the extended matrices

$$\overline{A} := \begin{pmatrix} -\mathbf{c}^t \\ A \end{pmatrix} \qquad \overline{\mathbf{b}}_\varepsilon := \begin{pmatrix} -\alpha - \varepsilon \\ \mathbf{b} \end{pmatrix} .$$

Then (**) is equivalent to $\overline{A}\mathbf{x} \leq \overline{\mathbf{b}}_\varepsilon$, and (*) is the special case $\varepsilon = 0$.
Fix $\varepsilon > 0$. We apply the FARKAS Lemma (in the variant of Proposition 1.15(3)) to obtain a non-negative vector $\overline{\mathbf{z}} = (z_0, \mathbf{z}) \geq 0$ such that

$$\overline{\mathbf{z}}^t\overline{A} \geq 0 \qquad \text{but} \quad \overline{\mathbf{z}}^t\overline{\mathbf{b}}_\varepsilon < 0 .$$

This implies

$$\mathbf{z}^t A \geq z_0\mathbf{c}^t \qquad \text{and} \quad \mathbf{z}^t\mathbf{b} < z_0(\alpha + \varepsilon) \qquad \mathbf{z} \geq 0, z_0 \geq 0 .$$

Further, applying the FARKAS Lemma for $\varepsilon = 0$, we see that there is no such $\overline{\mathbf{z}}$, hence our chosen $\overline{\mathbf{z}} = (z_0, \mathbf{z})$ must satisfy $\mathbf{z}^t\mathbf{b} \geq \mathbf{z}_0\alpha$ (otherwise $\overline{\mathbf{z}}$ would be a certificate that (*) has no solution!). So

$$z_0\alpha \leq \mathbf{z}^t\mathbf{b} < z_0(\alpha + \varepsilon) .$$

As $z_0 \geq 0$ this can only be true for $z_0 > 0$. Hence, for $\mathbf{y} := \frac{1}{z_0}\mathbf{z}$ we obtain

$$\mathbf{y}^t A \geq \mathbf{c}^t \qquad \mathbf{y}^t\mathbf{b} < \alpha + \varepsilon .$$

So $\mathbf{y}$ is a feasible solution of (D) of value less than $\alpha + \varepsilon$ for any chosen ε. By the weak duality theorem, however, the value is at least α. Hence, (D) is bounded, feasible and therefore has an optimal solution. Its value is between α and $\alpha + \varepsilon$ for any $\varepsilon > 0$, so

$$\mathbf{b}^t\mathbf{y} = \alpha . \qquad \square$$

We can further characterize the connections between primal and dual solutions. Let $\mathbf{s}$ and $\mathbf{r}$ be the slack vectors for the primal and dual program:

$$\begin{aligned} \mathbf{s} &:= \mathbf{b} - A\mathbf{x} & &(\text{i.e.} \quad A\mathbf{x} \leq \mathbf{b} \Leftrightarrow \mathbf{s} \geq 0) \\ \mathbf{r}^t &:= \mathbf{y}^t A - \mathbf{c}^t & &(\text{i.e.} \quad \mathbf{y}^t A \geq \mathbf{c}^t \Leftrightarrow \mathbf{r} \geq 0) \end{aligned}$$

Then

$$\mathbf{y}^t\mathbf{s} + \mathbf{r}^t\mathbf{x} = \mathbf{y}^t(\mathbf{b} - A\mathbf{x}) + (\mathbf{y}^t A - \mathbf{c}^t)\mathbf{x} = \mathbf{y}^t\mathbf{b} - \mathbf{c}^t\mathbf{x},$$

so

$$\mathbf{y}^t\mathbf{s} + \mathbf{r}^t\mathbf{x} = 0 \quad \Longleftrightarrow \quad \mathbf{y}^t\mathbf{b} = \mathbf{c}^t\mathbf{x}. \tag{Δ}$$

complementary slackness theorem

***Theorem 3.9* (complementary slackness).** *Let both programs (P) and (D) be feasible. Then feasible solutions $\overline{\mathbf{x}}$ and $\overline{\mathbf{y}}$ of (P) and (D) are both optimal if and only if*

(1) for each $i \in [m]$ one of s_i and $\overline{y}_i$ is zero, and
(2) for each $j \in [n]$ one of r_j and $\overline{x}_j$ is zero,

or, in a more compact way

$$\overline{\mathbf{y}}^t\mathbf{s} = 0 \qquad \textit{and} \qquad \mathbf{r}^t\overline{\mathbf{x}} = 0 .$$

Proof. The Duality Theorem states that $\overline{\mathbf{x}}$ and $\overline{\mathbf{y}}$ are optimal for their programs if and only if $\mathbf{c}^t\mathbf{x} = \mathbf{y}^t\mathbf{b}$. ($\Delta$) then implies that this is the case if and only if $\overline{\mathbf{y}}^t\mathbf{s} + \mathbf{r}^t\overline{\mathbf{x}} = 0$. By non-negativity of $\overline{\mathbf{y}}, \overline{\mathbf{x}}, \mathbf{s}, \mathbf{r}$ this happens if and only if the conditions in the theorem are satisfied. ☐

So if for some optimal solution $\overline{\mathbf{x}}$ some constraint is not satisfied with equality, then the corresponding dual variable is zero, and vice versa. We can rephrase this in the following useful way. Let $\overline{\mathbf{x}}$ be a feasible solution of

$$\max(\mathbf{c}^t\mathbf{x} \mid A\mathbf{x} \le \mathbf{b},\ \mathbf{x} \ge \mathbf{0}).$$

Then $\overline{\mathbf{x}}$ is optimal if and only if there is $\overline{\mathbf{y}}$ with

$$\overline{\mathbf{y}}^t A \ge \mathbf{c}^t \qquad\qquad \overline{\mathbf{y}} \ge \mathbf{0}$$

such that

$$\begin{aligned} \overline{x}_j > 0 \quad &\Longrightarrow \quad \overline{\mathbf{y}}^t\mathbf{a}^j = c_j \\ (A\overline{\mathbf{x}})_i < b_i \quad &\Longrightarrow \quad \overline{y}_i = 0\,, \end{aligned} \tag{3.1}$$

where $\mathbf{a}^j$ is the j-th column of A. So given some feasible solution $\overline{\mathbf{x}}$ we can set up the system (3.1) of linear equations and solve for $\overline{\mathbf{y}}$. If the solution exists and is unique, then $\overline{\mathbf{x}}$ and $\overline{\mathbf{y}}$ are optimal solutions of the primal and dual program. We will see later that a solution to (3.1) is always unique if $\overline{\mathbf{x}}$ is a **basic** feasible solution.

Example 3.10. (1) We consider the polytope $P := \{\mathbf{x} \mid A\mathbf{x} \le \mathbf{b}, \mathbf{x} \ge \mathbf{0}\} = \operatorname{conv}(V)$ with

$$A := \begin{pmatrix} 1 & 1 & 1 \\ -2 & 1 & 2 \end{pmatrix} \qquad \mathbf{b} := \begin{pmatrix} 5 \\ 2 \end{pmatrix} \qquad V := \begin{pmatrix} 0 & 0 & 0 & 1 & 2 & 5 \\ 0 & 0 & 2 & 4 & 0 & 0 \\ 0 & 1 & 0 & 0 & 3 & 0 \end{pmatrix}.$$

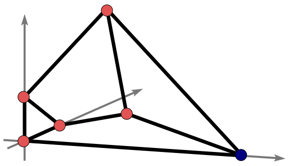


Figure 3.1

The polytope is shown in Figure 3.1. If we choose $\mathbf{c}^t := (1, -1, -2)$ as objective function, then the last column of V is the optimal solution $\overline{\mathbf{x}}$. This is the blue vertex in the figure. The corresponding dual optimal solution is $\overline{\mathbf{y}} = (1, 0)$. We compute the slack vectors

$$\mathbf{s} := \mathbf{b} - A\overline{\mathbf{x}} = \begin{pmatrix} 0 \\ 12 \end{pmatrix} \qquad\qquad \mathbf{r}^t := \overline{\mathbf{y}}^t A - \mathbf{c}^t = \begin{pmatrix} 0 & 2 & 3 \end{pmatrix}.$$

Then

$$\mathbf{y}^t\mathbf{s} + \mathbf{r}^t\mathbf{x} = \begin{pmatrix} 1 & 0 \end{pmatrix}\begin{pmatrix} 0 \\ 12 \end{pmatrix} + \begin{pmatrix} 0 & 2 & 3 \end{pmatrix}\begin{pmatrix} 5 \\ 0 \\ 0 \end{pmatrix} = 0.$$

(2) We consider the linear program $\max(\mathbf{c}^t\mathbf{x} \mid A\mathbf{x} \le \mathbf{b}, \mathbf{x} \ge \mathbf{0})$ with

$$A := \begin{pmatrix} 1 & 1 & 1 & 1 & 1 \\ -2 & 1 & 1 & -2 & -3 \\ -1 & 2 & 1 & 0 & -1 \end{pmatrix} \qquad \mathbf{b} := \begin{pmatrix} 5 \\ 2 \\ 4 \end{pmatrix}$$

$$\mathbf{c}^t := \begin{pmatrix} -1 & 3 & 1 & 1 & -1 \end{pmatrix}.$$

We are given the following potential optimal solution $\overline{\mathbf{x}}$ and compute the primal slack vector $\mathbf{s} := \mathbf{b} - A\overline{\mathbf{x}}$ for this solution:

$$\overline{\mathbf{x}} := \begin{pmatrix} 0 \\ 2 \\ 0 \\ 3 \\ 0 \end{pmatrix} \qquad\qquad \mathbf{s} := \begin{pmatrix} 0 \\ 6 \\ 0 \end{pmatrix}.$$

We want to use (3.1) to check that $\overline{\mathbf{x}}$ is indeed optimal. The system consists of three linear equations

$$\begin{pmatrix} y_1 & y_2 & y_3 \end{pmatrix}\begin{pmatrix} 1 \\ 1 \\ 2 \end{pmatrix} = 3\,, \qquad \begin{pmatrix} y_1 & y_2 & y_3 \end{pmatrix}\begin{pmatrix} 1 \\ -2 \\ 0 \end{pmatrix} = 1\,, \qquad \overline{y}_2 = 0\,.$$

This system has the unique solution $\overline{\mathbf{y}}^t = (1, 0, 1)$. This is a feasible solution of $\min(\mathbf{y}^t\mathbf{b} \mid \mathbf{y}^t A \ge \mathbf{c}^t, \mathbf{y} \ge \mathbf{0})$, and $\mathbf{c}^t\overline{\mathbf{x}} = 9 = \overline{\mathbf{y}}^t\mathbf{b}$, so both $\overline{\mathbf{x}}$ and $\overline{\mathbf{y}}$ are optimal for their programs. ◊

We want to discuss a geometrical interpretation of the duality theorem and complementary slackness. Let

$$\max(\mathbf{c}^t\mathbf{x} \mid A\mathbf{x} \le \mathbf{b}) \tag{P}$$
$$\min(\mathbf{y}^t\mathbf{b} \mid \mathbf{y}^t A = \mathbf{c}^t,\, \mathbf{y} \ge 0) \tag{D}$$

be a pair of dual programs for $A \in \mathbb{R}^{m\times n}$, $\mathbf{c} \in (\mathbb{R}^n)^*$, $\mathbf{b} \in \mathbb{R}^m$. The inequalities of the primal program define a polyhedron $P := \{\mathbf{x} \mid A\mathbf{x} \le \mathbf{b}\}$. Let $\overline{\mathbf{x}}$ and $\overline{\mathbf{y}}$ be optimal solutions of the primal and dual program. Complementary slackness tells us that

$$\overline{\mathbf{y}}^t(\mathbf{b} - A\overline{\mathbf{x}}) = \mathbf{0}\,. \tag{$\star$}$$

Hence, $\overline{\mathbf{y}}$ is non-zero only at those entries, at which $A\overline{\mathbf{x}} \le \mathbf{b}$ is tight. Let B be the set of row indices of A at which $\mathbf{a}_j\overline{\mathbf{x}} = b_j$. Let A_B the $(|B| \times n)$-sub-matrix of A spanned by these rows, and $\overline{\mathbf{y}}_B$ be the corresponding selection of entries of $\overline{\mathbf{y}}$. Note that by ($\star$) this contains all non-zero entries of $\overline{\mathbf{y}}$. The dual program states that $\mathbf{c}^t$ is contained in the cone spanned by the rows of A_B, and the dual solution $\overline{\mathbf{y}}$ gives a set of coefficients to represent $\mathbf{c}^t$ in this set of generators.

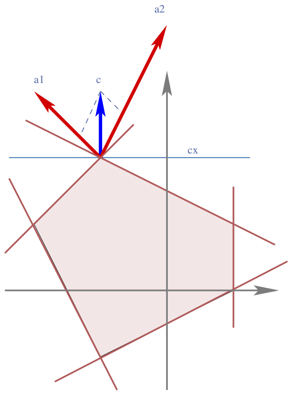


Figure 3.2

Example 3.11. Let A, $\mathbf{b}$, and $\mathbf{c}$ be given by

$$A := \begin{pmatrix} -1 & 1 \\ 1 & 2 \\ -2 & -1 \\ 1 & -2 \\ 1 & 0 \end{pmatrix} \qquad \mathbf{b} := \begin{pmatrix} 3 \\ 3 \\ 3 \\ 1 \\ 1 \end{pmatrix} \qquad \mathbf{c}^t := \begin{pmatrix} 0 & 1 \end{pmatrix}$$

Then

$$\overline{\mathbf{x}} := \begin{pmatrix} -1 \\ 2 \end{pmatrix} \qquad \overline{\mathbf{y}} := \begin{pmatrix} 1/3 & 1/3 & 0 & 0 & 0 \end{pmatrix}$$

are primal and dual solution. The primal solution is tight on the first two inequalities $-x_1 + x_2 \le 2$ and $x_1 + 2x_2 \le 4$. The corresponding functionals satisfy

$$1/3\begin{pmatrix} -1 & 1 \end{pmatrix} + 1/3\begin{pmatrix} 1 & 2 \end{pmatrix} = \begin{pmatrix} 0 & 1 \end{pmatrix}$$

See Figure 3.2 for an illustration. ◊

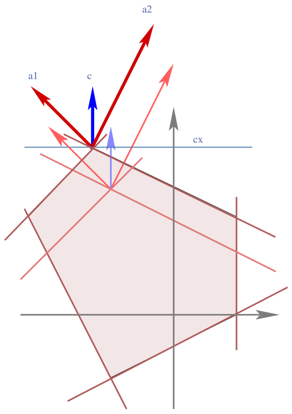


Figure 3.3

Using this geometric interpretation we can discuss the influence of small changes of $\mathbf{b}$ to the optimal value of the linear program. We assume for this that $m \ge n$, $|B| = n$ and $\operatorname{rank}(A_B) = n$. Hence, A_B is an invertible matrix and $\overline{\mathbf{x}} = A_B^{-1}\mathbf{b}_B$. Let $\Delta \in \mathbb{R}^m$ be the change in $\mathbf{b}$. If Δ is small enough, then the optimal solution of $A\mathbf{x} \le \mathbf{b} + \Delta$ will still be tight at the inequalities in B. So $\overline{\mathbf{x}}' = A_B^{-1}(\mathbf{b} + \Delta)_B$. However, using the duality theorem and complementary slackness we can compute the new optimal value without computing the new optimal solution $\overline{\mathbf{x}}'$. We have seen above that the non-zero coefficients of the dual solution are the coefficients of a representation of $\mathbf{c}^t$ in the rows of A_B. By our assumption, A_B stays the same, so the program with modified right hand side has the same dual solution $\overline{\mathbf{y}}^t$. By the duality theorem the new optimal value is $\overline{\mathbf{y}}^t(\mathbf{b} + \Delta)$. Further, changing right hand sides b_i that correspond to inequalities that are not tight for our optimal solution do not affect the optimal value (again, as long as the set B of tight inequalities stays the same). See Figure 3.3 for an example.

The crucial problem in these considerations is that we don't have a good criterion to decide whether a change Δ to $\mathbf{b}$ changes the set B or not. In Section 6 we will see that we can nevertheless efficiently exploit part this idea to compute optimal values of linear programs with variations in the right hand side.

Faces of Polyhedra 4

In this chapter we define faces of polyhedra and examine their relation to the interior and exterior description of a polyhedron. The main theorem of this section is a refined version of the MINKOWSKI-WEYL-Duality. We will explicitly characterize the necessary inequalities for the exterior and the necessary generators for the interior description. As intermediate results we obtain a characterization of the sets of optimal solutions of a linear program and of all linear functionals that lead to the same optimal solution.

We introduce some new notation to simplify the statements. Let $A \in \mathbb{R}^{m\times n}$ and $\mathbf{b} \in \mathbb{R}^m$. For subsets $I \subseteq [m]$ ($J \subseteq [n]$) we write A_{I*} (A_{*J}) for the matrix obtained from A by deleting all rows (columns) with index not in I (J). Similarly, $\mathbf{b}_I$ is the vector obtained from $\mathbf{b}$ by deleting all coordinates not in I. If $I = \{i\}$, then we write A_{i*} instead of $A_{\{i\}*}$, and b_i instead of $\mathbf{b}_{\{i\}}$.

Definition 4.1. *Let $A \in \mathbb{R}^{m\times n}$, $\mathbf{b} \in \mathbb{R}^m$ and $P := \{\mathbf{x} \mid A\mathbf{x} \le \mathbf{b}\}$. For any subset $F \subseteq P$ we define the* **equality set** *of F to be*

equality set

$$\mathrm{eq}(F) := \{i \in [m] \mid A_{i*}\mathbf{x} = b_i \text{ for all } \mathbf{x} \in F\}.$$

An inequality $A_{i*}\mathbf{x} = b_i$ is an **implicit equality** *if $i \in \mathrm{eq}(P)$. A point $\mathbf{x} \in P$ is an* **(relative) interior point** *of P if $A_{J*}\mathbf{x} < \mathbf{b}_J$ for $J := [m] - \mathrm{eq}(P)$. $\mathbf{x} \in P$ is a* **boundary point** *if it is not an interior point. The* **boundary** *∂P and the* **interior** *P° of P are the sets of all boundary and interior points, respectively.*

implicit equality
interior point
boundary point
boundary
interior

Lemma 4.2. *Any non-empty polyhedron $P := \{\mathbf{x} \mid A\mathbf{x} \le \mathbf{b}\}$ has an interior point.*

Proof. Let $I := \mathrm{eq}(P)$ and $J := [m] - I$. For any $j \in J$ there is some $\mathbf{x}_j \in P$ such that $A_{j*}\mathbf{x}_j < b_j$. Define $\mathbf{x} := \frac{1}{|J|}\sum_{j\in J} x_j$. Then $\mathbf{x}$ satisfies $A_{J*}\mathbf{x} < \mathbf{b}_J$. □

In particular, if $P \neq \varnothing$ is an affine space then $\mathrm{eq}(P) = [m]$ and any point $\mathbf{x} \in P$ is an interior point, i.e. $P = P^\circ$. In this case, $\partial P = \varnothing$. We will see later that this characterizes affine spaces.

Definition 4.3. *Points $\mathbf{a}_1, \ldots, \mathbf{a}_k \in \mathbb{R}^n$ are said to be* **affinely dependent** *if there are $\lambda_1, \ldots, \lambda_k$, $\sum \lambda_i = 0$, not all $\lambda_i = 0$, such that $\sum \lambda_i \mathbf{a}_i = 0$, and* **affinely independent** *otherwise.*

affinely (in)dependent

In other words, points $\mathbf{a}_1, \ldots, \mathbf{a}_k \in \mathbb{R}^n$ are *affinely* independent if and only if

$$\begin{pmatrix} 1 \\ \mathbf{a}_1 \end{pmatrix}, \ldots, \begin{pmatrix} 1 \\ \mathbf{a}_k \end{pmatrix} \in \mathbb{R}^{n+1}$$

are *linearly* independent. The dimension of the affine hull of k affinely independent points is $k - 1$.

Proposition 4.4. *Let $P := \{\mathbf{x} \mid A\mathbf{x} \le \mathbf{b}\}$ be a polyhedron and $J := \mathrm{eq}(P)$. Then*

$$\mathrm{aff}(P) = \{\mathbf{x} \mid A_{J*}\mathbf{x} = \mathbf{b}_J\}.$$

In particular, $\dim(P) = n - \mathrm{rank}(A_{J*})$.

Proof. Let $\mathbf{p}_1,\dots,\mathbf{p}_r \in P$ and $\lambda_1,\dots,\lambda_r$ with $\sum\lambda_i = 1$. Then $A_{J*}(\sum_{i=1}^r \lambda_i \mathbf{p}_i) = \sum_{i=1}^r \lambda_i A_{J*}\mathbf{p}_i = \mathbf{b}_J$. Hence, $\operatorname{aff}(P) \subseteq \{\mathbf{x} \mid A_{J*}\mathbf{x} = \mathbf{b}_J\}$.
Now suppose $\mathbf{z}$ satisfies $A_{J*}\mathbf{z} = \mathbf{b}_J$. If $\mathbf{z} \in P$ then there is nothing to prove, as $P \subset \operatorname{aff}(P)$. So assume $\mathbf{z} \notin P$. Pick an interior point $\mathbf{x} \in P$. Then the line through $\mathbf{p}$ and $\mathbf{z}$ contains at least one other point of P. Hence, the whole line is contained in the affine hull. □

full-dimensional

A polyhedron $P \subseteq \mathbb{R}^n$ is **full-dimensional** if $\dim(P) = n$. The previous proposition implies that this holds if and only if $\operatorname{eq}(P) = \varnothing$.

We will now show that we can define a finer combinatorial structure on the boundary points of a polyhedron by intersecting the boundary of P with certain affine hyperplanes.

Definition 4.5. *Let $P := \{\mathbf{x} \mid A\mathbf{x} \le \mathbf{b}\} \subseteq \mathbb{R}^n$ be a polyhedron and $\mathbf{c}^t \in \mathbb{R}^n$, $\delta \in \mathbb{R}$. The hyperplane $H_{\mathbf{c},\delta} := \{\mathbf{x} \mid \mathbf{c}^t\mathbf{x} = \delta\}$ is a*

valid hyperplane
supporting hyperplane

1. **valid hyperplane** *if $\mathbf{c}^t\mathbf{x} \le \delta$ for all $\mathbf{x} \in P$, and*
2. *a* **supporting hyperplane** *if additionally $\mathbf{c}^t\mathbf{x} = \delta$ for at least one $\mathbf{x} \in P$.*

face of a polyhedron
proper face

$F \subseteq P$ is a **face** *of P if either $F = P$ or $F = P \cap H$ for a valid hyperplane. If $F \ne P$ then F is a* **proper face***.*

dimension

Different hyperplanes may define the same face. Faces are the intersection of P with an affine space, hence faces are polyhedra themselves. The **dimension** of a face F of P is its dimension as a polyhedron.

Proposition 4.6. *Let P be a polyhedron. The set F of optimal solutions of a linear program $\max(\mathbf{c}^t\mathbf{x} \mid \mathbf{x} \in P)$ for some $\mathbf{c}^t \in (\mathbb{R}^n)^*$ is a face of P.*

Proof. Let F be the set of all optimal solutions. By definition, any optimal solution $\overline{\mathbf{x}} \in F$ of the linear program satisfies $\mathbf{c}^t\mathbf{x} \le \mathbf{c}^t\overline{\mathbf{x}}$ for all $\mathbf{x} \in P$. Hence, the hyperplane is valid. □

We denote the face defined by the previous proposition with $\operatorname{face}_{\mathbf{c}}(P)$. Note that $\operatorname{face}_{\mathbf{0}}(P) = P$, and $\operatorname{face}_{\mathbf{c}}(P) = \varnothing$ if the linear programming problem is unbounded for $\mathbf{c}^t$. In the last chapter we have seen that the objective function is in the cone defined by those inequalities that are tight at the optimal solution. We will see that this gives another representation for faces of a polyhedron.

Proposition 4.7. *Let $P := \{\mathbf{x} \mid A\mathbf{x} \le \mathbf{b}\}$ be a polyhedron and $I \subseteq [m]$. Then $F := \{\mathbf{x} \in P \mid A_{I*}\mathbf{x} = \mathbf{b}_I\}$ is a face of P.*

Proof. If $I = \varnothing$ then $F = P$ is a face. So assume $I \ne \varnothing$, and define $\mathbf{c}^t := \sum_{i \in I} A_{i*}$ and $\delta := \sum_{i \in I} b_i$. For any $\mathbf{x} \notin F$ at least one of the inequalities in $A_{I*}\mathbf{x} \le \mathbf{b}_I$ is strict, and thus

$$\mathbf{c}^t\mathbf{x} \quad \begin{cases} = \delta & \text{if } \mathbf{x} \in F \\ < \delta & \text{otherwise .} \end{cases}$$

Hence, $H := \{\mathbf{x} \mid \mathbf{c}^t\mathbf{x} = \delta\}$ is a valid hyperplane and $F = P \cap H$. □

Hence, any subset $I \subset [m]$ defines a face $\operatorname{face}_I(P) := \{\mathbf{x} \in P \mid A_{I*}\mathbf{x} = \mathbf{b}_I\}$ of P. The argument in the proof shows that more generally *any* conic combination of the rows of A_{I*} with strictly positive coefficients defines the same face $\operatorname{face}_I(P)$.

Example 4.8. Consider the polyhedron $P := \{\mathbf{x} \mid A\mathbf{x} \le \mathbf{b}\} = \operatorname{conv}(V)$ with

$$A := \begin{pmatrix} -1 & 0 \\ 0 & -1 \\ 1 & -1 \\ 1 & 0 \\ -1 & 2 \end{pmatrix} \qquad b := \begin{pmatrix} 0 \\ 0 \\ 1 \\ 2 \\ 2 \end{pmatrix} \qquad V := \begin{pmatrix} 0 & 1 & 2 & 2 & 0 \\ 0 & 0 & 1 & 2 & 1 \end{pmatrix}$$

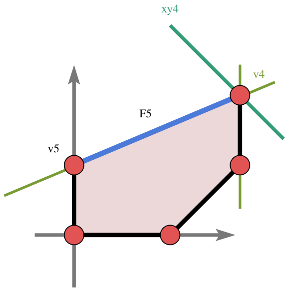


Figure 4.1

Then

$$F_5 := \{(x,y) \in P \mid -x+2y=2\}$$
$$v_4 := \{(x,y) \in P \mid x-y=4\} \qquad = \{(x,y) \in P \mid x=2, -x+2y=2\}\,.$$

$$\operatorname{face}_{\{1,5\}}(P) = \begin{pmatrix} 0 \\ 1 \end{pmatrix} = v_5 \qquad \operatorname{eq}\left(\begin{pmatrix} 2 \\ 2 \end{pmatrix}\right) = \{4,5\} = v_4\,. \qquad \Diamond$$

Theorem 4.9. *Let $P := \{\mathbf{x} \mid A\mathbf{x} \le \mathbf{b}\}$ be a polyhedron and $\mathbf{c}^t \in (\mathbb{R}^n)^*$, $\delta \in \mathbb{R}$. If*

$$F \;:=\; \{\mathbf{x} \mid \mathbf{c}^t\mathbf{x} = \delta\} \cap P$$

is a non-empty face of P then there is $I \subseteq [m]$ and $\boldsymbol{\lambda} \in \mathbb{R}^{|I|}$, $\boldsymbol{\lambda} \ge \mathbf{0}$ such that

$$F = \operatorname{face}_I(P) \qquad \boldsymbol{\lambda}^t A_{I*} = \mathbf{c}^t \qquad \boldsymbol{\lambda}^t \mathbf{b}_I = \delta\,.$$

Proof. $H := \{\mathbf{x} \mid \mathbf{c}^t\mathbf{x} = \delta\}$ is a supporting hyperplane of P. Hence, the linear programming problem $\max(\mathbf{c}^t\mathbf{x} \mid A\mathbf{x} \le \mathbf{b})$ is bounded and feasible, and F is the set of optimal solutions. By duality

$$\max(\mathbf{c}^t\mathbf{x} \mid A\mathbf{x} \le \mathbf{b}) = \min(\mathbf{y}^t\mathbf{b} \mid \mathbf{y}^t A = \mathbf{c}^t, \mathbf{y} \ge \mathbf{0})\,.$$

Let $\overline{\mathbf{y}}$ be an optimal solution of the dual program, and $I := \{i \in [m] \mid \overline{y}_i > 0\}$. The complementary slackness theorem implies

$$\begin{aligned} \mathbf{x} \in F \quad &\Leftrightarrow \quad \mathbf{c}^t\mathbf{x} - \delta = 0 \text{ and } \mathbf{x} \in P \quad \Leftrightarrow \quad \overline{\mathbf{y}}^t(A\mathbf{x} - \mathbf{b}) = 0 \text{ and } \mathbf{x} \in P \\ &\Leftrightarrow \quad \overline{\mathbf{y}}_I^t(A_{I*}\mathbf{x} - \mathbf{b}_I) = 0 \text{ and } \mathbf{x} \in P \quad \Leftrightarrow \quad A_{I*}\mathbf{x} = \mathbf{b}_I \text{ and } \mathbf{x} \in P\,, \end{aligned}$$

where the third equivalence follows from $\overline{y}_i = 0$ for $i \notin I$, and the forth from $y_i > 0$ for $i \in I$. Let $\boldsymbol{\lambda} = \overline{\mathbf{y}}_I$. Then $\boldsymbol{\lambda}^t A_{I*} = \mathbf{c}^t$, $\boldsymbol{\lambda}^t \mathbf{b}_I = \delta$, and $F = \operatorname{face}_I(P)$. □

We collect some important consequences of this theorem.

Corollary 4.10. *Let $P = \{\mathbf{x} \in \mathbb{R}^n \mid A\mathbf{x} \le \mathbf{b}\}$ for $A \in \mathbb{R}^{m\times n}$, $\mathbf{b} \in \mathbb{R}^m$ and $F \subseteq P$ non-empty. Then F is a face of P if and only if $F = \operatorname{face}_I(P)$ for some $I \subseteq [m]$.* □

Corollary 4.11. *Let P be a polyhedron and F a face.*

(1) $\operatorname{face}_{\operatorname{eq}(F)}(P) = F$ for any proper face $F \neq \emptyset$.
(2) If $G \subseteq F$, then G is a face of F if and only if it is a face of P.
(3) $\dim(F) \le \dim(P) - 1$ for any proper face F of P.
(4) P has at most $2^m + 1$ faces.

Proof. The first three are trivial, for the forth observe that there are 2^m subsets of $[m]$ that can possibly define a non-empty face, and $F = \emptyset$. □

Definition 4.12. *Let F be a proper face of a polyhedron P. The* **normal cone** *of F is*

normal cone

$$\mathcal{N}_F = \left\{\mathbf{c}^t \mid F \subseteq \operatorname{face}_{\mathbf{c}}(P)\right\}\,.$$

Let $F = \operatorname{face}_I(P)$. By Theorem 4.9, $\mathcal{N}_F$ is finitely generated by the rows of A_{I*}. If F, G are faces of P, and $F \subseteq G$, then $\mathcal{N}_G \subseteq \mathcal{N}_F$. $\mathcal{N}_P$ is the span of all linear functionals in $\operatorname{eq}(P)$. Hence, if P is full dimensional then $\mathcal{N}_P = \{\mathbf{0}\}$. The collection of all normal cones is an example for a more general structure characterized in the following definition.

Definition 4.13. *A* **fan** *in $\mathbb{R}^n$ is a finite collection $\mathcal{F} = \{C_1, \ldots, C_r\}$ of non-empty polyhedral cones such that*

fan

(1) Every non-empty face of a cone in $\mathcal{F}$ is in $\mathcal{F}$, and

(2) the intersection of any two cones in $\mathscr{F}$ is a face of both.

complete fan
pointed fan

A fan is **complete** *if* $\bigcup_{C \in \mathscr{F}} C = \mathbb{R}^n$*. It is* **pointed** *if* $\{\mathbf{0}\}$ *is a cone in* $\mathscr{F}$*.*

normal fan

The normal cones $\mathscr{N}_F$ of all proper faces F of P satisfy these two conditions, hence, they form a fan, the **normal fan** $\mathscr{F}_P$ of P. It is pointed if and only if P is full dimensional. Proposition 3.4 states that every linear functional defines a bounded feasible linear program if and only if P is not empty and $\operatorname{rec}(P) = \{\mathbf{0}\}$. Hence, the normal fan of P is complete if and only if P is a polytope. See Figure 4.2 for the normal fan a triangle. The normal cones of the three edges are the generators of the normal cones.

(ir)redundant constraint
(ir)redundant system

Definition 4.14. *Let $P := \{\mathbf{x} \mid A\mathbf{x} \leq \mathbf{b}\}$ be a polyhedron and $i \in [m]$. $A_{i*}\mathbf{x} \leq b_i$ is a* **redundant constraint***, if it can be removed from the system without changing P, and* **irredundant** *otherwise. If all constraints in $A\mathbf{x} \leq \mathbf{b}$ are irredundant, then the system is* **irredundant***, and* **redundant** *otherwise.*

Observe that redundancy is a property of the inequality system $A\mathbf{x} \leq \mathbf{b}$, not of the polyhedron. Redundant inequalities may become irredundant if some other redundant inequality is removed from the system. Clearly, any system $A\mathbf{x} \leq \mathbf{b}$ can be made irredundant by succesively removing redundant inequalities.

Example 4.15. Consider $P := \left\{ \begin{pmatrix} x \\ y \\ z \end{pmatrix} \,\middle|\, x, y, z \geq 0, x + y + 2z \geq 0, x + y + z = 1 \right\}$.
Both $z \geq 0$ and $x + y + 2z \geq 1$ are redundant, but removing one makes the other irredundant. See Figure 4.3. ◊

facet

Definition 4.16. *A proper non-empty face of P is a* **facet** *if it is not strictly contained in any other proper face.*

Example 4.17. The facets in Example 4.15 are given by the inequalities $x \geq 0$, $y \geq 0$, and $z \geq 0$.
Consider $P := \operatorname{conv}\left(\binom{0}{0}, \binom{1}{0}, \binom{0}{1}\right)$. This is a triangle with three facets

$$\operatorname{conv}\left(\binom{0}{0}, \binom{1}{0}\right) \qquad \operatorname{conv}\left(\binom{0}{0}, \binom{0}{1}\right) \qquad \operatorname{conv}\left(\binom{1}{0}, \binom{0}{1}\right)$$ ◊

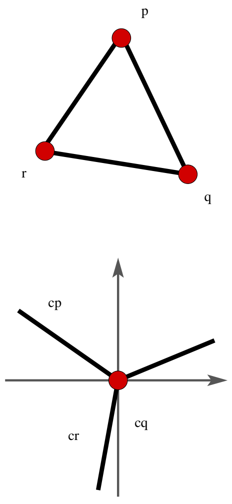


Figure 4.2

Theorem 4.18. *Let $P := \{\mathbf{x} \mid A\mathbf{x} \leq \mathbf{b}\}$ and $E := \operatorname{eq}(P)$, $I := [m] - E$, and assume that $A_{I*}\mathbf{x} \leq \mathbf{b}_I$ is irredundant. Then F is a facet of P if and only if $F = \operatorname{face}_i(P)$ for some $i \in I$.*
The number f of facets of P satisfies $f \leq |I|$ $(\leq m)$ with equality if and only if $A_{I*}\mathbf{x} \leq \mathbf{b}_I$ is irredundant.

Proof. "⇒": Let $F := \operatorname{face}_J(P)$ be a facet for some $J \subseteq I$ and choose $j \in J$. Then $j \notin E$, so $A_{j*}\mathbf{x} \leq b_j$ is not an implicit equality. Hence, $F' := \operatorname{face}_j(P)$ is a proper face of P, and $F \subseteq F'$. But F is maximal by assumption, so $F = F'$.
"⇐": Let $i \in I$ and $I' := I - \{i\}$. Let $\mathbf{x}$ be an interior point of P. Then $\mathbf{x}$ satisfies

$$A_{I*}\mathbf{x} < \mathbf{b}_I \qquad \text{(and} \quad A_{E*}\mathbf{x} = \mathbf{b}_E\ ) .$$

But $\mathbf{x} \notin F$, so F is a proper face. By assumption, the system $A_{I*}\mathbf{x} \leq \mathbf{b}_I$ is irredundant, so there is $\mathbf{y} \in \operatorname{aff}(P)$ that satisfies

$$A_{I'*}\mathbf{y} \leq \mathbf{b}_{I'} \qquad\qquad A_{i*}\mathbf{y} > b_i .$$

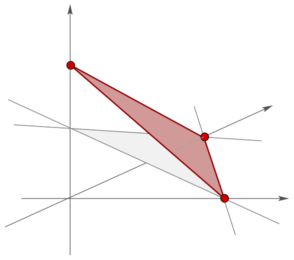


Figure 4.3

Hence, we can find $\lambda \in \mathbb{R}$, $0 < \lambda < 1$ such that $\mathbf{z} := \lambda\mathbf{x} + (1 - \lambda)\mathbf{y}$ satisfies

$$A_{i*}\mathbf{z} = b_i, \qquad A_{I'*}\mathbf{z} < \mathbf{b}_{I'} \qquad A_{E*}\mathbf{z} = \mathbf{b}_E .$$

See Figure 4.4. This implies that $\mathbf{z} \in F$, but not in any other face of P. Hence, F is a facet. □

We note some consequences of this theorem.

Corollary 4.19. *The dimension of a facet F is* $\dim(F) = \dim(P) - 1$.

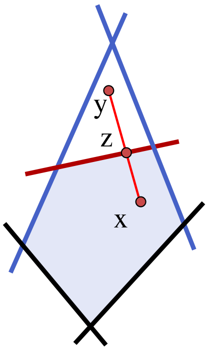


Figure 4.4

Proof. Let $P = \{\mathbf{x} \mid A\mathbf{x} \leq \mathbf{b}\}$for an irredundant system $A\mathbf{x} \leq \mathbf{b}$, and $J := \operatorname{eq}(P)$. Then $F = \operatorname{face}_i(P)$ for some $i \in [m] - J$. Let $J' := J + \{i\}$. Then $J' = \operatorname{eq}(F)$ and

$$\dim(F) = n - \operatorname{rank}(A_{J'*}) = n - (1 + \operatorname{rank}(A_{J*})) \qquad \square$$

Corollary 4.20. *If $P := \{\mathbf{x} \in \mathbb{R}^n \mid A\mathbf{x} \leq \mathbf{b}\} \subset \mathbb{R}^n$ satisfies $\dim(P) = n$ and $A\mathbf{x} \leq \mathbf{b}$ is irredundant, then $A\mathbf{x} \leq \mathbf{b}$ is unique up to scaling some of the inequalities with a positive scalar.* $\square$

The assumption that set $J = \operatorname{eq}(P) = \varnothing$, i.e. that P is full-dimensional is essential for the corollary. Otherwise we can add linear combinations of the implicit equalities to any inequality without changing the set of solutions. However, if we require that for $I := [m] - J$ and any $j \in J$, $i \in I$ the rows A_{i*} and A_{j*} are orthogonal, then the set A_{I*} is again unique up to scaling. This follows essentially from the fact that we can decompose a polyhedron into a Minkowski sum of a pointed polyhedron and a linear space (see (2.3)).

Corollary 4.21. *(1) Any proper non-empty face is the intersection of some facets.*
(2) A polyhedron has no proper faces if and only if it is an affine space.

Proof. (1) Any proper non-empty face is of the form $F = \operatorname{face}_I(P)$ for some $I \subseteq [m]$, $I \neq \varnothing$. Then $F = \bigcap_{i \in I} \operatorname{face}_i(P)$.
(2) $P = \{\mathbf{x} \mid A\mathbf{x} \leq \mathbf{b}\}$ for some $A \in \mathbb{R}^{m \times n}$ and $\mathbf{b} \in \mathbb{R}^m$ is an affine space if and only if $\operatorname{eq}(P) = [m]$. $\square$

We have seen that a system is irredundant if and only if each inequality corresponds to a facet of the polyhedron. The following proposition gives an easier check for this.

Proposition 4.22. *Let $P := \{\mathbf{x} \mid A\mathbf{x} \leq \mathbf{b}\}$ for $A \in \mathbb{R}^{m \times n}$, $\mathbf{b} \in \mathbb{R}^m$ be a polyhedron, and $J := \operatorname{eq}(P)$, $I := [m] - J$. The system $A_{I*}\mathbf{x} \leq \mathbf{b}$ is irredundant if and only if for any $i, j \in I$, $i \neq j$ there is a point $\mathbf{x} \in P$ such that*

$$A\mathbf{x} \leq \mathbf{b} \qquad A_{i*}\mathbf{x} = b_i \qquad A_{j*}\mathbf{x} < b_2\,.$$

Proof. If the system is irredundant, then we can choose $\mathbf{x} \in \operatorname{face}_i(P) \setminus \operatorname{face}_j(P)$. Conversely, assume that $A_i\mathbf{x} \leq b_i$ is a redundant inequality. Then $\operatorname{face}_i(P)$ is contained in a facet $\operatorname{face}_j(P)$ for some $j \in I$. Hence,

$$\{\mathbf{x} \mid A\mathbf{x} \leq \mathbf{b}, A_{i*}\mathbf{x} = b_i\} \subseteq \{\mathbf{x} \mid A\mathbf{x} \leq \mathbf{b}, A_{j*}\mathbf{x} = b_j\}\,.$$

This implies that for i, j there is no $\mathbf{x}$ satisfying the assumption of the proposition. $\square$

Faces of P can be ordered by inclusion. This gives a partially ordered set, the **face poset** $\mathscr{L}_P$. It completely captures the combinatorial properties of P. For a polytope P this is even an Eulerian lattice, the **face lattice** of P. Two polyhedra P and Q are **combinatorially equivalent** if their face lattices are isomorphic as posets. Note that this equivalence relation is much weaker than the affine equivalence defined in Section 2.

face poset
face lattice
combinatorially equivalent

Theorem 4.18 completely characterizes the facets of a polyhedron: They correspond to the constraints in the exterior description of that polyhedron, and we need precisely one inequality for each facet of P. With the next theorems we want to study the *minimal* faces of a polyhedron. Similar to the relation of facets to the exterior description of a polyhedron we will construct a close connection to the interior description of P.

Definition 4.23. *A face of a polyhedron P is* **minimal** *if there is no non-empty face G of P with $G \neq F$ but $G \subset F$.*

minimal face

Theorem 4.24. *Let $P = \{\mathbf{x} \mid A\mathbf{x} \le \mathbf{b}\}$ be a polyhedron and $L := \operatorname{lineal}(P)$.*

(1) $F = \operatorname{face}_I(P)$ for some $I \subseteq [m]$ is minimal if and only if

$$F = \{\mathbf{x} \in \mathbb{R}^n \mid A_{I*}\mathbf{x} = \mathbf{b}_I\} \quad (= \operatorname{aff}(F)).$$

(2) Any minimal face of P is a translate of L.

Proof. (1) A face of P has no proper faces if and only if it is an affine space. Hence, if $F = \{\mathbf{x} \in \mathbb{R}^n \mid A_{I*}\mathbf{x} = \mathbf{b}_I\}$, then it is minimal. Conversely, let F be minimal. We can write F as

$$F = \{\mathbf{x} \mid A_{I*}\mathbf{x} = \mathbf{b}_I,\ A_{J*}\mathbf{x} \le \mathbf{b}_J\}$$

for some $I, J \subset [m]$. We may assume that all inequalities in J are irredundant. But then $J = \varnothing$, as F has no proper faces.

(2) Let $F = \{\mathbf{x} \in \mathbb{R}^n \mid A_{I*}\mathbf{x} = \mathbf{b}_I\}$ be a minimal face. It suffices to prove that the lineality space of P and F coincide. Clearly

$$\{\mathbf{x} \in \mathbb{R}^n \mid A_{I*}\mathbf{x} = \mathbf{0}\} \supseteq \{\mathbf{x} \in \mathbb{R}^n \mid A\mathbf{x} = \mathbf{0}\} = \operatorname{lineal}(P)\,,$$

with equality if $\operatorname{rank}(A_{I*}) = \operatorname{rank}(A)$. Assume that the ranks differ. Then we can find $j \in [m] - I$ such that A_{j*} is not in the span of A_{I*}. But this means

$$\{\mathbf{x} \mid A_{I*}\mathbf{x} = \mathbf{b}_I, A_{j*}\mathbf{x} \le b_j\} \subsetneq \{\mathbf{x} \mid A_{I*}\mathbf{x} = \mathbf{b}_I\}\,,$$

so $\operatorname{face}_{I \cup \{j\}}(P)$ would be a proper face of F. □

Corollary 4.25. *Let $C := \{\mathbf{x} \mid A\mathbf{x} \le \mathbf{0}\}$ be a cone. Then $L := \operatorname{lineal}(C)$ is the unique minimal face of C.* □

Let $P = \{\mathbf{x} \mid A\mathbf{x} \le \mathbf{b}\}$ and

$$n_0 := \dim \operatorname{lineal} P = n - \operatorname{rank}(A)\,.$$

vertex

By Theorem 4.24 each minimal face of P has dimension n_0. In particular, all minimal faces have the same dimension. Minimal faces of a *pointed* polyhedron P are called **vertices**. They are 0-dimensional.

Example 4.26. (1) The polytope

$$P := \left\{ \begin{pmatrix} x \\ y \\ z \end{pmatrix} \,\middle|\, z \ge 0, x + z \le 2, y + z \le 2, -x + z \le 0, -y + z \le 0 \right\}$$

has 5 vertices

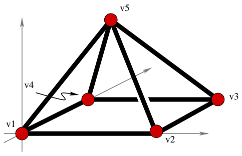


Figure 4.5

$$v_1 := \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix} = \operatorname{face}_{\{1,4,5\}}(P), \qquad v_2 := \begin{pmatrix} 2 \\ 0 \\ 0 \end{pmatrix} = \operatorname{face}_{\{1,2,5\}}(P),$$

$$v_3 := \begin{pmatrix} 2 \\ 2 \\ 0 \end{pmatrix} = \operatorname{face}_{\{1,2,3\}}(P), \qquad v_4 := \begin{pmatrix} 0 \\ 2 \\ 0 \end{pmatrix} = \operatorname{face}_{\{1,3,4\}}(P),$$

$$v_5 := \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} = \operatorname{face}_{\{2,3,4,5\}}(P)\,.$$

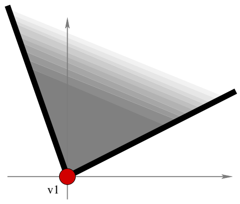


Figure 4.6

The last vertex is already defined by any three of the inequalities 2, 3, 4, and 5. See Figure 4.5.

(2) $C_1 := \left\{ \begin{pmatrix} x \\ y \end{pmatrix} \,\middle|\, -3x - y \le 0, x - 2y \le 0 \right\} = \operatorname{cone}\begin{pmatrix} -1 & 2 \\ 3 & 1 \end{pmatrix}$

The minimal face of C_1 is the apex $a := \begin{pmatrix} 0 \\ 0 \end{pmatrix}$, see Figure 4.6

(3) $C_2 := \left\{\left(\begin{smallmatrix} x \\ y \\ z \end{smallmatrix}\right) \,\middle|\, x+y-z \leq 0, -x-y-z \leq 0\right\} = \operatorname{cone}\left(\begin{smallmatrix} -1 & 1 & -1 & 1 \\ 1 & -1 & -1 & 1 \\ 0 & 0 & 1 & 1 \end{smallmatrix}\right)$

The minimal face of C_2 is the line spanned by $(-1, 1, 0)$. See Figure 4.7. ◊

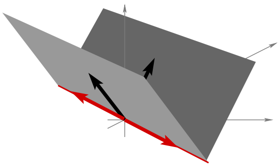


Figure 4.7

With the next theorems we obtain further characterizations of vertices of a polyhedron.

Theorem 4.27. *Let* $P = \{\mathbf{x} \mid A\mathbf{x} \leq \mathbf{b}\}$ *be a polyhedron and* $\mathbf{v} \in P$, $I = \operatorname{eq}(\mathbf{v})$. *Then the following are equivalent:*

(1) $\mathbf{v}$ *is a vertex of* P

(2) $\operatorname{rank}(A_{I*}) = n$

(3) There are no $\mathbf{x}_1, \mathbf{x}_2 \in P$ *such that there is* $0 < \lambda < 1$ *with* $\mathbf{v} = \lambda\mathbf{x}_1 + (1-\lambda)\mathbf{x}_2$.

Proof. (1) ⇒ (3) If $\mathbf{v}$ is a vertex, then there is $\mathbf{c} \in \mathbb{R}^n$, $\delta \in \mathbb{R}$ such that

$$\mathbf{c}^t\mathbf{v} = \delta \qquad\qquad \mathbf{c}^t\mathbf{x} < \delta \quad \text{for all } \mathbf{x} \in P,\, \mathbf{x} \neq \mathbf{v}\,.$$

Assume that there were $\mathbf{x}_1, \mathbf{x}_2 \in P$ and $0 < \lambda < 1$ such that $\mathbf{v} = \lambda\mathbf{x}_1 + (1-\lambda)\mathbf{x}_2$. Then

$$\delta = \mathbf{c}^t\mathbf{v} = \lambda\mathbf{c}^t\mathbf{x}_1 + (1-\lambda)\mathbf{c}^t\mathbf{x}_2 < \lambda\delta + (1-\lambda)\delta = \delta\,,$$

a contradiction.

(3) ⇒ (2) Let $J := [m] - I$. Then $\mathbf{v}$ is an interior point of $\{\mathbf{x} \mid A_{I*}\mathbf{x} = \mathbf{b}_I, A_{J*}\mathbf{x} \leq \mathbf{b}_J\}$. If $\operatorname{rank}(A_{I*}) < n$, then there is $\mathbf{y} \neq \mathbf{0}$ with $A_{I*}\mathbf{y} = \mathbf{0}$. So for $\varepsilon > 0$ small enough, $\mathbf{y}_\pm := \mathbf{v} \pm \varepsilon\mathbf{y} \in P$, and $\mathbf{v} = \frac{1}{2}(\mathbf{y}_+ + \mathbf{y}_-)$.

(2) ⇒ (1) The affine space defined by A_{I*} is 0-dimensional. □

In the following Section 6 we develop the simplex method to compute the optimal value of a linear program, which uses the standard form of a linear program as defined in Definition 3.1. Hence, we want to consider polyhedra given as $P := \{A\mathbf{x} = \mathbf{b}, \mathbf{x} \geq \mathbf{0}\}$ for some $A \in \mathbb{R}^{m \times n}$, $\mathbf{b} \in \mathbb{R}^m$ and their faces. We translate the previous theorem into this setting.

Corollary 4.28. *Let* $P = \{\mathbf{x} \mid A\mathbf{x} = \mathbf{b},\, \mathbf{x} \geq \mathbf{0}\}$, $\mathbf{v} \in P$ *and* $I := \{i \mid v_i > 0\}$. *Then*

$$\mathbf{v} \text{ is a vertex of } P \iff \operatorname{rank}(A_{*I}) = |I|$$
$$(\iff \text{the columns of } A_{*I} \text{ are linearly independent})\,.$$

Proof. Let $K := [n] - I$ and

$$C := \begin{pmatrix} A \\ -A \\ -\mathrm{Id}_n \end{pmatrix} \qquad\qquad \mathbf{d} := \begin{pmatrix} \mathbf{b} \\ -\mathbf{b} \\ \mathbf{0} \end{pmatrix}.$$

Then $P = \{\mathbf{x} \mid C\mathbf{x} \leq \mathbf{d}\}$. Using the system of linear inequalities $C\mathbf{x} \leq \mathbf{d}$ we define the equality set $J := \operatorname{eq}(\mathbf{v})$ of $\mathbf{v}$. Then $J = [2m] + (2m + K)$, so

$$\operatorname{rank}(C_{J*}) = \operatorname{rank}\begin{pmatrix} A \\ -A \\ \mathrm{Id}_{K*} \end{pmatrix} = |K| + \operatorname{rank}(A_{*I})$$

The claim now follows from the previous theorem. □

Proposition 4.29. *For a nonempty polyhedron* $P = \{\mathbf{x} \mid A\mathbf{x} \leq \mathbf{b}\}$ *the following are equivalent:*

(1) P *is pointed.*

(2) $\operatorname{rank}(A) = n$.

(3) $\operatorname{rec}(P)$ *is pointed.*

(4) Any non-empty face of P *is pointed.*

Proof. (1) $\Longleftrightarrow$ (2): The lineality space of P is $\text{lineal}(P) = \{\mathbf{x} \mid A\mathbf{x} = \mathbf{0}\}$. This is $\{\mathbf{0}\}$ if and only if $\text{rank}(A) = n$.
(3) and (4) are clearly equivalent to (2). $\square$

Corollary 4.30. *(1) If $P = \{\mathbf{x} \mid A\mathbf{x} = \mathbf{b},\, \mathbf{x} \geq \mathbf{0}\} \neq \varnothing$ then it is pointed.*
(2) If $P = \{\mathbf{x} \mid A\mathbf{x} \leq \mathbf{b},\, \mathbf{x} \geq \mathbf{0}\} \neq \varnothing$ then it is pointed.

Proof. For the first claim observe that $P = \{\mathbf{x} \mid C\mathbf{x} \leq \mathbf{d}\}$ for $C = \begin{pmatrix} A \\ -A \\ \text{Id}_n \end{pmatrix}$, and $\text{rank}(C) = n$, so P is pointed by the previous proposition. The proof for the second statement is similar. $\square$

Corollary 4.31. *If P is a pointed polyhedron and $\mathbf{x}$ is a bounded optimal solution of $\max(\mathbf{c}^t\mathbf{x} \mid \mathbf{x} \in P)$, then there is an optimal solution that is a vertex of P.*

Proof. The set of optimal solutions is a face F of P. F pointed, so it contains a vertex. $\square$

Corollary 4.32. *Linear programs in standard or canonical form have a finite optimal solution if and only if they have an optimal vertex.* $\square$

We can now prove a first refinement of the Affine MINKOWSKI Theorem 2.7.

Theorem 4.33. *Let $P = \{\mathbf{x} \in \mathbb{R}^n \mid A\mathbf{x} \leq \mathbf{b}\}$ be a polyhedron with r minimal faces $F_1, \ldots, F_r$. Choose $\mathbf{v}_i \in F_i$, $1 \leq i \leq r$. Then*

$$P = \text{conv}(\mathbf{v}_1, \ldots, \mathbf{v}_r) + \text{rec}(P)$$

Proof. Let $Q := \text{conv}(\mathbf{v}_1, \ldots, \mathbf{v}_r) + \text{rec}(P)$. Then clearly $Q \subseteq P$. By the Affine WEYL Theorem, there is a matrix $C \in \mathbb{R}^{s \times n}$ and $\mathbf{d} \in \mathbb{R}^s$, such that $Q = \{\mathbf{x} \mid C\mathbf{x} \leq \mathbf{d}\}$.
Suppose there is $\mathbf{z} \in P \setminus Q$. Then we can find $j \in [s]$ such that $\mathbf{c}_j^t\mathbf{z} > d_j$ for $\mathbf{c}_j^t := C_{j*}$. Hence

$$d_j = \max(\mathbf{c}_j^t\mathbf{x} \mid \mathbf{x} \in Q) < \mathbf{c}_j^t\mathbf{z} \leq \max(\mathbf{c}_j^t\mathbf{x} \mid \mathbf{x} \in P)\,.$$

If the maximum on the right hand side is infinite, then by Proposition 3.4 there is a vector $\mathbf{y} \in \text{rec}(P)$ such that $\mathbf{c}^t\mathbf{y} > 0$. But $\text{rec}(P) = \text{rec}(Q)$, so $\mathbf{y} \in \text{rec}(Q)$. Using the proposition again we obtain that also $\max(\mathbf{c}^t\mathbf{x} \mid \mathbf{x} \in Q)$ would is unbounded. This is a contradiction, so the maximum $\delta := \max(\mathbf{c}^t\mathbf{x} \mid \mathbf{x} \in P) > d_j$ is finite.
Let F_i be some minimal face of $\text{face}_{\mathbf{c}_j}(P)$. By assumption $\mathbf{v}_i \in F_i$. Hence, $\mathbf{c}^t\mathbf{v}_i = \delta > d_j$, so that $\mathbf{v}_i \notin Q$. This is a contradiction. So $P \subseteq Q$. $\square$

Given a decomposition of a polyhedron into a polytope and a cone, the previous theorem constructs a minimal set of generators for the polytope. In the rest of this chapter we want to achieve a similar set of generators for the cone part. We will see later that we need to look at faces of dimension $n_0 + 1$, i.e. one level above the minimal faces in the face poset. So let us first explore their structure.

Proposition 4.34. *Let $P = \{\mathbf{x} \mid A\mathbf{x} \leq \mathbf{b}\}$ be a polyhedron, and F a face of P of dimension $\dim(F) = n_0 + 1$. Then there are $I, J \subseteq [m]$ with*

$$F = \{\mathbf{x} \in \mathbb{R}^n \mid A_{J*}\mathbf{x} = \mathbf{b}_J,\, A_{I*}\mathbf{x} \leq \mathbf{b}_I\}\,, \qquad \text{rank}(A_{J*}) = \text{rank}(A) - 1\,, \qquad |I| \leq 2$$

Proof. F has a representation of the form

$$F = \{\mathbf{x} \mid A_{J*}x = \mathbf{b}_J,\, A_{I*}\mathbf{x} \leq \mathbf{b}_I\}$$

for some subsets $I, J \subseteq [m]$. We may assume that I is as small as possible, so that $J = \text{eq}(F)$. Let

$$H := \{\mathbf{x} \mid A_{J*}\mathbf{x} = \mathbf{b}_J\} = \text{aff}(F)\,,$$

and for any $i \in I$

$$J_i := J \cup \{i\} \qquad\qquad H_i := \{\mathbf{x} \mid A_{i*}\mathbf{x} = \mathbf{b}_i, A_{J_i*}\mathbf{x} = \mathbf{b}_{J_i}\}\,.$$

By our assumption any row in A_{I*} is not in the span of the rows of A_{J*}, so

$$\operatorname{rank}(A_{J_i*}) = \operatorname{rank}(A_J*) + 1 = n - \dim(F) + 1 = n - n_0 \text{for } i \in I.$$

and all H_i are minimal faces of P. Hence, all H_i are translates of the lineality space of P. Their intersections with H are parallel affine subspaces of dimension $\dim(H) - 1$ (they are all parallel to the lineality space). Choose a complement ℓ of H_i in H. See Figure 4.8. Then ℓ is one-dimensional and intersects each H_i in a single point. By convexity, there are at most two intersection points where ℓ could enter or leave P. So, as I is irredundant, there can be at most two such subspaces H_i. □

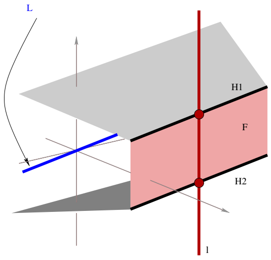


Figure 4.8

By this proposition any face F of dimension $\dim(F) = n_0 + 1$ has at most two facets, which itself are minimal faces of P. Hence, we can write F in the form

$$F = \operatorname{lineal}(P) + e$$

where e is either a segment or a ray. If P is pointed, then F is called an **edge** if e is a segment, and **extremal ray** otherwise. Two minimal faces of P are **adjacent** if they are contained in a common edge. If P is not pointed, then F is called a **minimal proper face**.

edge
extremal ray
adjacent minimal faces
minimal proper face

Example 4.35. Consider the polyhedron P and the cones C_1 and C_2 defined in the previous Example 4.26.

(1) The set $\{x \mid -z = 0, -y + z = 0, -x + z \le 0, x + z \le 2\}$ defines an edge of P.

(2) Both $\mathbb{R}_+\begin{pmatrix} -1 \\ 3 \end{pmatrix}$ and $\mathbb{R}_+\begin{pmatrix} 2 \\ 1 \end{pmatrix}$ are extremal rays (and minimal proper faces) of C_1.

(3) The proper minimal faces of C_2 are

$$\begin{aligned} F_1 &:= \{x \mid x + y - z = 0, -x - y - z \le 0\} \\ F_2 &:= \{x \mid x + y - z \le 0, -x - y - z = 0\}\,. \end{aligned} \qquad \Diamond$$

Let P be a polyhedron and $\mathbf{u}, \mathbf{v}$ adjacent vertices of P. Then Proposition 4.34 says that there are $I, J \subseteq [m]$ such that

$$\mathbf{u} = \operatorname{face}_I(P) \qquad\qquad \mathbf{v} = \operatorname{face}_J(P) \qquad\qquad |I \setminus J| = |J \setminus I| = 1\,.$$

Remark 4.36. There is a similar description for faces of codimension 2. These are always intersections of exactly 2 facets. There is no similar easy description for faces of dimension d with

$$n_0 + 2 \le d \le \dim(P) - 3\,. \qquad \Diamond$$

With this preparation the next theorem constructs a generating set for a cone.

Theorem 4.37. *Let* $C = \{\mathbf{x} \mid A\mathbf{x} \le \mathbf{0}\} \subseteq \mathbb{R}^n$ *and* $L := \operatorname{lineal}(C)$. *Let* $G_1, \dots, G_s$ *be the proper minimal faces of* C. *Choose* $\mathbf{y}_i \in G_i - L$ *for* $1 \le i \le s$. *Then*

$$C = \operatorname{cone}(\mathbf{y}_1, \dots, \mathbf{y}_s) + L\,.$$

Proof. Let $n_0 := \dim(L)$, $d := \dim(C)$ and $D := \operatorname{cone}(\mathbf{y}_1, \dots, \mathbf{y}_s) + L$. By Proposition 4.34 we can write the minimal proper face G_1 in the form

$$G_1 = \{\mathbf{x} \mid A_{J*}\mathbf{x} = 0,\ A_{i*}\mathbf{x} \le 0\}$$

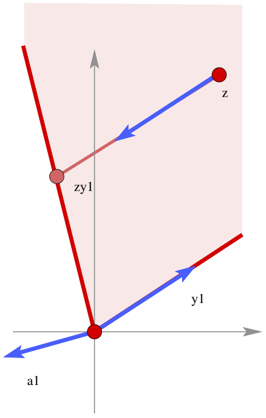


Figure 4.9

for some index $i \in [m]$. This implies $A_{i*}\mathbf{y}_1 < 0$, as $\mathbf{y}_1 \notin L$.

We prove the theorem now by induction on $k := d - n_0$. The claim is trivial if $k = 0$, so suppose $k > 0$ and write C as

$$C = \{\mathbf{x} \mid A_{K*}\mathbf{x} \le 0,\ A_{E*}\mathbf{x} = 0\}.$$

with $E = \operatorname{eq}(C)$ and K irredundant. By Theorem 4.18 the facets of C are in one-to-one correspondence with the inequalities in K. Any proper minimal face of a facet of C is also a proper minimal face of C. Hence, by induction, each facet of C is contained in D. Let $\mathbf{z} \in C$ and

$$\lambda := \max(\mu \mid \mathbf{z} - \mu\mathbf{y}_1 \in C).$$

As $A_{i*}\mathbf{y}_1 < 0$ and $A\mathbf{z} \le 0$, λ exists and is finite. For at least one $j \in K$ we have $A_{j*}(\mathbf{z} - \lambda\mathbf{y}_1) = 0$, so $\mathbf{z} - \lambda\mathbf{y}_1$ is contained in a facet of C, and hence in D. But this implies

$$\mathbf{z} = \lambda\mathbf{y}_1 + (\mathbf{z} - \lambda\mathbf{y}_1) \in D. \qquad \square$$

Combining this with the refined affine Minkowski-Theorem, we obtain

Theorem 4.38. *Let $P := \{\mathbf{x} \mid A\mathbf{x} \le \mathbf{b}\}$ and $L := \operatorname{lineal}(P)$. Let*

- *$F_1, \ldots, F_r$ be the minimal faces of P and*
- *$G_1, \ldots, G_s$ the minimal proper faces of $\operatorname{rec}(P)$.*

Choose

$$\mathbf{x}_i \in F_i \quad \text{for } 1 \le i \le r \qquad\qquad \mathbf{y}_j \in G_j - L \quad \text{for } 1 \le j \le s$$

and an arbitrary basis $\mathbf{z}_1, \ldots, \mathbf{z}_t$ of L. Then

$$P = \operatorname{conv}(\mathbf{x}_1, \ldots, \mathbf{x}_r) + \operatorname{cone}(\mathbf{y}_1, \ldots, \mathbf{y}_s) + \operatorname{lin}(\mathbf{z}_1, \ldots, \mathbf{z}_t).$$

and

- *$\mathbf{x}_1, \ldots, \mathbf{x}_r$ are unique up to permutation and adding vectors from L,*
- *$\mathbf{y}_1, \ldots, \mathbf{y}_s$ are unique up to permutation, adding vectors from L and scaling with a positive factor.* $\square$

We want to show that this representation is minimal. Let a polyhedron P be given by

$$P \;=\; \operatorname{conv}(X) \;+\; \operatorname{cone}(Y) \;+\; \operatorname{lin}(Z) \qquad (*)$$

for matrices $X \in \mathbb{R}^{n \times r}, Y \in \mathbb{R}^{n \times s}, Z \in \mathbb{R}^{n \times t}$. We may assume that $\operatorname{cone}(Y)$ is pointed. Then

$$L := \operatorname{lineal}(P) = \operatorname{lin}(Z) \qquad \text{and} \qquad \operatorname{rec}(P) = \operatorname{cone}(Y) + L.$$

Proposition 4.39. *Let F be a proper face of P given by $(*)$. Then there are $I \subseteq [r]$, $J \subseteq [s]$ such that*

$$F = \operatorname{conv}(X_{*I}) + \operatorname{cone}(Y_{*J}) + \operatorname{lin}(Z).$$

Proof. By definition, there is $\mathbf{c}^t \in \mathbb{R}^n$ and $\delta \in \mathbb{R}$ such that

$$F = P \cap \{\mathbf{x} \mid \mathbf{c}^t\mathbf{x} = \delta\},$$

and $\delta = \max(\mathbf{c}^t\mathbf{x} \mid \mathbf{x} \in P)$. So

$$\mathbf{c}^t Z = \mathbf{0} \qquad\qquad \mathbf{c}^t Y \le \mathbf{0} \qquad\qquad \mathbf{c}^t X \le \delta.$$

Let $$I := \{i \in [r] \mid \mathbf{c}^t X_{*i} = \delta\} \qquad J := \{j \in [s] \mid \mathbf{c}^t Y_{*j} = \mathbf{0}\}$$

and $$\overline{F} := \operatorname{conv}(X_{*I}) + \operatorname{cone}(Y_{*J}) + \operatorname{lin}(Z)\,.$$

Then $\overline{F} \subset F$. Any $\mathbf{u} \in F$ can be decomposed into $\mathbf{u} = \mathbf{x} + \mathbf{y} + \mathbf{z}$ for

$$\mathbf{x} \in \operatorname{conv}(X) \qquad \mathbf{y} \in \operatorname{cone}(Y) \qquad \mathbf{z} \in \operatorname{lin}(Z)\,.$$

So there are $\boldsymbol{\lambda} \in \mathbb{R}^r$, $\boldsymbol{\mu} \in \mathbb{R}^s$, $\boldsymbol{\eta} \in \mathbb{R}^t$ such that

$$\mathbf{u} = X\boldsymbol{\lambda} + Y\boldsymbol{\mu} + Z\boldsymbol{\eta}\,, \qquad \text{and} \qquad \delta = \mathbf{c}^t\mathbf{u} = \mathbf{c}^t X\boldsymbol{\lambda} + \mathbf{c}^t Y\boldsymbol{\mu} + \mathbf{c}^t Z\boldsymbol{\eta}\,.$$

Thus, $\mathbf{c}^t Y\boldsymbol{\mu} = 0$ and $\mathbf{c}^t X\boldsymbol{\lambda} = \delta$. This can only hold if $\mu_k = 0$ for $k \notin J$ and $\lambda_k = 0$ for $k \notin I$. So $\mathbf{x} \in \operatorname{conv}(X_{*I})$ and $\mathbf{y} \in \operatorname{cone}(Y_{*J})$, which implies $\mathbf{u} \in \overline{F}$. ☐

Let F be a face of a polyhedron P represented in the form (*). Minimal faces of F are minimal faces of P, and minimal proper faces of rec(F) are minimal proper faces of rec(P), so the above Proposition 4.39 implies that

(1) for any minimal face F of P there is $i \in [r]$ with $X_{*i} \in F$, and
(2) for any minimal proper face G of rec(P) there is $j \in [s]$ such that $Y_{*j} \in G$.

Now minimal faces of a polyhedron are disjoint, and minimal proper faces of a cone intersect only in the lineality space. So columns of X (Y) corresponding to different minimal (proper) faces must be different. Hence, any representation of a polyhedron in the form (*) contains a representation as defined in Theorem 4.38. In particular, if $P = \operatorname{conv}(\mathbf{x}_1, \dots, \mathbf{x}_r)$, then $\{\mathbf{x}_1, \dots, \mathbf{x}_r\}$ contains all vertices of P.

A polyhedron P is pointed if lineal(P) = $\{0\}$. So P is pointed if and only if P has a vertex. In particular, all polytopes are pointed. Recall that we can decompose any polyhedron P into the Minkowski sum

$$P \;=\; L \;+\; P'$$

where $L :=$ lineal(P) is a linear space and P' is a pointed polyhedron obtained by intersecting P with a complement W of lineal(P), i.e. $P' \;=\; P \cap W$. So P has a representation

$$P \;=\; L \;+\; \operatorname{conv}(\mathbf{x}'_1, \dots, \mathbf{x}'_r) \;+\; \operatorname{cone}(\mathbf{y}'_1, \dots, \mathbf{y}'_s)\,,$$

where $\mathbf{x}'_1, \dots, \mathbf{x}'_r$ are the vertices of P' and $\mathbf{y}'_1, \dots, \mathbf{y}'_r$ are generators of the extremal rays of rec(P').

Proposition 4.40. *Any pointed polyhedron $P = \{A\mathbf{x} \le \mathbf{b}\}$ has a minimal representation as*

$$P = \operatorname{conv}(\mathbf{v}_1, \dots, \mathbf{v}_r) + \operatorname{cone}(\mathbf{y}_1, \dots, \mathbf{y}_k)\,,$$

where $\mathbf{v}_1, \dots, \mathbf{v}_r$ are the vertices of P and $\mathbf{y}_1, \dots \mathbf{y}_k$ are generators the extreme rays of rec(P)*, which are unique up to scaling any of the $\mathbf{y}_i$ with a positive real number.* ☐

Complexity of Rational Polyhedra 5

In this chapter we want to look at algorithms, their running time behavior, at the complexity of the interior and exterior description of polyhedra, and the complexity of solutions of linear programs. Here, we have to give up the choice of $\mathbb{R}$ as our ground field, as numbers in $\mathbb{R}$ do not have a finite coding length. The usual choice here is the field $\mathbb{Q}$ of rational numbers. So from now on, all numbers, vectors and matrices have entries from $\mathbb{Q}$. Clearly, if

$$P = \{\mathbf{x} \mid A\mathbf{x} \le \mathbf{b}\} = \operatorname{conv}(X) + \operatorname{cone}(Y) + \operatorname{lin}(Z)\,,$$

then

$$\begin{array}{rcl} A, \mathbf{b} \text{ rational} & \Longrightarrow & X, Y, Z \text{ can chosen to be rational} \\ X, Y, Z \text{ rational} & \Longrightarrow & A, \mathbf{b} \text{ can chosen to be rational}\,. \end{array}$$

The most successful distinction in complexity over the last 50 years has been between problems that can be solved in polynomial time in the size of the input, and those for which such an algorithm is not known. The first type of problems is subsumed in the class $\mathscr{P}$ of polynomially solvable problems. Often, one sets these problems against the so called $\mathscr{NP}$-complete problems.

Clearly, except for very simple problems, the *running time* of an algorithm will depend on the size of the input. So we have to fix some measure for the input. Implementations on a computer usually use the binary encoding, which we will choose in the following.

Definition 5.1. *The* **size** *or* **coding length** $\langle s\rangle$ *of an object* s *is the minimum length of a* 0-1*-string that represents it.*

size
coding length

$z \in \mathbb{Z}$	$\langle z\rangle = 1 + \lceil \log_2(\lvert z\rvert + 1)\rceil$
$r \in \mathbb{Q}$	$r = \frac{a}{b}$, $a, b \in \mathbb{Z}$, $b \ge 0$ with $\gcd(a, b) = 1$, then $\langle r\rangle = 1 + \lceil \log_2(\lvert a\rvert + 1)\rceil + \lceil \log_2(\lvert b\rvert + 1)\rceil$
$\mathbf{v} \in \mathbb{Q}^n$	$\langle \mathbf{v}\rangle = \sum \langle \mathbf{v}_i\rangle$
$A \in \mathbb{Q}^{m\times n}$	$\langle A\rangle = \sum \langle a_{ij}\rangle$.

Table 5.1: A list of some common objects and their sizes

There are many other common choices for such a measure. However, most of them are "polynomially equivalent", so changing the measure does not change "polynomial solvability" of a problem, which is what we are interested in (the notable exception of this is the **unary encoding** of numbers). A list of some common objects and their sizes that we will use is in Table 5.1.

unary encoding

Now that we can measure the size of our input data, we can define the complexity of solving a *problem*. Usually one distinguishes between two different types of problems.

decision problem
feasible solution
instance
yes-instance
no-instance
optimization problem
objective function
minimization problem
maximization problem

Definition 5.2. *(1) In* **a decision problem** Δ *we have a set* S_I *of* **feasible solutions** *for every* **instance** $I \in \Delta$ *of the problem* Δ*. We have to decide whether* S_I *contains a solution that satisfies given properties. If there is such a solution, then* I *is a* ***yes*-instance***, otherwise a* ***no*-instance***.*

(2) In an **optimization problem** Π*, we again have a set* S_I *of* **feasible solutions** *for every* **instance** $I \in \Pi$ *of the problem* Π*, but additionally we have an* **objective function** $c_I : S_I \longrightarrow \mathbb{Q}$ *associated with the instance. The task is to find an optimal solution in* S_I *with respect to this function, so either a solution* $\overline{x} \in S_I$ *with* $c_I(\overline{x}) \leq c_I(x)$ *for all* $x \in S_I$ *in a* **minimization problem***, or* $c_I(\overline{x}) \geq c_I(x)$ *for all* $x \in S_I$ *in a* **maximization problem***.*

Example 5.3. Examples of decision problems are

(1) Does a system of linear inequalities $A\mathbf{x} \leq \mathbf{b}$ have a solution?
(2) Does a graph have a Hamiltonian cycle?
(A **Hamiltonian cycle** is a cycle in the graph that visits each vertex exactly once)

Examples for an optimization problem are

(1) Find an optimal solution of $\max(\mathbf{c}^t\mathbf{x} \mid A\mathbf{x} \leq \mathbf{b})$ or determine that the system is infeasible.
(2) Find a shortest path between two nodes in a graph. ◊

algorithm

An **algorithm** $\mathscr{A}$ is a finite list of instructions that perform operations on some data. $\mathscr{A}$ *solves* a problem Λ if, for any given input y of I, it determines in finite time an *output* $z \in S_I$, or stops without an output, if there is no solution.

running time

An important characteristic of an algorithm is its **running time**. This should clearly not depend on the specific hard- or software used, so we need an intrinsic measure.

running time function

Definition 5.4. *The* **running time function** *of an algorithm* $\mathscr{A}$ *for a problem* Λ *is a function* $f_{\mathscr{A}} : \mathbb{N} \longrightarrow \mathbb{N}$ *such that*

$$f_{\mathscr{A}}(s) = \max_{y \text{ with } \langle y \rangle \leq s} (\text{number of elementary operations performed by } \mathscr{A} \text{ on the input } y)$$

elementary operation

Elementary operations *are*

addition, subtraction, multiplication, division, comparison, and assignment.

polynomial (time)
polynomial-time solvable

An algorithm is said to be a **polynomial-time algorithm***, or just* **polynomial***, if its running time function* f *satisfies* $f = \mathscr{O}(g)$ *for a polynomial* g*. A problem* Λ *is* **polynomial-time solvable** *if there is a polynomial algorithm that solves it.*

Landau-*symbol*

The notation $\mathscr{O}$ for the estimate of f in the definition is one of the **LANDAU-symbols**:

$$f = \mathscr{O}(g) \qquad \exists C \in R, \exists n_0 \in \mathbb{N} : \quad f(n) \leq Cg(n) \text{ for all } n \geq n_0 ,$$

which is useful here as we are only interested in the asymptotic dependence of the running time, so we want to neglect coefficients and lower order terms.

In the definition of an algorithm we assumed that the elementary operations can be done in constant (unit) time. This clearly is a simplification, as the time needed for the multiplication of two numbers does depend on their size, but it is a polynomial time operation, so this choice does not affect polynomial-time solvability.

complexity class

Decision problems are categorized into several **complexity classes** with respect to their (asymptotic) running time.

Definition 5.5. *(1)* $\mathscr{P}$ *is the set of all decision problems* Δ *that are polynomial-time solvable.*

certificate

(2) $\mathscr{N}\mathscr{P}$ *is the collection of all decision problems* Δ *that have an associated problem* $\Delta' \in \mathscr{P}$ *such that for any yes-instance* $I \in \Delta$ *there is a* **certificate** $C(I)$ *of size polynomial in the size of* I *such that* $(I, C(I))$ *is a yes-instance of* Δ'*.*

(3) co-$\mathcal{NP}$ *is the collection of all problems that have a polynomial-time checkable certificate for any no-instance.*

An example of an $\mathcal{NP}$-problem is the question whether a given graph is Hamiltonian. A positive answer can be certified by giving a Hamiltonian cycle, and the correctness of this can be checked in polynomial time. Note, that there is no known certificate for the opposite answer whether a given graph is *not* Hamiltonian, so it is unknown whether the problem is in co-$\mathcal{NP}$.

Problems in $\mathcal{NP} \cap$ co-$\mathcal{NP}$ are those for which both a positive and a negative answer have a certificate that can be checked in polynomial time. Such problems are also called **well-characterized**. We will see some examples later. Clearly, we have $\mathcal{P} \subseteq \mathcal{NP} \cap$ co-$\mathcal{NP}$. There are only very few problems known that are in $\mathcal{NP} \cap$ co-$\mathcal{NP}$, but not known to be in $\mathcal{P}$. One of the most famous and important questions in complexity theory is whether

well-characterized

$$\mathcal{P} \neq \mathcal{NP} \ .$$

It is widely believed that the class $\mathcal{NP}$ is much larger than $\mathcal{P}$. However, there is no proof known. Any answer, positive or negative, to this would have far reaching consequences:

(1) If $\mathcal{P} = \mathcal{NP}$, then it may either imply a completely new type of polynomial-time algorithm, or it may prove that the concept of *polynomial-time* does not serve as a good characterization for algorithms.
(2) If $\mathcal{P} \neq \mathcal{NP}$, then this will likely give insight why some problems seem to be so much harder than others.

Definition 5.6. *A problem Δ is said to be* **(polynomial-time) reducible** *to a problem Δ' if there is a polynomial-time algorithm that returns, for any instance $I \in \Delta$ an instance $I' \in \Delta'$ such that*

(polynomial-time) reducible

$$I \text{ is a yes-instance} \iff I' \text{ is a yes-instance}.$$

This implies, that if Δ is reducible to Δ' and

(1) $\Delta' \in \mathcal{P}$, then also $\Delta \in \mathcal{P}$.
(2) $\Delta \in \mathcal{NP}$, then also $\Delta' \in \mathcal{NP}$.

A problem Δ is **$\mathcal{NP}$-complete** if any other problem $\Delta' \in \mathcal{NP}$ can be reduced to Δ. If Δ is reducible to Δ' and Δ' is $\mathcal{NP}$-complete, then so is Δ. $\mathcal{NP}$-complete problems are in some sense the hardest $\mathcal{NP}$-problems. There in fact do exist $\mathcal{NP}$-complete problems, e.g. the TRAVELING SALESMAN PROBLEM (TSP) or integer linear programming.

$\mathcal{NP}$-complete

Let Π be an optimization problem. We can associate a decision problem Δ to Π as follows:

$$\text{Given } r \in \mathbb{Q} \text{ and an instance } I, \text{ is there } z \in S_I \text{ with } c_I(z) \geq r\ ?$$

An optimization problem Π is **$\mathcal{NP}$-hard**, if the associated decision problem is $\mathcal{NP}$-complete. Now we analyze the complexity of linear systems of equations and inequalities.

$\mathcal{NP}$-hard

Lemma 5.7. *Let $A \in \mathbb{Q}^{n \times n}$ of size s. Then $\langle \det(s) \rangle \leq 2s$.*

Proof. Let $A = (\frac{p_{ij}}{q_{ij}})_{ij}$ and $\det(A) = \frac{p}{q}$ for completely reduced fractions $\frac{p_{ij}}{q_{ij}}$, $\frac{p}{q}$ and $q_{ij}, q > 0$. We can assume that $n \geq 2$ as otherwise $\det(A) = A$ and

$$\sum \langle p_{ij} \rangle \geq n\,,$$

as otherwise $\det(A) = 0$. Clearly, we have

$$q + 1 \leq \prod_{i,j=1}^{n} (q_{ij} + 1) \leq 2^{\sum \log_2(q_{ij}+1)} \leq 2^{\sum \lceil \log_2(q_{ij}+1) \rceil} \leq 2^{s-n}$$

By the definition of the determinant via the Laplace formula

$$|\det(A)| \leq \prod_{i,j=1}^{n} (|p_{ij}| + 1)$$

Further we have that

$$|p| + 1 = |\det(A)|\, q + 1 \leq \prod_{i,j=1}^{n} (|p_{ij}| + 1)\,(q_{ij} + 1) \leq 2^{\sum \log_2(|p_{ij}|+1) + \log_2(q_{ij}+1)}$$
$$\leq 2^{\sum \lceil \log_2(|p_{ij}|+1) + \log_2(q_{ij}+1) \rceil} \leq 2^s - n^2 \,.$$

Combining this gives $\langle \det(A) \rangle = 1 + \lceil \log_2(|p| + 1) \rceil + \lceil \log_2(q + 1) \rceil < 2s$. ☐

***Corollary 5.8**. Let $A \in \mathbb{Q}^{m\times n}$, $\mathbf{b} \in \mathbb{Q}^m$, and assume that the rows of the matrix $(A|\mathbf{b})$ have size $\leq \varphi$. If $A\mathbf{x} = \mathbf{b}$ has a solution, then it has one of size at most $4n^2\varphi$.*

Proof. We may assume that $\operatorname{rank} A = m$. Then we can reorder the columns of A such that $\overline{A} := A_{*[m]}$ is non-singular.

Then $x := (\overline{\mathbf{x}}, 0)$ for $\overline{\mathbf{x}} := \overline{A}^{-1}\mathbf{b}$ is a solution of the system, and by CRAMER's rule, the entries of $\overline{\mathbf{x}}$ are

$$\overline{x}_i = \frac{\det(\overline{A}_j)}{\det(\overline{A})} \,.$$

where $\overline{A}_j$ is the matrix obtained by replacing the j-th column of $\overline{A}$ by $\mathbf{b}$. Now the size of these matrices is at most $m\varphi$, so their determinants have size at most $2m\varphi$, and hence,

$$\langle \overline{x}_i \rangle \leq 4m\varphi.$$

So $\langle \mathbf{x} \rangle \leq 4mn\varphi \leq 4n^2\varphi$. ☐

This implies that the problem Λ_1 of deciding whether a rational system of linear equations $A\mathbf{x} = \mathbf{b}$ has a solution is well-characterized:

(1) If $A\mathbf{x} = \mathbf{b}$ has a solution, then it has one of size polynomially bounded by the sizes of A and $\mathbf{b}$, and given such an $\mathbf{x}$, we can check in polynomial time that $A\mathbf{x} = \mathbf{b}$. So $\Lambda_1 \in \mathcal{NP}$.

(2) If $A\mathbf{x} = \mathbf{b}$ has no solution, then there is $\mathbf{y}$ such that $\mathbf{y}^t A = \mathbf{0}$ and $\mathbf{y}^t\mathbf{b} \neq 0$. Again, we can assume that $\mathbf{y}$ has size polynomial in that of A and $\mathbf{b}$, and the correctness can be verified in polynomial time. So $\Lambda_1 \in$ co-$\mathcal{NP}$.

In fact, this problem is in $\mathcal{P}$, as can be shown by proving that GAUSSIAN elimination is in $\mathcal{P}$.

***Lemma 5.9**. If a system of linear inequalities $A\mathbf{x} \leq \mathbf{b}$ has a solution, then it has one of size polynomially bounded in the size of A and $\mathbf{b}$.*

Proof. Let $A \in \mathbb{Q}^{m\times n}$, $P := \{\mathbf{x} \mid A\mathbf{x} \leq \mathbf{b}\}$ and $I \subset [m]$ such that $\{\mathbf{x} \mid A_{I*}\mathbf{x} = \mathbf{b}_I\}$ is a minimal face of P. By Corollary 5.8, it contains a vector of size polynomially bounded in the sizes of A_{I*} and $\mathbf{b}_I$. These are in turn bounded by the sizes of A and $\mathbf{b}$. ☐

This implies that the problem Λ_2 of deciding whether a system of rational linear inequalities $A\mathbf{x} \leq \mathbf{b}$ has a solution is in $\mathcal{NP}$.

By FARKAS Lemma $A\mathbf{x} \leq \mathbf{b}$ has no solution if and only if there is a vector $\mathbf{y} \in \mathbb{Q}^m$ such that $\mathbf{y} \geq \mathbf{0}$, $\mathbf{y}^t A = \mathbf{0}$ and $\mathbf{y}^t\mathbf{b} < 0$. Again, we can choose this vector to be polynomially bounded in the sizes of A and $\mathbf{b}$, so $\Lambda_2 \in$ co-$\mathcal{NP}$. In total, Λ_2 is well-characterized. This implies the following proposition.

Proposition 5.10. *Let $A \in \mathbb{Q}^{m\times n}$, $\mathbf{b} \in \mathbb{Q}^m$, $\mathbf{c} \in \mathbb{Q}^n$, and $\delta \in \mathbb{Q}$. Then the following problems are well-characterized:*

1. *Decide whether $A\mathbf{x} \le \mathbf{b}$ has a solution.*
2. *Decide whether $A\mathbf{x} = \mathbf{b}$ has a nonnegative solution.*
3. *Decide whether $A\mathbf{x} \le \mathbf{b}$, $\mathbf{c}^t\mathbf{x} > \delta$ has a solution.*
4. *Decide whether $\max(\mathbf{c}^t\mathbf{x} \mid A\mathbf{x} \le \mathbf{b})$ is infeasible, bounded feasible or unbounded feasible.*

Proof. (1) follows directly from the above, the arguments for (2) and (3) are similar, for (4) use the dual program and the duality theorem. □

Further, it is easy to see that the problems (1), (2) and (3) of the previous proposition are polynomially equivalent, which means that, if one can solve any of the problems in polynomial time, then one can solve the others in polynomial time.

The four problems are in fact in $\mathscr{P}$, which follows from KACHIYAN's *ellipsoid method* for solving linear programs. However, this is currently not a practical algorithm. The simplex method, which we will meet in the next chapter, is still both numerically more stable and faster. However, the simplex method is not a polynomial time algorithm.

Definition 5.11. *Let $P \subseteq \mathbb{Q}^n$ be a rational polyhedron.*

1. *The* **facet complexity** *of P is the smallest number φ such that $\varphi \ge n$ and there exists a rational system $Ax \le b$ that describes P and each inequality has size bounded by φ.*
2. *The* **vertex complexity** *of P is the smallest number ν such that $\nu \ge n$ and there exists rational matrices X and Y such that $P = \operatorname{conv}(X) + \operatorname{cone}(Y)$ and each row of X and Y has size bounded by ν.*

facet complexity

vertex complexity

Proposition 5.12. *Let $P \subset \mathbb{Q}^n$ be a rational polyhedron with facet complexity φ and vertex complexity ν. Then there is a constant $C > 0$ such that*

$$\nu \le Cn^2\varphi \qquad\qquad \varphi \le Cn^2\nu .$$

Proof. (1) $\nu \le 4n^2\varphi$: Let $P = \{Ax \le b\}$. The claim is a direct consequence of the representation of P as

$$P = \operatorname{conv}(X) + \operatorname{cone}(Y) + \operatorname{lin}(Z) .$$

where all rows of B, C and D are solutions of some subsystem of $A\mathbf{x} \le \mathbf{b}$, which are, by Corollary 5.8, bounded by $4n^2\varphi$.

(2) $\varphi \le 4n^2\nu$:
Assume $P = \operatorname{conv}(X) + \operatorname{cone}(Y)$ for rational matrices X and Y whose rows have size at most ν.
Suppose first that $\dim P = n$. Then each facet of P is determined by a linear equation of the form (where ξ is the variable)

$$\det \begin{bmatrix} 1 & \xi^t \\ \mathbf{1}_{|I|} & X_{I*} \\ \mathbf{0}_{|J|} & Y_{J*} \end{bmatrix} = 0 .$$

Expanding by the first row we obtain

$$\sum_{i=1}^{n} (-1)^i \left(\det(D_i)\right) \xi_i = -\det(D_0) ,$$

for the $(n \times n)$-sub-matrices D_i obtained by deleting the first row and the i-th column. D_0 has size at most $2n(\nu+1)$, and each D_i has size at most $2n\nu$. Therefore, the equation and the corresponding inequality for the facet have size at most $n \cdot (2n(\nu+1) + 2n\nu \le 6n^2\nu$.

So assume $\dim P < n$. By adding a subset of $\{0, e_1, \ldots, e_n\}$, where 0 is the zero vector and e_j are the standard unit vectors, to the polyhedron, one can obtain a new polyhedron Q with $\dim Q = n$ such that P is a face of Q. By the previous argument, this gives us $n - \dim P$ equations of size at most $6n^2\nu$ defining the affine hull.
Further, we can choose $n - \dim(P)$ coordinate directions so that the projection P' is full-dimensional and affinely equivalent to P'. P' has vertex complexity at most ν. By the previous argument, we may describe P' by linear inequalities of size at most $6(n - \dim(P))^2\nu$. Extend these inequalities by zeros to obtain inequalities valid for P. Together with the equations for the affine hull, we obtain a description of the required size. ☐

A direct corollary of this theorem concerns the size of solutions of linear programs.

Corollary 5.13. *Let $A \in \mathbb{Q}^{m\times n}$, $\mathbf{b} \in \mathbb{Q}^m$, $\mathbf{c}^t \in \mathbb{Q}^n$. Assume that each entry of A, $\mathbf{b}$, and $\mathbf{c}$ is bounded by s, and that the optima of*

$$\delta := \max(\mathbf{c}^t\mathbf{x} \mid A\mathbf{x} \le \mathbf{b}, \mathbf{x} \ge \mathbf{0}) = \min(\mathbf{y}^t\mathbf{b} \mid \mathbf{y}^t A \ge \mathbf{c}, \mathbf{y} \ge \mathbf{0})$$

are finite. Then $\langle\delta\rangle$ is bounded by n and s, and both the primal and the dual program have an optimal solution of size bounded by n and s.

Proof. Each linear inequality of $A\mathbf{x} \le \mathbf{b}$ has size bounded by $t := (n+1)s$, so the facet complexity is at most t. The optimum of the primal program is attained at a vertex, so it is bounded by $4n^2t$. The bound on the dual solution follows similar. Both imply a polynomial bound on the optimal value. ☐

Refining this, you could even prove that if $\delta = p/q$ is a completely reduced fraction, then q is bounded by n and s. So given any lower bound on δ, you could prove that the solution space of a rational linear program is finite.

The Simplex Algorithm 6

In this section we will introduce the **Simplex algorithm** to solve a linear programming problem

Simplex algorithm

$$\max(\mathbf{c}^t\mathbf{x} \mid \mathbf{x} \in P) \tag{6.1}$$

over a polyhedron $P = \{\mathbf{x} \mid A\mathbf{x} \leq \mathbf{b}\}$. This algorithm was first considered by George Dantzig in 1947, based on work for the U.S. military, and is still one of the most successful algorithms for solving a linear program. Similar ideas have been developed independently by Kantorovich already in 1939. The main idea is to enumerate vertices of a polyhedron in a clever way, namely to proceed from one vertex to another with better objective value along an edge.

Many different variants and improvements have been presented since then, and we will consider several important in the following section. The Simplex algorithm is not a polynomial time algorithm, just because there are polyhedra with an exponential number of vertices in the number of facets, and you may have to enumerate all of them. Nevertheless, it has been proved by Kachiyan in 1979 that linear programming is in $\mathscr{P}$. To prove this, he developed the **Ellipsoid method**. In 1984 Karmarkar presented another polynomial time method, the **interior point method**.

ellipsoid method

interior point method

However, the simplex algorithm is still the fastest method for many "real world problems", as degeneracies leading to an exponential running time only rarely occur. There are up to now no useful implementations of the Ellipsoid method, but in recent time some useful implementations of the interior point method have been released. Still, the simplex algorithm is the most flexible method and can easily be adapted to take inherent additional structure of a problem into account. In this text we will concentrate on the simplex algorithm and only briefly mention the main ideas in other methods.

We introduce some standard notation of linear programming. The algorithm is often discussed in a purely algebraic setting, and many notions used for linear programs reflect this. We will introduce several along our development of the algorithm to make the connection to standard textbooks. Let

$$\begin{array}{lrcl} \text{maximize} & \mathbf{c}^t\mathbf{x} & & \\ \text{subject to} & A\mathbf{x} & = & \mathbf{b} \\ & \mathbf{x} & \geq & \mathbf{0} \end{array} \tag{6.2}$$

for $A \in \mathbb{R}^{m\times n}$, $\mathbf{b} \in \mathbb{R}^m$, $\mathbf{c} \in \mathbb{R}^n$ be a linear program in standard form. We have seen in Section 3 that we can transform any linear program into this form, so it suffices to find an algorithm for problems of this type. Geometrically, we want to maximize a linear function $\mathbf{c}^t$ over the polyhedron

$$P := \{\mathbf{x} \mid A\mathbf{x} = \mathbf{b},\ \mathbf{x} \geq \mathbf{0}\}.$$

If some row of A is linearly dependent from the others, than we can discard it without changing the feasible region. Hence, we can in the following assume rank $A = m \leq n$. This reduction can be done efficiently using GAUSSian elimination. We assume further that $P \neq \varnothing$. This is a restriction (we exclude infeasible programs), and we see later how we can detect empty polyhedra.

Given $P \neq \varnothing$, Corollary 4.30(1) tells us that P is pointed, i.e. it has a vertex. From Proposition 4.6 we know that the set of optimal solutions is a face F of P. As P has a vertex also F has a vertex, and we can in the following restrict our search for an optimal solution to vertices of P. It is standard in linear programming to write A_S for the matrix A_{*S}, $S \subseteq [m]$. We will follow this notation in this and the next section.

Corollary 4.28 states that a vertex $\mathbf{v}$ of P corresponds to a solution of

$$A_B \mathbf{v}_B = \mathbf{b}\,, \qquad \mathbf{v}_N = \mathbf{0} \qquad \text{and} \qquad \mathbf{v} \geq \mathbf{0}$$

where $N \subset [n]$ has size $|N| \geq n - \operatorname{rank} A = n - m$, $B := [n] \setminus J$ and the columns of A_B are linearly independent. By adding columns to B we can always assume that $|B| = m$ and A_B is non-singular. We might have a choice here, so the assignment of a set B to a vertex $\mathbf{v}$ need not be unique.

Conversely, let $B \subseteq [m]$ with $|B| = m$ and A_B non-singular and $N := [m] \setminus B$. Then B determines a unique solution

$$\mathbf{v}_B := A_B^{-1}\mathbf{b} \qquad \mathbf{v}_N = \mathbf{0} \tag{6.3}$$

of $A\mathbf{x} = \mathbf{b}$ which is a vertex of P if $\mathbf{v}_B \geq 0$. So the linear program $\max(\mathbf{c}^t\mathbf{x} \mid A\mathbf{x} = \mathbf{b}, \mathbf{x} \geq 0)$ has an optimal solution of the form (6.3) if it is bounded. This motivates the following definition.

basis
basic and non-basic variables
basic solution
basic feasible solution
(non-)degenerate
feasible basis

Definition 6.1. *A subset $B \subset [n]$ is called a* **basis** *if A_B is non-singular. Variables with index in B are called* **basic***, all others are* **non-basic***.*

A **basic solution** *is a solution $\mathbf{x}$ such that $x_j = 0$ for $j \notin B$. If $\mathbf{x}$ is also feasible, then it is a* **basic feasible solution***. A basis B is called* **feasible***, if the corresponding $\mathbf{x}$ is a basic feasible solution. A basic solution is* **non-degenerate** *if no entry of $\mathbf{x}_B$ is zero, and* **degenerate** *otherwise.*

So we have the following relations:

$$\text{vertex} \quad \overset{\text{unique}}{\longleftrightarrow} \quad \text{basic feasible solution} \quad \overset{\text{not unique}}{\longleftrightarrow} \quad \text{feasible basis}$$

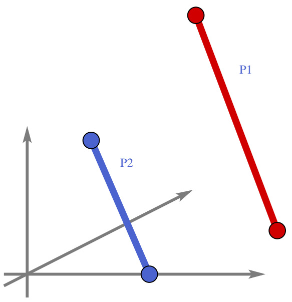


Figure 6.1

Example 6.2. (1) Consider the following matrix $A \in \mathbb{R}^{2\times 3}$, and right hand side vectors $\mathbf{b}_1, \mathbf{b}_2 \in \mathbb{R}^2$, for the basis $B = (1,2)$:

$$A := \left(\begin{smallmatrix} 1 & 1 & 0 \\ 4 & -1 & 5 \end{smallmatrix}\right) \qquad A_B = \left(\begin{smallmatrix} 1 & 1 \\ 4 & -1 \end{smallmatrix}\right) \qquad \mathbf{b}_1 = \left(\begin{smallmatrix} 8 \\ 12 \end{smallmatrix}\right) \qquad \mathbf{b}_1 = \left(\begin{smallmatrix} 3 \\ 12 \end{smallmatrix}\right).$$

This leads to the basic solutions

$$\mathbf{x}_1 = \left(\begin{smallmatrix} 4 \\ 4 \\ 0 \end{smallmatrix}\right) \qquad \mathbf{x}_2 = \left(\begin{smallmatrix} 3 \\ 0 \\ 0 \end{smallmatrix}\right).$$

The first solution is non-degenerate, the second degenerate. The polytopes defined by these inequalities are

$$P_1 := \operatorname{conv}\left(\begin{smallmatrix} 4 & 0 \\ 4 & 8 \\ 0 & 4 \end{smallmatrix}\right) \quad \text{and} \quad P_2 := \operatorname{conv}\left(\begin{smallmatrix} 0 & 3 \\ 3 & 0 \\ 3 & 0 \end{smallmatrix}\right)$$

respectively. See Figure 6.1.

(2) Consider the linear program

$$\begin{array}{rcrcl} x & + & y & \leq & 1 \\ 2x & + & y & \leq & 2 \\ & & x, y & \geq & 0 \end{array}$$

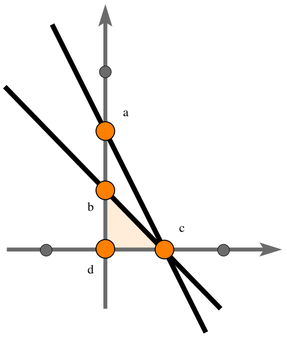


Figure 6.2

The set of solutions is shown in Figure 6.2. The polytope has four vertices $\mathbf{v}_1, \mathbf{v}_2$, $\mathbf{v}_3$, and $\mathbf{v}_4$. Introducing slack variables r and s, we obtain

$$\begin{array}{rcrcrcrcl} x & + & y & + & r & & & = & 1 \\ 2x & + & y & & & + & s & = & 2 \\ & & & & & x, y, & r, s & \geq & 0 \end{array} \,,$$

B	A_B	x_B	x_N	x	vertex	
$(1,2)$	$\begin{pmatrix}1&1\\2&1\end{pmatrix}$	$(x,y)=(1,\ 0)$	(r,s)	$(1,0,\ 0,0)$	$\mathbf{v}_3$	degenerate
$(1,3)$	$\begin{pmatrix}1&1\\2&\end{pmatrix}$	$(x,r)=(1,\ 0)$	(y,s)	$(1,0,\ 0,0)$	$\mathbf{v}_3$	degenerate
$(1,4)$	$\begin{pmatrix}1&\\2&1\end{pmatrix}$	$(x,s)=(1,\ 0)$	(y,r)	$(1,0,\ 0,0)$	$\mathbf{v}_3$	degenerate
$(2,3)$	$\begin{pmatrix}1&1\\1&\end{pmatrix}$	$(y,r)=(2,-1)$	(x,s)	$(0,2,-1,0)$	$\mathbf{v}_1$	infeasible
$(2,4)$	$\begin{pmatrix}1&\\1&1\end{pmatrix}$	$(y,s)=(1,\ 1)$	(x,r)	$(0,1,\ 0,1)$	$\mathbf{v}_2$	
$(3,4)$	$\begin{pmatrix}1&\\&1\end{pmatrix}$	$(r,s)=(1,\ 1)$	(x,y)	$(0,0,\ 1,1)$	$\mathbf{v}_4$	

Table 6.1: Bases and solutions for Example 2

and in matrix form

$$A = \begin{pmatrix}1&1&1&0\\2&1&0&1\end{pmatrix} \qquad\qquad b = \begin{pmatrix}1\\2\end{pmatrix}$$

The bases of this matrix are shown in Table 6.1 In this example each (2×2)-minor is a basis (which is not true in general!). The vertex $\mathbf{v}_3$ corresponds to three different bases, so it is degenerate. All other basic solutions are non-degenerate. The basis $B = (2,3)$ with solution $\mathbf{v}_1$ is infeasible.

(3) The previous example showed that redundant inequalities lead to degenerate solutions of a linear program. The following two examples demonstrate two other types of degeneracies. We'll see later that there are no other.

  (a) Consider

$$A := \left(\begin{smallmatrix}1&1&0\\0&0&1\end{smallmatrix}\right) \qquad\qquad b := \left(\begin{smallmatrix}1\\0\end{smallmatrix}\right).$$

  The corresponding linear program has two bases

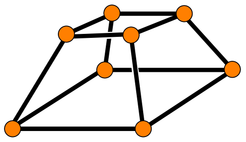
    (i) $B_1 := (1,3)$ with basic solution $x = \left(\begin{smallmatrix}1\\0\\0\end{smallmatrix}\right)$, and

    (ii) $B_2 := (2,3)$ with basic solution $x = \left(\begin{smallmatrix}0\\1\\0\end{smallmatrix}\right)$.

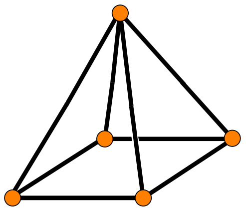
  Both solutions are degenerate, but the basis is in both cases uniquely determined.

  (b) The degeneracies in the previous two examples may seem artificial, as they originate from a redundant description of the polyhedron. We can get rid of them by removing a row (redundant inequality) or column (redundant variable). In dimension 2 any degeneracy is of this type. However, starting with dimension 3 there are degeneracies that are inherent to the geometry of the problem. See Figure 6.3 for an example. The polytope in (a) has no degenerate vertices, while the top vertex in (b) is degenerate (it lies in the intersection of more than 3 hyperplanes).

  Polytope $P \subseteq \mathbb{R}^n$ with the property that at most n hyperplanes meet in a common point are called **simple**. ◊



(a): non-degenerate



(b): degenerate

Figure 6.3

simple polytope

We have seen that any bounded feasible linear program in standard form assumes its optimal value at a vertex of the polyhedron. This is a finite number, so the following is a valid algorithm to compute the optimal value:

(1) compute the finite number of bases,

(2) check whether they are feasible, and in that case
(3) compute the corresponding basic feasible solution and its value.

The maximum over these numbers is the desired optimum. However, there may be up to $\binom{m}{n}$ different feasible bases, which is far too many to enumerate in reasonable time (consider e.g. the case $n = 2m$. Then $\binom{m}{n}$ grows roughly like 4^m, which is exponential in n). So we need a better idea to obtain the optimal value of a linear program.

The basic idea for the approach that we discuss in the rest of this chapter is to move along edges of the polyhedron from one vertex (feasible basis) to an adjacent one with a better objective value, until we reach the optimum. To start this procedure we need an initial feasible solution, i.e. a vertex of the polyhedron. We postpone finding such a vertex and assume first it is given to us and develop a method to move to a better vertex. The outlined approach a priori does not guarantee that we do not enumerate all bases (and there are examples where one does!). Yet, at least experimentally, it seems to be a far superior method than pure enumeration.

In the following, B will always denote a (not necessarily feasible) basis and $N := [n]-B$. To simplify notation and computations, from now on we think of B as a vector in $\mathbb{N}^m$, and similarly $N \in \mathbb{N}^{n-m}$. In particular, B is ordered, and for $B = (1,3,2)$ the matrix A_B contains the columns 1, 3, and 2 in that order. For $1 \leq i \leq n$ we let $B(i)$ be the entry of B at position i, and similarly for N. We will still use the notation $s \in B$ when appropriate.

Example 6.3. This will be our running example for the algorithm. We omit 0's in the matrix to emphasize the positions of the non-zero elements. Let

$$A := \begin{bmatrix} 1 & & 1 & 3 & -1 \\ & 1 & & 1 & \\ & 1 & 1 & 2 & -2 \end{bmatrix} \qquad \mathbf{b} := \begin{bmatrix} 5 \\ 1 \\ 3 \end{bmatrix}$$

Then $B := (1,2,3)$ and $B' := (1,4,3)$ are feasible bases, and

$$A_B = \begin{bmatrix} 1 & & 1 \\ & 1 & \\ & 1 & 1 \end{bmatrix} \qquad A_{B'} = \begin{bmatrix} 1 & 3 & 1 \\ & 1 & \\ & 2 & 1 \end{bmatrix}.$$

◊

Now suppose we are given a basic feasible solution $\overline{\mathbf{x}}$ for some basis B, i.e.

$$\overline{\mathbf{x}}_b \geq \mathbf{0} \qquad \overline{\mathbf{x}}_N = \mathbf{0}.$$

Multiplying a system of linear equations with a non-singular matrix does not change the set of solutions. By assumption, A_B is a non-singular matrix, so we can multiply $A\mathbf{x} = \mathbf{b}$ with A^{-1} on both sides to obtain

$$A_B^{-1}\mathbf{b} \;=\; A_B^{-1}A\mathbf{x} \;=\; \mathbf{x}_B + A_B^{-1}A_N\mathbf{x}_N\,.$$

This immediately implies $\overline{\mathbf{x}}_B = A_B^{-1}\mathbf{b}$, as $\overline{\mathbf{x}}_N = \mathbf{0}$. We can achieve this transformation with Gaussian elimination.

Example 6.3 continued. Using the basis $B = (1,2,3)$ we have

$$A_B^{-1} = \begin{bmatrix} 1 & 1 & -1 \\ & 1 & \\ & -1 & 1 \end{bmatrix} \qquad A_B^{-1}A = \begin{bmatrix} 1 & & & 2 & 1 \\ & 1 & & 1 & \\ & & 1 & 1 & -2 \end{bmatrix}$$

$$\overline{\mathbf{x}}_B = A_B^{-1}\mathbf{b} = \begin{bmatrix} 3 \\ 1 \\ 2 \end{bmatrix} \geq \mathbf{0}$$

◊

Remark 6.5. Sometimes it is easy to find an initial feasible basis for a linear program. Consider the program in canonical form

$$\max(\mathbf{c}^t\mathbf{x} \mid A\mathbf{x} \leq \mathbf{b},\, x \geq \mathbf{0})$$

for a vector $\mathbf{b}$ that satisfies $\mathbf{b} \geq \mathbf{0}$. If we transform this program into standard form,

$$\max(\mathbf{c}^t\mathbf{x} \mid A\mathbf{x} + \mathrm{Id_m}\mathbf{s} = \mathbf{b},\, \mathbf{x}, \mathbf{s} \geq \mathbf{0})\,,$$

then $\overline{\mathbf{x}} := \mathbf{0}$, $\overline{\mathbf{s}} := \mathbf{b}$ is a basic feasible solution. ◊

A key observation for the algorithm is that the set of optimal solution of our linear program is invariant under adding a scalar multiple of a row of A to the objective functional $\mathbf{c}^t$, i.e. for some $i \in [m]$ and $\lambda \in \mathbb{R}$ the linear programs

$$\max(\mathbf{c}^t\mathbf{x} \mid A\mathbf{x} = \mathbf{b}\,,\, ,\mathbf{x} \geq 0)$$

and

$$\max((\mathbf{c}^t + \lambda A_{i*})\mathbf{x} \mid A\mathbf{x} = \mathbf{b}\,,\, \mathbf{x} \geq 0)$$

have the same optimal solutions. The optimal value of the program is not invariant, it changes by λb_i. A very convenient transformation of the objective function $\mathbf{c}^t$ is

$$\overline{\mathbf{c}}^t := \mathbf{c}^t - \mathbf{c}^t A_B^{-1} A\,.$$

Then

$$\overline{\mathbf{c}}^t_B = \mathbf{c}^t_B - \mathbf{c}^t_B A_B^{-1} A_B = \mathbf{0} \qquad \text{and} \qquad \overline{\mathbf{c}}^t_N = \mathbf{c}^t_N - \mathbf{c}^t_B A_B^{-1} A_N\,,$$

so $\overline{\mathbf{c}}^t\overline{\mathbf{x}} = 0$ for our basic feasible solution. The corresponding value of the original program is $\mathbf{c}^t_B A_B^{-1}\mathbf{b}$. The functional $\overline{\mathbf{c}}^t$ is the **reduced cost** (of the basis B).

reduced cost

Example 6.3 continued.

$$\text{Let} \quad \mathbf{c}^t := \begin{bmatrix} 1 & 0 & 1 & 4 & 1 \end{bmatrix}, \quad \text{then} \quad \overline{\mathbf{c}}^t = \begin{bmatrix} 0 & 0 & 0 & 1 & 2 \end{bmatrix}$$ ◊

It is common and very convenient to write the reduced costs $\overline{\mathbf{c}}^t$ and the transformed matrix $A_B^{-1}A$ and its corresponding right hand side $A_B^{-1}\mathbf{b}$ in the following concise tableau form.

$$T_B := \left[\begin{array}{c|c} -\mathbf{c}^t_B A_B^{-1}\mathbf{b} & \overline{\mathbf{c}}^t \\ \hline A_B^{-1}\mathbf{b} & A_B^{-1}A \end{array}\right] \tag{6.4}$$

The offset $-\mathbf{c}^t_B A_B^{-1}\mathbf{b}$ induced by the change in the objective function is placed in the upper right corner. This corresponds to the objective value of the basic solution $\overline{\mathbf{x}}_B := A_B^{-1}\mathbf{b}$ corresponding to the basis B, as $\overline{\mathbf{x}}_N = 0$. Hence, if $\mathbf{x}$ is an optimal basic solution, then $\mathbf{c}^t_B A_B^{-1}\mathbf{b}$ is the corresponding optimal value.

***Definition 6.7**. T_B is the* **simplex tableau** *associated to the basis B.*

simplex tableau

Example 6.3 continued. We have chosen $B = (1, 2, 3)$ as a basis. So

$$T_B = \left[\begin{array}{r|rrrrr} -5 & & & & 1 & 2 \\ \hline 3 & 1 & & & 2 & 1 \\ 1 & & 1 & & 1 & \\ 2 & & & 1 & 1 & -2 \end{array}\right]$$

◊

The simplex algorithm works on these tableaus. Roughly, in each step the algorithm checks whether the current basic solution is optimal, and if not, determines a new basis of the linear program corresponding to a basic solution with larger optimal value, and updates the tableau accordingly. In the following we work out one step of this algorithm.

First we need a criterion to decide whether our current basic feasible solution $\overline{\mathbf{x}}$ for a feasible basis B is optimal, and if not, a way to improve the basis. $\overline{\mathbf{c}}^t$ and $\mathbf{c}^t$ differ only by a fixed linear combination of the rows of A, so we can consider $\overline{\mathbf{c}}^t$ in the following. As $\overline{\mathbf{c}}^t_B = \mathbf{0}$, a change in a coordinate of $\mathbf{x}_B$ does not affect the objective value. So at least one of the non-basic variables $\mathbf{x}_N$ must change (which then possibly also affects basic variables). As $\mathbf{x} \geq \mathbf{0}$ for all feasible solutions, it must increase from zero to some positive value.

First assume that $\overline{\mathbf{c}}^t_N \leq 0$. Let $\mathbf{z} \in P$. Then $\overline{\mathbf{c}}^t_N \mathbf{z}_N \leq 0$ and hence

$$\mathbf{c}^t\mathbf{z} = \overline{\mathbf{c}}^t\mathbf{z} + \mathbf{c}^t_B A_B^{-1} A\mathbf{z} = \overline{\mathbf{c}}^t_N \mathbf{z}_N + \mathbf{c}^t_B A_B^{-1}\mathbf{b} \leq \mathbf{c}^t_B A_B^{-1}\mathbf{b} = \mathbf{c}^t_B \overline{\mathbf{x}}_B = \mathbf{c}^t\overline{\mathbf{x}},$$

which implies that $\overline{\mathbf{x}}$ is an optimal solution. Hence, we have found a simple termination condition for our algorithm: If the reduced costs are non-positive, then the corresponding basic solution is optimal.

Now assume that there is $r \in N$ such that $\overline{c}_r > 0$. Let r' be such that $N(r') = r$. $\mathbf{e}^{(r)} \in \mathbb{R}^n$ be the r-th unit vector. Then there is $\lambda \geq 0$ such that

$$\overline{\mathbf{x}}_B \geq \lambda\, A_B^{-1} A\, \mathbf{e}^{(r)}. \tag{6.5}$$

If $\overline{\mathbf{x}}$ is non-degenerate, then $\overline{\mathbf{x}}_B = A_B^{-1}\mathbf{b} > \mathbf{0}$ and we can choose $\lambda > 0$. Define a vector $\mathbf{x}^\lambda$ by

$$\mathbf{x}^\lambda_B := \overline{\mathbf{x}}_B - \lambda\, A_B^{-1} A_N\, \mathbf{e}^{(r)}_N, \qquad \mathbf{x}^\lambda_N := \lambda\, \mathbf{e}^{(r)}_N. \tag{6.6}$$

Then $\mathbf{x}^\lambda \geq \mathbf{0}$ and

$$A\mathbf{x}^\lambda = A_B\mathbf{x}^\lambda_B + A_N\mathbf{x}^\lambda_N = A_B\overline{\mathbf{x}}_B - A_B\,\lambda\, A_B^{-1}A_N\,\mathbf{e}^{(r)}_N + A_N\,\lambda\,\mathbf{e}^{(r)}_N = A_B\overline{\mathbf{x}}_B = \mathbf{b},$$

so $\mathbf{x}^\lambda \in P$ and

$$\begin{aligned}\mathbf{c}^t\mathbf{x}^\lambda \; &= \mathbf{c}^t_B\overline{\mathbf{x}}_B - \mathbf{c}^t_B\,\lambda\, A_B^{-1}A_N\,\mathbf{e}^{(r)}_N + \mathbf{c}^t_N\,\lambda\,\mathbf{e}^{(r)}_N = \\ &\quad \mathbf{c}^t\overline{\mathbf{x}} + (\mathbf{c}^t_N - \mathbf{c}^t_B A_B^{-1}A_N)\,\lambda\,\mathbf{e}^{(r)}_N \;=\; \mathbf{c}^t\overline{\mathbf{x}} + \lambda\,\overline{c}_r \geq \mathbf{c}^t\overline{\mathbf{x}}.\end{aligned}$$

So we can change our solution $\overline{\mathbf{x}}$ by increasing the r-th entry of $\overline{\mathbf{x}}$ from 0 to λ and simultaneously decreasing $\overline{\mathbf{x}}_B$ by $\lambda\, A_B^{-1}A\,\mathbf{e}^{(r)}$. The new objective value differs from the old by $\lambda\overline{c}_r$. If $\overline{\mathbf{x}}$ is non-degenerate, then we can choose $\lambda > 0$, so $\lambda\overline{c}_r > 0$, and the objective value really increases by a strictly positive amount.

Example 6.3 continued. In our example, $\overline{c}_4 = 1 > 0$, so we can choose $r = 4$. If we choose $\lambda = 1/2$, then

$$\mathbf{x}^\lambda_B = \begin{bmatrix} 3 \\ 1 \\ 2 \end{bmatrix} - \frac{1}{2}\begin{bmatrix} 2 \\ 1 \\ 1 \end{bmatrix} = \begin{bmatrix} 2 \\ 1/2 \\ 3/2 \end{bmatrix} \qquad \mathbf{x}^\lambda_N = \begin{bmatrix} 1/2 \\ 0 \end{bmatrix}.$$

The objective value increases by $\lambda\overline{c}_4 = 1/2 \cdot 1 = 1/2$. ◊

The condition

$$\mathbf{x}^\lambda_B \;=\; \overline{\mathbf{x}}_B - \lambda\, A_B^{-1}A_N\,\mathbf{e}^{(r)}_N \;\geq\; \mathbf{0} \qquad \Longleftrightarrow \qquad \overline{\mathbf{x}}_B \;\geq\; \lambda\, A_B^{-1}A_N\,\mathbf{e}^{(r)}_N \tag{6.7}$$

tells us how large we can choose λ in this process. If $A_B^{-1}A_N\,\mathbf{e}^{(r)}_N \leq \mathbf{0}$, then any $\lambda > 0$ satisfies (6.7), and our linear program is unbounded. If at least one entry of $A_B^{-1}A_N\,\mathbf{e}^{(r)}_N$

is positive, then (6.7) gives a proper restriction for λ. We want to determine the maximal λ so that $\mathbf{x}_B^\lambda \geq 0$. To simplify the notation we introduce $\overline{\mathbf{b}} := A_B^{-1}\mathbf{b}$ and $\overline{A} := A_B^{-1}A$ and define

$$\lambda := \min\left(\frac{\overline{b}_i}{\overline{A}_{ir}} \;\middle|\; i \in [m], \overline{A}_{ir} > 0 \right) \geq 0\,, \tag{6.8}$$

Then (6.7) is satisfied, but any $\lambda' > \lambda$ would violate it. Choose a row s' at which this minimum is attained and set $s := B(s')$. Let B' be the vector obtained from B by replacing the s'-th entry with r (i.e. s with r).

By our choice of s' we have $\overline{A}_{s'r} > 0$. Hence, using Gaussian elimination, we can transform the r-th column of $\overline{A}$ into the unit vector $\mathbf{e}^{(s)} \in \mathbb{R}^m$. Formally, these elimination steps can be achieved by multiplication of $\overline{A}$ with an invertible matrix $\Lambda_{BB'}^{-1}$ from the left,

$$M \;=\; \Lambda_{BB'}^{-1} A_B^{-1} A\,. \tag{6.9}$$

By construction we have $(A_{s'*})_B = \mathbf{e}^{(s)}$, so $M_{B'} = \mathrm{Id}_\mathrm{m}$. Now $M_{B'} = (\Lambda_{BB'}^{-1}A_B^{-1})A_{B'}$, so $A_{B'}^{-1} = \Lambda_{BB'}^{-1}A_B^{-1}$ and B' is a basis. Hence, B' is a feasible basis.

For later use we note the exact form of the transformation matrix Λ used in the proof. We can rewrite the equation $\Lambda_{BB'}A_{B'}^{-1} = A_B^{-1}$ in the form $A_{B'} = A_B\Lambda_{BB'}$, so $\Lambda_{BB'}$ is a unit matrix in which the s-th column is replaced by the r-th column of A_BA, i.e. the solution of the linear system of equations $A\mathbf{d} = A_{*r}$. Its inverse $\Lambda_{BB'}^{-1}$ is then easily seen to be a unit matrix in which the s-th column is replaced by a vector λ defined via

$$\lambda_j := \begin{cases} \frac{1}{\overline{A}_{s'r}} & j = s' \\ -\frac{\overline{A}_{jr}}{\overline{A}_{s'r}} & \text{otherwise} \end{cases}$$

We fix some more common linear programming terminology.

Definition 6.10. *The index s is the* **leaving index***, the index r the* **entering index** *in this step of the algorithm. The corresponding variables of $\overline{\mathbf{x}}$ are the* **leaving** *and* **entering variable***. The element $\overline{A}_{s'r}$ is the* **pivot element** *of this step.*

leaving/entering index
leaving/entering variable
pivot element

Example 6.3 continued. If we choose $j = 4$, then $\overline{c}_4 = 1 > 0$. and (6.8) tells us that we can choose $\lambda := \min(3/2, 1, 2) = 1$. Then

$$\mathbf{x}_B^\lambda = \begin{bmatrix} 3 \\ 1 \\ 2 \end{bmatrix} - 1 \cdot \begin{bmatrix} 2 \\ 1 \\ 1 \end{bmatrix} = \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix} \qquad\qquad \mathbf{x}_N^\lambda = \begin{bmatrix} 1 \\ 0 \end{bmatrix}.$$

The objective value increases by 1, and the new basis is $B' = (1, 4, 3)$. ◊

Note that existence of a pivot element $\overline{A}_{s'r} > 0$ does not necessarily mean that the program is bounded. We just cannot detect this at the current step of the algorithm.

There is one case we have not dealt with yet. If $\overline{c}_r > 0$ for some $r \in N$ and $A_B^{-1}A_{*r}$ has positive entries, but $\mathbf{x}$ is degenerate, then λ may be 0. We can still move from the basis B to the new basis B' by exchanging r with s, but the objective value does not increase. This can indeed occur, and we postpone a discussion of a possible solution to the end of this chapter.

To finish the step of the algorithm we have to update the tableau to the new basis. Given B' we can of course compute a new simplex tableau by inverting $A_{B'}$. However, on a computer inverting a matrix is expensive and numerically unstable, so we want to avoid this.

We have already seen above that we can move from $A_B^{-1}A$ and $A_{B'}^{-1}A$ by multiplying with $\Lambda_{BB'}^{-1}$ from the left. Now

$$A_{B'}^{-1}A \;=\; \Lambda_{BB'}^{-1}A_B^{-1}A \qquad\qquad \text{implies} \qquad\qquad A_{B'}^{-1}\mathbf{b} = \Lambda_{BB'}^{-1}A_B^{-1}\mathbf{b}\,.$$

Hence, the lower part of the tableau can be obtained with Gaussian elimination using only the s'-th row, which we apply simultaneously both to the matrix and the left column.

It remains to update the reduced costs. Consider the vector

$$\tilde{\mathbf{c}}^t := \overline{\mathbf{c}} - \mu \overline{A}_{s'*}$$

for some $\mu \in \mathbb{R}$. The entries $\overline{A}_{s'r}$ and $\overline{c}_r$ are positive, and the only non-zero entry among those in B is $\overline{A}_{s's}$. Hence, we can find $\mu > 0$ such that $\tilde{\mathbf{c}}^t_{B'} = \mathbf{0}$. The following proposition implies that then already $\tilde{\mathbf{c}}^t = \mathbf{c}^t - \mathbf{c}^t_{B'} A^{-1}_{B'} A$, i.e. $\tilde{\mathbf{c}}^t$ are the reduced cost for the new basis.

Proposition 6.12. *If B is a basis and $\boldsymbol{\lambda}^t \in \mathbb{R}^n$ satisfies $(\mathbf{c}^t - \boldsymbol{\lambda}^t A)_B = \mathbf{0}$, then $\boldsymbol{\lambda}^t = \mathbf{c}^t_N A^{-1}_B$.*

Proof. Suppose we have $\boldsymbol{\lambda}, \boldsymbol{\mu} \in \mathbb{R}^n$ such that $\mathbf{0} = (\mathbf{c}^t - \boldsymbol{\lambda}^t A)_B = (\mathbf{c}^t - \boldsymbol{\mu}^t A)_B$. Subtracting the two equations gives $\mathbf{0} = (\boldsymbol{\lambda}^t - \boldsymbol{\mu}^t) A_B$. As B is a basis, A_B is non-singular, so the only solution of $\mathbf{y}^t A_B = \mathbf{0}$ is $\mathbf{y} = \mathbf{0}$. Hence $\boldsymbol{\lambda} = \boldsymbol{\mu}$. □

A similar reasoning as before shows that extending this operation on the left column gives the new objective value. Thus we have obtained the complete simplex tableau in the new basis. This completes the step in the algorithm.

The crucial property of the simplex tableau for this procedure to work is that

$$\overline{\mathbf{c}}^t_B = \mathbf{0} \qquad \text{and} \qquad \overline{\mathbf{x}}_N = \mathbf{0}, \tag{6.10}$$

i.e. the non-zero entries of $\overline{\mathbf{c}}$ and $\overline{\mathbf{x}}$ have disjoint support. This implied both the simple optimality test and the way to choose a leaving and entering index. The final update of the tableau is necessary as after a change of basis (6.10) is violated.

Summarizing, we have now developed the following rule to transform a tableau of one basis into the the tableau of another basis.

pivot element

> Choose a **pivot element** $\overline{A}_{s'r} > 0$ in the tableau.
>
> Perform Gaussian elimination by row operations in such a way that the r-th column becomes the $(m+1)$-dimensional unit vector with a 1 at position $(s+1)$ (which is the position of the s'-th row of A).

Example 6.3 continued. We have chosen $B = (1, 2, 3)$ as a basis. So

$$T_B = \left[\begin{array}{c|ccccc} -5 & & & & 1 & 2 \\ \hline 3 & 1 & & & 2 & 1 \\ 1 & & 1 & & \boxed{1} & \\ 2 & & & 1 & 1 & -2 \end{array}\right]$$

We want to bring $r = 4$ into the basis. The corresponding column is

$$\overline{A}_{*r} = \begin{bmatrix} 2 \\ 1 \\ 1 \end{bmatrix} \qquad \text{and} \qquad \overline{b} := \begin{bmatrix} 3 \\ 1 \\ 2 \end{bmatrix}.$$

To determine λ_0 we have to find the minimum of $\frac{3}{2}, 1$ and 2. The minimum of these numbers is $\lambda_0 = 1$, and it is attained for the second row, so $s = 2$. We mark the chosen pivot element in the current tableau by a box around that entry. Performing one step of the simplex algorithm gives the new tableau

$$A^{-1}_{B'} = \begin{bmatrix} 1 & -1 & -1 \\ & 1 & \\ & -2 & 1 \end{bmatrix} \qquad T_{B'} = \left[\begin{array}{c|ccccc} -6 & & -1 & & & 2 \\ \hline 1 & 1 & -2 & & & 1 \\ 1 & & 1 & & 1 & \\ 1 & & -1 & 1 & & -2 \end{array}\right]$$

for the new basis $B' = (1, 4, 3)$. ◊

Repeating this step until the reduced cost become non-positive now computed the optimal solution, which completes the simplex algorithm in the case that we are already given a feasible initial basis of our linear program. This algorithm is commonly called the **phase II** of the simplex algorithm. The following table summarizes the different steps.

phase II

THE SIMPLEX ALGORITHM (PHASE II)

INPUT: $A \in \mathbb{R}^{m\times n}$, $\mathbf{b} \in R^m$, and $\mathbf{c} \in \mathbb{R}^n$.
$P := \{\mathbf{x} \mid A\mathbf{x} = b,\ \mathbf{x} \geq 0\} \neq \varnothing$
B feasible basis.

OUTPUT: An optimal solution of

$$\max(\mathbf{c}^t\mathbf{x} \mid A\mathbf{x} = \mathbf{b},\ \mathbf{x} \geq 0) = \max(\mathbf{c}^t\mathbf{x} \mid \mathbf{x} \in P)$$

if it exists, and the information that the system is unbounded, otherwise.

ALGORITHM:

(1) Compute the simplex tableau T_B for the basis B.
(2) (OPTIMALITY TEST) If $\overline{\mathbf{c}} \leq \mathbf{0}$, then B is optimal and we stop.
(3) (ENTERING COLUMN) If not, then we choose an index r, such that $\overline{c}_r > 0$.
(4) (BOUNDEDNESS) If the corresponding column of $A_B^{-1}A$ is non-positive, then the program is unbounded, and we stop.
(5) (PIVOT ELEMENT) Otherwise, choose an index s' such that $\frac{\overline{b}_r}{\overline{A}_{s'r}} \geq 0$ is minimal. Let $s := B(s')$.
(6) (EXCHANGE) Exchange r and s in both vectors B and N.
(7) (UPDATE) Update the tableau T_B to the new basis using Gaussian elimination.
(8) Return to (2).

The algorithm in this form only works if in each step the basic feasible solution is non-degenerate. There is still choice in the algorithm as we have not yet fixed how we choose the entering and leaving index. We will see later that we can consistently deal with non-degenerate solutions by fixing a rule how to choose the entering and leaving variable. Obvious choices are to take always the smallest index, or the one with the largest increase in the objective function, or to choose at random.

Example 6.3 continued. We finish our example and compute an optimal solution. Given the tableau for the basis B' taking $r = 5$ is only choice. Also for the pivot element $s' = 1$ is the only possible choice, as all other entries in the last column are non-positive. This gives $\lambda_0 = \frac{1}{1} = 1$. Using $B'(1) = 1$ we obtain the new basis $B'' = (5, 4, 3)$, $N'' = (2, 1)$ with associated tableau

$$A_{B''}^{-1} = \begin{bmatrix} 1 & -1 & -1 \\ & 1 & \\ 2 & -4 & 1 \end{bmatrix} \qquad T_{B''} = \left[\begin{array}{r|rrrrr} -8 & -2 & 3 & & & \\ \hline 1 & 1 & -2 & & & 1 \\ 1 & & \boxed{1} & & 1 & \\ 3 & 2 & -5 & 1 & & \end{array}\right]$$

The only choices in the new tableau are $r = 2$ and $s' = 2$. The corresponding pivot element is marked with a box in the above tableau. So we obtain a new basis $B''' := (5, 2, 3)$, $N''' := (4, 1)$ with tableau

$$A_{B'''}^{-1} = \begin{bmatrix} -1 & & 1 \\ & 1 & \\ -2 & 1 & 1 \end{bmatrix} \qquad T_{B'''} := \left[\begin{array}{r|rrrrr} -11 & -2 & & & -3 & \\ \hline 3 & 1 & & & 2 & 1 \\ 1 & & 1 & & 1 & \\ 8 & 2 & & 1 & 5 & \end{array}\right]$$

Now $\bar{c} \leq 0$, and we have reached the optimal value 11 for the program. The associated solution is

$$\bar{\mathbf{x}}_{B'''} = \begin{bmatrix} 3 \\ 1 \\ 8 \end{bmatrix}. \qquad \text{Reordering the entries and extending gives} \qquad \bar{\mathbf{x}} = \begin{bmatrix} 0 \\ 1 \\ 8 \\ 0 \\ 3 \end{bmatrix}$$

◊

Before we discuss degenerate solutions we want to find a way to compute an initial basic feasible solution for the linear programming problem

$$\max(\mathbf{c}^t\mathbf{x} \mid A\mathbf{x} = \mathbf{b},\, \mathbf{x} \geq \mathbf{0}) \;=\; \max(\mathbf{c}^t\mathbf{x} \mid \mathbf{x} \in P)\,, \tag{P}$$

which we left open above. We may assume that $\mathbf{b} \geq 0$, as multiplication of an equation in $A\mathbf{x} = \mathbf{b}$ by -1 does not change the solution set. We consider the following auxiliary program:

$$\delta := \max(-\mathbf{1}^t\mathbf{z} \mid \mathbf{z} + A\mathbf{x} = \mathbf{b},\, \mathbf{x}, \mathbf{z} \geq \mathbf{0})\,. \tag{AP}$$

This has the obvious basic feasible solution

$$B := (1, \ldots, m) \qquad\qquad \bar{\mathbf{z}} := \mathbf{b} \qquad\qquad \bar{\mathbf{x}} := \mathbf{0}\,,$$

so we can apply the phase II developed above. If (P) has a feasible solution $\bar{\mathbf{x}}$, then $(0, \bar{\mathbf{x}})$ is a feasible solution of (AP) with objective value $\delta = 0$. As $\delta \leq 0$ for all feasible $(\mathbf{z}, \mathbf{x})$, we have

$$\bar{\mathbf{x}} \text{ feasible for (P)} \quad \Longleftrightarrow \quad (0, \bar{\mathbf{x}}) \text{ is optimal for (AP) with optimal value } \delta = 0.$$

Now (AP) is bounded, so by duality there is an optimal solution $(\bar{\mathbf{z}}, \bar{\mathbf{x}})$ with optimal value $\bar{\delta} \leq 0$ and a basis $\bar{B} \subseteq \{1, \ldots, m+n\}$. We distinguish the two cases $\bar{\delta} < 0$ and $\bar{\delta} = 0$.

(1) $\delta < 0$: Then P is infeasible, i.e. $P = \varnothing$, as any $\mathbf{x} \in P$ would be optimal for (AP) with value 0.

(2) $\bar{\delta} = 0$: This implies $\bar{\mathbf{z}} = 0$. As $\bar{\mathbf{x}} \geq \mathbf{0}$ and $A\bar{\mathbf{x}} = \mathbf{b}$, $\bar{\mathbf{x}}$ is feasible for (P). Further, $I := \{i \mid \bar{x}_i > 0\} \subseteq \bar{B}$ and $|I| \leq m$, so we can extend I to a feasible basis B of (P) with basic feasible solution $\bar{\mathbf{x}}$. B is unique if $\bar{\mathbf{x}}$ is non-degenerate.

phase I

Solving this auxiliary program to obtain an initial feasible basis to start phase II of the simplex algorithm is commonly called **phase I** of the simplex algorithm.

Example 6.15. Let

$$A := \begin{pmatrix} 2 & 3 & 1 \\ 1 & 2 & 1 \end{pmatrix} \qquad \mathbf{b} := \begin{pmatrix} 5 \\ 3 \end{pmatrix} \qquad \mathbf{c}^t := \begin{pmatrix} 1 & 4 & 2 \end{pmatrix}.$$

We compute the reduced costs $\bar{\mathbf{d}}^t$ for the auxiliary program as

$$\begin{aligned} \bar{\mathbf{d}}^t &:= \begin{pmatrix} -1 & -1 & 0 & 0 & 0 \end{pmatrix} + \begin{pmatrix} 1 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 & 2 & 3 & 1 \\ 0 & 1 & 1 & 2 & 1 \end{pmatrix} \\ &= \begin{pmatrix} 0 & 0 & 3 & 5 & 2 \end{pmatrix}. \end{aligned}$$

So applying phase II of the simplex algorithm to the auxiliary problem gives the following tableaus. We again mark the chosen pivot elements in each step with a box.

$$\widetilde{T}_{(1,2)} := \left[\begin{array}{c|ccccc} 8 & & & 3 & 5 & 2 \\ \hline 5 & 1 & & 2 & 3 & 1 \\ 3 & & 1 & 1 & 2 & \boxed{1} \end{array}\right]$$

$$\widetilde{T}_{(1,5)} := \left[\begin{array}{c|ccccc} 2 & & -2 & 1 & 1 & \\ \hline 2 & 1 & -1 & \boxed{1} & 1 & \\ 3 & & 1 & 1 & 2 & 1 \end{array}\right]$$

$$\widetilde{T}_{(3,5)} := \left[\begin{array}{c|ccccc} & -1 & -1 & & & \\ \hline 2 & 1 & -1 & 1 & 1 & \\ 1 & -1 & 2 & & 1 & 1 \end{array}\right]$$

This tableau is optimal with basis $(3,5)$.

Removing the auxiliary variables from the tableau we obtain the feasible basis $B = (1,3)$ for the original problem with feasible basic solution $\overline{\mathbf{x}}_B^t = \begin{pmatrix} 2 & 0 & 1 \end{pmatrix}$. ◊

The final tableau of the phase I contains almost all necessary information to write down the initial tableau for the phase II of the algorithm.

(1) The first column contains $A_B^{-1}\mathbf{b}$.
(2) The next m columns contain A_B^{-1}, the rest is $A_B^{-1}A$.
(3) The reduced costs are $\overline{\mathbf{c}}^t = \mathbf{c}^t - \mathbf{c}_B^t A_B^{-1} A$ and the current objective value is $-\mathbf{c}_B A_B^{-1} b$.

Example 6.15 continued. We have

$$A_B^{-1} = \begin{pmatrix} 1 & -1 \\ -1 & 2 \end{pmatrix} \qquad A_B^{-1}A = \begin{pmatrix} 1 & 1 & \\ & 1 & 1 \end{pmatrix} \qquad A_B^{-1}\mathbf{b} = \begin{pmatrix} 2 \\ 1 \end{pmatrix}$$

To start the simplex algorithm on the original program we have to compute the reduced costs for the original cost function in this basis.

$$\overline{\mathbf{c}}^t := \begin{pmatrix} 1 & 4 & 2 \end{pmatrix} + \begin{pmatrix} 1 & 2 \end{pmatrix} \begin{pmatrix} 1 & 1 & 0 \\ 0 & 1 & 1 \end{pmatrix} = \begin{pmatrix} 0 & 1 & 0 \end{pmatrix}.$$

So we obtain the following sequence of tableaus from the simplex algorithm.

$$T_{(1,3)} := \left[\begin{array}{c|ccc} -4 & & 1 & \\ \hline 2 & 1 & 1 & \\ 1 & & \boxed{1} & 1 \end{array}\right] \qquad T_{(1,2)} := \left[\begin{array}{c|ccc} -5 & & & -1 \\ \hline 1 & 1 & & -1 \\ 1 & & 1 & 1 \end{array}\right]$$

so $\overline{\mathbf{x}}^t = \begin{pmatrix} 1 & 1 & 0 \end{pmatrix}$ is an optimal solution with value 5. ◊

This example was particularly nice as the optimal solution of the auxiliary program is non-degenerate, which implies that the basis only contains rows of the original matrix. This allowed us to obtain the initial tableaux for the original problem by deleting the auxiliary columns. For a degenerate solution there might remain auxiliary variables in the basis. The restriction of the solution $(\mathbf{z}, \mathbf{x})$ to $\mathbf{x}$ is still a basic feasible solution for the original problem, and we can compute the corresponding basis and transform the matrix accordingly.

The following observation tells us how to avoid this. If s is the auxiliary index in the final basis B, and s' the corresponding row (i.e. $B(s') = s$), then $\overline{b}_{s'} = 0$ in $\overline{\mathbf{b}} := A_B^{-1}\mathbf{b}$. So we can safely multiply that row by -1 without changing the linear program. So to remove an auxiliary variable from the basis we can just look at the row corresponding to it, take any non-zero element corresponding to an original variable and do a pivot step on this element. This replaces the auxiliary variable by an original one in the basis. Observe that there must be a non-zero element as the original matrix A has full row rank.

Example 6.17.

$$A := \begin{pmatrix} 1 & 1 & 1 \\ 1 & 1 & -1 \end{pmatrix} \qquad \mathbf{b} := \begin{pmatrix} 6 \\ 3 \end{pmatrix} \qquad \mathbf{c}^t := \begin{pmatrix} 2 & 1 & 3 \end{pmatrix}.$$

Then

$$\tilde{T}_{(1,2)} := \left[\begin{array}{c|ccccc} 9 & & & 3 & 2 & \\ \hline 6 & 1 & & 2 & 1 & 1 \\ 3 & & \boxed{1} & 1 & -1 & \end{array}\right] \qquad \tilde{T}_{(1,3)} := \left[\begin{array}{c|ccccc} 0 & & -3 & & -1 & 3 \\ \hline 0 & 1 & -2 & & \boxed{-1} & 3 \\ 3 & & 1 & 1 & 1 & -1 \end{array}\right]$$

This leads to the feasible basis $\overline{\mathbf{x}}^t = \begin{pmatrix} 3 & 0 & 0 \end{pmatrix}$ that we can use to start phase I of the algorithm. However, the solution is degenerate and the variable z_1 is still in the basis. Feasible bases for the original problem are (up to permuting the entries) $B = (1,2)$ and $B = (1,3)$.

We do one more pivot step to remove z_1 from the basis.

$$\tilde{T}_{(4,3)} := \left[\begin{array}{c|ccccc} 0 & -1 & -1 & & & \\ \hline 0 & -1 & 2 & & 1 & -3 \\ 3 & 1 & -1 & 1 & & 2 \end{array}\right]$$

We obtain the basis $B = (2,1)$ with associated basic solution solution $\mathbf{x}_B = \begin{pmatrix} 0 \\ 3 \end{pmatrix}$ for the original problem. Now B does not contain auxiliary variables anymore, so we obtain an initial tableaux of the original problem by deleting the auxiliary columns and recomputing the reduced costs.

$$T_{(2,1)} := \left[\begin{array}{c|ccc} -6 & & & 2 \\ \hline 0 & & 1 & -3 \\ 3 & 1 & & \boxed{2} \end{array}\right] \qquad T_{(2,3)} := \left[\begin{array}{c|ccc} -9 & -2 & & \\ \hline 9/2 & 3/2 & 1 & \\ 3/2 & 1/2 & & 1 \end{array}\right]$$

so

$$\overline{\mathbf{x}} = \begin{pmatrix} 0 \\ 9/2 \\ 3/2 \end{pmatrix}$$

is an optimal solution of the problem. ◊

two-step method

Combining this phase I to obtain an initial feasible basic solution with the phase II described above is commonly called the **two-step method** for linear programs. We summarize the complete simplex algorithm.

THE SIMPLEX ALGORITHM

INPUT: $A \in \mathbb{R}^{m\times n}$, $\mathbf{b} \in R^m$, and $\mathbf{c} \in \mathbb{R}^n$.

OUTPUT: An optimal solution of

$$\max(\mathbf{c}^t\mathbf{x} \mid A\mathbf{x} = \mathbf{b},\ \mathbf{x} \geq 0)$$

if it exists, and the information that the system is unbounded or infeasible otherwise.

(1) Reduce the matrix to rank $A = m \geq n$.

(2) If $m = n$, then use Gaussian elimination to compute the unique solution $\overline{\mathbf{x}}$ and check whether $\overline{\mathbf{x}} \geq 0$. If so, output $\overline{\mathbf{x}}$, otherwise the program is infeasible. STOP.

(3) Solve the auxiliary linear program

$$\tilde{\delta} = \max(-\mathbf{1}^t\mathbf{z} \mid \mathbf{z} + A\mathbf{x} = \mathbf{b},\ \mathbf{x}, \mathbf{z} \geq \mathbf{0}).$$

If $\tilde{\delta} < 0$ then the program is infeasible. STOP.

(4) Start phase II with the initial tableau obtained in phase I.

Theorem 6.18. *If all basic feasible solutions are non-degenerate, then the simplex algorithm stops after a finite number of steps.*

Proof. There are only a finite number of bases. In each step the objective value strictly increases, so the algorithm visits each basis at most once. □

We have not fixed a rule to decide which variable we choose as entering and leaving variable if there is more than one possibility. The assumption that all solutions are non-degenerate is essential in the above theorem as long as we don't specify such a rule.

We never return to a basis with lower objective value. So, if the above method does not terminate, then it must stay at a degenerate basic solution. As there are only finitely many bases it must cycle between bases $B_1, \ldots, B_k$ defining this basic solution. This effect is called **cycling**. The following example shows that this can indeed occur also in small problems.

cycling

Example 6.19. (This example is taken from Chvátal's book [Chv83])
Consider the linear program $\max(\mathbf{c}^t x \mid A\mathbf{x} \le \mathbf{b}, \mathbf{x} \ge \mathbf{0})$ for

$$A := \begin{bmatrix} 1/2 & -11/2 & -5/2 & 9 \\ 1/2 & -3/2 & -1/2 & 1 \\ 1 & & & \end{bmatrix} \quad \mathbf{b} := \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} \quad \mathbf{c}^t := \begin{bmatrix} 10 & -57 & -9 & -24 \end{bmatrix}.$$

Introducing slack variables we arrive at the simplex tableau:

$$\left[\begin{array}{c|ccccccc} 0 & 10 & -57 & -9 & -24 & & & \\ \hline 0 & \boxed{1/2} & -11/2 & -5/2 & 9 & 1 & & \\ 0 & 1/2 & -3/2 & -1/2 & 1 & & 1 & \\ 1 & 1 & & & & & & 1 \end{array}\right]$$

Now we fix rules for the two potential choices in each round of the algorithm:

(1) If there is more than one $r \in N$ with $\overline{c}_r > 0$ then we choose that with largest $\overline{c}_r$.
(2) If there is more then one element in the chosen column that could serve as a pivot element, then we choose the one corresponding to the basic variable with smallest index.

In the following we indicate the current basis by labeling the rows of the tableaux.

$$\begin{array}{c} \\ 1 \\ 6 \\ 7 \end{array} \left[\begin{array}{c|ccccccc} & 0 & 53 & 41 & -204 & -20 & & \\ \hline & 1 & -11 & -5 & 18 & 2 & & \\ & & \boxed{4} & 2 & -8 & -1 & 1 & \\ 1 & & 11 & 5 & -18 & -2 & & 1 \end{array}\right]$$

$$\begin{array}{c} \\ 1 \\ 2 \\ 7 \end{array} \left[\begin{array}{c|ccccccc} & & & 29/2 & -98 & -27/4 & -53/4 & \\ \hline & 1 & & \boxed{1/2} & -4 & -3/4 & -11/4 & \\ & & 1 & 1/2 & -2 & -1/4 & 1/4 & \\ 1 & & & -1/2 & 4 & 3/4 & -11/4 & 1 \end{array}\right]$$

$$\begin{array}{c} \\ 3 \\ 2 \\ 7 \end{array} \left[\begin{array}{c|ccccccc} & -29 & & & 18 & 15 & -93 & \\ \hline & 2 & & 1 & \boxed{-8} & -3/2 & 11/2 & \\ & -1 & 1 & & 2 & 1/2 & -5/2 & \\ 1 & 1 & & & & & & 1 \end{array}\right]$$

$$\begin{array}{c} \\ 3 \\ 4 \\ 7 \end{array} \left[\begin{array}{c|ccccccc} & -20 & -9 & & & 21/2 & -141/2 & \\ \hline & -2 & 4 & 1 & & \boxed{1/2} & -9/2 & \\ & -1/2 & 1/2 & & 1 & 1/4 & -5/4 & \\ 1 & 1 & & & & & & 1 \end{array}\right]$$

$$\begin{array}{c} \\ 5 \\ 4 \\ 7 \end{array} \left[\begin{array}{c|ccccccc} & 22 & -93 & -21 & & & 24 & \\ \hline & -4 & 8 & 2 & & 1 & -9 & \\ & 1/2 & -3/2 & -1/2 & 1 & & \boxed{1} & \\ 1 & 1 & & & & & & 1 \end{array}\right]$$

$$\begin{array}{c} \\ 5 \\ 6 \\ 7 \end{array} \left[\begin{array}{c|ccccccc} & 10 & -57 & -9 & -24 & & & \\ \hline & 1/2 & -11/2 & -5/2 & 9 & 1 & & \\ & 1/2 & -3/2 & -1/2 & 1 & & 1 & \\ 1 & 1 & & & & & & 1 \end{array}\right]$$

At this point, we are back at the original tableau, so the simplex algorithm does not terminate if we use these rules to choose entering and leaving variable. ◊

To avoid cycling, we need an additional parameter that we associate to each pair of a basis and its corresponding objective value, which *increases* even when the objective value does not. As we have observed already we still have the selection rules for entering and leaving variable at our hand to achieve this. We discuss in the following several strategies for this tasks, and prove that some combinations avoid cycling. This will prove that the simplex algorithm with these selections rules also terminates after a finite number of steps even for degenerate solutions.

pivot strategies

These **pivot strategies** clearly influence the path of the algorithm along the graph of the polyhedron, so they also influence the running time of the algorithm. We have seen in the example that not all pivot strategies avoid cycling, and experimentally strategies that avoid cycling have a longer average running time than others. As (again experimentally) cycling is rare in practical problems, a common choice is to use a strategy that promises to be good on the given problem, and switch to another if cycling is detected. We need some preparation.

lexicographically positive

Definition 6.20. *A non-zero vector* $\mathbf{x} \in \mathbb{R}^n$ *is* **lexicographically positive** *if its first non-zero entry is positive. We write* $\mathbf{x} \succ 0$.

lexicographically larger

A vector $\mathbf{y}$ *is* **lexicographically larger** *than* $\mathbf{x}$*, written as* $\mathbf{y} \succ \mathbf{x}$*, if* $\mathbf{y} - \mathbf{x} \succ \mathbf{0}$.

The lexicographic order $\succ$ defines a total order on $\mathbb{R}^n$. For $S \subset \mathbb{R}^n$, let lex-min(S) be the lexicographically smallest vector in S.

Here are some common choices for the entering index r.

smallest variable index

(1) **smallest variable index**: Choose the smallest index r of the variables with $\overline{c}_r > 0$ (observe that this need not be the first non-zero entry of $\overline{c}_N$).

steepest ascend

(2) **steepest ascend**: choose the column index r with the largest $\overline{c}_r$.

largest improvement

(3) **largest improvement**: choose the column index r that induces the largest improvement on the objective value (i.e. such that the λ defined in (6.8) is maximal for this column).

The first rule is very easy to implement, while the two other rules are computationally much more involved. Yet, one would expect that using them leads to a smaller number of rounds in the algorithm. Given the entering index r here are some strategies to choose the leaving index $s = B(s')$.

(1) **lexicographic rule**: Choose the leaving index s' so that

lexicographic rule

$$\frac{1}{\overline{A}_{s'r}}\overline{A}_{s'*} = \text{lex-min}\left(\frac{1}{\overline{A}_{ir}}\overline{A}_{i*} \,\middle|\, \overline{A}_{ir} > 0\right).$$

(2) **smallest index**: Choose the smallest possible s'.

smallest index

(3) **smallest variable index**: Choose the smallest s.

smallest variable index

Theorem 6.21. *Using the lexicographic rule no basis can be repeated. In particular, the simplex algorithm with this rule cannot cycle (independent of the rule to choose an entering index).*

Proof. Assume that the initial feasible basis B satisfies $B = [m]$, i.e. $A = (A_B\, A_N)$. Then all rows of $\overline{A} := A_B^{-1}A$ are lexicographically positive.

Now assume that at some intermediate step of the algorithm all rows of $\overline{A} = A_B^{-1}A$ are lexicographically positive, and assume that the algorithm chooses the column r. Let $J := \{j \in [m] \mid \overline{A}_{jr} > 0\}$, and let s' be the row chosen by the lexicographic rule. Let B' be the updated basis. The rows of $A' := A_{B'}^{-1}A$ are

$$j \in J \;:\quad \overline{A}'_{j*} = \overline{A}_{j*} - \frac{\overline{A}_{jr}}{\overline{A}_{s'r}}\overline{A}_{s'*} \succ 0 \qquad\qquad \text{by the choice of } s'$$

$$j \notin J \;:\quad \overline{A}'_{j*} = \overline{A}_{j*} - \frac{\overline{A}_{jr}}{\overline{A}_{s'r}}\overline{A}_{s'*} = \overline{A}_{j*} + \left|\,\frac{\overline{A}_{jr}}{\overline{A}_{s'r}}\,\right|\overline{A}_{s'*} \succ 0\,.$$

So the rows of $\overline{A}'$ are lexicographically positive. The reduced costs change by

$$\overline{\mathbf{c}}' \;=\; \overline{\mathbf{c}} \;-\; \frac{\overline{c}_r}{\overline{A}_{s'r}}\overline{A}_{s'*}\,,$$

so $\overline{\mathbf{c}}' \prec \overline{\mathbf{c}}$ and the reduced costs strictly decrease in each step. This implies that the tableau for any basis is unique, so no basis can occur twice during the algorithm. ☐

Another rule that guarantees finiteness of the simplex algorithm is

> **Bland's rule**: In each step, choose the smallest index for entering the basis, and then the smallest index for leaving the basis.

Bland's rule

Experiments suggest however that the lexicographic rule or Bland's rule are not a good choice. They tends to create many more pivoting steps than other rules on average.

We have already observed in the introduction that in the worst-case, none of the above rules is "good" in the sense that it guarantees polynomial running time. For all common selection rules there are linear programs so that the simplex algorithm with this particular rules runs through all vertices of the polyhedron, and there are polytopes with an exponential number of vertices in the number of facets. It is an open problem whether there are (yet undiscovered) selection rules that guarantee a polynomial running time. For a randomized approach for the selection of entering and leaving index it has been proved that the expected run time is sub-exponential.

Finally we note without proof that there is no *optimal* pivot strategy in the sense that it does not effect worst case performance. There is a huge amount of literature about which rules may be better in which cases (expected, on average, ...).

The simplex tableau is a nice and useful tool for solving a linear program by hand. For implementations, there is far too much unnecessary information contained in the tableau, and we compute lots of unnecessary information. In particular, in each tableau we store a unit matrix. We will discuss a way to reduce both stored data and computation time. However, we can only touch on this topic, as there are far too many different ways to attack this problem in different settings (i.e. for sparse systems, many redundant constraints, ...). For a more thorough treatment the reader is e.g. referred to [Chv83].

Now assume that we are given a linear program in standard form with matrix $A \in \mathbb{R}^{m\times n}$, right hand side $\mathbf{b} \in \mathbb{R}^m$, and cost vector $\mathbf{c}^t \in \mathbb{R}^n$ together with a feasible basis B and its corresponding basic solution $\overline{\mathbf{x}}_B$. We consider one iteration of the simplex algorithm.

In the first step we have to determine the entering index, i.e. the index of a positive entry in $\mathbf{c}^t - \mathbf{c}^t_B A_B^{-1} A$. We can compute this vector in two steps by solving the system of linear equations $\mathbf{y}^t A_B = \mathbf{c}^t_B$ and then computing $\overline{\mathbf{c}}^t_N := \mathbf{c}^t_N - \mathbf{y}^t A_N$. Depending on our strategy for choosing the entering index it might suffice to compute only a part of $\overline{\mathbf{c}}^t$. If the basic solution is not already optimal then we find an index r such that $\overline{c}_r > 0$.

Given the entering index we have to determine the leaving index. For this we need the current basic solution $\overline{\mathbf{x}}_B$ and the r-th column of the matrix $A_B^{-1}A$. The latter can be computed from $A_B\mathbf{d} = A_{*r}$, and the minimal quotient λ as in (6.7) determines the entering index s. The new basic solution is now $\overline{\mathbf{x}} - \lambda\mathbf{d}$ corresponding to the basis obtained from B by replacing r with s.

revised simplex method

The simplex method in this form is often called the **revised simplex method**, whereas the tableau form described earlier is the **standard simplex method**. In each iteration we have to solve two systems of equations with m equations instead of Gaussian elimination on an $(m \times n)$ system, and we basically have no update step. Given the matrix A_B we need only one new column of A in each iteration, namely the column corresponding to the entering index. This is particularly advantageous if the matrix A is too large to keep it in memory, or if columns are only generated when needed.

The efficiency of this revised method mainly depends on the fast solution of the two systems of equations, one to determine the entering, and one to determine the leaving index. Note that the constraint matrix A_B in each iteration only differs in one column from the matrix used in the previous iteration. So usually, these systems are not solved from scratch, but some method is implemented to reuse part of the computations. We sketch one such method.

Let $A_{B_1}, A_{B_2}, \ldots, A_{B_k}$ be the sequence of basis matrices obtained in the iterations of the simplex algorithm. From our discussion above we know that two consecutive such matrices differ only by multiplication with the matrix $\Lambda_{B_{k-1}B_k}$ defined in (6.9),

$$A_{B_k} = A_{B_{k-1}}\Lambda_{B_{k-1}B_k} .$$

Doing one initial Gaussian elimination we can assume that $A_{B_1} = \mathrm{Id_m}$. Then

$$A_{B_k} = \Lambda_{B_1B_2}\Lambda_{B_2B_3}\cdots\Lambda_{B_{k-1}B_k} ,$$

and we can solve the system $\mathbf{y}^t A_{B_k} = \mathbf{c}^t_B$ by successively solving

$$\mathbf{u}^t_{k-1}\Lambda_{B_{k-1}B_k} = \mathbf{c}^t_B \qquad \mathbf{u}^t_{k-2}\Lambda_{B_{k-2}B_{k-1}} = \mathbf{u}^t_{k-1} \qquad \ldots \qquad \mathbf{y}^t\Lambda_{B_1B_2} = \mathbf{u}^t_1 .$$

Though this might look like a stupid idea, note that the systems in the above chain are extremely easy to solve, as the matrices differ only in one column from the unit matrix. Thus, in each system only one entry has to be computed. Similarly, we can solve the second system $A_{B_k}\mathbf{d} = A_{*r_k}$, where r_k is the entering index in the k-th step, with such a chain,

$$\Lambda_{B_1B_2}\mathbf{v}_1 = A_{*r_k} \qquad \Lambda_{B_2B_3}\mathbf{v}_2 = A_{*r_k}\mathbf{v}_1 \qquad \ldots \qquad \Lambda_{B_{k-1}B_k}\mathbf{d} = \mathbf{v}_{k-1} .$$

There is one problem left with this approach. If we have done a large number of pivot steps in the algorithm then despite their simplicity the time to solve all these systems might become larger then the time to solve the original system from scratch. However, given a decomposition of a matrix A into a product of upper or lower triangular matrices gives a similarly simple computation of a solution of a system of linear equations. So at some point it might be advisable to replace the product

$$\Lambda_{B_1B_2}\Lambda_{B_2B_3}\cdots\Lambda_{B_{k-1}B_k} \qquad (6.11)$$

by a decomposition into upper or lower triangular matrices. Doing this with the first basis matrix also spares the initial matrix inversion to obtain $A_{B_1} = \mathrm{Id_m}$. As there are many different ways to obtain a decomposition into triangular matrices there is much choice in the actual implementation of this approach. One also needs to decide at which point the sequence of $\Lambda_{B_{i-1}B_i}$ is replaced with such a decomposition. These choices, and the previous choices for strategies to choose the entering and leaving index greatly influence the running time of the simplex algorithm. See standard texts on Linear Programming for more on this.

Let us finally discuss an economic interpretation of the simplex method in connection with the slackness theorem. We use the following general setting. Any variable x_j counts units of some object or service provided, and a row of A corresponds to a resource used in the production of the x_j. An entry c_j of the cost vector gives a profit obtained for each unit of the variable x_j, and the column A_{*j} of A collects different resources needed to obtain one unit of x_j. Given a basic feasible solution $\overline{\mathbf{x}}_B$, the slackness theorem tells us that we can check optimality of this solution by

solving $\mathbf{y}^tA_B = \mathbf{c}_B^t$ and checking $\mathbf{y}^tA_N \geq \mathbf{c}_n$.

The system of equations is precisely the first system we have to solve in the revised simplex method, and the system of inequalities is the one we have to consider to find the entering index. If it is satisfied, then the solution is optimal, and otherwise any index corresponding to a column that violates it can be taken as an entering index (well, almost, as for degenerate solutions $\overline{\mathbf{x}}$ the system of equations considered in the complementary slackness theorem may be smaller than the above).

Solving $\mathbf{y}^tA_B = \mathbf{c}_B^t$ can be seen as assigning a value to each resource such that the return $c_B^t\overline{x}_B$ of the current solution matches the value of the amount $A_B\overline{\mathbf{x}}_b$ of the resources needed to obtain $\overline{x}_B$, i.e. $c_B^t\overline{\mathbf{x}}_B = \mathbf{y}^tA_B\overline{\mathbf{x}}$. In this interpretation $\mathbf{y}$ is sometimes called the **shadow price** of the resources. Evaluating $\mathbf{y}^tA_N$ gives the current shadow prices of the variables that are currently not in the basis. As $\overline{\mathbf{x}}_N = \mathbf{0}$ these correspond to goods or services currently not produced. This process is sometimes called **pricing out** of the non-basic variables. If non of these variables costs less in the shadow prices the it would return according to the objective function (i.e. $\mathbf{y}^tA_{*j} \geq \overline{c}_j$ for each non-basic index $j \in N$) then the current solution is optimal (The converse does not necessarily hold if $\overline{\mathbf{x}}$ is degenerate. Though there is an optimal solution that satisfies this we might need additional steps in the simplex algorithm to obtain it). Otherwise we can pick a variable r whose cost is less than its return, i.e. some r such that $\mathbf{y}^tA_{*r} < \overline{c}_r$.In this case it is advisable to substitute the production of x_r for some of the variables in the basis. The second system of equations $A_B\mathbf{d} = A_{*r}$ precisely computes for us a way to substitute one unit of the entering variable x_r for a mix of the current basic variables. The maximal amount of the new variable we can put into our basic solution is given by the parameter λ of (6.7). In this process exactly one of the currently produced objects is removed from the basis, i.e. its production is stopped. Observe that the $\mathbf{y}^t$ computed in the first system of linear equations is nothing but the dual solution corresponding to $\overline{\mathbf{x}}$. By the duality theorem it will be feasible if and only if the primal solution is optimal. We will meet this last observation again in the next section.

shadow price

pricing out

Duality and Post-Optimization 7

In this chapter we introduce variants of the simplex method using information on the dual solution contained in the simplex tableau, and we discuss the effect of changes in the input data on the optimal solution, and ways to quickly regenerate an optimal solution. The following discussions just touch the surface on these topics. Using as much information as possible contained in a tableau, and being able to quickly cope with varying input is important in many real-world applications of linear optimization. So there is an extensive literature dealing with these questions.

We have seen at the end of the last chapter that the simplex tableau also contains information about the dual solution. In the following we want to explore how we can use this information for solving linear programs. We modify our definition of a simplex tableau to support this.

Let a linear program $\max(\mathbf{c}^t\mathbf{x} \mid A\mathbf{x} = \mathbf{b},\, \mathbf{x} \geq \mathbf{0})$ for $A \in \mathbb{R}^{m\times n}$, $\mathbf{b} \in \mathbb{R}^m$, and $\mathbf{c}^t \in \mathbb{R}^n$ be given. We assume $\operatorname{rank} A = m$. If $B \subseteq [n]$ is a feasible basis, then

$$T_B := \left[\begin{array}{c|c} -\mathbf{c}_B^t A_B^{-1}\mathbf{b} & \overline{\mathbf{c}}^t \\ \hline A_B^{-1}\mathbf{b} & A_B^{-1}A \end{array}\right]$$

is the simplex tableau associated to B. We have already observed that $A_B^{-1}A$ always contains a unit matrix in the rows corresponding to B, and the reduced cost $\overline{\mathbf{c}}^t$ are 0 in the entries of the basis. Hence, this part of the tableau does not contain any information about the linear program. However, we used it to keep track of the current basis when transforming a tableau into a new basis. We can record this information in a more compact form by removing the unit matrix and $\overline{\mathbf{c}}_B^t$ from the tableau and labeling the rows of $A_B^{-1}A_N$ with the corresponding basis index, and the columns of $A_B^{-1}A_N$ with the corresponding non-basis index, so if $B = (b_1, \ldots, b_m)$ and $N = (n_1, \ldots, n_{n-m})$, then we can write the tableau in the form

$$\begin{array}{c|c|c} & & n_1 \quad \cdots \quad n_{n-m} \\ \hline & -\mathbf{c}_B^t A_B^{-1}\mathbf{b} & \overline{\mathbf{c}}_N^t \\ \hline \begin{array}{c} b_1 \\ \vdots \\ b_m \end{array} & A_B^{-1}\mathbf{b} & A_B^{-1}A \end{array}$$

Example 7.1. Reconsider the linear program given in Example 6.3:

$$A := \begin{pmatrix} 1 & & 2 & 3 & -1 \\ & 1 & & 1 & \\ & 1 & 1 & 2 & -2 \end{pmatrix} \qquad \mathbf{b} := \begin{pmatrix} 5 \\ 1 \\ 3 \end{pmatrix} \qquad \mathbf{c}^t := \begin{pmatrix} 1 & 0 & 1 & 4 & 1 \end{pmatrix}.$$

The basis $B = (1, 2, 3)$ is primal feasible with

$$A_B^{-1}A_N = \begin{pmatrix} 2 & 1 \\ 1 & \\ 1 & -2 \end{pmatrix} \qquad A_B^{-1}\mathbf{b} = \begin{pmatrix} 3 \\ 1 \\ 2 \end{pmatrix} \qquad \overline{\mathbf{c}}_N^t = \begin{pmatrix} 1 & 2 \end{pmatrix}.$$

So the reduced tableau is

$$\begin{array}{c r|rr}
 & & 4 & 5 \\
 & -5 & 1 & 2 \\ \cline{2-4}
1 & 3 & 2 & 1 \\
2 & 1 & 1 & \\
3 & 2 & 1 & -2
\end{array}$$

◊

reduced simplex tableau

The simplex tableau in this form is called the **reduced simplex tableau** associated to the basis B. We need to modify our update rules for this tableau. Let $\overline{A} := A_B^{-1}A_N$. If $r = N(r')$ enters the basis and $s = B(s')$ leaves it, then the column below the index r (the r'-th column of $\overline{A}$) is the pivot column, the row with index s (i.e. the s'-th row of $\overline{A}$ is the pivot row, and the element $\overline{A}_{s'r'}$ at their intersection is the pivot element. By construction, this is positive. In the original update step we used Gauss elimination to transform the r'-th row into a unit vector. We still do this Gauss elimination for all columns except the r'-th (but including the right hand side $\overline{\mathbf{x}} = A_B^{-1}\mathbf{b}$), and in the r'-th column we enter the s-th column of the original tableau, so at the position (i, r') for $i \neq s'$ we write $-\overline{A}_{jr'}/\overline{A}_{s'r'}$, and at position (s', r') we put ${}^1/\overline{A}_{s'r'}$.

Example 7.1 continued. We finish the example in the new notation and obtain the following sequence of tableaus. The pivot elements are marked with a box.

$$\begin{array}{c r|rr}
 & & 4 & 5 \\
 & -5 & 1 & 2 \\ \cline{2-4}
1 & 3 & 2 & 1 \\
2 & 1 & \boxed{1} & \\
3 & 2 & 1 & -2
\end{array}$$

$$\begin{array}{c r|rr}
 & & 2 & 5 \\
 & -6 & -1 & 2 \\ \cline{2-4}
1 & 1 & -2 & \boxed{1} \\
4 & 1 & 1 & \\
3 & 1 & -1 & -2
\end{array}$$

$$\begin{array}{c r|rr}
 & & 2 & 1 \\
 & -8 & 3 & -2 \\ \cline{2-4}
5 & 1 & -2 & 1 \\
4 & 1 & \boxed{1} & \\
3 & 3 & -5 & 2
\end{array}$$

$$\begin{array}{c r|rr}
 & & 4 & 1 \\
 & -11 & -3 & -2 \\ \cline{2-4}
5 & 3 & 2 & 1 \\
2 & 1 & 1 & \\
3 & 8 & 5 & 2
\end{array}$$

◊

Note that we can write down a simplex tableau even if the current basis is not feasible, so we don't require this in the following. Feasibility of a basis B (the set of row labels) can then be read off from the tableau. B is feasible if and only if the first column in the lower part is non-negative. A pivot step that replaces some index $s \in B$ in the basis by some index $r \in N$ not in the basis is allowed as long as the set $B' := (B \setminus \{r\}) \cup \{s\}$ is still a basis. This is if and only if the pivot element $A_{s'r'} \neq 0$, where r' and s' are the column and the row of the current constraint matrix $A_B^{-1}A_N$ that correspond to the variable indices r and s.

If we consider linear programs and their duals it is more convenient to use the canonical form of the program as this emphasizes the symmetry between a program and its dual. So for the following considerations we look at a pair

$$\max(\mathbf{c}^t\mathbf{x} \mid A\mathbf{x} \leq \mathbf{b}, \quad \mathbf{x} \geq \mathbf{0}) \tag{P}$$
$$\min(\mathbf{y}^t\mathbf{b} \mid \mathbf{y}^t A \geq \mathbf{c}^t, \;\; \mathbf{y} \geq \mathbf{0}). \tag{D}$$

of dual linear programs for $A \in \mathbb{R}^{m\times n}$, $\mathbf{b} \in \mathbb{R}^m$, and $\mathbf{c}^t \in \mathbb{R}^n$. Note that in this form we cannot assume rank $A = m$ (and indeed this is rarely satisfied). We will see later that we can easily transfer our considerations to a linear program in standard form and its dual. (P) and (D) exhibit the particularly nice feature of programs in canonical form that they one can be transformed into the other via

$$\begin{aligned} A &\longleftrightarrow -A^t \\ \mathbf{b} &\longleftrightarrow -\mathbf{c}^t \end{aligned}$$

and
$$\mathbf{c} \longleftrightarrow -\mathbf{b}^t\,.$$

To apply the simplex method we nevertheless need programs in standard form, so we transform both programs in a symmetric way by introducing slack variables (and taking the negative transpose of the dual). We introduce $\mathbf{u}$ as slack variables for the primal, and $\mathbf{v}$ for the dual. So our programs are

$$\begin{aligned} &\max(\;\;\mathbf{c}^t\mathbf{x} \mid \quad\;\; \mathbf{u} + A\mathbf{x} = \;\;\mathbf{b},\;\; \mathbf{x},\mathbf{u} \geq \mathbf{0}) \\ &\max(-\mathbf{b}^t\mathbf{y} \mid -A^t\mathbf{y} + \;\;\mathbf{v} = -\mathbf{c},\;\; \mathbf{y},\mathbf{v} \geq \mathbf{0})\,. \end{aligned}$$

Note that the primal program has as many slack variables as the dual has variables, and vice versa. We have the following obvious basic solutions (not necessarily feasible):

for the primal: $B := (\; 1, \;\; \dots, \;\; m \;)$ $\qquad (\overline{\mathbf{u}}, \overline{\mathbf{x}}) := (\mathbf{b}, \;\; \mathbf{0})$

for the dual: $N := (\; m+1, \;\; \dots, \;\; n+m \;)$ $\qquad (\overline{\mathbf{y}}, \overline{\mathbf{v}}) := (\mathbf{0}, -\mathbf{c})$

Writing down the reduced tableaus for the two programs using these initial bases and $B = (B_1, \dots, B_m)$, $N = (N_1, \dots, N_n)$ gives

$$\begin{array}{c|c|ccc|} & & N_1 & \dots & N_n \\ \hline & -\mathbf{c}_B^t\mathbf{b} & & \mathbf{c}^t & \\ \hline B_1 & & & & \\ \vdots & \mathbf{b} & & A & \\ B_m & & & & \\ \hline \end{array}$$

for the primal, and

$$\begin{array}{c|c|ccc|} & & B_1 & \dots & B_m \\ \hline & \mathbf{c}_B^t\mathbf{b} & & -\mathbf{b}^t & \\ \hline N_1 & & & & \\ \vdots & -\mathbf{c} & & -A^t & \\ N_n & & & & \\ \hline \end{array}$$

for the dual program.

Looking at this we can see that the dual tableau is just the negative transpose of the primal. The criteria for feasibility and optimality obtained in the previous chapter state that

(1) the primal tableau is feasible if $\mathbf{b} \geq \mathbf{0}$, and the dual tableau is feasible if $\mathbf{c} \leq \mathbf{0}$,
(2) the primal tableau is optimal if $\mathbf{c} \leq \mathbf{0}$, and the dual tableau is optimal if $\mathbf{b} \geq \mathbf{0}$.

primal feasible
dual feasible

Definition 7.3. *The basis B is called* **primal feasible** *if $\mathbf{b} \geq \mathbf{0}$ and dual feasible if $\mathbf{c}^t \leq \mathbf{0}$.*

Hence, we can read off primal and dual feasibility and optimality from either tableau. If both tableaus are primal feasible, then the current basis is already optimal. We want to show that this property is preserved in the simplex algorithm if we do the same steps in both tableaus. Let $s \in B$ be an index that we can replace with some index $r \in N$, i.e. $A_{sr} \neq 0$. Then we can replace the dual basis index r with the index s, as $(A^t)_{rs} = A_{sr} \neq 0$. We go back to standard tableaus to compute the transformed tableaus. For this we define

$$K := \begin{pmatrix} \mathrm{Id}_\mathrm{m} & A \end{pmatrix} \qquad\qquad \overline{K} := \begin{pmatrix} -A^t & \mathrm{Id}_\mathrm{n} \end{pmatrix}$$

Doing a pivot step amounts to multiplication of K and $\overline{K}$ with an invertible matrix Λ and $\overline{\Lambda}$ from the left to obtain new matrices $K' := \Lambda K$ and $\overline{K}' := \overline{\Lambda K}$. We obtain new bases B' for the primal and N' for the dual. As we have exchanged r for s in B and s for r in N we still have the property that $B' = [n+m] \setminus N'$. The special form of K and $\overline{K}$ implies $K\overline{K}^t = -A^t + A = 0$. Multiplication with Λ and $\overline{\Lambda}$ does not change this, we still have $K'(\overline{K}')^t = \Lambda K \overline{K}^t \overline{\Lambda}^t = 0$. Hence,

$$K'_B(\overline{K}'_B)^t + K'_N(\overline{K}'_N)^t = 0 \qquad\Leftrightarrow\qquad (\overline{K}'_B)^t(\overline{K}'_N)^{-t} = (K'_B)^{-1}K'_N$$

The new matrices in the reduced tableau are $(K'_B)^{-1}K'_N$ for the primal and $(K'_N(\overline{K}'_B)^{-1})^t$ for the dual, so the negative transpose property is preserved for the two matrices. This is easily seen to extend to $\mathbf{b}$ and $\mathbf{c}^t$. This implies the following general relation between bases of the two programs:

$$B \text{ basis of the primal} \quad\Longleftrightarrow\quad N := [n+m] \setminus B \text{ basis of the dual program.}$$

If one of the tableaus is optimal with primal optimal basis B, then we can read off both the primal and dual optimal solution from the tableau. Note however, that we have implicitly solved the linear programs in standard form, so setting

$$\begin{array}{llll} & \overline{\mathbf{x}}'_B := \mathbf{b} & \overline{\mathbf{x}}'_N := \mathbf{0} & \text{for the primal} \\ \text{and} & \overline{\mathbf{y}}'_N := -\mathbf{c} & \overline{\mathbf{y}}'_B := \mathbf{0} & \text{for the dual} \end{array}$$

gives the solutions of the problems in standard form, and we have to project $\overline{\mathbf{x}}'$ to the last n variables to obtain the primal, and $\overline{\mathbf{y}}'$ to the first m variables to obtain the dual solution of the original programs in standard form.

If we use the algorithm with reduced simplex tableaus as described at the beginning of this section, then we can now start the simplex algorithm on either tableau, provided the associated primal or dual solution is feasible. However, we have seen above that the dual tableau is just the negative transpose of the primal, so we can interpret the iterations of the simplex algorithm on the dual tableau also in the primal tableau. In this description the dual simplex does the same steps as the primal, but we

(1) exchange row and column operations, scale the pivot column so that the pivot element changes to -1, and write down the negative of the scaling factors in the pivot row,
(2) choose an index r with negative entry in $\mathbf{b}$ as the entering row, and
(3) the leaving index s among those with negative entry in the r-th row.

This simplex algorithm for the dual program operating on the primal tableau is commonly called the **dual simplex algorithm**. As a primal pivot step in the primal tableau results in the same tableau as the corresponding dual step in the dual tableau, we can even drop the first change, and just do a normal primal pivot.

dual simplex algorithm

Example 7.4. Consider the linear program $\max(\mathbf{c}^t\mathbf{x} \,|\, A\mathbf{x} \leq \mathbf{b}\,,\ \mathbf{x} \geq \mathbf{0})$ with

$$A := \begin{pmatrix} -1 & -1 \\ -2 & -1 \\ 2 & 1 \end{pmatrix} \qquad \mathbf{b} := \begin{pmatrix} -3 \\ -5 \\ 8 \end{pmatrix} \qquad \mathbf{c}^t := \begin{pmatrix} -3 & -2 \end{pmatrix}$$

the obvious primal basis $B = [3]$ of the linear program $\max(\mathbf{c}^t\mathbf{x} \mid \mathbf{y} + A\mathbf{x} = \mathbf{b},\ \mathbf{x}, \mathbf{y} \geq \mathbf{0})$ with $\overline{\mathbf{x}} := \mathbf{b}$ is not feasible, so we try the dual program $\max(-\mathbf{b}^t\mathbf{y} \mid -A^t\mathbf{y} + \mathbf{x} = -\mathbf{c},\ \mathbf{x}, \mathbf{v} \geq \mathbf{0})$. This has a basic solution $\overline{\mathbf{y}} = -\mathbf{c} \geq 0$ with basis $N = (4, 5)$ which is feasible. We set up the tableau:

$$\begin{array}{c|c|cc}
 & & 4 & 5 \\
 & 0 & -3 & -2 \\ \cline{2-4}
1 & -3 & -1 & -1 \\
2 & -5 & \boxed{-2} & -1 \\
3 & 8 & 2 & 1
\end{array}$$

We have to find a row with $b_i < 0$, so we can take either the first or second row. We choose the second, i.e. $s' = 2$ and $s := B(s') = 2$. Among the negative entries of that row we have to find one that minimizes the quotient ${}^{c_j}/_{A_{s'j}}$. As ${}^{-3}/_{-2} < {}^{-2}/_{-1}$ this is $r' = 1$, i.e. $r = N(r') = 4$. So after the first step the tableau is

$$\begin{array}{c|c|cc}
 & & 2 & 5 \\
 & {}^{15}/_{2} & -{}^{3}/_{2} & -{}^{1}/_{2} \\ \cline{2-4}
1 & -{}^{1}/_{2} & -{}^{1}/_{2} & \boxed{-{}^{1}/_{2}} \\
4 & {}^{5}/_{2} & -{}^{1}/_{2} & {}^{1}/_{2} \\
3 & 3 & 1 & 0
\end{array}$$

Note that we can obtain the tableau in two ways. If we want to do the dual pivot step, then we add $-{}^5/_2$ of the second column to the first, $-{}^1/_2$ of the second column to the third, scale the second column by ${}^1/_2$ so that the framed entry is -1, and replace the second row with the negative of scaling factors, so with ${}^5/_2$, $-{}^1/_2$ and ${}^1/_2$. For the primal step we add $-{}^3/_2$ of the third row to the first, $-{}^1/_2$ of the third to the second, ${}^1/_2$ of the third to the forth, scale the third row by $-{}^1/_2$, and replace the second column with the scaling factors. Either way gives the above tableau.

The first row of the new tableau still contains a negative entry, so we are not yet optimal. For the next picot step we have to take $s' = 1$ and $r' = 2$, so $s = 1$ and $r = 5$. The new tableau is

$$\begin{array}{c|c|cc}
 & & 2 & 1 \\
 & 8 & -1 & -1 \\ \cline{2-4}
5 & 1 & 1 & -2 \\
4 & 2 & -1 & 1 \\
3 & 3 & 1 & 0
\end{array}$$

This tableau is optimal. so the optimal dual solution is $\overline{\mathbf{y}}^t = (1,\, 1)$, and the optimal primal solution is $\overline{\mathbf{x}} = \begin{pmatrix} 2 \\ 1 \end{pmatrix}$ with optimal value -8. ◊

We can use the dual algorithm to replace our *phase I* used in the previous chapter to obtain an initial feasible basis for a linear program in canonical form. In this step, the objective function is irrelevant, we are only interested in a vertex of the polyhedron $\{\mathbf{x} \mid A\mathbf{x} \leq \mathbf{b}, \mathbf{x} \geq \mathbf{0}\}$. But feasibility of the dual solution hinges only on $\mathbf{c}^t$. If this is negative, then the trivial initial dual solution is feasible. So to obtain a feasible solution we can replace $\mathbf{c}^t$ with an arbitrary non-zero cost vector $\mathbf{d}^t$ that satisfies $\mathbf{d}^t \leq \mathbf{0}$. A suitable choice would e.g. be $\mathbf{d}^t := -\mathbf{1}^t$. We can then solve the dual problem with the above method, and obtain an optimal dual and primal solution for the modified problem. The primal solution is in particular feasible, so we can use this to solve the original problem. The only thing we have to change in the last tableau of the auxiliary problem is the cost function.

Example 7.5. Consider the linear program $\max(\mathbf{c}^t\mathbf{x} \mid A\mathbf{x}\ le\ \mathbf{b},\ \mathbf{x} \geq \mathbf{0})$ in canonical form with

$$A := \begin{pmatrix} 2 & 1 \\ -1 & 1 \\ -3 & -1 \end{pmatrix} \qquad \mathbf{b} := \begin{pmatrix} 12 \\ 6 \\ -6 \end{pmatrix} \qquad \mathbf{c}^t := \begin{pmatrix} 1 & 3 \end{pmatrix}.$$

The initial reduced tableau is

$$\begin{array}{cr|rr}
 & & 4 & 5 \\
 & 0 & 1 & 3 \\ \hline
1 & 12 & 2 & 1 \\
2 & 6 & -1 & 1 \\
3 & -6 & -3 & -1
\end{array}$$

Hence, neither the primal nor the dual solution is optimal. We replace the current cost function by $\mathbf{d}^t := \begin{pmatrix} -1 & -1 \end{pmatrix}$ to obtain the tableau

$$\begin{array}{cr|rr}
 & & 4 & 5 \\
 & 0 & -1 & -1 \\ \hline
1 & 12 & 2 & 1 \\
2 & 6 & -1 & 1 \\
3 & -6 & -3 & -1
\end{array}$$

Now the dual basis is feasible, and we apply the dual simplex algorithm. For this, we have to choose a row with negative first entry, so we have to take $s' = s = 3$. Then among the columns with negative entry in this row we choose the one with the smallest quotient. In our case, there is only one such column, and we have to take $r' = 1$, $r = 4$. If we do the pivot step we obtain the tableau

$$\begin{array}{cr|rr}
 & & 3 & 5 \\
 & 2 & -1/3 & -4/3 \\ \hline
1 & 8 & 2/3 & 5/3 \\
2 & 8 & -1/3 & 2/3 \\
4 & 2 & -1/3 & -1/3
\end{array}$$

This tableau is dual optimal, and we read off the dual and primal solutions to be

$$\mathbf{x} := \begin{pmatrix} 2 \\ 0 \end{pmatrix} \qquad \mathbf{y}^t := \begin{pmatrix} 0 & 0 & -1/3 \end{pmatrix}$$

As we have only changed the cost function, the primal solution is also feasible for the original problem. Observe that the tableau also tells us the values of the slack variables.

Indeed, if we plug in the primal solution into $A\mathbf{x} \leq \mathbf{b}$ we obtain

$$\begin{pmatrix} 12 \\ 6 \\ -6 \end{pmatrix} - \begin{pmatrix} 2 & 1 \\ -1 & 1 \\ -3 & -1 \end{pmatrix} \begin{pmatrix} 2 \\ 0 \end{pmatrix} = \begin{pmatrix} 8 \\ 8 \\ 0 \end{pmatrix}$$

reflecting the fact that the first two slack variables of the primal are 8, and the last one is not in the basis, so it is 0. We can now set up a tableau for the original function using our feasible solution by replacing the cost function

$$\begin{array}{c} \\ \\ 1 \\ 2 \\ 4 \end{array} \left[\begin{array}{r|rr} & 3 & 5 \\ -2 & 4/3 & 10/3 \\ \hline 8 & 2/3 & 5/3 \\ 8 & -1/3 & 2/3 \\ 2 & -1/3 & -1/3 \end{array} \right]$$

We can solve this program with the primal simplex algorithm. ◊

Similarly, we could modify the right hand side $\mathbf{b}$ to obtain a feasible primal solution, solve the primal problem to obtain a feasible dual solution for this modified problem, and then solve the original dual problem to obtain optimal primal and dual solutions of the original problem.

In the previous chapter we have considered a primal problem in the standard form $\max(\mathbf{c}^t\mathbf{x} \mid A\mathbf{x} = \mathbf{b},\ \mathbf{x} \geq \mathbf{0})$. Here, the dual problem is $\min(\mathbf{y}^t\mathbf{b} \mid \mathbf{y}^t A \geq \mathbf{c}^t)$, and the duality between the two problems is less obvious than in the formulation (P) and (D) used for the dual simplex algorithm. However, the reduced tableau obtained at the beginning of the chapter shows how to transform such a problem into our format. Given a basis B of the program we can rewrite the constraints as

$$\begin{aligned} A\mathbf{x} = \mathbf{b},\ \mathbf{x} \geq \mathbf{0} \quad &\Longleftrightarrow \quad A_N\mathbf{x}_N + A_B\mathbf{x}_B = \mathbf{b},\ \mathbf{x}_N, \mathbf{x}_B \geq \mathbf{0} \\ &\Longleftrightarrow \quad A_B^{-1}A_N\mathbf{x}_N + \mathbf{x}_B = \mathbf{b},\ \mathbf{x}_N, \mathbf{x}_B \geq \mathbf{0} \\ &\Longleftrightarrow \quad A_B^{-1}A_N\mathbf{x}_N \leq \mathbf{b},\ \mathbf{x}_N \geq \mathbf{0}. \end{aligned}$$

The cost function in this basis is

$$\mathbf{c}^t\mathbf{x} := \mathbf{c}_B^t\mathbf{x}_B + \mathbf{c}_N^t\mathbf{x}_N = \mathbf{c}_B^t A_B^{-1}\mathbf{b} + \mathbf{c}_N^t\mathbf{x}_N = k + \mathbf{c}_N^t\mathbf{x}_N$$

for $k := \mathbf{c}_B^t A_B^{-1}\mathbf{b}$ constant. So we can rewrite our linear program as

$$\max(\mathbf{c}_N^t\mathbf{x}_N \mid A_B^{-1}A_N\mathbf{x}_N \leq \mathbf{b},\ \mathbf{x}_N \geq \mathbf{0}). \tag{CP}$$

The program has the same optimal solutions (the omitted variables $\mathbf{x}_B$ are the slack of the obtained solution), and the optimal value differs by the constant k. (CP) is a linear program in canonical form that yields exactly the reduced tableau associated to the basis B that we have defined for the original problem. With this transformed problem we can now apply the primal or dual simplex algorithm as given above.

Example 7.6. Let

$$A := \begin{pmatrix} 2 & 1 & -3 & -4 & -2 \\ -3 & -1 & 4 & 5 & 4 \end{pmatrix} \quad \mathbf{b} := \begin{pmatrix} -6 \\ 7 \end{pmatrix} \quad \mathbf{c}^t := \begin{pmatrix} 1 & 3 & -5 & -9 & 2 \end{pmatrix}$$

Then $B = (1, 2)$ is dual feasible and

$$A_B^{-1} = \begin{pmatrix} -1 & -1 \\ 3 & 2 \end{pmatrix} \qquad \overline{A} = \begin{pmatrix} 1 & & -1 & -1 & -2 \\ & 1 & -1 & -2 & 2 \end{pmatrix} \qquad \overline{\mathbf{b}} = \begin{pmatrix} -1 \\ -4 \end{pmatrix}$$

$$\overline{\mathbf{c}}^t = \begin{pmatrix} 0 & 0 & -1 & -2 & -2 \end{pmatrix} \qquad \overline{\mathbf{y}}^t = \begin{pmatrix} 8 & 5 \end{pmatrix}.$$

So B is not optimal. The corresponding tableau is

$$\begin{array}{cr|rrr}
 & & 3 & 4 & 5 \\
 & 13 & -1 & -2 & -2 \\ \cline{2-5}
1 & -1 & \boxed{-1} & -1 & -2 \\
2 & -4 & 1 & -2 & 2
\end{array}$$

Using this algorithm we obtain the following sequence of tableaus:

$$\begin{array}{cr|rrr}
 & & 1 & 4 & 5 \\
 & 14 & -1 & -1 & 0 \\ \cline{2-5}
3 & 1 & -1 & 1 & 2 \\
2 & -3 & \boxed{-1} & -1 & 4
\end{array}$$

$$\begin{array}{cr|rrr}
 & & 2 & 4 & 5 \\
 & 13 & 0 & -1 & -4 \\ \cline{2-5}
3 & 4 & 2 & -1 & -2 \\
1 & 3 & 1 & -1 & -4
\end{array}$$

This tableau is optimal as lower part of the first column is non-negative. The linear program has optimal value 17 and primal basic feasible solution

$$\overline{\mathbf{x}}^t := \begin{pmatrix} 3 & 0 & 4 & 0 & 0 \end{pmatrix}. \qquad \Diamond$$

primal-dual methods

In the simplex methods we have discussed so far we have only used the duality theorem in the observation that a primal and dual feasible basis must be optimal. In the end of Chapter 3 we have also proved the complementary slackness theorem and have seen that we can sometimes produce a dual solution from a primal one by solving a system of linear equations. Solving equations is much easier than solving inequalities, so we now want to exploit this also for our simplex algorithm. This leads to the so called **primal-dual methods**. We consider again

$$\max(\mathbf{c}^t\mathbf{x} \mid A\mathbf{x} = \mathbf{b}, \mathbf{x} \geq \mathbf{0}) = \max(\mathbf{c}^t\mathbf{x} \mid \mathbf{x} \in P) \tag{P}$$

$$\min(\mathbf{y}^t\mathbf{b} \mid \mathbf{y}^t A \geq \mathbf{c}^t) = \min(\mathbf{y}^t\mathbf{b} \mid \mathbf{y} \in P^*). \tag{D}$$

As before we may assume that $\mathbf{b} \geq \mathbf{0}$. Now assume that we have a dual feasible solution $\overline{\mathbf{y}}$ and some vector $\overline{\mathbf{x}} \geq \mathbf{0}$ with the property

$$\overline{x}_i > 0 \quad \Longrightarrow \quad \overline{y}^t A_{*i} \; = \; c_i \qquad \text{for all } 1 \leq i \leq n. \tag{7.1}$$

The complementary slackness theorem tells us that if $A\overline{\mathbf{x}} = \mathbf{b}$ then $\overline{\mathbf{x}}$ is primal feasible and optimal. Clearly, if $\mathbf{b} = \mathbf{0}$, then $\mathbf{x} = \mathbf{0}$ satisfies (7.1), so it is optimal. Otherwise we have that $\mathbf{1}^t\mathbf{b} > 0$.

Let $\qquad J := \{j \mid \overline{\mathbf{y}}^t A_{*j} = c_j\}$

and $\qquad \tilde{A} := A_{*J}.$

Consider the linear program

$$\min(\mathbf{1}^t\mathbf{z} \mid \mathrm{Id}_{\mathrm{m}}\mathbf{z} + \tilde{\mathrm{A}}\mathbf{u} = \mathbf{b},\, \mathbf{u}, \mathbf{z} \geq \mathbf{0}). \tag{7.2}$$

The first m columns (which are an identity matrix) are always a feasible basis for this linear program with basic feasible solution

$$\overline{\mathbf{z}} := \mathbf{b}, \qquad\qquad \overline{\mathbf{u}} := \mathbf{0}.$$

It has objective value $\delta = \mathbf{1}^t\mathbf{z} = \mathbf{1}^t\mathbf{b} > 0$. The program is bounded below by 0. The duality theorem implies that there is an optimal solution $(\overline{\mathbf{z}}, \overline{\mathbf{u}})$ with optimal value $\overline{\delta} := \mathbf{1}^t\overline{\mathbf{z}} \geq 0$. We distinguish the two cases $\overline{\delta} = 0$ and $\overline{\delta} > 0$.

If $\overline{\delta} = 0$, then $\overline{\mathbf{z}} = 0$, and

$$\overline{\mathbf{x}}_J := \overline{\mathbf{u}} \qquad\qquad \overline{\mathbf{x}}_{[n]-J} := 0$$

is a primal feasible solution that satisfies (7.1), hence it is optimal.

Otherwise $\overline{\delta} > 0$. Let $\overline{\mathbf{w}}$ be the dual optimal solution corresponding to (7.2). Then

$$\overline{\delta} = \overline{\mathbf{w}}^t\mathbf{b}, \qquad\qquad \overline{\mathbf{w}} \leq \mathbf{1}, \qquad\qquad \overline{\mathbf{w}}^t\tilde{A} \leq 0.$$

By construction,

$$\overline{\mathbf{y}}^t A_{*k} > c_k \qquad \text{for all } k \notin J.$$

Hence, there is $\varepsilon > 0$ such that

$$(\overline{\mathbf{y}}^t - \varepsilon\overline{\mathbf{w}}^t)A \geq \mathbf{c}^t,$$

so $\overline{\mathbf{y}}' := \overline{\mathbf{y}} - \varepsilon\overline{\mathbf{w}}$ is a dual feasible solution. Its objective value is

$$(\overline{\mathbf{y}}')^t b = \overline{\mathbf{y}}^t b - \varepsilon\overline{\mathbf{w}}^t\mathbf{b} = \overline{\mathbf{y}}^t\mathbf{b} - \varepsilon\overline{\delta} < \overline{\mathbf{y}}^t\mathbf{b},$$

so $\overline{\mathbf{y}}'$ has a better objective value than $\overline{\mathbf{y}}$. We may repeat these steps with this improved dual solution. This may, however, not lead to a finite algorithm, as we cannot a priori control the improvement $\varepsilon\overline{\delta}$. We need some argument that ensures that this value is bounded below in each step by some strictly positive lower bound to guarantee finiteness. Often, this can be guaranteed by giving some interpretation of the restricted primal program as a simplified version of the original one. We'll later see an application of this method.

At the end of this chapter, and at the end of our discussion of the simplex method to solve linear programs, we want to discuss the influence of changes to the input data on an optimal solution obtained by some simplex method. There are four main scenarios that we will distinguish:

(1) A change in the right hand side **b**,
(2) a change in the cost function **c**,
(3) addition of a new variable to the program, and
(4) addition of a new constraint to the program.

Such considerations are quite important in real-world applications, as the input data might change according to changes in the market situation (e.g. prices raise, or supply decreases), or measurement of some input is subject to uncertainty and we want to adjust the solution or just estimate the stability of our solution.

Let again $\max(\mathbf{c}^t\mathbf{x} \mid A\mathbf{x} \leq \mathbf{b}, \mathbf{x} \geq \mathbf{0})$ be a linear program in canonical form, and assume that we have an optimal basis B with optimal solution $\overline{\mathbf{x}}$ and value δ. Then B is primal and dual feasible. This implies in particular that the associated reduced costs $\mathbf{c}^t$ are non-positive. Let $\overline{\mathbf{y}}$ be the dual optimal solution corresponding to $\overline{\mathbf{x}}$.

We start with a change in the right Let $\mathbf{b}' := \mathbf{b} + \Delta$ be a new right hand side for some $\Delta \in \mathbb{R}^m$. $\overline{\mathbf{y}}$ is a feasible solution of

$$\min(\mathbf{y}^t(\mathbf{b} + \Delta) \mid \mathbf{y}^t A \geq \mathbf{c}^t, \mathbf{y} \geq \mathbf{0}).$$

This implies that $\overline{\mathbf{y}}^t\mathbf{b}$ is an upper bound for the optimal value δ' of the new program, so that

$$\delta' \leq \overline{\mathbf{y}}^t(\mathbf{b} + \Delta) = \overline{\mathbf{y}}^t b + \overline{\mathbf{y}}^t\Delta = \delta + \overline{\mathbf{y}}^t\Delta. \tag{7.3}$$

We may view $\overline{\mathbf{y}}$ as a measure for the sensitivity of the optimum value of the original program for a perturbation of $\mathbf{b}$. The new basic solution for the basis B is

$$\overline{\mathbf{x}}'_B := A_B^{-1}\mathbf{b}' = A_B^{-1}(\mathbf{b} + \Delta) = \overline{\mathbf{x}}_B + A_B^{-1}\Delta.$$

The reduced costs $\overline{\mathbf{c}}^t = \mathbf{c}^t - \mathbf{c}_B^t A_B^{-1} A$ of the current basic solution do not change, so the basis B remains to be dual feasible.

- If $\overline{\mathbf{x}}' \geq \mathbf{0}$, then B is still primal feasible, so it is optimal. This is the case if and only if

  $$-A_B^{-1}\Delta \leq \overline{\mathbf{x}}_B$$

  which is clearly satisfied in a small region around the origin for non-degenerate solutions $\overline{\mathbf{x}}$. In particular, we have equality in (7.3). We have met this case already at the end of Chapter 3.
- If $\overline{\mathbf{x}}' \not\geq \mathbf{0}$, then we can use this basis as the input for the dual simplex algorithm and obtain a new basic solution.

These considerations also show that the set of all right hand sides b for which $\overline{\mathbf{x}}$ is an optimal solution is the polyhedral cone

$$C_{\mathbf{b}} := \{\mathbf{x} \mid A_B^{-1}\mathbf{x} \leq 0\}\,.$$

Now we look at how changes in the cost function $\mathbf{c}$ affect the optimal value. Assume that we change $\mathbf{c}$ to $\mathbf{c}' := \mathbf{c} + \Delta$. Then

$$\begin{aligned}(\overline{\mathbf{c}}')^t &= \mathbf{c}^t + \Delta^t - (\mathbf{c}_B + \Delta_B)^t A_B^{-1} A \\ &= \overline{\mathbf{c}}^t + \Delta^t - \Delta_B^t A_B^{-1} A\end{aligned}$$

and
$$(\mathbf{c}^t + \Delta)^t \overline{\mathbf{x}} = \mathbf{c}_B^t A_B^{-1}\mathbf{b} + \Delta_B^t A_B^{-1}\mathbf{b}$$

As a primal solution $\overline{\mathbf{x}}$ is optimal if and only if the reduced costs are non-positive, our solution remains optimal for the new program if and only if

$$\overline{\mathbf{c}}^t + \Delta^t - \Delta_B^t A_B^{-1} A \leq \mathbf{0}\,.$$

Otherwise, some entry of $(\overline{\mathbf{c}}')^t$ is positive, and we can replace the old by the new reduced costs in the simplex tableau and restart the primal simplex algorithm.

Now we examine how we can deal with the case that a new variable is added to our program. This adds a new column $\mathbf{a} = A_{*(n+1)}$ to the matrix A and a new entry c_{n+1} to the cost function. The basis B is still a feasible basis. We may compute the new reduced costs via

$$\overline{c}_{n+1} = c_{n+1} - \mathbf{c}_B^t A_B^t \mathbf{a} = c_{n+1} - \overline{\mathbf{y}}^t \mathbf{a}\,,$$

where $\overline{\mathbf{y}}$ is the dual optimal solution of the original program. If $\overline{c}_{n+1} \leq 0$, then $\overline{\mathbf{x}}$ is still optimal, otherwise we can add $\overline{c}_{n+1}$ and $\overline{A}_{*(n+1)} = A_B^{-1}\mathbf{a}$ at the end of the tableau and restart the simplex algorithm. The first step puts x_{n+1} into the basis.

Finally, we look at the addition of a new constraint to our linear program. Let $\mathbf{a}_{m+1}^t \in (\mathbb{R}^n)^*$ and $b_{m+1} \in \mathbb{R}$. We add the constraint $\mathbf{a}_{m+1}^t\mathbf{x} = b_{m+1}$ as $(m+1)$-th row to the matrix A. If it is linear dependent from the other equations, then we could discard it, so we assume that it is linear independent and the rank of A increases to $m+1$. We distinguish the two cases whether $\overline{\mathbf{x}}$ satisfies $\mathbf{a}_{m+1}^t\overline{\mathbf{x}} = b_{m+1}$ or not.

Assume first that the basic solution $\overline{\mathbf{x}}$ satisfies $\mathbf{a}_{m+1}^t\overline{\mathbf{x}} = b_{m+1}$. Then $\overline{\mathbf{x}}$ is also optimal for the extended program. We just add some non-basic column to the basis. $\overline{\mathbf{x}}$ becomes a degenerate solution.

Now assume that the basic solution $\overline{\mathbf{x}}$ does not satisfy $\mathbf{a}_{m+1}^t\overline{\mathbf{x}} = b_{m+1}$. Then we introduce a new slack variable x_{n+1} and change the new equation to

$$\mathbf{a}_{m+1}^t\mathbf{x} + \varepsilon x_{n+1} = b_{m+1}$$

where we take $\varepsilon := 1$ if $\mathbf{a}_{m+1}^t \overline{\mathbf{x}} > b_{m+1}$ and $\varepsilon := -1$ otherwise. We extend the basis by the new last column and put x_{n+1} with value $b_{m+1} - \mathbf{a}_{m+1}^t \overline{\mathbf{x}}$ into the basis. The new basic solution is primal infeasible, as $\overline{x}_{n+1} < 0$. However, we have not changed the reduced costs, so they are still non-positive. So the new basis is dual feasible.

We can start the dual simplex algorithm, which will choose some entry of the new constraint as pivot element. If there is none, i.e. if this row is non-negative, then the dual problem is unbounded. This means that the intersection of the added hyperplane with the original polyhedron is empty, and the problem becomes infeasible.

Otherwise, the dual algorithms computes an optimal solution $\tilde{\mathbf{x}}$. As we have added an artificial variable x_{n+1}, we still need to argue that there is an optimal solution for which $x_{n+1} = 0$. If this is not already satisfied, then we can transform our optimal solution in the following way. Let

$$\eta := \frac{\tilde{x}_{n+1}}{\tilde{x}_{n+1} - \overline{x}_{n+1}} .$$

Then $\eta > 0$, as $\overline{x}_{n+1} < 0$. Define

$$\hat{\mathbf{x}} := \eta \overline{\mathbf{x}} + (1 - \eta) \tilde{\mathbf{x}} .$$

Then $\hat{\mathbf{x}} \geq 0$ as

$$\begin{aligned} \hat{x}_{n+1} &= \frac{\tilde{x}_{n+1}}{\tilde{x}_{n+1} - \overline{x}_{n+1}} \overline{x}_{n+1} + \tilde{x}_{n+1} - \frac{\tilde{x}_{n+1}}{\tilde{x}_{n+1} - \overline{x}_{n+1}} \tilde{x}_{n+1} \\ &= \frac{1}{\tilde{x}_{n+1} - \overline{x}_{n+1}} \left(\tilde{x}_{n+1} \overline{x}_{n+1} - \tilde{x}_{n+1} \tilde{x}_{n+1} + \tilde{x}_{n+1} \tilde{x}_{n+1} - \overline{x}_{n+1} \tilde{x}_{n+1} \right) \\ &= 0 . \end{aligned}$$

The dual solution $\overline{\mathbf{y}}$ corresponding to $\overline{\mathbf{x}}$ is dual feasible, which implies that $\mathbf{c}^t \overline{\mathbf{x}} = \overline{\mathbf{y}}^t \mathbf{b} \geq \mathbf{c}^t \tilde{\mathbf{x}}$ by weak duality. Thus, the objective value satisfies

$$\mathbf{c}^t \hat{\mathbf{x}} = \eta \mathbf{c}^t \overline{\mathbf{x}} + (1 - \eta) \tilde{\mathbf{x}} \geq \lambda \mathbf{c}^t \tilde{\mathbf{x}} + (1 - \eta) \mathbf{c}^t \tilde{\mathbf{x}} = \mathbf{c}^t \tilde{\mathbf{x}} ,$$

so $\hat{\mathbf{x}}$ is optimal.

There are other important changes to the input data that occur in real-world applications which we will not discuss in this text. In particular, we have not discussed how changing a column or row of the matrix A influences the optimal solution. Efficient handling of all these changes are quite important for any practical purposes. You can find extensive discussions on such problems in the literature.

Integer Polyhedra 8

For the rest of the course we restrict to **rational polyhedra**, that is, we only look at polyhedra $P = \{\mathbf{x} \mid A\mathbf{x} \leq \mathbf{b}\}$ such that $A \in \mathbb{Q}^{m\times n}$ and $\mathbf{b} \in \mathbb{Q}^m$. By scaling inequalities with an appropriate factor, we may even assume that $A \in \mathbb{Z}^{m\times n}$ and $\mathbf{b} \in \mathbb{Z}^m$.

rational polyhedra

All vertices of P and all generators of rec P are given as the solution $\mathbf{z}$ of a subsystem $A_{I*}\mathbf{z} = \mathbf{b}_I$ of $A\mathbf{x} \leq b$. From Cramer's rule we may deduce that any solution of a rational system of equations is rational. Hence, using Theorem 4.38, we see that P can be written as $P = \operatorname{conv}(X) + \operatorname{cone}(Y)$ for rational matrices X and Y, and $\operatorname{cone}(Y) = \operatorname{rec}(P)$.

Remark 8.1. We can scale a generator of a cone C by an arbitrary positive factor without changing C. Hence, we can even assume that all generators of rec(P) are integral, i.e. Y is an integral matrix.

Using methods similar to those presented in Chapter 5, one can prove that the absolute value of all entries of Y is bounded by the maximum size of a sub-determinant of A. ◊

Definition 8.2. *Let $A \in \mathbb{Q}^{m\times n}$, $\mathbf{b} \in \mathbb{Q}^m$, $\mathbf{c} \in \mathbb{Q}^n$. The* **integer linear programming problem** *(in standard form) is the problem to determine*

integer linear programming problem (standard form)

$$\delta := \max(\mathbf{c}^t\mathbf{x} \mid A\mathbf{x} = \mathbf{b},\ \mathbf{x} \geq \mathbf{0},\ \mathbf{x} \in \mathbb{Z}^n)\,. \tag{IP}$$

The weak duality estimate

$$\max(\mathbf{c}^t\mathbf{x} \mid A\mathbf{x} = \mathbf{b},\ \mathbf{x} \geq \mathbf{0},\ \mathbf{x} \in \mathbb{Z}^n) \ \leq\ \min(\mathbf{y}^t\mathbf{b} \mid \mathbf{y}^t A \geq \mathbf{c}^t,\ \mathbf{y} \in \mathbb{Z}^m)\,.$$

remains valid for integer linear programs. However, we usually do not have equality in this relation, i.e. the (strong) duality theorem fails. The **LP-relaxation** of (IP) is the linear program

LP-relaxation

$$\delta_{lp} := \max(\mathbf{c}^t\mathbf{x} \mid A\mathbf{x} = \mathbf{b},\ \mathbf{x} \geq \mathbf{0})\,. \tag{LP}$$

The optimal solution of (LP) is the **LP-optimum** of (IP). Clearly, δ_{lp} is an upper bound for δ, but this bound may be quite far away from the true solution.

LP-optimum

Example 8.3. (1) Consider the linear program $\max(x \mid 2x \leq 1,\ x \geq 0\ x \in \mathbb{Z})$ with dual $\min(y \mid 2y \geq 1\ y \geq 0\ y \in \mathbb{Z})$. The primal optimum is 0, while the dual optimum is 1. Without the integrality constraint, both programs are feasible with solution $x = y = 1/2$ and optimal value $1/2$.

(2) Let

$$A := \begin{pmatrix} 1 & 1 \\ -n+2 & 1 \\ n & 1 \end{pmatrix} \qquad \mathbf{b} := \begin{pmatrix} 1 \\ 1 \\ n \end{pmatrix} \qquad \mathbf{c}^t := \begin{pmatrix} 0 & -1 \end{pmatrix}$$

The integer optimum is $(0,1)$, but the LP-optimum is $(1/2, -n/2)$. See Figure 8.1. ◊

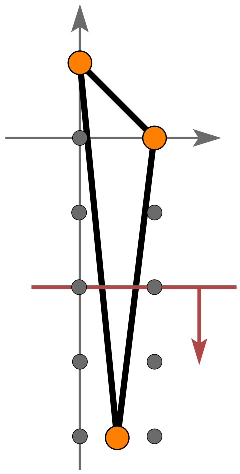


Figure 8.1

The second example suggests that, if the two solutions differ, then the matrix A must have large entries. This is in fact true, one can give bounds on the difference between an integer solution and the solution of the corresponding relaxation in terms of the size of A and $\mathbf{b}$, see e.g. Schrijver's book [Sch86].

Remark 8.4. Note that for integer linear programs we also have to be careful with transformations between different forms of the program. Consider $\max(x \mid -x \le -{}^1\!/_2,\ \le {}^3\!/_2,\ x \ge 0,\ x \in \mathbb{Z})$. This linear program in canonical form as the optimal solution $x = 1$. However, if we introduce slack variables y and z to write

$$-x + y = -{}^1\!/_2 \qquad x + z = {}^3\!/_2 \qquad x, y, z \ge 0$$

then we cannot require y, z to be integral (as the solution $x = 1$ corresponds to $y, z = {}^1\!/_2$), so the new program is not an integral linear program. We discuss transformations that preserve integer points later. ◊

integer hull

Definition 8.5. *The* **integer hull** *of P is*

$$P_I := \operatorname{conv}(P \cap \mathbb{Z}^n)\,.$$

integer polyhedron
0/1-polytope

A rational polyhedron P is an **integer polyhedron** *if $P = P_I$. If all vertices of an integer polytope are in $\{0,1\}^n$, then it is a* **0/1-polytope**. *See Figure 8.2 for an example.*

integer part
floor
fractional part
ceiling

We need some notation for the next theorem. Let $\alpha \in \mathbb{R}$. Then the **integer part** (or **floor**) $\lfloor \alpha \rfloor$ of α is the largest integer a smaller than α. The **fractional part** of α is $\{\alpha\} := \alpha - \lfloor \alpha \rfloor$. The **ceiling** $\lceil \alpha \rceil$ of α is the smallest integer larger than α.

Theorem 8.6. *Let P be a rational polyhedron. Then P_I is a polyhedron. If $P_I \neq \varnothing$, then $\operatorname{rec}(P) = \operatorname{rec}(P_I)$.*

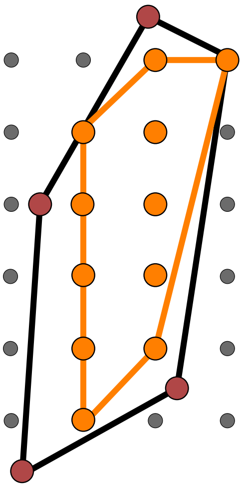


Figure 8.2

Proof. If P is a polytope, then $P \cap \mathbb{Z}^n$ is finite, so P_I is a polytope. If P is a cone, then $P = P_I$ by Remark 8.1.

So assume $P = Q + C$ for a polytope Q and a cone C. Let $\mathbf{y}_1, \ldots, \mathbf{y}_s \in \mathbb{Z}^m$ be generators of C and define

$$\Pi := \left\{ \sum_{i=1}^{s} \lambda_i \mathbf{y}_i \;\middle|\; 0 \le \lambda_i \le 1,\ 1 \le i \le s \right\}.$$

Π is a polytope (a parallelepiped), so $Q + \Pi$ and $(Q + \Pi)_I$ are polytopes. We will show that

$$P_I = (Q + \Pi)_I + C\,. \tag{8.1}$$

"$\subseteq$": Let $\mathbf{x} \in P \cap \mathbb{Z}^n$. Then there are $\mathbf{q} \in Q$ and $\mathbf{y} \in C$ such that

$$\mathbf{x} = \mathbf{q} + \mathbf{y}, \qquad \text{and} \qquad \mathbf{y} = \textstyle\sum_{i=1}^{s} \lambda_i \mathbf{y}_i$$

for some $0 \le \lambda_i \le 1$, $1 \le i \le s$. We can split the coefficients $\lambda_i \{\lambda_i\} + \lfloor \lambda \rfloor$ into their fractional and integral part. We define

$$\mathbf{z} := \textstyle\sum_{i=1}^{s} \{\lambda_i\} \mathbf{y}_i \in \Pi \qquad \text{and} \qquad \mathbf{y}' := \textstyle\sum_{i=1}^{s} \lfloor \lambda_i \rfloor \mathbf{y}_i \in C \cap \mathbb{Z}^n\,,$$

so $\mathbf{y} = \mathbf{y}' + \mathbf{z}$. Hence, we can write $\mathbf{x}$ in the form

$$\mathbf{x} = \mathbf{q} + \mathbf{y} = \mathbf{q} + \mathbf{y}' + \mathbf{z} \qquad \Longrightarrow \qquad \mathbf{x} - \mathbf{y}' = \mathbf{q} + \mathbf{z}\,.$$

The right hand side $\mathbf{x} - \mathbf{y}'$ of the second equation is in $\mathbb{Z}^n$. So also the vector on the left side of the equation is integral. This implies $\mathbf{q} + \mathbf{z} \in (Q + \Pi)_I$. As $\mathbf{y}' \in C$, we obtain

$$\mathbf{x} = (\mathbf{q} + \mathbf{z}) + \mathbf{y}' \in (Q + \Pi)_I + C\,.$$

"$\supseteq$": This follows directly from the following chain:

$$(Q + \Pi)_I + C \subseteq P_I + C = P_I + C_I \subseteq (P + C)_I = P_I\,.$$

Finally, if $P_I \neq \varnothing$, then the recession cone of P_I is uniquely defined by the decomposition in (8.1). Hence, $\operatorname{rec}(P) = C = \operatorname{rec}(P_I)$. □

So in principle, an integer linear program is just a linear program over the integer hull of a polyhedron. However, to use this approach we need an exterior description of P_I. This is in general difficult (non-polynomial) to compute. It is slightly easier (but still not polynomial) to decide whether $P_I = \varnothing$. We will see a method to compute the integer hull of arbitrary rational polytopes in Chapter 13.

A central problem for integer polyhedra is the characterization of those rational polyhedra that satisfy $P = P_I$. The next proposition gives several important criteria for this. In the Chapters 10 and 12 we will give a characterization of two important families of polyhedra with integral vertices.

Proposition 8.7. *Let P be a rational polyhedron. Then the following are equivalent.*

(1) $P = P_I$,

(2) each nonempty face of P contains an integral point,

(3) each minimal face of P contains an integral point,

(4) the linear program $\max(\mathbf{c}^t\mathbf{x} \mid \mathbf{x} \in P)$ *has an integral optimal solution for each* $\mathbf{c} \in \mathbb{R}^n$ *for which the maximum is finite.*

Proof.

(1) ⇒ (2) Let $\mathbf{c} \in \mathbb{R}^n$ and $\delta \in \mathbb{R}$ define a nonempty face F of P, i.e.

$$F = \{\mathbf{x} \mid \mathbf{c}^t\mathbf{x} = \delta\} \cap P \neq \varnothing .$$

Let $\mathbf{x} \in F$. As $P = P_I$, $\mathbf{x}$ can be written as a convex combination of integer vectors

$$\mathbf{x} = \textstyle\sum_{i=1}^k \lambda_i \mathbf{x}_i \qquad \text{for } \mathbf{x}_i \in P \cap \mathbb{Z}^n,\ \lambda_i \geq 0,\ 1 \leq i \leq k,\ \sum_{i=1}^s \lambda_i = 1.$$

Then $$\delta = \mathbf{c}^t\mathbf{x} = \mathbf{c}^t \textstyle\sum_{i=1}^k \lambda_i \mathbf{x}_i = \sum_{i=1}^k \lambda_i \mathbf{c}^t \mathbf{x}_i .$$

Hence, $\mathbf{c}^t\mathbf{x}_i \leq \delta$ for all $1 \leq i \leq k$ implies $\mathbf{c}^t\mathbf{x}_i = \delta$, so $\mathbf{x}_i \in F$ for $1 \leq i \leq k$.

(2) ⇒ (3) This is just a specialization of the assumption.

(3) ⇒ (4) If $\max(\mathbf{c}^t\mathbf{x} \mid \mathbf{x} \in P)$ is finite, then the set of optimal solutions defines a face F of P. Any face of P contains a minimal face of P, so also F contains a minimal face G of P. G contains an integral point by assumption.

(4) ⇒ (3) Let $H := \{\mathbf{x} \mid \mathbf{c}^t\mathbf{x} = \delta\}$ be a valid hyperplane that defines a minimal face F of P. Then $\delta = \max(\mathbf{c}^t\mathbf{x} \mid \mathbf{x} \in P)$, so there is some integral vector in F by assumption.

(3) ⇒ (1) Let $F_1, \ldots, F_k$ be the minimal faces of P. We can choose an integral point $\mathbf{x}_i$ in each. Let $Q := \operatorname{conv}(\mathbf{x}_1, \ldots, \mathbf{x}_k)$ and $C := \operatorname{rec}(P)$ the recession cone of P. Then $Q = Q_I$ and $C = C_I$. By our representation theorem for polyhedra, Theorem 4.38, we have $P = Q + C$. This implies

$$P = Q + C = Q_I + C_I \subseteq (Q + C)_I = P_I \subseteq P .$$

so that $P = P_I$. □

In particular, if P is pointed, then $P = P_I$ if and only if every vertex of P is integral.

Corollary 8.8. *If P is an integer polyhedron, then the integer linear programming problem over P can be solved with the simplex algorithm.* □

We have seen in Proposition 2.11 that affine images of polyhedra are polyhedra, and any function that maps polyhedra to polyhedra must be affine. Clearly the same is true for rational polyhedra and rational affine maps (i.e. affine maps $\mathbf{x} \mapsto M\mathbf{x} + \mathbf{b}$ for a rational matrix $M \in \mathbb{Q}^{m \times n}$ and a rational vector $\mathbf{b} \in \mathbb{Q}^n$). In the previous chapters we have sometimes used affine transformations to obtain a nice representation of a polyhedron. For example, up to an affine map we can assume that a polyhedron is full-dimensional. Implicitly, we also perform such a transformation when we solve systems using GAUSSian elimination. All row operations preserve the space spanned by the rows of the matrix, which ensures us that the solution we find in the affine space spanned by the matrix in row echelon form is also a point in the original space. Further, if we are given rational polyhedron $P \subseteq \mathbb{R}^n$, a linear functional $\mathbf{c}^t \in \mathbb{R}^n$, and a rational affine map $\varphi : \mathbb{R}^n \to \mathbb{R}^n$, then the optimal solutions of $\max((\varphi^*\mathbf{c})^t\mathbf{x} \mid \mathbf{x} \in \varphi(P))$ are the affine images of the optimal solutions of $\max(\mathbf{c}^t\mathbf{x} \mid \mathbf{x} \in P)$. Both properties do not hold anymore if we restrict to integer polyhedra and integral optimal solutions to integer linear programs. In general, the affine image of an integer polyhedron is not integral anymore (shift a segment with integral end points by some non-integral amount). The number of points in $|P \cap \mathbb{Z}^n|$ for a rational polyhedron P is not invariant under affine maps (Think e.g. of scaling a polyhedron. For any polytope P there is some $\varepsilon > 0$ such that $|\varepsilon P \cap \mathbb{Z}^n| \leq 1$.). Hence, finding an integral solution in an affine transform of a polyhedron doesn't tell us much about integral solutions in the original polyhedron.

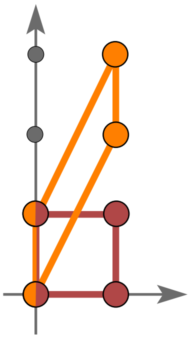


Figure 8.3

In the following we want to determine the set of transformations that preserve integrality and the information about integer points in a polyhedron, i.e. transformations that preserve the set $\mathbb{Z}^d$. Such transformations are called **unimodular**. Clearly, this is a subset of the affine transformations, so we can write such a map as $\mathbf{x} \mapsto U\mathbf{x} + \mathbf{b}$ for some $U \in \mathbb{Q}^{m \times n}$ and $\mathbf{b} \in \mathbb{Q}^n$. A translation by a vector $\mathbf{b}$ preserves $\mathbb{Z}^n$ if and only if $\mathbf{b} \in \mathbb{Z}^n$. Hence, we have to characterize linear transformations U such that $U\mathbb{Z}^n = \mathbb{Z}^n$. This is given by the following proposition.

unimodular transformation

Proposition 8.9. *Let $A \in \mathbb{Z}^{m \times m}$. Then $\det(A) = \pm 1$ if and only if the solution of $A\mathbf{x} = \mathbf{b}$ is integer for any $\mathbf{b} \in \mathbb{Z}^m$.*

Proof. "⇒": By Cramer's rule, the entries of $\mathbf{x}$ are $x_i = \pm\det(A_i)$, where A_i is the matrix obtained from A by replacing the i-th column with $\mathbf{b}$.

"⇐": If $|\det A| > 1$, then $0 < |\det A^{-1}| < 1$, so A^{-1} contains a non-integer entry a_{ij}. If $\mathbf{e}_j \in \mathbb{Z}^m$ is the j-th unit vector, then $A\mathbf{x} = \mathbf{e}_j$ has no integer solution. □

unimodular matrix

unimodularly equivalent

Definition 8.10. *A transformation $U \in \mathbb{Z}^{m \times m}$ is **unimodular** if $|\det U| = 1$.*

*Two polyhedra P and Q are **unimodularly equivalent** if there is a unimodular transformation φ that maps P to Q.*

Obviously the inverse U^{-1} and the product UV of unimodular transformation U and V are again unimodular. Figure 8.3 shows two 4-gons that are unimodularly equivalent via the unimodular transformation $\left(\begin{smallmatrix} 1 & 0 \\ 2 & 1 \end{smallmatrix}\right)$.

lattice

lattice preserving transformations

Remark 8.11. Unimodular transformations are more general those transformations that are **lattice preserving**. A **lattice** Λ is a discrete subgroup of $\mathbb{R}^n$. It can always be written in the form

$$\Lambda = \mathbb{Z}\mathbf{v}_1 + \mathbb{Z}\mathbf{v}_2 + \cdots + \mathbb{Z}\mathbf{v}_k$$

for a finite number of linearly independent vectors $\mathbf{v}_1, \ldots, \mathbf{v}_k \in \mathbb{R}^n$. A non-singular transformation T is lattice preserving if $T(\Lambda) = \Lambda$. $\mathbb{Z}^n$ is the special lattice spanned by the standard unit vectors. In the same way as for $\mathbb{Z}^n$ one could also consider integer polyhedra w.r.t. to any other lattice. ◊

The equivalent of the row echelon form of linear algebra will then be the HERMITE normal form, and we will see in the next theorem that any rational matrix can be transformed to such a matrix using certain unimodular column operations.

Hermite *normal form*

Definition 8.12. *A matrix $A = (a_{ij})_{ij} \in \mathbb{R}^{m \times n}$ with $\operatorname{rank}(A) = m$ is in **HERMITE normal form**, if the following conditions are satisfied:*

1. *$A_{*J} = 0$ for $J = [n] \setminus [m]$.*
2. *$A_{*[m]}$ is non-singular, lower-triangular, and non-negative.*
3. *$a_{ii} > a_{ij}$ for $j < i$.*

Example 8.13. The first matrix is in HERMITE normal form, the other two are not.

$$\begin{bmatrix} 5 & 0 & 0 & 0 \\ 3 & 4 & 0 & 0 \\ 2 & 1 & 4 & 0 \end{bmatrix} \qquad \begin{bmatrix} 5 & 0 & 0 \\ 3 & 4 & 0 \\ 2 & 7 & 4 \end{bmatrix} \qquad \begin{bmatrix} 5 & 0 & 0 & 0 \\ 3 & 4 & 0 & 0 \\ 6 & 1 & 7 & 4 \end{bmatrix} \qquad \Diamond$$

Similarly to the column operations used for GAUSS elimination we consider the following three **elementary unimodular column operations** to transform a matrix into HERMITE normal form:

elementary unimodular column operations

(1) exchange two columns of A,
(2) multiply a column by -1,
(3) add an integral multiple of a column to another column.

Similarly we can define **elementary unimodular row operations**. Elementary unimodular column operations on a matrix A are given by right multiplication of A with a unimodular transformation. Similarly, row operations are given by left multiplication.

elementary unimodular row operations

Theorem 8.14. *Let $A \in \mathbb{Q}^{m\times n}$. Then there is a unimodular matrix $U \in \mathbb{Z}^{m\times m}$ such that AU is in* HERMITE *normal form.*

Proof. W.l.o.g. we may assume that A is integral (scale back in the end). We use elementary column operations to transform A into HERMITE normal form. The unimodular transformation matrix U is then given by multiplication of the corresponding transformation matrices.

Suppose that we have, using elementary column operations, achieved the following situation:

$$A' := \begin{bmatrix} H & 0 \\ B & C \end{bmatrix}$$

for an integral upper triangular matrix H and some integral matrices B and C (we may start this process by choosing $C = A$).

Let $\gamma_1, \ldots, \gamma_k$ be the entries of the first row of C. Using (2) we may assume that $\gamma_i \geq 0$ for all i. As C has full row rank, not all γ_i are zero. We apply the following two steps to the first row of C:

(1) Using (1), we can reorder the columns of C so that $\gamma_1 \geq \gamma_2 \geq \ldots \neq \gamma_k$.
(2) If $\gamma_2 \neq 0$, then subtract the second column from the first and repeat.

This process terminates as in each step the total sum of all entries in the first row of C strictly decreases. Then $\gamma_2 = \gamma_3 = \ldots = \gamma_k = 0$ and the number of rows in upper triangular form in A' has increased by one. Repeating this, we obtain

$$A'' := \begin{bmatrix} H & 0 \end{bmatrix}$$

with an upper triangular integral matrix H, and all diagonal entries of H are positive. If still

$$h_{ij} < 0 \qquad \text{or} \qquad h_{ij} > h_{ii} \qquad \text{for some } i > j$$

then we can add an integral multiple of the i-th column to the j-column to ensure

$$0 \leq h_{ij} < h_{ii}\,.$$

If we apply this procedure from top to bottom, then a once corrected entry of H is not affected by corrections of later entries. This finally transforms A into HERMITE normal form. □

Remark 8.15. For a general lattice Λ spanned by some vectors $\mathbf{v}_1, \ldots, \mathbf{v}_k \in \mathbb{R}^n$ it is not obvious how we can decide whether a given vector $\mathbf{v}$ is in the lattice or not.

The HERMITE normal form solves this problem. If A is the matrix with columns $\mathbf{v}_1, \ldots, \mathbf{v}_k$ and if B is the non-singular matrix obtained from the HERMITE normal form, then $\mathbf{v}$ is in the lattice if and only if $B^{-1}\mathbf{v}$ is integral. This is immediate from the fact that the columns of A and B span the same lattice. ◊

The HERMITE normal form of a matrix $A \in \mathbb{Q}^{m \times n}$ with $\operatorname{rank}(A) = m$ is in fact unique, but we don't need this in the following. This is an easy consequence from the observation that the lattice spanned by the columns is invariant under elementary operations.

Remark 8.16. It follows that if a system $Avx = vb$ has at least one integral solution $\overline{\mathbf{x}}$ then the set of all integral solutions forms a lattice in that hyperplane, i.e. there are integral solutions $\mathbf{x}_1, \ldots, \mathbf{x}_k$ such that any integral solution is of the form

$$\overline{\mathbf{x}} + \lambda_1 \mathbf{x}_1 + \cdots + \lambda_k \mathbf{x}_k$$

for $\lambda_1, \ldots, \lambda_k \in \mathbb{Z}$. ◊

***Theorem 8.17* (Integral Alternative Theorem).** *Let $A\mathbf{x} = \mathbf{b}$ be a rational system of linear inequalities.*

Then either this system has an integer solution $\overline{\mathbf{x}}$ or there is a rational vector $\mathbf{y}$ such that $\mathbf{y}^t A$ is integral but $\mathbf{y}^t \mathbf{b}$ is not.

Proof. If $\overline{\mathbf{x}}$ is an integral solution of $A\mathbf{x} = \mathbf{b}$ and $\mathbf{y}^t A$ is integral, then also $\mathbf{y}^t \mathbf{b} = \mathbf{y}^t A \mathbf{x}$ is integral. So at most one of the two possibilities can hold.

Now assume that $\mathbf{y}^t \mathbf{b}$ is integral whenever $\mathbf{y}^t A$ is integral. If $A\mathbf{x} = \mathbf{b}$ had no solution, then the FARKAS Lemma implies that there is a vector $\mathbf{z} \in \mathbb{Q}^m$ such that $\mathbf{z}^t A = \mathbf{0}$, but $\mathbf{z}^t \mathbf{b} > 0$. By Scaling $\mathbf{z}$ with any positive factor does not change this property, so we may assume $\mathbf{z}^t \mathbf{b} = 1/2$. This is a contradiction to our assumption. Hence $A\mathbf{x} = \mathbf{b}$ has at least one solution, and we can assume that the rows of A are linearly independent.

Both statements in the theorem are invariant under elementary unimodular column operations. So by Theorem 8.14 we can assume that A has the form $A = [H \;\; 0]$ for a lower triangular matrix H. Now $H^{-1}A = [\mathrm{Id}_{\mathrm{m}} \;\; 0]$ is integral, so $H^{-1}b$ is integral by our assumption (apply it to all row vectors of H^{-1} separately). But

$$[H \;\; 0] \begin{bmatrix} H^{-1}b \\ 0 \end{bmatrix} = b$$

So $x = \begin{bmatrix} H^{-1}b \\ 0 \end{bmatrix}$ is an integral solution. □

***Proposition 8.18**. Let $A \in \mathbb{Z}^{m \times n}$ with $\operatorname{rank}(A) = m$. The following are equivalent:*

1. *The greatest common divisor of the sub-determinants of A or order m is one.*
2. *$A\mathbf{x} = \mathbf{b}$ has an integral solution $\mathbf{x}$ for each integral vector $\mathbf{b}$.*
3. *For each $\mathbf{y}$, if $\mathbf{y}^t A$ is integral, then $\mathbf{y}$ is integral.*

Proof. All three claims are invariant under unimodular column operations. Hence, we may assume that A is in HERMITE normal form $A = (B \;\; 0)$. But then, all three claims are equivalent to B being a unit matrix. □

***Proposition 8.19**. Let $P = \{\mathbf{x} \mid A\mathbf{x} \leq \mathbf{b}\}$ for rational A and $\mathbf{b}$. Then P is integral if and only if each rational supporting hyperplane of P contains an integral point.*

Proof. Assume first that P is integral. Any rational supporting hyperplane intersects P in a face F that contains a minimal face. So the claim follows from Proposition 8.7.

Now suppose that every rational supporting hyperplane contains an integral point. We may assume that A and b are integral. Let F be a minimal face of P and $I := \operatorname{eq}(F)$. So

$$\begin{aligned} F &= \{\mathbf{x} \mid A_{I*}\mathbf{x} = \mathbf{b}_I\} \cap P \\ &= \{\mathbf{x} \mid A_{I*}\mathbf{x} = \mathbf{b}_I\} \qquad \text{(by minimality of } F) . \end{aligned}$$

If F does not contain an integer point, then there is a rational $\mathbf{y}$ such that

$$\mathbf{c}^t := \mathbf{y}^t A_{I*} \in \mathbb{Z}^n \qquad \text{but} \qquad \delta := \mathbf{y}^t \mathbf{b}_I \notin \mathbb{Z} ,$$

by the integral alternative theorem, Theorem 8.17. Adding a vector $\mathbf{y}' \in \mathbb{Z}^m$ to $\mathbf{y}$ does not change this property. Hence, we may assume that $\mathbf{y} > 0$. Now let $\mathbf{z} \in P$.

Then $$\mathbf{c}^t\mathbf{z} \; = \; \mathbf{y}^t A_{I*}\mathbf{z} \; \leq \; \mathbf{y}^t\mathbf{b}_I \; = \; \delta$$

with equality if $\mathbf{z} \in P$ (here the inequality needs $\mathbf{y} > 0$). So $H := \{\mathbf{z} \mid \mathbf{c}^t\mathbf{z} = \delta\}$ is a supporting hyperplane. But $\mathbf{c} \in \mathbb{Z}^n$ and $\delta \notin \mathbb{Z}$ implies that H does not contain an integer point. This contradicts our assumption. □

A vector $\mathbf{x} = (x_1, \ldots, x_n) \in \mathbb{Z}^n$ is **primitive** if $\gcd(x_1, \ldots, x_n) = 1$. Any integral vector can be transformed into a primitive one by dividing each entry with the common g.c.d. of all entries.

primitive vector

Corollary 8.20. *Let $A \in \mathbb{Q}^{m\times n}$ and $\mathbf{b} \in \mathbb{Q}^m$. Then the following are equivalent:*

(1) There is an integral optimal solution to $\max(\mathbf{c}^t\mathbf{x} \mid A\mathbf{x} \leq \mathbf{b})$ *for each* $\mathbf{c}^t \in \mathbb{Q}^n$ *for which that maximum is finite.*

(2) $\max(\mathbf{c}^t\mathbf{x} \mid A\mathbf{x} \leq \mathbf{b})$ *is integral for each integral vector* $\mathbf{c}^t \in \mathbb{Z}^n$ *for which the maximum is finite.*

Proof. (1) ⇒ (2) If the maximum is attained for an integral vector $\mathbf{x}$ and $\mathbf{c}^t$ is integral, then the maximum is an integer.

(2) ⇒ (1) Let $H := \{\mathbf{x} \mid \mathbf{c}^t\mathbf{x} = \delta\}$ be a rational supporting hyperplane of $P := \{\mathbf{x} \mid A\mathbf{x} \leq \mathbf{b}\}$ for a primitive integral $\mathbf{c}^t$ (i.e. the g.c.d. of the entries is 1). Then $\delta = \max(\mathbf{c}^t\mathbf{x} \mid A\mathbf{x} \leq \mathbf{b})$, so $\delta \in \mathbb{Z}$. By Theorem 8.17 we know that H contains an integral point.

H was arbitrary, so any rational supporting hyperplane of P contains an integral point. Proposition 8.19 implies that P is integral, so $\max(\mathbf{c}^t\mathbf{x} \mid A\mathbf{x} \leq b)$ has an integral optimal solution whenever it is finite. □

Let C be a rational cone. Scaling the generators with a positive factor does not change the cone, so we can assume that all generators are integral and primitive. Any integral point in the cone is a rational conic combination of these generators. However, it is in general not true that any integral point in the cone is an integral conic combination of the generators.

Example 8.21. Let $C \subseteq R^2$ be the cone spanned by $\binom{1}{1}$ and $\binom{-1}{1}$. Then $\binom{0}{1}$ is in the cone, but cannot be written as

$$\binom{0}{1} \; = \; \lambda\binom{1}{1} + \mu\binom{-1}{1}$$

for integral $\lambda, \mu \geq 0$. ◊

We have to add further generators if we want to obtain a set of vectors that generate each integer point in the cone by an integral conic combination.

Definition 8.22. *Let $\mathscr{H} := \{\mathbf{h}_1, \ldots, \mathbf{h}_t\} \subset \mathbb{Q}^n$ be a finite set. $\mathscr{H}$ is a* **Hilbert basis** *of $C := \operatorname{cone}(\mathbf{h}_1, \ldots, \mathbf{h}_t)$ if every integral vector in C is an integral conic combination of $\mathbf{h}_1, \ldots, \mathbf{h}_t$. $\mathscr{H}$ is an* **integral Hilbert basis***, if all $\mathbf{h}_i$ are integral.*

Hilbert basis

integral Hilbert basis

All Hilbert bases that we consider in this course will be integral. So when we speak of a Hilbert basis, then we implicitly mean that it is integral.

Example 8.23. Let $k \in \mathbb{N}$. Consider the cone

$$C := \operatorname{cone}\left(\begin{pmatrix} 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 1 \\ k \end{pmatrix}\right)$$

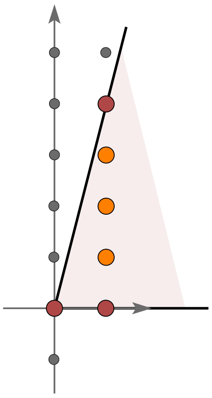


Figure 8.4

Then

$$\mathscr{H} := \left\{\begin{pmatrix} 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 1 \\ 1 \end{pmatrix}, \begin{pmatrix} 1 \\ 2 \end{pmatrix}, \dots, \begin{pmatrix} 1 \\ k \end{pmatrix}\right\}$$

is a HILBERT basis of C, see Figure 8.4. This is minimal in the sense that no subset of $\mathscr{H}$ is a HILBERT basis and any other HILBERT basis must contain $\mathscr{H}$. Hence, computation of a HILBERT basis cannot be polynomial in the size of a generating set of the cone. ◊

Although the previous example shows that a HILBERT basis can be large compared to the set of rays of a cone, it is still finite. The following theorem shows that this is always true.

Theorem 8.24. *Let $C \subseteq \mathbb{R}^m$ be a rational cone. Then C is generated by an integral* HILBERT *basis $\mathscr{H}$. If C is pointed, then there is a unique minimal* HILBERT *basis contained in every other* HILBERT *basis of the cone.*

Proof. Let $\mathbf{y}_1, \dots, \mathbf{y}_k$ be primitive integral vectors that generate C, and define the parallelepiped

$$\Pi := \left\{\sum_{i=1}^k \lambda_i \mathbf{y}_i \mid 0 \le \lambda_i \le 1,\ 1 \le i \le k\right\}.$$

as in the proof of Theorem 8.6. Let $\mathscr{H} := \Pi \cap \mathbb{Z}^n$. Observe that $\mathbf{y}_1, \dots, \mathbf{y}_k \in \mathscr{H}$, so $\mathscr{H}$ generates C. We will prove that $\mathscr{H}$ is a Hilbert basis of C.
Let $\mathbf{x} \in C \cap \mathbb{Z}^n$ be any integral vector in C. Then there are $\eta_1, \dots, \eta_k \ge 0$ such that $\mathbf{x} = \sum_{i=1}^k \eta_i \mathbf{y}_i$. We can rewrite this as

$$\mathbf{x} = \sum_{i=1}^k (\lfloor \eta_i \rfloor + \{\eta_i\})\mathbf{y}_i, \qquad \text{so that} \qquad \mathbf{x} - \sum_{i=1}^k \lfloor \eta_i \rfloor \mathbf{y}_i = \sum_{i=1}^k \{\eta_i\}\mathbf{y}_i.$$

The left side of this equation is integral. Hence, also the right side is integral. But

$$\mathbf{h} := \sum_{i=1}^k \{\eta_i\}\mathbf{y}_i \in \Pi,$$

so $\mathbf{h} \in \Pi \cap \mathbb{Z}^n = \mathscr{H}$. This implies that $\mathbf{x}$ is a integral conic combination of points in $\mathscr{H}$. So $\mathscr{H}$ is a Hilbert basis.
Now assume that C is pointed. Then there is $\mathbf{b}^t \in \mathbb{R}^n$ such that

$$\mathbf{b}^t\mathbf{x} > 0 \qquad \text{for all} \qquad \mathbf{x} \in C \setminus \{0\}.$$

Let $\quad K := \{\mathbf{y} \in C \cap \mathbb{Z}^m \mid \mathbf{y} \ne 0,\ \mathbf{y} \text{ not a sum of two other integral vectors in } C\}.$

Then $K \subseteq \mathscr{H}$, so K is finite. Assume that K is not a Hilbert basis. Then there is $\mathbf{x} \in C$ such that $\mathbf{x} \notin \mathbb{N}K$. Choose an $\mathbf{x}$ such that $\mathbf{b}^t\mathbf{x}$ is as small as possible. Since $\mathbf{x} \notin K$, there must be $\mathbf{x}_1, \mathbf{x}_2 \in C$ such that $\mathbf{x} = \mathbf{x}_1 + \mathbf{x}_2$. But

$$\mathbf{b}^t\mathbf{x}_1 \ge 0, \qquad \mathbf{b}^t\mathbf{x}_2 \ge 0, \qquad \mathbf{b}^t\mathbf{x} \ge 0 \qquad \text{and} \qquad \mathbf{b}^t\mathbf{x} = \mathbf{b}^t\mathbf{x}_1 + \mathbf{b}^t\mathbf{x}_2,$$

so $\quad \mathbf{b}^t\mathbf{x}_1 \le \mathbf{b}^t\mathbf{x}, \qquad \mathbf{b}^t\mathbf{x}_2 < \mathbf{b}^t\mathbf{x}.$

By our choice of $\mathbf{x}$ we get $\mathbf{x}_1, \mathbf{x}_2 \in \mathbb{N}K$, so that $\mathbf{x} \in \mathbb{N}K$, a contradiction. ☐

Remark 8.25. (1) The minimal Hilbert basis of a non-pointed cone is not unique:

$$\{\pm\mathbf{e}_1, \pm\mathbf{e}_2\} \qquad\qquad \{\mathbf{e}_1, \mathbf{e}_2, -\mathbf{e}_1 - \mathbf{e}_2\}$$

are both Hilbert bases of $\mathbb{R}^2$. See Figure 8.5.

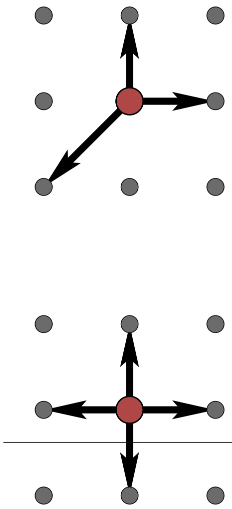


Figure 8.5

(2) Every vector in a minimal Hilbert basis is primitive. ◊

Remark 8.26. Combining Theorems 8.6 and 8.24 we obtain that for any rational polyhedron P there are integral vectors $\mathbf{x}_1, \ldots, \mathbf{x}_t$ and $\mathbf{y}_1, \ldots, \mathbf{y}_s$ such that

$$P \cap \mathbb{Z}^n = \left\{ \sum_{i=1}^t \lambda_i \mathbf{x}_i + \sum_{j=1}^s \mu_j \mathbf{y}_j \;\middle|\; \sum_{i=1}^t \lambda_i = 1,\, \lambda_i, \mu_j \in \mathbb{Z} \text{ for } 1 \le i \le t,\, 1 \le j \le s \right\}$$

This gives a parametric solution for linear Diophantine inequalities. ◊

The Matching Polytope 9

Let $G = (V, E)$ be an undirected graph. G is **bipartite**, if we can write V as the disjoint union of two non-empty sets A and B (the **color classes** of G) such that no edge has both end points in the same set.

Definition 9.1. *A* **matching** *in G is a subset $M \subset E$ of the edges so that no two edges are incident. The* **size** *of a matching M is $|M|$.*

A matching is **maximal** *if it is not contained in a matching of larger size, and it is* **maximum** *if it has the largest size among all matchings.*

A matching is **perfect***, if any vertex is incident to one edge of M.*

It is easy to see that a maximal matching covers at least half of the vertices, but P_4 shows that this may be tight. Clearly, $|V|$ must be even if G contains a perfect matching.

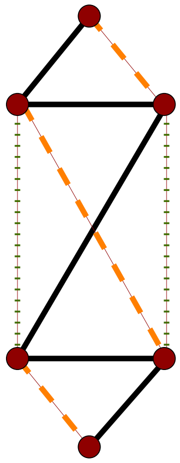


Figure 9.1

Example 9.2. Let $V := \{a, b, c, d, e, f\}$ and

$$E := \{(a, b), (a, c), (b, c), (b, d), (b, e), (c, d), (c, e), (d, f), (e, f)\}.$$

Then $M := \{(b, d), (c, e)\}$ is a maximal matching, and $M := \{(a, b), (c, d), (e, f)\}$ is a maximum and perfect matching. See also Figure 9.1. ◊

The **characteristic vector** $\chi^M \in \{0, 1\}^{|E|}$ of a matching M is the vector defined by

$$\chi_e^M := \begin{cases} 1 & \text{if } e \in M \\ 0 & \text{otherwise}. \end{cases}$$

Definition 9.3. *Let $G = (V, E)$ be a graph. The* **perfect matching polytope** *$P_{pm}(G)$ of G is*

$$P_{pm}(G) := \text{conv}(\chi^M \mid M \textit{ is a perfect matching in } G).$$

and the **matching polytope** *$P_{pm}(G)$ of G is*

$$P_m(G) := \text{conv}(\chi^M \mid M \textit{ is a matching in } G).$$

All characteristic vectors are 0/1-vectors, hence, they form a subset of the vertex set of the cube $[0, 1]^{|E|}$. This implies that any characteristic vector of a (perfect) matching is a vertex of the (perfect) matching polytope. The origin and the standard unit vectors of $\mathbb{R}^{|E|}$ are characteristic vectors of matchings, so $\dim(P_m(G) = |E|$. The dimension of $P_{pm}(G)$ is certainly smaller than $|E|$, as all points are contained in the hyperplane $\{\mathbf{x} \mid \mathbf{1}^t\mathbf{x} = {}^{|V|}/_2\}$.

Let $\omega : E \longrightarrow \mathbb{R}$ be some **weight function** on the edges of G. The **weight** of a matching M is $\sum_{e \in M} w(e)$. The problem of finding a maximum weight (perfect) matching is the linear program

$$\text{maximize} \quad \sum_{e \in E} \omega(e) x_e \qquad \text{subject to} \quad \mathbf{x} \in P_m(G) \quad (\mathbf{x} \in P_{pm}(G)).$$

In particular, choosing unit weights would return a maximum matching. This could be solved via linear programming if we had an inequality description of the polytope.

We want to derive such a description now. It is due to Jack Edmonds (1965). It is sometimes claimed that this discovery initiated the use of polyhedral methods in combinatorial optimization. We start with bipartite graphs, as they are much easier to handle. For $U \subseteq V$, let

$$\delta(U) := \{e \in E \mid |e \cap U| = 1\}$$

be the set of all edges that have exactly one end point in U. If $U = \{v\}$, then we write $\delta(v)$ instead of $\delta(\{v\})$. Further, for $\mathbf{x} \in \mathbb{R}^{|E|}$ and $S \subseteq E$ let $\mathbf{x}(S) := \sum_{e \in S} x_e$.

Clearly, any convex combination x of characteristic vectors of perfect matchings satisfies the following sets of constraints.

$$x_e \geq 0 \qquad \text{for all } e \in E \qquad (9.1a)$$

$$\mathbf{x}(\delta(v)) = 1 \qquad \text{for all } v \in V \qquad (9.1b)$$

The next theorem says that these inequalities even suffice if G is bipartite.

Theorem 9.4. *If G is bipartite then* (9.1) *give a complete description of the perfect matching polytope $P_{pm}(G)$.*

Proof. Let $Q := \{\mathbf{x} \in \mathbb{R}^{|E|} \mid \mathbf{x} \text{ satisfies } (9.1)\}$. Q is clearly contained in the 0/1-cube, so it is bounded and a polytope. Further, $P_{pm}(G) \subseteq Q$, so we need to prove $Q \subseteq P_{pm}(G)$. Let $\mathbf{x}$ be a vertex of Q, $E_{\mathbf{x}} := \{e \mid x_e > 0\}$, and $F := (V, E_{\mathbf{x}})$ the sub-graph defined by F. Suppose that F contains a cycle C. As G is bipartite, C must have even length, and C can be split into the disjoint union $C = M \mathbin{\dot\cup} N$ of two matchings in G.

For small $\varepsilon > 0$, both

$$\mathbf{y}_+ := \mathbf{x} + \varepsilon(\chi^M - \chi^N) \qquad \text{and} \qquad \mathbf{y}_- := \mathbf{x} - \varepsilon(\chi^M - \chi^N)$$

satisfy (9.1). Hence, $\mathbf{x} = \frac{1}{2}(\mathbf{y}_+ + \mathbf{y}_-)$ is a convex combination of two different vectors representing $\mathbf{x}$, so $\mathbf{x}$ is not a vertex (by Theorem 4.27). This contradicts the choice of $\mathbf{x}$, so F contains no cycles.

Hence F is a forest, so any connected component of F contains at least one node of degree 1. It follows from (9.1b) that all nodes have degree 1. Thus, $E_{\mathbf{x}}$ is a perfect matching. □

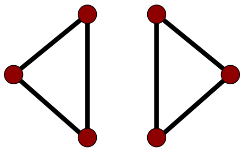


Figure 9.2

Here is an alternative proof of the theorem, that uses the characterization of the Birkhoff polytope.

Proof. As before, let $Q := \{\mathbf{x} \in \mathbb{R}^{|E|} \mid \mathbf{x} \text{ satisfies (9.1a) and (9.1b)}\}$. Then Q is a polytope that contains $P_{pm}(G)$.

Let X, Y be the two colour classes of G, and let $\mathbf{x}$ be a point of Q. By (9.1b) we obtain

$$|X| = \sum_{v \in X} \mathbf{x}(\delta(v)) = \sum_{v \in Y} \mathbf{x}(\delta(v)) = |Y|.$$

so $|X| = |Y| =: n$. Define a matrix $A = (a_{ij})_{ij}$ by $a_{ij} = x_e$ if there is the edge e between the i-th vertex of X and the j-th vertex of Y, and $a_{ij} = 0$ otherwise.

By (9.1), A is a doubly stochastic matrix. By Exercise 12.4, it is a convex combination of permutation matrices. By non-negativity, these permutation matrices have zero entries whenever A has a zero entry. So any permutation matrix in this combination correspond to a perfect matching in G. Hence $\mathbf{x}$ is a convex combination of characteristic vectors of perfect matchings in G. □

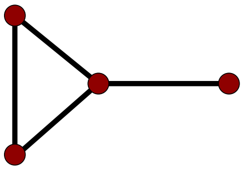


Figure 9.3

Example 9.5. (1) For a general graph the first two sets of inequalities are not enough: Let G be the union of two two disjoint triangles, i.e. the graph

$$G = (V, E) \quad V = \{A, B, C, a, b, c\} \quad E = \{(A, B), (B, C), (A, C), (a, b), (b, c), (a, c)\}.$$

Then $\mathbf{x} = \frac{1}{2}(1, 1, 1, 1, 1, 1)$ is a solution of the first two sets of inequalities, but clearly G does not have a perfect matching, so $P = \varnothing$. See Figure 9.2.

(2) The converse of the theorem is not true. Let

$$G := (V, E) \qquad V := \{a, b, c, d\} \qquad E := \{(a, b), (b, c), (c, d), (d, b)\}.$$

See Figure 9.3. Then $P_{pm}(G) = \{(1, 0, 1, 0)\}$ is characterized by (9.1). ◊

Using this theorem one can derive a description of the matching polytope $P_m(G)$, which is defined to be the convex hull of all characteristic vectors χ^M of all matching $M \subset E$ in G. Consider the set of inequalities

$$x_e \geq 0 \qquad \text{for all } e \in E \tag{9.2a}$$

$$\mathbf{x}(\delta(v)) \leq 1 \qquad \text{for all } v \in V. \tag{9.2b}$$

Clearly any convex combination of incidence vectors of matchings satisfies these inequalities.

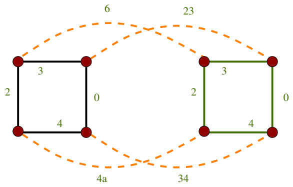


Figure 9.4

Proposition 9.6. *The matching polytope $P_m(G)$ of a graph G is determined by (9.2a) and (9.2b) if and only if G is bipartite.*

Proof. "⇒": If G is not bipartite then G has an odd cycle C. Let $x_e := \frac{1}{2}$ if $e \in C$, and $x_e = 0$ otherwise. Then x satisfies (9.2) but is not in the matching polytope.

"⇐": Define a new graph $H = (W, F)$ by taking two disjoint copies G_1 and G_2 of G and connecting corresponding vertices in the two copies G_1 and G_2 by an edge. Then H is bipartite if G is bipartite.

The first graph in Figure 9.4 shows H if G is a square. For any $\mathbf{x} \in \mathbb{R}^{|E|}$ construct a point in $\mathbf{y} \in \mathbb{R}^{|F|}$ by

(1) $y_e = x_e$ if e is an edge in G_1 or G_2, and

(2) $y_e = 1 - \mathbf{x}(\delta(u_1))$ if $e = (u_1, u_2)$ is an edge between corresponding nodes in the two copies G_1 and G_2.

By this construction we have

$$\mathbf{x} \text{ satisfies } (9.2) \qquad \Longrightarrow \qquad \mathbf{y} \text{ satisfies } (9.1). \qquad \square$$

Hence, $\mathbf{y}$ is a convex combination of perfect matchings in the graph H, which implies that $\mathbf{x}$ is a convex combination of matchings in G. Figure 9.5 shows the decomposition into perfect matchings for the graph H of Figure 9.4.

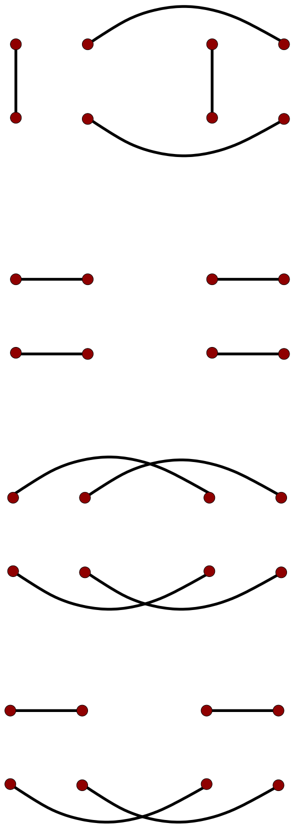


Figure 9.5

Now we head for general graphs. These are much more complicated to handle. Again we first deal with perfect matchings. The general case follows from this with almost the same proof as for Proposition 9.6.

Consider the following system of linear inequalities.

$$\mathbf{x}(\delta(U)) \geq 1 \qquad \text{for all } U \subseteq V \text{ such that } |U| \text{ is odd.} \tag{9.3}$$

It is easy to see that these inequalities are satisfied for any perfect matching in a graph G. The following theorem tells us that together with the previous inequalities (9.1) they also suffice. We will later see in Chapter 13 that the additional inequalities arise naturally from a construction of the integer hull of a general polyhedron.

***Theorem 9.7* (Edmonds 1965).** *Let $G = (V, E)$ with $|V|$ even. Then its perfect matching polytope is*

$$P_{pm}(G) = \{\mathbf{x} \mid \mathbf{x} \textit{ satisfies } (9.1) \textit{ and } (9.3)\}.$$

Proof. Let Q be the polytope determined by (9.1) and (9.3). Then we clearly have $P_{pm}(G) \subseteq Q$.

Suppose that the converse is not true. Among those graphs G for which $Q \neq P_{pm}(G)$ choose one with $|V|+|E|$ as small as possible. We may assume that $|V|$ is even, as otherwise $\mathbf{x}(V) \geq 1$ by the third condition, which implies $Q = \varnothing = P_{pm}(G)$. The same argument also shows that each connected component of G has an even number of vertices. Let $\mathbf{y} \in Q$ be a vertex of Q that is not in $P_{pm}(G)$.

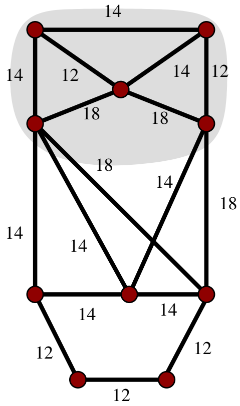


Figure 9.6

Claim 1. *$0 < y_e < 1$ for all $e \in E$.*

Proof. Assume $y_e = 0$. Let

$$E' := E - \{e\} \qquad G' := (V, E') \qquad Q := \{\mathbf{x} \in \mathbb{R}^{|E'|} \mid \mathbf{x} \text{ satisfies (9.1) and (9.3) for } G'\}\,.$$

Then the projection of $\mathbf{y}$ onto $\mathbb{R}^{|E'|}$ is in Q', but not in $P_{pm}(G')$, so G' would be a smaller counter-example.

If $y_e = 1$, then either e is an isolated edge or there is an incident edge f. In the first case we obtain a smaller counterexample by removing e and its end points. In the second case $y_f = 0$ by (9.2b), and we could remove f by the previous argument. ⋄

By (9.1b) each node has degree at least 2. This implies $|E| \geq |V|$ by double counting. If $|E| = |V|$, then G is a collection of cycles, and by (9.3) each cycle has even length. So G would be a bipartite graph. However, in this case the theorem follows from Theorem 9.4, so G is not be a counter-example. Hence we can assume that $|E| > |V|$.

By the choice of $\mathbf{y}$ there must be $|E|$ linearly independent constraints among the inequalities in (9.1) and (9.3) that are satisfied with equality.

By Claim 1 no inequality in (9.1a) is satisfied with equality. As $|E| > |V|$, at least one of the inequalities in (9.3) is satisfied with equality. Hence, there is an odd subset U of V such that $\mathbf{y}(\delta(U)) = 1$. Let $\overline{U} = V - U$. Then

$$1 \;=\; \mathbf{y}(\delta(U)) \;=\; \mathbf{y}(\delta(\overline{U}))\,.$$

We can restrict to odd subsets of size $3 \leq |U| \leq |V| - 3$, as otherwise the equation is already contained in the set (9.1b)).

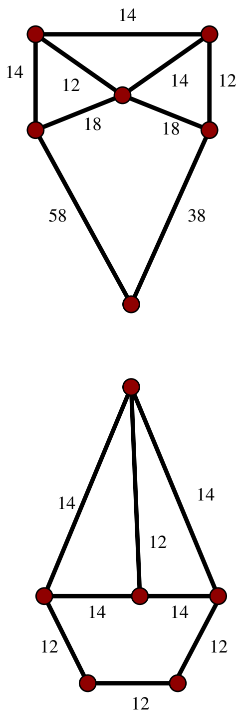


Figure 9.7

Let $H_1 := (V_1, E_1) := G/\overline{U}$. So H_1 is a graph with vertex set $V_1 := U \cup \{\overline{u}\}$ and edge set

$$E_1 := \{e \in E \mid e \subseteq U\} \cup \{(v, \overline{u}) \mid v \in U \text{ and there is } w \in \overline{U} \text{ such that } (v,w) \in E\}\,.$$

Define a projection $\mathbf{y}^{(1)}$ of $\mathbf{x}$ onto the edge set of H_1 by

$$y^{(1)}_e := y_e \qquad \text{for } e \subseteq U \qquad\qquad y^{(1)}_{(v,\overline{u})} := \sum_{w \in \overline{U},\,(v,w) \in E} y_{(v,w)} \qquad \text{for } v \in U\,.$$

Figure 9.6 shows an example of a graph G and a subset U of the vertices. Figure 9.7 shows its two projections H_1 and H_2.

Claim 2. *$\mathbf{y}^{(1)}$ satisfies (9.1) and (9.3) for H_1.*

Proof. By definition, $\mathbf{y}^{(1)} \geq 0$, so (9.1a) is satisfied. Now look at the inequalities (9.1b). We distinguish two cases, vertices $v \in U$ and $\overline{u}$.

$$v \in U: \qquad \mathbf{y}^{(1)}(\delta(v)) = \sum_{w \in U,\,(v,w) \in E} y^{(1)}_{(v,w)} + \sum_{w \in \overline{U}} y^{(1)}_{(v,w)} = \mathbf{y}(\delta(v)) = 1$$

$$\mathbf{y}^{(1)}(\delta(\overline{u})) = \sum_{\substack{v \in U\, w \in \overline{U} \\ (v,w) \in E}} y_{(v,w)} = \mathbf{y}(\delta(U)) = 1\,.$$

by the particular choice of U.

Let $W \subseteq U \cup \{\overline{u}\}$, $|W|$ odd. Again, we distinguish two cases, $\overline{u} \notin W$ and $\overline{u} \in W$.

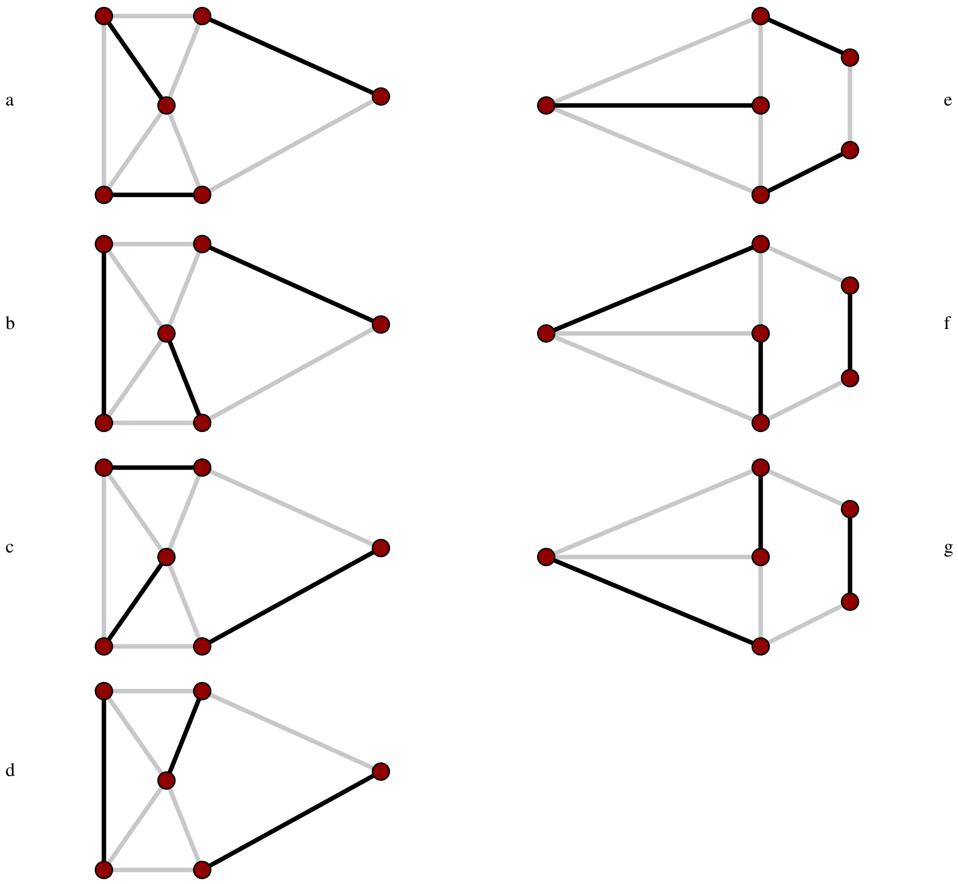


Figure 9.8: The decomposition of $y^{(1)}$ and $y^{(2)}$ into perfect matchings.

If $\overline{w} \notin W$, then by the definition of the values of $\mathbf{y}^{(1)}$ on the edges connected to $\overline{u}$. $\mathbf{y}^{(1)}(\delta(W)) = \mathbf{y}(\delta(W)) \geq 1$.

If $\overline{u} \in W$, then let $W' := W - \{\overline{u}\}$ and $W'' := U - W'$. As $|W|$ is odd, $|W'|$ is even, and $|W''|$ is odd. Thus, $\delta(W)$ is the set of all edges between $W'$and W'', and all edges between $\overline{u}$ and W''. Hence, by our definition of the values of $\mathbf{y}^{(1)}$ on the edges,

$$\mathbf{y}^{(1)}(\delta(W)) = \mathbf{y}(\delta(W'')) \geq 1\,.$$

This proves that $y^{(1)}$ also satisfies (9.3). So in total, $\mathbf{y}^{(1)}$ satisfies all inequalities in (9.1) and (9.3) with respect to the graph H_1. ◇

Now H_1 has less vertices and edges than G, so it is not a counterexample, so $\mathbf{y}^{(1)}$ is in the perfect matching polytope $P_{pm}(H_1)$ of H_1.

Similarly define the graph $H_2 := (V_2, E_2) := G/U$. The projection $\mathbf{y}^{(2)}$ of $\mathbf{y}$ onto $\mathbb{R}^{|E_2|}$ satisfies (9.1) and (9.3), so $\mathbf{y}^{(2)} \in P_{pm}(H_2)$.

So there are perfect matchings $M_1^{(1)}, \ldots, M_k^{(1)}$ of H_1 and $N_1^{(2)}, \ldots, N_l^{(2)}$ of H_2 together with coefficients $\lambda_1^{(1)}, \ldots, \lambda_k^{(1)}$ and $\lambda_1^{(2)}, \ldots, \lambda_l^{(2)}$ such that

$$\mathbf{y}^{(1)} = \sum_{i=1}^{k} \lambda_i^{(1)} \chi^{M_i^{(1)}} \qquad\qquad \mathbf{y}^{(2)} = \sum_{j=1}^{l} \lambda_j^{(2)} \chi^{N_j^{(2)}}\,.$$

See Figure 9.8 for the perfect matchings obtained for the example shown in Figures 9.6 and 9.7. y is rational, so by multiplying both equations with a suitable factor, and repeating matchings if necessary, we can assume that

$$\mathbf{y}^{(1)} = \frac{1}{p}\sum_{i=1}^{p}\chi^{M_i^{(1)}} \qquad\qquad \mathbf{y}^{(2)} = \frac{1}{p}\sum_{i=1}^{p}\chi^{N_i^{(2)}}.$$

for some $p \in \mathbb{N}$. We can even assume that p is chosen large enough that py_e is integer for all $e \in E$.

We can lift a perfect matching $M_i^{(1)}$ to a matching M_i in G that covers all edges in $\overline{U}$ and exactly one vertex in U (so it contains exactly one edge from $\delta(U)$). Let $v \in \overline{U}$ and $f := (v, \overline{u})$. Then by construction

$$py_f^{(1)} = \sum_{w \in \overline{U}} py_{(v,w)},.$$

Hence, we can choose the liftings M_i in such a way that any edge $e \in \delta(U)$ is contained in exactly py_e of them. Similarly, we can lift the matchings $N_j^{(2)}$ to matchings N_j in such a way that any edge $e \in \delta(U)$ is contained in exactly py_e of them. See Figure 9.9 for the liftings in our example..

Thus, we can pair the lifted matchings according to the edge in $\delta(U)$ they contain. Relabel the matchings from 1 to p so that M_k and N_j contain the same edge from $\delta(U)$. Then $L_k := M_k \cup N_k$ is a perfect matching in G and

$$\mathbf{y} = \frac{1}{p}\sum_{k=1}^{p}\chi^{L_k}.$$

This implies that $\mathbf{y} \in P_{pm}(G)$, in contradiction to our assumption. So $Q = p_{pm}(G)$ □

Recall that $E(U)$ for some $U \subseteq V$ is the set of all edges e with both end points in U,

$$E(U) := \{e \in e \mid e \subseteq U\}.$$

The characteristic vector of a matching in G clearly satisfies

$$x_e \geq 0 \qquad \text{for all } e \in E \tag{9.4a}$$

$$\mathbf{x}(\delta(v)) \leq 1 \qquad \text{for all } v \in V \tag{9.4b}$$

$$\mathbf{x}(E(U)) \leq \frac{1}{2}(|U| - 1) \qquad \text{for all } U \subseteq V \text{ such that } |U| \text{ is odd.} \tag{9.4c}$$

The next theorem also shows that these inequalities suffice. Its proof is completely analogous to the proof of Proposition 9.6.

Theorem 9.8*. The matching polytope $P_m(G)$ of an undirected graph $G = (V, E)$ is given by*

$$P_{pm}(G) = \{\mathbf{x} \mid \mathbf{x} \textit{ satisfies } (9.4)\}.$$

Proof. Each vector in the matching polytope satisfies the inequalities (9.4).

We need to prove that they suffice. Define a graph $H = (W, F)$ by taking two disjoint copies $G_1 = (V_1, E_1)$ and $G_2 = (V_2, E_2)$ of G and connect corresponding vertices by an edge. For subsets $X_1 \subseteq V_1$ let $X_2 \subseteq V_2$ be the corresponding set in G_2 and vice versa.

For any $\mathbf{x} \in \mathbb{R}^{|E|}$ we construct a point $\mathbf{y} \in \mathbb{R}^{|F|}$ by

(1) $y_e = x_e$ if e is an edge in G or G', and

(2) $y_e = 1 - \mathbf{x}(\delta(u_1))$ if $e = (u_1, u_2)$ is an edge between corresponding vertices in the two copies G_1 and G_2.

Assume that x satisfies (9.4). Then

(1) $\mathbf{y} \geq 0$.

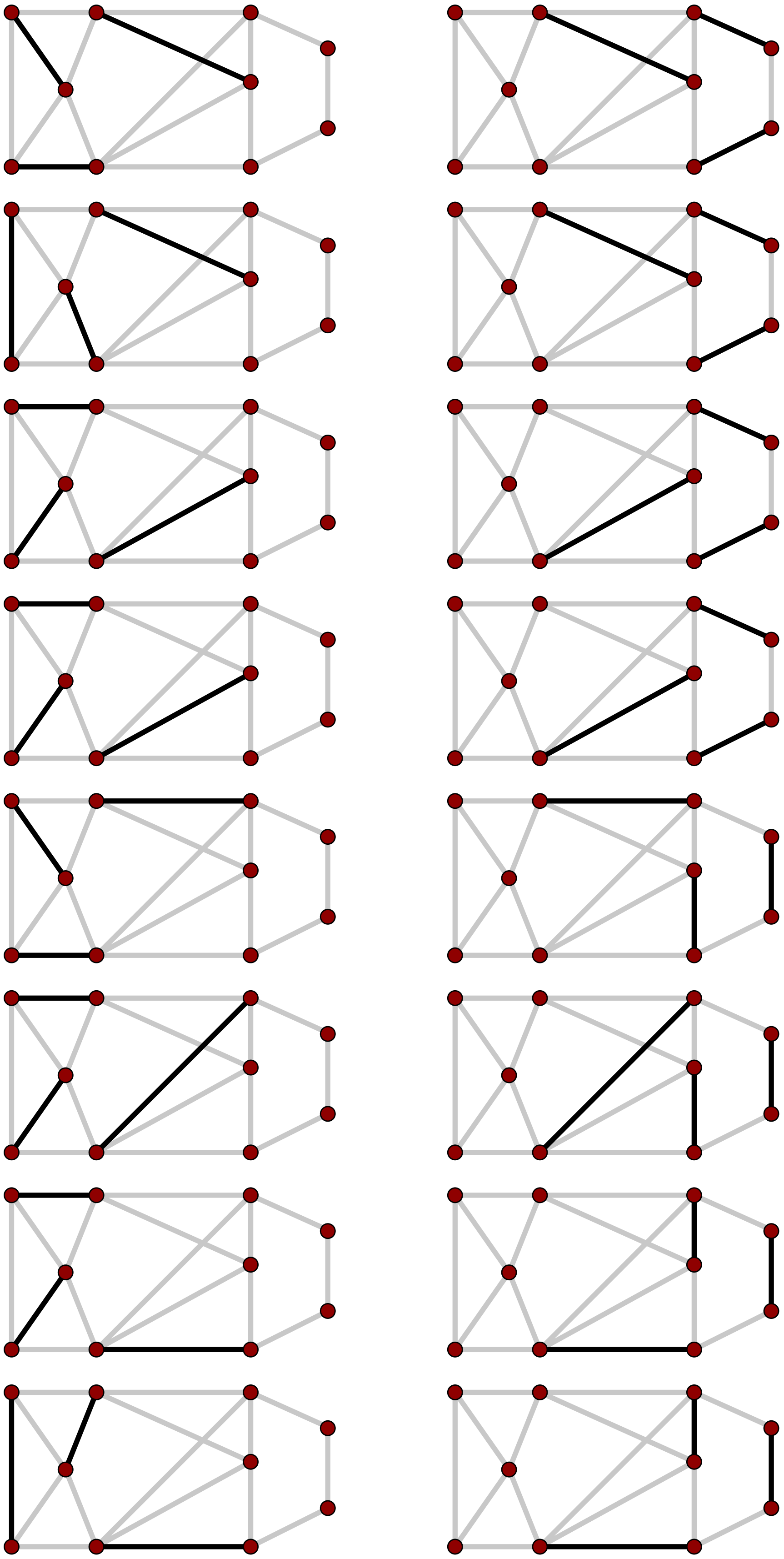


Figure 9.9: The lifted and paired matchings of the projected graphs

(2) For each $v_1 \in V_1$, $\mathbf{y}(\delta_H(v_1)) = \mathbf{y}(\delta_H(v_2)) = 1$.

(3) Let $U = X_1 \cup Y_2$ be an odd subset of $W = V_1 \cup V_2$ with $X_1, Y_1 \subseteq V_1$. Then clearly

$$\mathbf{y}(\delta_H(U)) \geq \mathbf{y}(\delta_H(X_1 - Y_1)) + \mathbf{y}(\delta_H(Y_2 - X_2)).$$

So we may assume that X_1, Y_1 are disjoint.
One of the sets X_1, Y_1 must be odd. So we may assume that this is X_1. Then we can as well assume that $Y_1 = \emptyset$.
So it suffices to show that $\mathbf{y}(\delta_H(U)) \geq 1$ for all odd $U \subseteq V_1$. Now

$$|U| = \sum_{v \in U} \mathbf{y}(\delta_H(v)) = \mathbf{y}(\delta_H(U)) + 2\mathbf{y}(E_H(U))$$

so that

$$\mathbf{y}(\delta(U)) = |U| - 2\mathbf{y}(E_H(U)) \geq |U| - 2(\frac{1}{2}(|U| - 1)) = 1.$$

Hence, $\mathbf{y}$ satisfies (9.2), so it is in the perfect matching polytope of H, so $\mathbf{x}$ is in the matching polytope of G. ☐

Unimodular Matrices 10

In this chapter we will look at the special class of unimodular matrices. The nice property of such matrices is that polyhedra that have a unimodular constraint matrix (and integral right hand side) are integral. We will see in the next chapter that many graph problems can be described by a linear program with a (totally unimodular) constraint matrix, and that this property, together with the duality theorem, gives particularly nice and simple proofs of some important theorems in graph theory. In particular, we will prove the MaxCut-MinFlow-Theorem, and MENGER's Theorem.

Definition 10.1. *Let $A \in \mathbb{Z}^{m\times n}$ be an integral matrix. Let B be a square sub-matrix B of A. Then* $\det(B)$ *is a* **minor** *of A. It is a* **maximal minor** *if $B \in \mathbb{Z}^{s\times s}$ for $s = \min(m,n)$.*

(1) $A \in \mathbb{Z}^{m\times n}$ is **totally unimodular** *if all minors are either 0 or ± 1.*
(2) A is **unimodular** *if $\operatorname{rank} A = \min(m,n)$ and each maximal minor is 0 or ± 1.*

minor
maximal minor
totally unimodular
unimodular

Observe that this is a generalization of our previous definition of a unimodular transformation to the case of non-square matrices. Let $A \in \mathbb{Z}^{m\times n}$ be totally unimodular. The following operations preserve total unimodularity:

(1) taking the transpose,
(2) adding a rwo or column that is a unit vector,
(3) deleting a row or column that is a unit vector,
(4) multiplying a row or column with -1,
(5) interchanging two rows or columns,
(6) duplicating a row or column.

Proposition 10.2. *A matrix $A \in \mathbb{Z}^{m\times n}$ is totally unimodular if and only if $(\,\mathrm{Id_m}\ \ \mathrm{A}\,)$ is unimodular.*

Proof. "⇒": Let B be a maximal sub-matrix of $(\mathrm{Id_m}\ \mathrm{A})$. Let $J \subseteq [m]$ be the set of indices of columns from $\mathrm{Id_m}$ in B. Let B' be the matrix obtained from B by deleting the first $|J|$ columns and the rows with index in J. Using the Laplace formula for determinants we obtain $\varepsilon \in \{0,1\}$ with

$$\det(B) = (-1)^{\varepsilon} \det(B').$$

By total unimodularity of A, $\det(B') \in \{0, \pm 1\}$.
"⇐": Let B be a square sub-matrix of A and let J be the row indices missed by B. Let B' be the maximal square sub-matrix obtained by taking the complete columns of A corresponding to columns of B and the unit vectors e_j, for $j \in J$, from $\mathrm{Id_m}$. Again using the Laplace formula, there is $\varepsilon \in \{0, \pm 1\}$, such that

$$\det(B) = (-1)^{\varepsilon} \det(B') \in \{0, \pm 1\}\,. \qquad \square$$

Theorem 10.3. *Let $P = \{\mathbf{x} \mid A\mathbf{x} \le \mathbf{b}\}$ for $A \in \mathbb{Z}^{m\times n}$ and $\mathbf{b} \in \mathbb{Z}^m$. If A is totally unimodular, then P is integral.*

Proof. Let $I \subseteq [m]$ correspond to a minimal face F of P. We may assume that I is a minimal system defining F, so that A_{I*} has full rank. By reordering the columns of

A_{I*} we may assume that $A_{I*} = (\,U \;\; V\,)$ for a non-singular matrix U. Since A is totally unimodular, $\det(U) = \pm 1$ and

$$\mathbf{x} = \begin{pmatrix} U^{-1}\mathbf{b}_I \\ 0 \end{pmatrix} \in \mathbb{Z}^n$$

is integral. Hence, the face F contains an integral point. ☐

Corollary 10.4. *Let $A \in \mathbb{Z}^{m\times n}$, $\mathbf{b} \in \mathbb{Z}^m$, and $\mathbf{c} \in \mathbb{Z}^n$. Assume that A is totally unimodular. Then*

$$\max(\mathbf{c}^t\mathbf{x} \mid A\mathbf{x} \le \mathbf{b}) = \min(\mathbf{y}^t\mathbf{b} \mid \mathbf{y}^t A = \mathbf{c}^t,\ \mathbf{y} \ge \mathbf{0})$$

with integral optimal solutions.

Proof. Integrality of the primal problem is immediate from the previous Theorem 10.3. We can represent the dual program using the matrix

$$\overline{A} := \begin{bmatrix} A^t \\ -A^t \\ -\mathrm{Id_m} \end{bmatrix}$$

which is also totally unimodular. So also the dual has an integral optimal solution. The equality is just the LP-duality relation. ☐

Theorem 10.5. *Let $A \in \mathbb{Z}^{m\times n}$ and $\operatorname{rank}(A) = m$. Then $P = \{\mathbf{x} \mid A\mathbf{x} = \mathbf{b}, \mathbf{x} \ge \mathbf{0}\}$ is integral for any $\mathbf{b} \in \mathbb{Z}^m$ if and only if A is unimodular.*

Proof. Suppose first that A is unimodular and that $\mathbf{b} \in \mathbb{Z}^m$. By Corollary 4.30 (1), P is pointed, so any minimal face is a vertex. Let $\mathbf{x}$ be a vertex of P. Let $I := \{i \mid x_i > 0\}$. By our characterization of vertices of polyhedra in standard form, Theorem 4.28, the columns of A_{*I} are linearly independent. Hence, we can extend I to I' such that $B := A_{*I'}$ is a maximal non-singular square sub-matrix, and $\mathbf{x} = B^{-1}\mathbf{b}$. But $\det(B) = \pm 1$, so $\mathbf{x}$ is integral by Lemma 8.9.

Now suppose that P is integral whenever $\mathbf{b}$ is integral, and let B be a maximal non-singular sub-matrix. We will use the characterization of Lemma **??** to show that $\det B = \pm 1$. So we need to show that for any $\mathbf{z} \in \mathbb{Z}^m$ also $B^{-1}\mathbf{z}$ is integral. Let $\mathbf{z} \in \mathbb{Z}^m$ and choose $\mathbf{y} \in \mathbb{Z}^m$ such that $\mathbf{x} := \mathbf{y} + B^{-1}\mathbf{z} \ge \mathbf{0}$. Let

$$\mathbf{b} := B\mathbf{x} = B\mathbf{y} + \mathbf{z} \in \mathbb{Z}^m\,.$$

We may extend $\mathbf{x}$ with zeros to a vector $\overline{\mathbf{x}}$ so that $A\overline{\mathbf{x}} = B\mathbf{x} = \mathbf{b}$. So $\overline{\mathbf{x}} \in P$, as by construction $\overline{\mathbf{x}} \ge \mathbf{0}$.

Let $I := \{i \in [n] \mid \overline{x}_i > 0\}$. The columns of A_{*I} are a subset of those of B, so $\operatorname{rank} A_{*I} = |I|$. Theorem 4.28 now implies that $\mathbf{x}$ is a vertex of P. By assumption, P is integral, so $B^{-1}\mathbf{z} = \mathbf{x} - \mathbf{y} \in \mathbb{Z}^m$. ☐

This theorem implies integrality also for other representations of polyhedra.

Proposition 10.6 **(Theorem of Hoffman and Kruskal, 1956).** *Let $A \in \mathbb{Z}^{m\times n}$. The following statements are equivalent:*

(1) A is totally unimodular.
(2) $P := \{\mathbf{x} \mid A\mathbf{x} \le \mathbf{b}, \mathbf{x} \ge \mathbf{0}\}$ is integral for any $\mathbf{b} \in \mathbb{Z}^m$.
(3) $P := \{\mathbf{x} \mid \mathbf{a} \le A\mathbf{x} \le \mathbf{b}, \mathbf{c} \le \mathbf{x} \le \mathbf{d}\}$ is integral for any $\mathbf{a}, \mathbf{b}, \mathbf{c}, \mathbf{d} \in \mathbb{Z}^m$.

Proof. (1) $\Longleftrightarrow$ (2): A is totally unimodular if and only if $(A \;\; \mathrm{Id_m})$ is unimodular, and by the previous Theorem 10.5, $(A \;\; \mathrm{Id_m})$ is unimodular if and only if

$$Q := \{(\mathbf{x}, \mathbf{y}) \mid A\mathbf{x} + \mathbf{y} = \mathbf{b},\ \mathbf{x}, \mathbf{y} \ge \mathbf{0}\}$$

is integral for any right hand side $\mathbf{b} \in \mathbb{Z}^m$. So we are done if we can show that

$$\mathbf{x} \text{ is a vertex of } P \quad \Longleftrightarrow \quad (\mathbf{x}, \mathbf{b} - A\mathbf{x}) \text{ is a vertex of } Q.$$

The polyhedron P is pointed, by Corollary 4.30(2). By our characterization of vertices, $\mathbf{x} \in P$ is a vertex of P if and only if $\mathbf{x}$ satisfies n linear independent inequalities of

$$\begin{bmatrix} A \\ -\mathrm{Id_m} \end{bmatrix} \mathbf{x} \leq \begin{bmatrix} \mathbf{b} \\ \mathbf{0} \end{bmatrix}$$

with equality. Let B be such a set of n inequalities. Let

$$\begin{aligned} I &:= B \cap [m], & J &:= \{j - m \mid j \in B \cap \{m+1, \ldots, m+n\}\}, \\ K &:= [n] \setminus J \quad \text{and} & \bar{I} &:= [m] \setminus I. \end{aligned}$$

Then $\mathbf{x}_J = \mathbf{0}$ and $A_{I*}\mathbf{x} = \mathbf{b}_I$, and the square sub-matrix A_{IK} has full rank $|I| = |K|$. This implies that also $(A \mid \mathrm{Id_m})_{\mathrm{K}\cup\bar{\mathrm{I}}}$ has full rank m. Define $\mathbf{y} \in \mathbb{Z}^m$ by

$$\mathbf{y}_I := \mathbf{0} \qquad \text{and} \qquad \mathbf{y}_{\bar{I}} := \mathbf{b}_{\bar{I}} - A_{\bar{I}*}\mathbf{x}.$$

Then $(\mathbf{x}, \mathbf{y})_{K\cup\bar{I}} = (\mathbf{x}_K, \mathbf{y}_{\bar{I}})$ and $\mathbf{x}, \mathbf{y})_{J\cup I} = \mathbf{0}$. By our characterization of vertices of polyhedra in standard form in Theorem 4.28, $(\mathbf{x}, \mathbf{y})$ is a vertex of Q, so $\mathbf{x}$ is integral. The converse direction is similar.

(1) $\Longleftrightarrow$ (3): This follows from a simple translation of P and rewriting the inequalities in the following form:

$$P = \{\mathbf{x} + \mathbf{c} \mid \mathbf{a} - A\mathbf{c} \leq A\mathbf{x} \leq \mathbf{b} - A\mathbf{c},\ \mathbf{0} \leq \mathbf{x} \leq \mathbf{d} - \mathbf{c}\} = \{\mathbf{x} \mid M\mathbf{x} \leq \mathbf{m}, \mathbf{x} \geq \mathbf{0}\}$$

for
$$M := \begin{pmatrix} A \\ -A \\ \mathrm{Id_m} \end{pmatrix} \qquad \mathbf{m} := \begin{pmatrix} \mathbf{b} - A\mathbf{c} \\ A\mathbf{c} - \mathbf{a} \\ \mathbf{d} - \mathbf{c} \end{pmatrix}.$$

Now M is unimodular if and only if A is unimodular, so we can use (2). □

A direct consequence of the second statement of the previous proposition is the following characterization of total unimodularity by the integrality of solutions of a linear program.

Corollary 10.7. *Let $A \in \mathbb{Z}^{m\times n}$. Then A is totally unimodular, if and only if for all $\mathbf{b} \in \mathbb{Z}^m$, $\mathbf{c} \in \mathbb{Z}^m$, the optimal solutions $\overline{\mathbf{x}}$ and $\overline{\mathbf{y}}$ of*

$$\max(\mathbf{c}^t\mathbf{x} \mid A\mathbf{x} \leq \mathbf{b},\ \mathbf{x} \geq \mathbf{0}) = \min(\mathbf{y}^t\mathbf{b} \mid \mathbf{y}^t A \geq \mathbf{c}^t,\ \mathbf{y} \geq \mathbf{0})$$

are integral, if the values are finite. □

With the next theorems we study criteria that ensure unimodularity of a matrix A. In particular we will see the incidence matrices of bipartite graphs and general directed graphs are always unimodular.

Theorem 10.8 **(Ghouila, Houri 1962).** *Let $A \in \mathbb{Z}^{m\times n}$. Then A is totally unimodular if and only if for each $J \subseteq [n]$ there is $\boldsymbol{\xi} \in \{0, \pm 1\}^n$ with $J = \{j \mid \xi_j \neq 0\}$ and*

$$A\boldsymbol{\xi} \in \{0, \pm 1\}^m.$$

Proof. "⇒": Let A be totally unimodular and $\chi \in \{0,1\}^n$ the characteristic vector of the chosen collection of columns of A. Consider the polyhedron

$$Q := \{\mathbf{x} \in \mathbb{R}^n \mid \lfloor \tfrac{1}{2}A\chi \rfloor \leq A\mathbf{x} \leq \lceil \tfrac{1}{2}A\chi \rceil,\ \mathbf{0} \leq \mathbf{x} \leq \chi\}.$$

Then $\frac{1}{2}\chi \in Q$, so $Q \neq \emptyset$. Further, Q is pointed by Theorem 4.30(2). So Q has a vertex $\mathbf{x}$. $\mathbf{x}$ is integral by assumption and $\mathbf{0} \le \mathbf{x} \le \chi$, so $\mathbf{x}$ is a 0/1-vector. Let

$$\boldsymbol{\xi} := \chi - 2\mathbf{x} \in \{0, \pm 1\}^m .$$

Then $\boldsymbol{\xi} \equiv \chi (\mod 2)$. Further, $A\boldsymbol{\xi} = A\chi - 2A\mathbf{x}$, so

$$A\chi - 2\lceil \tfrac{1}{2}A\chi \rceil \;\le\; A\boldsymbol{\xi} \;\le\; A\chi - 2\lfloor \tfrac{1}{2}A\chi \rfloor .$$

which implies that $A\boldsymbol{\xi} \in \{0, \pm 1\}^m$. So $\boldsymbol{\xi}$ is an incidence vector of a partition of the columns as required in the theorem.

"$\Leftarrow$": We use induction to prove that every $(k \times k)$-sub-matrix has determinant 0 or ± 1. If $k = 1$ then this follows from our assumption that each column has only entries in 0 and ± 1. In particular, all entries of A are 0 or ± 1.

Let $k \ge 2$ and B a non-singular sub-matrix of A with column vectors $\mathbf{b}_1, \ldots, \mathbf{b}_k$. By Cramer's rule, the entry $\widetilde{b}_{ij}$ at position (i, j) of the inverse B^{-1} of B is

$$\widetilde{b}_{ij} = \frac{\det B_{ij}}{\det B} .$$

where B_{ij} is the matrix obtained from B by replacing the j-th column with the i-th unit vector. Using the induction hypothesis and the Laplace rule we know that $\det B_{ij} \in \{0, \pm 1\}$. So

$$\overline{B} := (\det B) B^{-1}$$

has only entries in $\{0, \pm 1\}$. Let $\overline{\mathbf{b}}$ be the first column of $\overline{B}$ and $I := \{i \mid \overline{b}_i \neq 0\}$. By assumption there is $\boldsymbol{\xi} \in \{0, \pm 1\}^n$ such that $I = \{i \mid \xi_i \neq 0\}$ and $B\xi \in \{0, \pm 1\}^m$.

By construction, $B\overline{\mathbf{b}} = (\det B)\mathbf{e}_1$, where $\mathbf{e}_1$ is the first unit vector. So for $2 \le i \le k$ we get

$$0 = \mathbf{e}_i^t B \overline{\mathbf{b}} = \sum_{j \in I} \overline{b}_j b_{ij} .$$

As $b_{ij} \in \{0, \pm 1\}$ this implies that for all i the set $\{j \in I \mid b_{ij} \neq 0\}$ is even. Hence

$$B\boldsymbol{\xi} \;=\; \varepsilon \mathbf{e}_1 \qquad\qquad \text{for some } \varepsilon \in \{0, \pm 1\}$$

$\varepsilon = 0$ would imply that B is singular, so $\varepsilon = \pm 1$. Now $\frac{1}{\det B} B\overline{\mathbf{b}} = \mathbf{e}_1 = \pm B\xi$, so

$$\boldsymbol{\xi} = \pm \frac{1}{\det B} \overline{\mathbf{b}} .$$

But both $\boldsymbol{\xi}$ and $\overline{\mathbf{b}}$ have only entries in $\{0, \pm 1\}$, so $\det B = \pm 1$. ☐

Corollary 10.9. *Let M be an $(n \times m)$-matrix with entries in $\{0, \pm 1\}$ and the property that each column contains at most one 1 and at most one -1. Then M is totally unimodular.*

Proof. We give two proofs of this:

1. The matrix M is totally unimodular if and only if M^t is totally unimodular. By assumption, the sum of any collection of column vectors of M^t is a vector in $\{0, \pm 1\}^m$, so the claim follows from Theorem 10.8.
2. Let B a $(k \times k)$-sub-matrix. We prove the result by induction on k. The result is obvious if $k = 1$, so assume $k \ge 2$.
   If B contains a column with no non-zero entry then $\det B = 0$. If B contains a column with exactly one non-zero entry then the result follows from the Laplace formula applied to this column, and the induction hypothesis. Finally, if all columns of B have exactly two non-zero entries then the sum of the entries in each column is 0, so $\det B = 0$. ☐

Definition 10.10. *Let $G = (V, A)$ be a directed graph with n vertices $v_1, \dots, v_n$ and m edges $a_1, \dots, a_m$. The* **incidence matrix** *$D_G := (d_{ij}) \in \{0, \pm 1\}^{n \times m}$ of G is defined by*

incidence matrix, undirected graph

$$d_{ij} := \begin{cases} 1 & \text{if } a_j \text{ is an outgoing arc of } v_i \\ -1 & \text{if } a_j \text{is an incoming arc of } v_i \\ 0 & \text{otherwise} \end{cases}$$

Corollary 10.11. *The incidence matrix of a directed graph $G = (V, A)$ is totally unimodular.* ☐

Definition 10.12. *Let $G = (V, E)$ be an undirected graph with n vertices $v_1, \dots, v_n$ and m edges $e_1, \dots, e_m$. The* **incidence matrix** *$D_G := (d_{ij}) \in \{0, \pm 1\}^{n \times m}$ of G is defined by*

incidence matrix, directed graph

$$d_{ij} := \begin{cases} 1 & \text{if } a_j \text{ is incident to } v_i \\ 0 & \text{otherwise} \end{cases}$$

Theorem 10.13. *The incidence matrix of an undirected graph $G = (V, E)$ is totally unimodular if and only if the graph is bipartite.*

Proof. First assume that G is bipartite. By renumbering the vertices of the graph we may assume that D_G splits into

$$D_G = \begin{bmatrix} D_1 \\ D_2 \end{bmatrix}$$

where D_1 corresponds to the vertices in one color class and D_2 to the vertices in the other. So any column of D_1 and D_2 contains exactly one 1. Let B be an $(k \times k)$-submatrix of D_G. We prove the theorem by induction. If $k = 1$, then the claim is easy, so assume $k \geq 2$. We distinguish three cases:

(1) If B contains a column with no non-zero entry, then $\det(B) = 0$.
(2) If B contains a column with exactly one non-zero entry, then we can use the Laplace formula and the induction hypothesis.
(3) If all columns of B contain exactly two non-zero entries, then the sum of the rows contained in D_1 equals the sum of the rows contained in D_2. Hence, $\det B = 0$.

If G is not bipartite, then it contains an circuit of odd length k. Let B be the $(k \times k)$-matrix defined by this circuit. Then we can reorder the rows so that

$$B = \begin{bmatrix} 1 & 0 & \cdots & \cdots & 1 \\ 1 & 1 & & & 0 \\ & 1 & \ddots & & \vdots \\ & & \ddots & \ddots & \vdots \\ & & & 1 & 0 \\ & & & 1 & 1 \end{bmatrix}$$

Hence, $\det B = \pm 2$. ☐

Lemma 10.14. *Let B be a $m \times m$-matrix with entries in $\{0, \pm 1\}$. If $|\det B| > 2$, then B has a square sub-matrix C with $|\det C| = 2$.*

Proof. Consider the matrix $\overline{B} := (\, B \ \mathrm{Id_m} \,)$. We consider the following operations on the matrix $\overline{B}$:

(1) adding or subtracting a row to another row,
(2) multiplying a row with -1.

We transform $\overline{B}$ into a new matrix $\widetilde{B}$ using these operations in such a way that

(1) all entries of $\widetilde{B}$ are still in $\{0, \pm 1\}$,
(2) all unit vectors occur among the columns of $\widetilde{B}$,
(3) the first k columns of $\widetilde{B}$ are unit vectors.

Let $\widetilde{B}$ be such that k is maximal. Up to sign these operations do not alter the determinant of maximal sub-matrices. Reordering the first and second m columns of $\widetilde{B}$ allows us to assume that there is an $(m \times m)$-matrix B' such that

$$\overline{B}' = \begin{bmatrix} \mathrm{Id_k} & B' & 0 \\ 0 & & \mathrm{Id_l} \end{bmatrix}.$$

Up to sign the first m columns of $\widetilde{B}$ have the same determinant as B. $|\det B| \geq 2$ implies $k < m$.

If we cannot transform any further column among the first m to a unit vector without violating condition (1), then up to multiplication of rows with -1 there must be $I = \{i_1, i_2\}$ and $J = \{j_1, j_2\}$ such that

$$\widetilde{B}_{IJ} = \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix} \qquad \text{or} \qquad \widetilde{B}_{IJ} = \begin{bmatrix} 1 & -1 \\ -1 & -1 \end{bmatrix}.$$

Let $\widetilde{H}$ be the matrix with columns j_1, j_2 and all unit vectors except e_{i_1} and e_{i_2} and let H be the corresponding sub-matrix of $\overline{B}$. Then $|\det H| = |\det \widetilde{H}| = 2$, and as all columns of H corresponding to columns of $\overline{B}$ with index $> m$ are unit vectors, B must have a sub-matrix of determinant ± 2. □

Corollary 10.15. *Let $A \in \{0, \pm 1\}^{m \times n}$. Then A is totally unimodular if and only if no square sub-matrix has determinant ± 2.* □

Theorem 10.16. *Let $A \in \mathbb{Z}^{m \times n}$. Then A is totally unimodular if and only if each non-singular sub-matrix B of A has a row with an odd number of non-zero entries.*

Proof. "⇒": The number of non-zero entries in a row is odd if and only if the sum of the entries in that row is odd.

Assume that B is a $(k \times k)$-sub-matrix such that all row sums even. We have to show that B is singular. By Theorem 10.8, there is a vector $\xi \in \{\pm 1\}^k$ such that $B\xi \in \{0, \pm 1\}^k$. But even row sums imply $B\xi = 0$, so B is singular.

"⇐": Let B be a non-singular sub-matrix of A. By induction, any proper sub-matrix of B is unimodular, so if the claim fails, then $|\det B| \geq 2$. By Corollary 10.15, $\det B = \pm 2$.

Consider B as a matrix over GF(2). Its determinant is 0, so the columns of B are linearly dependent over GF(2). Hence, there is $\lambda \in \{0, 1\}^m$ such that $B\lambda$ has only even entries. Let $J := \{j \mid \lambda_j = 1\}$. Then the row sums of all rows of B_{*J} are even. Hence, by assumption, no maximal sub-matrix of B_{*J} is non-singular. So $\det B = 0$, a contradiction. □

Theorem 10.17. *Let $A \in \mathbb{R}^{m \times n}$ be a matrix of full row rank. Then the following are equivalent:*

(1) For each basis B of A the matrix $A_B^{-1}A$ is integral,
(2) for each basis B of A the matrix $A_B^{-1}A$ is totally unimodular,
(3) there is a basis B of A such that the matrix $A_B^{-1}A$ is totally unimodular.

Proof. We may assume that $A = [\mathrm{Id_m}\ \overline{\mathrm{A}}]$ as all three claims are invariant under multiplication of A by a non-singular matrix.

(1) ⇒ (2): It suffices to show that A is unimodular. Choosing $B = [m]$, i.e. $A = \mathrm{Id_m}$ shows that A is integral. For any other basis we obtain from $A_B^{-1}\mathrm{Id_m} = \mathrm{A_B^{-1}}$ that A_B^{-1} is integral. Hence $\det A_B = \pm 1$.

(2) ⇒ (3): By specialization.

(3) ⇒ (1): Let B be such a basis. Then $A_B^{-1}A$ is integral. For any other basis B', $A_{B'}^{-1}A = A_{B'}^{-1}A_B A_B^{-1}A$, and $A_{B'}^{-1}A_B$ is integral by the total unimodularity of A (its rows are the solution of $A_{B'}^{-1}\mathbf{x} = A_{*j}$ for $j \in B$). □

***Theorem 10.18**. Let $A \in \mathbb{Z}^{m\times n}$. The A is totally unimodular if and only if for each non-singular sub-matrix B and each vector $\mathbf{y}$ with entries in $\{0, \pm 1\}$, the g.c.d. of the entries of $\mathbf{y}^t B$ is one.*

Proof. Let B and $\mathbf{y}$ be as in the theorem. Let g be the g.c.d. of the entries of $\mathbf{y}^t B$. Then $\frac{1}{g}\mathbf{y}^t B$ is integral. If A is totally unimodular, then B^{-1} is integral, so $\frac{1}{g}\mathbf{y}^t = \frac{1}{g}\mathbf{y}^t BB^{-1}$ is integral, so $g = 1$.

Now assume that the g.c.d. is 1 for every combination of $\mathbf{y}$ and B. Then A has only entries in $\{0, \pm 1\}$. Let B be a non-singular sub-matrix of A. The common g.c.d. of the entries of $\mathbf{1}^t B$ is 1, so at least one entry is odd. Hence, B has a row with an odd number of non-zero entries. The claim now follows from Theorem 10.16. □

Remark 10.19. Let $G = (V, A)$ be a directed graph and $T = (V, A_0)$ a directed tree on the same vertex set as G. Let M be the following $(|A_0| \times |A|)$-matrix: For $a_0 \in A_0$ and $a = (u, v) \in A$ let P be the (unique) undirected path from u to v. The entry at (a_0, a) is

$$\begin{array}{rl} 1 & \text{if } a_0 \text{ occurs in } P \text{ in forward direction,} \\ -1 & \text{if } a_0 \text{ occurs in } P \text{ in backward direction,} \\ 0 & \text{if the path does not run through } a. \end{array}$$

This matrix is the **network matrix** of G and T. If we allow loops in the graph, then class of network matrices is closed under taking sub-matrices.

network matrix

In 1980, Seymour proved that all totally unimodular matrices can be built from network matrices and two other matrices by using eight types of operations. The decomposition can be done in polynomial time. This implies that the decision problem whether some given matrix A is totally unimodular is in $\mathscr{P}$. ◊

Applications of Unimodularity 11

In this chapter we want to present several applications of total unimodularity and integer linear programming to combinatorial optimization. We need the following notion.

characteristic vector

Definition 11.1. *Let $A \subseteq B$ be two sets. The* **characteristic vector** *of A in B is the vector $\chi^A = (\chi^A_j)_{j\in B} \in \{0,1\}^{|B|}$ defined by*

$$\chi^A_j := \begin{cases} 1 & \text{if } j \in A \\ 0 & \text{otherwise}. \end{cases}$$

matching

matching number

Let $G := (V, E)$ be an undirected graph with n vertices and m edges. A **matching** M in G as a subset of the edge set E such that no two edges in M are incident. The **matching number** of G is

$$\nu(G) := \max(|M| \mid M \text{ is a matching in } G).$$

Maximum Matching Problem

Clearly, $\nu(G) \le \frac{1}{2}|V|$. The Maximum Matching Problem asks for the largest cardinality of a matching in a graph (observe, that not every inclusion maximal matching realizes this number). We want to use integer linear programming to find this number, so we have to translate this into a geometric question.

Let D be the incidence matrix of G. Then $\mathbf{x} \in \mathbb{R}^m$ is the incidence vector of a matching in G if and only if

$$D\mathbf{x} \le \mathbf{1}, \qquad \mathbf{x} \ge \mathbf{0}, \qquad \mathbf{x} \text{ integer}.$$

Then the size of a maximum matching is just

$$\nu(G) = \max(\mathbf{1}^t\mathbf{x} \mid D\mathbf{x} \le \mathbf{1}, \mathbf{x} \ge \mathbf{0}, \mathbf{x} \in \mathbb{Z}^m). \tag{11.1}$$

Now we restrict to the case that G is bipartite. Then D is a totally unimodular matrix, so by Proposition 10.6 the polytope

$$P_m(G) := \{\mathbf{x} \mid D\mathbf{x} \le \mathbf{1}, \mathbf{x} \ge \mathbf{0})$$

defined by D is integral. Note that this is exactly the matching polytope of the bipartite graph G that we have studied already in Theorem 9.4. Hence, all basic optimal solutions of

$$\max(\mathbf{1}^t\mathbf{x} \mid D\mathbf{x} \le \mathbf{1}, \mathbf{x} \ge \mathbf{0}). \tag{11.2}$$

are integral and coincide with those of (11.1). So we can find a maximum matching in a bipartite graph with the simplex algorithm. Let us look at the dual linear program:

$$\min(\mathbf{y}^t\mathbf{1} \mid \mathbf{y}^t D \ge \mathbf{1}, \mathbf{y} \ge \mathbf{0}). \tag{11.3}$$

Again, total unimodularity of D implies that all optimal solutions of this program are integral, so (11.3) coincides with

$$\min(\mathbf{y}^t\mathbf{1} \mid \mathbf{y}^t D \ge \mathbf{1}, \mathbf{y} \ge \mathbf{0}, \mathbf{y} \in \mathbb{Z}^n). \tag{11.4}$$

Let $\overline{\mathbf{y}} \in \mathbb{R}^n$ be an optimal solution. No entry of $\overline{\mathbf{y}}$ is larger than 1, as reducing it to 1 yields a feasible solution with smaller objective value. Thus $\overline{\mathbf{y}} \geq \mathbf{0}$implies that $\overline{\mathbf{y}}$ is the incidence vector of some subset

$$W := \{v \mid y_v = 1\} \subseteq V\,.$$

vertex cover

vertex cover number

of the vertices of G. The condition $\overline{\mathbf{y}}^t D \geq \mathbf{1}$ implies that for each edge $e = (u, v) \in E$ at least one of the end points u and v is in W. Such a set is a **vertex cover** in G. See also Figure 11.1. The **vertex cover number** of G is

$$\tau(G) := \min(|W| \mid W \text{ is a vertex cover of } G)\,.$$

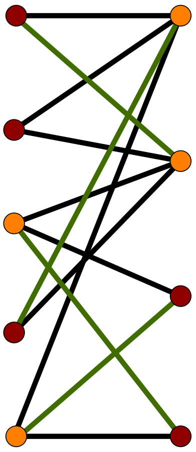


Figure 11.1

Any integral optimal solution of (11.4) defines a vertex cover of minimal size $\tau(G)$ in G. Clearly, for any graph, not necessarily bipartite, this number is related to the matching number via

$$\nu(G) \leq \tau(G)\,,$$

since any matching edge must be covered and no vertex can cover more than one matching edge. Already the graph K_3 shows that this inequality may be strict in general. See Figure 11.2.

However, if G is bipartite, then $\nu(G)$ and $\tau(G)$ are given by (11.1) and (11.4), or equivalently, by (11.2) and (11.3). The duality theorem implies that the values of these linear programs coincide. This proves the following classical result of König and Egerváry.

***Theorem 11.2* (König, Egerváry 1931).** *Let G be a bipartite graph. Then the maximum size of a matching in G equals the minimum size of a vertex cover in G:*

$$\nu(G) \;=\; \tau(G)\,. \qquad \square$$

vertex cover polytope

In the same way as for matchings we may also define the **vertex cover polytope** $P_{vc}(G) \subseteq \mathbb{R}^{|V|}$:

$$P_{vc}(G) := \operatorname{conv}(\chi^W \mid W \text{ is a vertex cover of } G)\,.$$

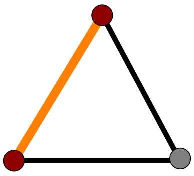


Figure 11.2

Here G need not be bipartite. However, for bipartite graphs we know from (11.4) that $P_{vc}(G)$ is defined by the inequalities

$$0 \leq y_v \leq 1 \qquad \text{for each } v \in V \qquad (11.5a)$$

$$y_u + y_v \geq 1 \qquad \text{for each } e = (u, v) \in E \qquad (11.5b)$$

In analogy to the case of the matching polytope of a bipartite graph, we can characterize bipartite graphs by the exterior description of the vertex cover polytope.

Theorem 11.3. *G is bipartite if and only if $P_{vc}(G)$ is determined by (11.5a) and (11.5b).*

Proof. We have already seen above that (11.5) suffice if G is bipartite. So suppose G contains an odd circuit C with n vertices. Define $\mathbf{y} := \frac{1}{2}\mathbf{1} \in \mathbb{Z}^n$. Then $\mathbf{y}$ satisfies (11.5). However, any vertex cover of G contains at least $\frac{n+1}{2}$ vertices of C, but $y(C) = \frac{n}{2}$. So $\mathbf{y} \notin P_{vc}(G)$. $\square$

perfect matching

k-regular graph

Theorem 11.2 has some nice consequences. Recall that a matching is **perfect** if each vertex is incident to a matching edge. A graph G is **k-regular** if all vertices have the same degree k, i.e. if all vertices are incident to exactly k edges.

***Corollary 11.4* (Frobenius' Theorem).** *A bipartite graph $G = (V, E)$ has a perfect matching if and only if each vertex cover has size at least $\frac{1}{2}|V|$.*

Proof. G has a perfect matching if and only if $\nu(G) \geq \frac{1}{2}|V|$. □

***Corollary 11.5* (König).** *Each k-regular bipartite graph $G = (V, E)$, $k \geq 1$, has a perfect matching.*

Proof. $|E| = \frac{1}{2}k|V|$. Any vertex covers exactly k edges, hence, any vertex cover contains at least $\frac{1}{2}|V|$ vertices. □

More generally, we can look for **weighted matchings** in a graph G. Given a **weight function** $\mathbf{w} \in \mathbb{Z}_+^m$ on the edges of G, the **weight** of a matching $M \subset E$ is

weighted matching
weight function
weight

$$\text{weight}(M) := \sum_{e \in M} w_e\,.$$

The **Maximum Weighted Matching Problem** is the task to find a matching of maximum weight. Expressed as an integer linear program, we want to solve

Maximum Weighted Matching Problem

$$\max(\mathbf{w}^t\mathbf{x} \mid D\mathbf{x} \leq \mathbf{1},\, \mathbf{x} \geq \mathbf{0},\, \mathbf{x} \in \mathbb{Z}^m)\,.$$

If G is bipartite, then D is totally unimodular, and we know that any optimal solution of the relaxed problem

$$\max(\mathbf{w}^t\mathbf{x} \mid D\mathbf{x} \leq \mathbf{1},\, \mathbf{x} \geq \mathbf{0})$$

is integral. Note however, that a maximum weighted matching need not have maximum cardinality. We obtain a min-max relation as before.

Proposition 11.6. *Let $G = (V, E)$ be a bipartite graph and $\mathbf{w} \in \mathbb{Z}_+^m$ a weight function on the edges.*

The maximum weight of a matching in G is equal to the minimum value of $\mathbf{1}^t\mathbf{f}$, where $\mathbf{f}$ ranges over all $\mathbf{f} \in \mathbb{Z}_+^n$ that satisfy

$$f_u + f_v \geq w_e \qquad \text{for all edges } e = (u, v) \in E\,.$$

Proof. Using total unimodularity of the incidence matrix D of G and the duality theorem this statement can be written as

$$\max(\mathbf{w}^t\mathbf{x} \mid D\mathbf{x} \leq \mathbf{1},\, \mathbf{x} \geq \mathbf{0}) \;=\; \min(\mathbf{f}^t\mathbf{1} \mid \mathbf{f}^t D \geq \mathbf{w}^t,\, \mathbf{f} \geq \mathbf{0})\,. \qquad \square$$

A vector $\mathbf{f}$ as in the proposition is called a **w-vertex cover**. The value $\mathbf{1}^t\mathbf{f}$ is the **size** of $\mathbf{f}$. So the proposition just says that the maximum weight of a matching in a bipartite graph equals the minimum size of a $\mathbf{w}$-vertex cover.

$\mathbf{w}$-vertex cover
size

Another closely related problem is the **Minimum Weight Perfect Matching Problem** (alternatively you can of course also search for the maximum). Given a graph G and a weight function $\mathbf{w} \in \mathbb{Z}^m$ on the edges, we want to find a perfect matching M in G of minimum weight

Minimum Weight Perfect Matching Problem

$$\text{weight}(M) := \sum_{e \in M} w_e\,.$$

So we want to solve

$$\min(\mathbf{w}^t\mathbf{x} \mid D\mathbf{x} = \mathbf{1},\, \mathbf{x} \geq \mathbf{0},\, \mathbf{x} \in \mathbb{Z}^m)\,.$$

If G is bipartite, then this coincides with

$$\min(\mathbf{w}^t\mathbf{x} \mid D\mathbf{x} = \mathbf{1},\, \mathbf{x} \geq \mathbf{0})\,.$$

The corresponding polyhedron $P_{pm}(G)$ is the perfect matching polytope of Theorem 9.4. Unimodularity and the duality theorem again immediately imply the following statement.

Proposition 11.7. *Let $G = (V, E)$ be a bipartite graph having a perfect matching and let $\mathbf{w} \in \mathbb{Q}^m$ be a weight function.*

The minimum weight of a perfect matching is equal to the maximum value of $\mathbf{1}^t\mathbf{f}$ taken over all $\mathbf{f} \in \mathbb{Z}^n$ such that

$$y_u + y_v \geq w_e \qquad \text{for each edge } e = (u, v) \in E \qquad \square$$

edge cover

edge cover number

In a similar way we can also derive another classical result of König about bipartite graphs. An **edge cover** in G is a subset $F \subseteq E$ of the edges of G such that for any vertex $v \in V$ there is an incident edge in F. The **edge cover number** of G is

$$\rho(G) := \min(|F| \mid F \text{ is an edge cover of } G).$$

Determining $\rho(G)$ is the MINIMUM EDGE COVER PROBLEM.

stable set

stable set number

A **stable set** in G is a subset $S \subseteq V$ of the vertices such that no two vertices in G are connected by an edge. The **stable set number** $\alpha(G)$ is

$$\alpha(G) := \max(|S| \mid S \text{ is a stable set in } G).$$

Finding $\alpha(G)$ is the MAXIMUM STABLE SET PROBLEM.

***Theorem 11.8* (König 1933).** *Let G be a bipartite graph. Then the maximum size of a stable set equals the minimum size of an edge cover in G.*

Proof. Let D be the incidence matrix of G, and let $S \subseteq V$. Then $\mathbf{y} \in \mathbb{R}^n$ is the characteristic vector of a stable set in G if and only if

$$\mathbf{y}^t D \leq \mathbf{1} \qquad \mathbf{y} \geq \mathbf{0} \qquad \mathbf{y} \text{ integer}.$$

A characteristic vector $\mathbf{x} \in \mathbb{R}^m$ defines an edge cover in G if and only if $\mathbf{x} \in \{0, 1\}^m$ and $D\mathbf{x} \geq \mathbf{1}$. Again, we can reformulate this to

$$D\mathbf{x} \geq \mathbf{1}, \qquad \mathbf{x} \geq \mathbf{0}, \qquad \mathbf{x} \text{ integer}.$$

The incidence matrix of G is totally unimodular by Theorem 10.13. So integer and linear optima coincide. Hence, using the duality theorem, the optimal value of the integer linear program

$$\max(\mathbf{1}^t\mathbf{y} \mid \mathbf{y}^t D \leq \mathbf{1}, \mathbf{y} \geq \mathbf{0}, \mathbf{y} \in \mathbb{Z}^n) \qquad \text{(MAX STABLE SET)}$$

equals the optimal value of the linear program

$$\min(\mathbf{1}^t\mathbf{x} \mid D\mathbf{x} \geq \mathbf{1}, \mathbf{x} \geq \mathbf{0}\, \mathbf{x} \in \mathbb{Z}^m), \qquad \text{(MIN EDGE COVER)}$$

where the first computes the size of a maximum stable set and the second the size of a minimum edge cover. $\square$

This theorem is *dual* to Theorem 11.2 in that it interchanges the rôle of vertices and edges, and minimum and maximum.

Proposition 11.9. *Let $G = (V, E)$ be a bipartite graph and $\mathbf{w} \in Z_+^m$ a weight function on the edges.*

The minimum weight of an edge cover in G is equal to the maximum of $\mathbf{1}^t\mathbf{f}$, where the maximum is taken over all $\mathbf{f} \in \mathbb{Z}_+^n$ that satisfy $f_u + f_v \leq w_e$ for all edges $e = (u, v) \in E$.

Proof. This is completely analogous to the proof of Proposition 11.6. Total unimodularity of D implies that the theorem is equivalent to

$$\max(\mathbf{w}^t\mathbf{x} \mid D\mathbf{x} \geq \mathbf{1}, \mathbf{x} \geq \mathbf{0}) = \min(\mathbf{f}^t\mathbf{1} \mid \mathbf{f}^t D \leq \mathbf{w}^t, \mathbf{f} \geq \mathbf{0}).$$

This equality is now just the duality theorem. $\square$

Now we consider a directed graph $G = (V,A)$ with n vertices and m arcs. For a set $U \subseteq V$ let

$$\delta^+(U) := \{(u,v) \mid u \in U,\ v \notin U\}$$

be the set of outgoing edges of U. For a vertex $v \in V$ we set $\delta^+(v) := \delta^+(\{v\})$. Similarly, we define the sets $\delta^-(U)$ and $\delta^-(v)$ of incoming edges. A subset $C \subseteq A$ is a **cut** in G if $C = \delta^+(U)$ for some $U \subseteq V$. C is **proper** if $\varnothing \neq U \neq V$. Let $c \in (\mathbb{R}_+ \cup \{\infty\})^m$ be a **capacity** on the arcs of G. The **capacity** of a cut C is

cut
proper cut
capacity

$$\mathrm{c}(C) \ := \ \sum_{a \in C} c_a \,.$$

Let $\mathbf{f} \in \mathbb{R}^A_+$ be a non-negative vector on the arcs of G. $\mathbf{f}$ satisfies the **flow conservation condition** at a vertex $v \in V$ if

flow conservation

$$\sum_{a \in \delta^-(v)} f_a \ = \ \sum_{a \in \delta^+(v)} f_a \,.$$

Let $s, t \in V$ be two vertices of G. $\mathbf{f}$ is a s-t**-flow** in G if

s-t-flow

(1) $\mathbf{f}$ satisfies flow conservation at all vertices $v \in V - \{s,t\}$ and
(2) $\mathrm{v}(\mathbf{f}) := \sum_{a \in \delta^+(s)} f_a - \sum_{a \in \delta^-(s)} f_a \ = \ \sum_{a \in \delta^-(t)} f_a - \sum_{a \in \delta^+(t)} f_a \ \geq \ 0\,.$

s is the **source** of the flow and t the **sink**. $\mathrm{v}(\mathbf{f})$ is the **value** of the flow $\mathbf{f}$. It is the net out-flow of the source (which equals the net in-flow of the sink). A flow is a **circulation** in G if flow conservation also holds at s and t. Let $\mathbf{c} \in (\mathbb{R}_+ \cup \{\infty\})^m$ be a capacity on the arcs of G. A flow $\mathbf{f}$ is **subject to c** if $\mathbf{f} \leq \mathbf{c}$. A **maximum** s-t**-flow** is an s-t-flow subject to $\mathbf{c}$ of maximum value. Finding such a flow is the MAXIMUM FLOW PROBLEM.

source, sink
value of a flow
circulation
flow subject to c
maximum s-t-flow
Maximum Flow Problem

We want to transform this into a linear program. Let D be the incidence matrix of G and D' the matrix obtained by deleting the rows corresponding to the vertices s and t. Then D' is totally unimodular by Corollary 10.9 and

$$\mathbf{f} \text{ is an } s\text{-}t\text{-flow subject to } \mathbf{c} \qquad \Longleftrightarrow \qquad D\mathbf{f} \ = \ \mathbf{0}, \ \ \mathbf{0} \leq \mathbf{f} \leq \mathbf{c}. \tag{11.6}$$

Let $\mathbf{w}$ be the row of D corresponding to s. Then $\mathbf{w}$ has a 1 at the position of an outgoing edge of s, and a -1 at the position of an incoming edge. Hence, if $\mathbf{f}$ is a s-t-flow on G, then $\mathrm{v}(\mathbf{f}) = \mathbf{w}^t\mathbf{f}$.

The MAXIMUM FLOW PROBLEM can now be written as

$$\max(\mathbf{w}^t\mathbf{f} \mid D'\mathbf{f} \ = \ \mathbf{0}, \ \ \mathbf{0} \leq \mathbf{f} \leq \mathbf{c})\,. \tag{11.7}$$

Although the matrix D' is totally unimodular, we cannot conclude integrality of an optimal solution, as the capacities may not be integer. However, the cost function $\mathbf{w}$ is integral, as it is given by a row of the integral matrix D. So we can pass to the dual program

$$\min(\mathbf{y}^t\mathbf{c} \mid \mathbf{y} \geq \mathbf{0} \text{ and there is } \mathbf{z} \text{ such that } \mathbf{y}^t + \mathbf{z}^t D' \geq \mathbf{w}^t)\,. \tag{11.8}$$

Here, the constraint matrix and right hand side are

$$M \ := \ \begin{pmatrix} -\mathrm{Id_m} & -\mathrm{Id_m} \\ -D' & 0 \end{pmatrix} \qquad\qquad \mathbf{b} \ := \ \begin{pmatrix} -\mathbf{w}^t \\ \mathbf{0} \end{pmatrix}.$$

By Theorem 10.3 the polyhedron $\{\mathbf{x} \mid \mathbf{x}^t M \leq \mathbf{b}^t\}$ is integral, so (11.8) has an integral optimal solution $(\overline{\mathbf{y}}, \overline{\mathbf{z}})$. We want to interpret this in the graph G.

Extend $\overline{\mathbf{z}}$ by $\overline{z}_s = -1$ and $\overline{z}_t = 0$. Then

$$\overline{\mathbf{y}}^t + \overline{\mathbf{z}}^t D \geq 0\,. \tag{11.9}$$

Define

$$U := \{v \in V \mid \bar{z}_v \leq -1\} \qquad \text{and} \qquad C := \delta^+(U)\,.$$

then $s \in U$ and $t \notin U$, so C is an s-t-cut in G. We want to compute the capacity of C. Let $a = (u, v) \in C$. By construction,

$$\bar{z}_u \leq -1 \qquad \text{and} \qquad \bar{z}_v \geq 0\,.$$

But (11.9) implies that $y_a + z_u - z_v \geq 0$, so $y_a \geq z_v - z_u \geq 1$. Hence

$$\mathrm{c}(C) \leq \mathbf{y}^t \mathbf{c} = \mathrm{v}(\mathbf{f}) \tag{11.10}$$

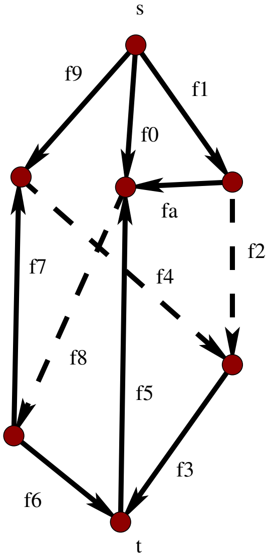


Figure 11.3

by duality. However, the capacity of any cut C in G is clearly an upper bound on the size of an s-t-flow in G, so we have equality in (11.10). This is the well known MaxFlow-MinCut Theorem. See Figure 11.3 for an example. The dashed edges in the graph are a cut. The arc labels $i|J$ denote the flow and the capacity of that arc.

***Theorem 11.10* (Ford & Fulkerson 1956).** *Let $G = (V, A)$ be a directed graph, $\mathbf{c} \geq \mathbf{0}$ a capacity function on the edges of G and s, t two vertices of G.*

The maximum value of an s-t-flow in G equals the minimum capacity of an s-t-cut in G. If all capacities are integer, then the optimal flow can be chosen to be integer.

Proof. The only missing piece is the integrality of $\mathbf{f}$ if $\mathbf{c}$ is integer. But in this case the value of

$$\max(\mathbf{w}^t \mathbf{f} \mid D'\mathbf{f} \;=\; \mathbf{0}, \mathbf{0} \;\leq\; \mathbf{f} \;\leq\; \mathbf{c}, \mathbf{f} \in \mathbb{Z}^m)$$

equals (11.7) by total unimodularity of D. □

s-t-path

Let s, t be vertices of a directed graph $G = (V, A)$. An **s-t-path** P is an alternating sequence

$$P = v_0 = s, a_0, v_1, a_1, v_2, \ldots, a_{k-2}, v_{k-1}, a_{k-1}, v_k = t\,,$$

arc disjoint path

It is **arc disjoint** if $a_i \neq a_j$ for all $0 \leq i < j \leq k-1$. Let $P_1, \ldots, P_k$ be s-t-paths in G, and $\chi_{P_1}, \ldots, \chi_{P_k} \in \{0, 1\}^m$ their incidence vectors. Then

$$\mathbf{f} := \chi_{P_1} + \cdots + \chi_{P_k}$$

is an integral flow of value k in G.

***Theorem 11.11* (Menger 1927).** *Let $G = (V, A)$ be a directed graph and s, t vertices of G. Then the maximum number of pairwise arc disjoint s-t-paths equals the minimum size of an s-t-cut.*

Proof. Define a capacity $\mathbf{c}$ on G by $c_a = 1$ for all $a \in A$. Then the size of a cut equals its capacity. Let C_0 be a cut in G of minimum size s. Let $P_1, \ldots, P_k$ be a set of arc disjoint s-t-paths. Then $s \geq k$ as C_0 contains at least one arc of each path.

By the MaxFlow-MinCut-Theorem, and as $\mathbf{c}$ is integral, there is an integral flow $\mathbf{f}$ of value s. The capacity constraints imply that $f_a \in \{0, 1\}$ for each $a \in A$.

We show by induction that $\mathbf{f}$ can be decomposed into s disjoint s-t-paths. Let $B := \{a \mid f_a = 1\}$. Clearly, if $\mathbf{f} \neq \mathbf{0}$, then flow conservation implies that B contains an s-t-path P. Hence, we can construct a new st-s-flow $\mathbf{f}'$ by

$$f' := f - \chi_P\,.$$

$\mathbf{f}'$ has value $s - 1$. Further, $f'_a = 0$ if $a \in P$. Repeating this, we obtain s arc disjoint s-t-paths in G. □

We finish this chapter with an example for the primal-dual-method for solving a linear program that we have seen towards the end of Chapter 6. There we started from the dual linear programs in the form

$$\begin{aligned}\max(\mathbf{c}^t\mathbf{x} \mid A\mathbf{x} = \mathbf{b},\, \mathbf{x} \geq \mathbf{0})\\ \min(\mathbf{y}^t\mathbf{b} \mid \mathbf{y}^t A \geq \mathbf{c}^t)\,,\end{aligned}$$

where $A \in \mathbb{R}^{m\times n}$, $\mathbf{b} \in \mathbb{R}^m$, and $\mathbf{c} \in \mathbb{R}^n$. We want to have a slightly different form here: We multiply $A, \mathbf{b}$ and $\mathbf{c}$ by -1 to obtain the dual pair of programs

$$\begin{aligned}\min(\mathbf{c}^t\mathbf{x} \mid A\mathbf{x} = \mathbf{b},\, \mathbf{x} \geq \mathbf{0})\\ \max(\mathbf{y}^t\mathbf{b} \mid \mathbf{y}^t A \leq \mathbf{c}^t)\,.\end{aligned} \tag{11.11}$$

The primal-dual method applies with almost no change also to these programs. We just have to reverse all inequality signs. Given a feasible dual solution $\overline{\mathbf{y}} \in \mathbb{R}^m$, and the set $J := \{j \mid \mathbf{y}^t A_{*j} = c_j\} \subset [m]$ we looked at the restricted primal program

$$\min(\mathbf{1}^t\mathbf{z} \mid \mathbf{z} + A_{*J}\mathbf{u} = \mathbf{b},\, \mathbf{u}, \mathbf{z} \geq \mathbf{0})\,.$$

and the corresponding dual program

$$\max(\mathbf{w}^t\mathbf{b} \mid \mathbf{w} \leq \mathbf{1},\, \mathbf{w}^t A_{*J} \leq \mathbf{0})\,.$$

Using the optimal solution $\overline{\mathbf{w}}$ of the dual program we showed that there is some $\varepsilon > 0$ such that $\overline{\mathbf{y}} + \varepsilon\overline{\mathbf{w}}$ is a dual feasible solution with better objective value.

Let us consider again the MAXFLOW-MINCUT problem that we have introduced above. We give a different representation as a linear program. Let $G := (V, A)$ be a directed graph with n vertices and m arcs. Let $\mathbf{d} \in \mathbb{Z}^V$ be defined by

$$d_v := \begin{cases} 1 & \text{if } v = t\\ -1 & \text{if } v = s\\ 0 & \text{otherwise}\,.\end{cases}$$

Let $\mathbf{f} \in \mathbb{Z}^A$ denote a flow in G, $\mathbf{c} \in \mathbb{R}^A$, $\mathbf{c} \geq \mathbf{0}$ the capacities on the arcs, and $w \in \mathbb{R}$ the flow value. Then a maximum flow can be computed by solving the linear program

$$\begin{aligned}&\max(w \mid D\mathbf{f} - \mathbf{d}w \leq \mathbf{0},\, \mathbf{0} \leq \mathbf{f} \leq \mathbf{c})\\ &\quad = \max(w \mid D\mathbf{f} - \mathbf{d}w \leq \mathbf{0},\, -\mathbf{f} \leq \mathbf{0},\, \mathbf{f} \leq \mathbf{c})\,.\end{aligned} \tag{11.12}$$

The inequality system $D\mathbf{f} - w\mathbf{d} \leq \mathbf{0}$ may seem to be a weaker condition that flow conservation, as it only requires that the in-flow of a vertex $v \neq s, t$ is larger than the out-flow. However, $\mathbf{b} := \mathbf{1}^t(D\mathbf{f} + w\mathbf{d}) = \mathbf{0}$, which implies that every entry of $\mathbf{b}$ is $\mathbf{0}$. So $D\mathbf{f} + w\mathbf{d} \leq \mathbf{0}$ implies $D\mathbf{f} + w\mathbf{d} = \mathbf{0}$ for any feasible solution $(\mathbf{f}, \mathbf{d})$. The additional row $\mathbf{d}$ of the constraint matrix can be seen as an additional arc from t to s that sends all flow back to the source. This way we have flow conservation at any vertex (i.e. we have a **circulation** in the graph), and we want to maximize the flow on the artificial arc.

circulation

We can use this linear program as the dual program in the pair (11.11). As $\mathbf{c} \geq \mathbf{0}$ we know that $\mathbf{f} = \mathbf{0}$, $w = 0$ is a feasible solution of this linear program, so we can start the primal-dual algorithm. Let now $(\mathbf{f}, w)$ be any dual feasible solution. In order to write down the restricted linear program we have to find all inequalities of the dual program that are satisfied with equality by our solution.

By our above consideration, the inequalities in $D\mathbf{f} + w\mathbf{d} \leq \mathbf{0}$ are always satisfies with equality. Let

$$U := \{a \in A \mid f_a = c_a\} \qquad\qquad L := \{a \in A \mid f_a = 0\}$$

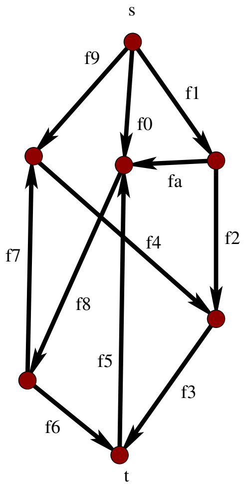


Figure 11.4

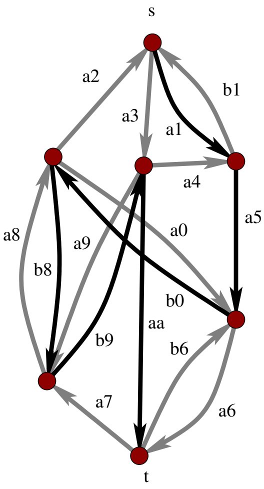


Figure 11.5

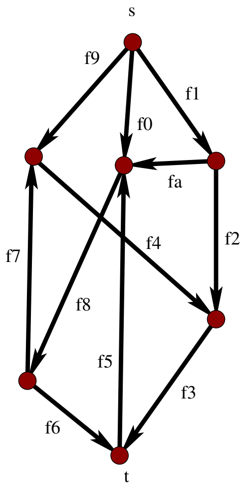


Figure 11.6

The dual of the restricted linear program for the solution $(\mathbf{f}, w)$ is

$$\max(z \mid D\mathbf{x} - z\mathbf{d} \le \mathbf{0},\ x_U \le 0,\ -x_L \le 0,\ x \le \mathbf{1},\ z \le 1)\,.$$

The constraint matrix of this linear program is totally unimodular, so we have an integral optimal solution $(\mathbf{x}, z)$. If $z = 0$, then our flow was optimal. Otherwise, $z = 1$, and we may interpret the dual program in the following way:

Let G^r be directed graph on the vertex set V obtained from G in the following way.

(1) for all arcs $a \in L$ we insert a into G^r,
(2) for all arcs $a = (u, v) \in U$ we insert (v, u) into G^r,
(3) for all arcs $a = (u, v) \in A - (U \cup L)$ we insert both (u, v) and (v, u) into G^r.

G^r is the **residual graph** of G corresponding to the flow $\mathbf{f}$. The linear program 11 finds a path P from s to t in G^r:

(1) We have an integral flow $\mathbf{x}$ that is ≤ 1 on each arc.
(2) Its value is 1, so $-1 \le x_a \le 1$ on all arcs.
(3) $0 \le x_a \le 1$ on arcs $a \in L$
(4) $-1 \le x \le 0$ on arcs in $a \in U$ (so $0 \le x_a \le 1$ on the reversed arcs in G^r).
(5) Flow conservation implies that at each vertex we have at most one incoming and one outgoing arc.

(we may create additional disconnected loops in the graph. They don't improve the flow.) The path P is an **augmenting path** in G^r. Figures 11.4-11.6 show an example. The first figure shows a graph with a feasible flow and capacities on the arcs. The second figure shows an augmenting path, and the last the new flow after adding the augmenting path.

The new dual feasible solution

$$\tilde{\mathbf{f}} := \mathbf{f} + \varepsilon\mathbf{x} \qquad\qquad \tilde{w} := w + \varepsilon z \tag{11.13}$$

increases the flow along this path by some $\varepsilon > 0$. We can determine the maximum possible ε by assigning capacities to the arcs in G^r:

(1) c_a on arcs originating from L or U.
(2) $c_a - f_a$ on the forward and f_a on the backward arc for all other arcs.

The minimum of the capacities along the augmenting path is the maximum ε. This is strictly positive by construction. Viewed in the graph, the flow update in (11.13) does the following. For each forward arc in the path P (i.e. one that has the same orientation as in G), we add ε to the flow on the arc. For each backward arc in the path we subtract ε. All other arcs stay unchanged. This is the Ford-Fulkerson algorithm for finding a maximum flow in a directed graph. The arguments above show that it is in fact an algorithm based on linear programming.

If $\mathbf{c}$ is integral, then $\varepsilon \ge 1$. In this case, the flow value increases by at least one in each step. As $\mathbf{1}^t\mathbf{c}$ is an upper bound for the maximum flow, this algorithm reaches the optimum after a finite number of steps. The same argument applies if $\mathbf{c}$ is rational, as we can scale the capacities to be integer. If $\mathbf{c}$ is not rational, then this algorithm in this form may fail to finish in a finite number of steps. However, choosing the augmenting path more carefully also guarantees finiteness for non-rational capacities (see any book on graph algorithms).

If $\mathbf{c}$ is integral, then we can also choose the flow to be integral, as we may start with the flow $\mathbf{f} = \mathbf{0}$ and then always increase the flow on the arcs by an integer. By an argument similar to the one in the proof of Menger's theorem, one can show that such an integral flow can always be decomposed into a set of paths.

There are more graph algorithms that can be seen to be a primal-dual linear programming algorithm. Among them is e.g. the Hungarian algorithm for finding a minimum weight perfect matching.

Total Dual Integrality

12

In the last two chapters we found a condition on the constraint matrix A that ensured the polyhedron $P := \{\mathbf{x} \mid A\mathbf{x} \le \mathbf{b}\}$ to be integral for any integral right hand side $\mathbf{b}$. Matrices with this property are exactly the totally unimodular matrices. We have seen that already this quite special class has many interesting applications.

In this chapter we want to study a more general relation. This time we want to consider which combinations of a constraint matrix A and a right hand side vector $\mathbf{b}$ lead to integer polyhedra. This naturally leads to the notion of totally dual integral systems of linear inequalities. We will see that this completely characterizes integrality of a polyhedron (however, as integer linear programming is not known to be in $\mathscr{P}$, there is no good algorithm that detects this). In the next chapter we will see that this leeds to a method to construct the integer hull of a polyhedron by cutting the polyhedron with affine hyperplanes that separate a fractional vertex from all integral points in the polyhedron.

***Definition 12.1* (Edmonds and Giles 1977).** *A rational system $A\mathbf{x} \le \mathbf{b}$ is* **totally dual integral (TDI)** *if*

totally dual integral
TDI

$$\min(\mathbf{y}^t\mathbf{b} \mid \mathbf{y}^tA = \mathbf{c}^t, \mathbf{y} \ge \mathbf{0}) = \max(\mathbf{c}^t\mathbf{x} \mid A\mathbf{x} \le \mathbf{b}) \qquad (12.1)$$

has an integral optimum dual solution $\overline{\mathbf{y}}$ for all integral $\mathbf{c}^t \in \mathbb{Z}^n$ for which the dual program is finite.

Adding any valid inequality to a TDI system preserves TDIness:

Proposition 12.2. *If $A\mathbf{x} \le \mathbf{b}$ is a TDI system and $\mathbf{c}^t\mathbf{x} \le \delta$ is a valid inequality for $P := \{\mathbf{x} \mid A\mathbf{x} \le \mathbf{b}\}$, then also $A\mathbf{x} \le \mathbf{b}$, $\mathbf{c}^t\mathbf{x} \le \delta$ is TDI.* ☐

Proposition 12.3. *Let $A\mathbf{x} \le \mathbf{b}$ be a TDI-system and $\mathbf{b} \in \mathbb{Z}^m$. Then $\max(\mathbf{c}^t\mathbf{x} \mid A\mathbf{x} \le \mathbf{b})$ has an integral optimal solution for any rational $\mathbf{c}^t$ for which the program is bounded.*

Proof. If $\mathbf{b}$ is integral, then $\min(\mathbf{y}^t\mathbf{b} \mid \mathbf{y}^tA = \mathbf{c}^t, \mathbf{y} \ge \mathbf{0})$ has an integral optimal value for all integral $\mathbf{c}^t$, as $A\mathbf{x} \le \mathbf{b}$ is TDI. Now the claim follows from Corollary 8.20. ☐

Corollary 12.4. *Let $A\mathbf{x} \le \mathbf{b}$ be a TDI system and $\mathbf{b} \in \mathbb{Z}^m$. Then $P := \{\mathbf{x} \mid A\mathbf{x} \le \mathbf{b}\}$ is an integral polyhedron.* ☐

Theorem 10.3 (or Corollary 10.4) now immediately implies the following characterization, which implies that TDI is a generalization of total unimodularity in the sense that it considers both the matrix A and the right hand side $\mathbf{b}$ instead of just the matrix A.

Corollary 12.5. *A rational system $A\mathbf{x} \le \mathbf{b}$ is TDI for each vector $\mathbf{b}$ if and only if A is totally unimodular.*

Proof. Total unimodularity of A implies that the polyhedron $P^* := \{\mathbf{y} \mid \mathbf{y}^tA = \mathbf{c}^t, \mathbf{y} \ge \mathbf{0}\}$ is integral for any integral $\mathbf{c}^t$. So $A\mathbf{x} \le \mathbf{b}$ is TDI. If $A\mathbf{x} \le \mathbf{b}$ is TDI for any integral $\mathbf{b}$, then $P := \{\mathbf{x} \mid A\mathbf{x} \le \mathbf{b}\}$ is integral for any $\mathbf{b}$, so A is totally unimodular. ☐

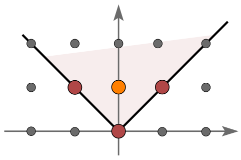


Figure 12.1

Example 12.6. TDI is a property of the inequality system $Ax \leq b$, not of the polyhedron it defines:

Consider $A_1 := \begin{pmatrix} -1 & -1 \\ 0 & -1 \\ 1 & -1 \end{pmatrix}$ and $A_2 := \begin{pmatrix} -1 & -1 \\ 1 & -1 \end{pmatrix}$

Then $A_1\mathbf{x} \leq \mathbf{0}$ and $A_2\mathbf{x} \leq \mathbf{0}$ define the same cone (see Figure 12.1), but the first is TDI and the second is not. One may observe that the rows of the first matrix form a Hilbert basis of the dual cone, while the rows of the second do not. We will see later that this indeed characterizes TDI. ◊

In the next chapter we will discuss an algorithm that computes the integer hull of a polyhedron. This is based on the fact that we can describe any polyhedron by a TDI system and that faces of a polyhedron described by such a system are again given by a TDI system. This will use the following theorem.

Theorem 12.7. *Let $A\mathbf{x} \leq \mathbf{b}$, $\mathbf{a}^t\mathbf{x} \leq \beta$ be a TDI system. Then also $A\mathbf{x} \leq \mathbf{b}$, $\mathbf{a}^t\mathbf{x} = \beta$ is TDI.*

Proof. Let $\mathbf{c}^t \in \mathbb{Z}^n$ such that

$$\max(\mathbf{c}^t\mathbf{x} \mid A\mathbf{x} \leq \mathbf{b},\, \mathbf{a}^t\mathbf{x} = \beta) = \max(\mathbf{c}^t\mathbf{x} \mid A\mathbf{x} \leq \mathbf{b},\, \mathbf{a}^t\mathbf{x} \leq \beta, -\mathbf{a}^t\mathbf{x} \leq -\beta)$$

is finite with optimal solution $\overline{\mathbf{x}}$. The dual of this program is

$$\min(\mathbf{y}^t\mathbf{b} + (z_+ - z_-)\beta \mid \mathbf{y}^t A + (z_+ - z_-)\mathbf{a}^t = \mathbf{c}^t,\, \mathbf{y} \geq \mathbf{0})\,. \tag{12.2}$$

We have to find an integral optimal solution of this program. Let $(\overline{\mathbf{y}}, \overline{z}_+, \overline{z}_-)$ be a (possibly fractional) optimal solution. Choose $k \in \mathbb{N}$ such that

$$\overline{z}_- \leq k \qquad \text{and} \qquad k \cdot \mathbf{a} \in \mathbb{Z}^n$$

and let $\hat{\mathbf{c}}^t := \mathbf{c}^t + k\mathbf{a}^t$, and $\overline{u} := \overline{z}_+ - \overline{z}_- + k$.

$\overline{\mathbf{x}}$ is also a feasible solution of

$$\max(\hat{\mathbf{c}}^t\mathbf{x} \mid A\mathbf{x} \leq \mathbf{b},\, \mathbf{a}^t\mathbf{x} \leq \beta) \tag{12.3}$$

and $(\overline{\mathbf{y}}, \overline{u})$ is a feasible solution of the dual program

$$\min(\mathbf{y}^t\mathbf{b} + u\beta \mid \mathbf{y}^t A + u\mathbf{a}^t = \hat{\mathbf{c}}^t, \mathbf{y} \geq \mathbf{0}, u \geq 0)\,. \tag{12.4}$$

The duality theorem implies that both programs (12.3) and (12.4) are bounded.
As $A\mathbf{x} \leq \mathbf{b}$, $\mathbf{a}^t\mathbf{x} \leq \beta$ is TDI, the system (12.4) has an integral optimal solution $(\tilde{\mathbf{y}}, \tilde{u})$.

Let $\tilde{z}_+ := \tilde{u}$ and $\tilde{z}_- := k$.

Then $\tilde{\mathbf{y}}^t A + (\tilde{z}_+ - \tilde{z}_-)a^t = \tilde{\mathbf{y}}^t A + \tilde{u}\mathbf{a}^t - k\mathbf{a}^t = \hat{\mathbf{c}}^t - k\mathbf{a}^t = \mathbf{c}^t$.

so $(\tilde{\mathbf{y}}, \tilde{z}_+, \tilde{z}_-)$ is an integral feasible solution of (12.2). Its objective value is

$$\tilde{\mathbf{y}}^t\mathbf{b} + (\tilde{z}_+ - \tilde{z}_-)\beta = \tilde{\mathbf{y}}^t\mathbf{b} + \tilde{u}\beta - k\beta \leq \overline{\mathbf{y}}^t\mathbf{b} + \overline{u}\beta - k\beta = \overline{\mathbf{y}}^t\mathbf{b} + (\overline{z}_+ - \overline{z}_-)\beta\,,$$

where the inequality follows as $(\overline{\mathbf{y}}, \overline{u})$ is feasible. $(\overline{\mathbf{y}}, \overline{z}_+, \overline{z}_-)$ is optimal, so we must have equality in this relation. Hence, $(\tilde{\mathbf{y}}, \tilde{z}_+, \tilde{z}_-)$ is an integral optimal solution of the dual program (12.2). □

Proposition 12.8. *Let $A \in \mathbb{Z}^{m\times n}$ and $\mathbf{b} \in \mathbb{Q}^m$. The system*

$$A\mathbf{x} \;\le\; \mathbf{b},\; \mathbf{x} \;\ge\; \mathbf{0} \qquad\qquad (\,A\mathbf{x} \;=\; \mathbf{b},\; \mathbf{x} \;\ge\; \mathbf{0}\,)$$

is TDI if and only if

$$\min(\mathbf{y}^t\mathbf{b} \mid \mathbf{y}^t A \ge \mathbf{c}^t,\, \mathbf{y} \ge \mathbf{0}) \qquad\qquad (\,\min(\mathbf{y}^t\mathbf{b} \mid \mathbf{y}^t A \ge \mathbf{c}^t)\,)$$

has an integral optimal solution for any integral $\mathbf{c}^t$ for which it is finite.

Proof. The system $A\mathbf{x} \le \mathbf{b}$, $\mathbf{x} \ge \mathbf{0}$ can be written as

$$\begin{pmatrix} A \\ -\mathrm{Id_m} \end{pmatrix} \mathbf{x} \le \begin{pmatrix} \mathbf{b} \\ \mathbf{0} \end{pmatrix}$$

So it is TDI if and only if for any integral $\mathbf{c}^t$

$$\min(\mathbf{y}^t\mathbf{b} \mid \mathbf{y}^t A - \mathbf{z}^t = \mathbf{c}^t,\, \mathbf{y}, \mathbf{z} \ge \mathbf{0}) \tag{12.5}$$

has an integral optimal solution $(\mathbf{y}, \mathbf{z})$ if the program is bounded. But this has an integral optimal solution if and only if

$$\min(\mathbf{y}^t\mathbf{b} \mid \mathbf{y}^t A \ge \mathbf{c}^t,\, \mathbf{y} \ge \mathbf{0})$$

has an integral optimal solution. The second claim is similar. □

Now we give a geometric interpretation of total dual integrality by relating it to Hilbert bases in the cones of the normal fan of a polyhedron.

Theorem 12.9. *Let $A \in \mathbb{Q}^{m\times n}$, $\mathbf{b} \in \mathbb{Q}^m$ and $P := \{\mathbf{x} \mid A\mathbf{x} \le \mathbf{b}\}$. The system $A\mathbf{x} \le \mathbf{b}$ is TDI if and only if for each face F of P the rows of $A_{\mathrm{eq}(F)*}$ are a Hilbert basis of $\mathrm{cone}(A_{\mathrm{eq}(F)*})$.*

Proof. Suppose that $A\mathbf{x} \le \mathbf{b}$ is TDI. Let F be a face of P, $I := \mathrm{eq}(F)$ and $J := [m] - I$. Let $\mathbf{c}^t \in \mathrm{cone}(A_{I*}) \cap \mathbb{Z}^n$. Then

$$\max(\mathbf{c}^t\mathbf{x} \mid A\mathbf{x} \le \mathbf{b}) \;=\; \min(\mathbf{y}^t\mathbf{b} \mid \mathbf{y}^t A = \mathbf{c}^t,\, \mathbf{y} \ge \mathbf{0}) \tag{12.6}$$

is optimally solved by any $\overline{\mathbf{x}} \in F$. By assumption, the minimum has an integral optimal solution $\overline{\mathbf{y}}$. The complementary slackness theorem, Theorem 3.9, implies that $\overline{\mathbf{y}}_J = 0$. Hence $\mathbf{c}^t = \overline{\mathbf{y}}_I A_{I*}$ is an integral conic combination of the generators of $\mathrm{cone}(A_{I*})$.

Suppose conversely that the value of the programs in (12.6) is finite for some $\mathbf{c}^t \in \mathbb{Z}^m$. Let F be the minimal face of P determined by $\mathbf{c}^t$ and $I := \mathrm{eq}(F)$. By the same argument as before there is some $\overline{\mathbf{y}}$ such that

$$\mathbf{c}^t = \overline{\mathbf{y}}_I A_{I*} \in \mathrm{cone}(A_{I*}) \qquad \text{and} \qquad \overline{\mathbf{y}}_J = \mathbf{0}\,.$$

However, $\overline{\mathbf{y}}_I \ge \mathbf{0}$ may now be fractional. Using the assumption there is a non-negative integral vector $\mathbf{z}$ such that $\mathbf{c}^t = \mathbf{z}^t A_{I*}$. Extending $\mathbf{z}$ with zeros to $\overline{\mathbf{z}}$ such that $\mathbf{z} = \overline{\mathbf{z}}_I$ we obtain for all $\mathbf{x} \in F$

$$\mathbf{c}^t = \lambda^t A \qquad \text{and} \qquad \overline{\mathbf{z}}^t\mathbf{b} = \overline{\mathbf{z}}^t A\mathbf{x} = \mathbf{c}^t\mathbf{x}\,.$$

So $\overline{\mathbf{z}}$ is an optimal solution of (12.6). As $\mathbf{c}^t$ was arbitrary, this implies that $A\mathbf{x} \le \mathbf{b}$ is TDI. □

As we may always choose the optimal solutions in the previous proof to lie in a minimal face of P we have even proved the stronger statement that it suffices to consider only minimal faces F in the theorem. Using this and choosing $\mathbf{b} = \mathbf{0}$ in the previous theorem we obtain

Corollary 12.10. *The rows of a rational matrix form a Hilbert basis if and only if $A\mathbf{x} \le \mathbf{0}$ is TDI.* □

minimally TDI

Definition 12.11. *Let $A\mathbf{x} \le \mathbf{b}$ be TDI. Then $A\mathbf{x} \le \mathbf{b}$ is* **minimally TDI** *if any proper subsystem of $A\mathbf{x} \le \mathbf{b}$ that defines the same polyhedron is not TDI.*

A TDI system $A\mathbf{x} \le \mathbf{b}$ is minimally TDI if and only if each constraint defines a supporting hyperplane of $P := \{\mathbf{x} \mid A\mathbf{x} \le \mathbf{b}\}$ and cannot be written as a non-negative integral linear combination of other inequalities in $A\mathbf{x} \le \mathbf{b}$.

Theorem 12.12. *Let P be a rational polyhedron.*

(1) Then there is a TDI system $A\mathbf{x} \le \mathbf{b}$ for an integral matrix A that defines P.
(2) $\mathbf{b}$ can be chosen to be integral if and only if P is integral.
(3) If P is full-dimensional, then there is a unique minimal TDI system defining P.

See Figure 12.2 for an example of a TDI-system defining a triangle.

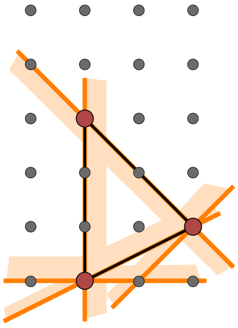


Figure 12.2

Proof. (1) Let F be a minimal face of P with normal cone C_F. By Theorem 8.24, the cone C_F has an integral Hilbert basis $\mathscr{H}_F := \{\mathbf{a}_1, \ldots, \mathbf{a}_r\}$. For $1 \le i \le r$ define

$$\beta_i := \max(\mathbf{a}_i^t \mathbf{x} \mid \mathbf{x} \in P)\,. \tag{12.7}$$

Then $\quad F \subseteq \{\mathbf{x} \mid \mathbf{a}_i^t \mathbf{x} = \beta_i\} \quad$ and $\quad P \subseteq \{\mathbf{x} \mid \mathbf{a}_i^t \mathbf{x} \le \beta_i\}$

for all i. Let $\mathscr{A}_F$ be the collection of inequalities $\mathbf{a}_i \mathbf{x} \le \beta_i$.

Let $A\mathbf{x} \le \mathbf{b}$ be the inequality system obtained from all $\mathscr{A}_F$, where F ranges over all minimal faces of P. Then $A\mathbf{x} \le \mathbf{b}$ determines P and is TDI by Theorem 12.9.

(2) If P is integral, then the β_i's in (12.7) are integral. So also $\mathbf{b}$ is integral. If conversely $\mathbf{b}$ is integral, then P is integral by Corollary 12.4.

(3) If P is full-dimensional, then each normal cone C_F is pointed. Hence, again by Theorem 8.24 it has a unique minimal integral Hilbert basis. Let $A\mathbf{x} \le \mathbf{b}$ be the inequality system obtained in the same way as above, but only using these minimal sets of generators. By Theorem 12.9, the system $A\mathbf{x} \le \mathbf{b}$ must be a subsystem of any TDI-system defining P. □

Corollary 12.13. *A rational polyhedron P is integral if and only if there is a TDI-system $A\mathbf{x} \le \mathbf{b}$ with integral $\mathbf{b}$ that defines it.* □

We shortly discuss the complexity of the following problem:

> Given a rational system of linear equations $Ax = b$, decide whether a rational system of linear equations has an integral solution. (12.8)

HERMITE normal forms can be used to show that this problem is well characterized. It is even in $\mathscr{P}$, as a transforming a matrix into HERMITE normal form can by done in polynomial time (see Schrijver's book [Sch86, Ch. 5]). Using the methods introduced in Chapter 5 one shows the following theorem.

Theorem 12.14. *Let $A \in \mathbb{Q}^{m\times n}$ with $\operatorname{rank}(A) = m$.*

(1) The HERMITE normal form of A has size polynomially bounded by the size of A.
(2) There is a unimodular matrix U such that AU is in HERMITE normal form and the size of U is polynomially bounded in the size of A.

Proof. We can assume that A is integral, as multiplication of a by a constant also multiplies the Hermit normal form by the same constant.

(1) Let $(B\ 0)$ be the HERMITE normal form of A for some lower triangular non-singular square matrix $B = (b_{ij})_{ij}$. The main idea of the proof is now the following observation. Let $j \in [m]$. The g.c.d. of all maximal minors of the first j rows of A is invariant under elementary column operations. Hence, this g.c.d. is the same in A and $(B\ 0)$. But in the latter it is just the product of the first j diagonal entries of B, as all other maximal minors are 0. Now B has its maximal entries on the diagonal, so the size of B is polynomially bounded by that of A.

(2) We may assume that $A = (A_1\ A_2)$ for some non-singular matrix A_1. Now consider the following matrix and its HERMITE normal form:

$$\begin{pmatrix} A_1 & A_2 \\ 0 & \mathrm{Id_m} \end{pmatrix} \qquad \text{with normal form} \qquad \begin{pmatrix} B & 0 \\ B_1 & B_2 \end{pmatrix}.$$

The sizes of B, B_1, B_2 are polynomially bounded in the size of A. Then also

$$U := \begin{pmatrix} A_1 & A_2 \\ 0 & \mathrm{Id_m} \end{pmatrix}^{-1} \cdot \begin{pmatrix} B & 0 \\ B_1 & B_2 \end{pmatrix}.$$

has bounded size and $AU = (B\ 0)$. ☐

Corollary 12.15. *If a rational system $A\mathbf{x} = \mathbf{b}$ of linear equations has an integral solution, then it has one of size polynomially bounded in the sizes of A and $\mathbf{b}$.*

Proof. We may assume that $\operatorname{rank}(A) = m$. Then there is a unimodular transformation U of polynomially bounded size (in that of A) such that $AU = (B\ 0)$ is in HERMITE normal form. Then

$$\bar{\mathbf{x}} := U \begin{pmatrix} B^{-1}b \\ 0 \end{pmatrix}$$

is an integral solution of $A\mathbf{x} = \mathbf{b}$ of bounded size. ☐

Using the integral alternative theorem, Theorem 8.17, this implies that the decision problem (12.8) is in $\mathcal{NP} \cap \text{co-}\mathcal{NP}$: We can assume that A has full rank and is given in HERMITE normal form $A = (B\ 0)$. Then a positive certificate is given by the previous corollary. By the integral alternative theorem, for a negative certificate we have to provide a rational vector $\mathbf{y}$ such that $\mathbf{y}^t A$ is integral, but $\mathbf{y}^t\mathbf{b}$ is not. By the proof of the corollary, some row of B^{-1} has this property if $A\mathbf{x} = \mathbf{b}$ does not have an integral solution.

Chvátal-Gomory Cuts

13

cutting plane

In this chapter we discuss the **method of cutting planes** to compute the integer hull of a polyhedron. The central idea of this approach is to successively **cut off** fractional minimal faces of the polyhedron by hyperplanes that separate the face from all integral points in the polyhedron. This general approach was developed around 1950 by GOMORY to use linear programming methods in integer linear programming. This has been quite successful, and there have been developed several similar approaches since. It has turned out that his idea is also of great theoretical interest. We will discuss in this final chapter one of the first approaches by GOMORY and CHVÁTAL.

We need some more notation. For any vector $\mathbf{b} \in \mathbb{R}^m$ let $\lfloor \mathbf{b} \rfloor$ and $\lceil \mathbf{b} \rceil$ denote the vector obtained by applying $\lfloor\,.\,\rfloor$ or $\lceil\,.\,\rceil$ to each component of $\mathbf{b}$. Let P_I denote the integer hull of a polyhedron, i.e. $P_I := \operatorname{conv}(P \cap \mathbb{Z}^n)$. Let $\mathbf{c} \in \mathbb{Z}^n$ and $\delta \in \mathbb{R}$ be given, and consider the rational half-space

$$H := \{\mathbf{x} \mid \mathbf{c}^t\mathbf{x} \le \delta\}\,.$$

Its integral closure H_I clearly satisfies

$$H_I \subseteq \overline{H} := \{\mathbf{x} \mid \mathbf{c}^t\mathbf{x} \le \lfloor \delta \rfloor\}\,.$$

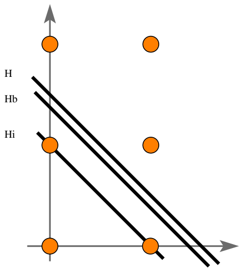


Figure 13.1

We have $H_I = \overline{H}$ if $\mathbf{c}^t$ is primitive, i.e. if $\gcd(c_1, \ldots, c_n) = 1$. Geometrically, H_I arises from H by shifting H along its normal vector $\mathbf{c}^t$ until it contains an integer point.

Example 13.1. Consider $\mathbf{c}^t := (3, 3)$ and $\delta := 9/2$. Then (see Figure13.1):

$$\overline{H} = \{\mathbf{x} \mid \mathbf{c}^t\mathbf{x} \le 4\} \qquad \text{and} \qquad H_I = \{\mathbf{x} \mid \mathbf{c}^t\mathbf{x} \le 3\}\,. \qquad \Diamond$$

elementary closure

Definition 13.2. *Let P be a polyhedron. The* **elementary closure** *of P is the set*

$$P^{(1)} := \bigcap_{\substack{H \text{ rational half-space}\\ \text{containing } P}} H_I\,.$$

Inductively we define $P^{(k)} := \left(P^{(k-1)}\right)^{(1)}$.

cutting planes

As $P_I \subseteq H_I$ for any half-space H whose boundary plane is valid, we know that $P_I \subseteq P^{(1)}$. It clearly suffices to consider supporting half-spaces in the intersection. The affine hyperplanes bounding the the half-spaces H_I are **cutting planes** of P. Repeating this procedure, we obtain a chain

$$P^{(1)} \supseteq P^{(2)} \supseteq P^{(3)} \supseteq \ldots \subseteq P^{(k)} \supseteq \ldots \supseteq P_I\,. \tag{13.1}$$

A priori, this does not seem to be a useful definition, as the intersection ranges over the infinite set of all supporting hyperplanes. However, the next theorem shows that in the case of an TDI system the intersection is already defined by a finite set of half-spaces. The approach of rounding down inequalities is one of the central ideas behind cutting plane methods to solve integer linear programs. The general method is the following: One successively solves linear programs over $P^{(k)}$ for increasing k until the obtained optimal solution is integral. The elementary closures need not be determined completely as one is only interested in the optimum with respect to one linear functional.

Theorem 13.3. *Let $P := \{\mathbf{x} \mid A\mathbf{x} \le \mathbf{b}\}$ be a polyhedron defined by a TDI system $A\mathbf{x} \le \mathbf{b}$ for an integral matrix $A \in \mathbb{Z}^{m\times n}$. Then $P^{(1)} = \{x \mid A\mathbf{x} \le \lfloor \mathbf{b} \rfloor\}$.*

Proof. Let $Q := \{\mathbf{x} \mid A\mathbf{x} \le \lfloor \mathbf{b} \rfloor\}$. If $P = \varnothing$, then $P^{(1)} = \varnothing$. Any $\mathbf{x}$ that satisfies $A\mathbf{x} \le \lfloor \mathbf{b} \rfloor$ also satisfies $A\mathbf{x} \le \mathbf{b}$, so $Q = \varnothing$, which proves the theorem in this case.
So we may assume that $P \ne \varnothing$. Each inequality in $A\mathbf{x} \le \mathbf{b}$ defines a rational half-space containing P, so

$$P^{(1)} \subseteq \{\mathbf{x} \mid A\mathbf{x} \le \lfloor \mathbf{b} \rfloor\} = Q\,.$$

We need to prove the reverse inclusion, i.e. we have to show that

$$Q \subseteq H_I$$

for any rational half-space H containing P. Let $H := \{\mathbf{x} \mid \mathbf{c}^t\mathbf{x} \le \delta\}$ be such a half-space. We may assume that $\mathbf{c}^t$ is a primitive integer vector, so that

$$H_I = \{\mathbf{x} \mid \mathbf{c}^t\mathbf{x} \le \lfloor \delta \rfloor\}\,.$$

By duality, and as P is contained in H, we have

$$\delta \ge \max(\mathbf{c}^t\mathbf{x} \mid A\mathbf{x} \le \mathbf{b}) \;=\; \min(\mathbf{y}^t\mathbf{b} \mid \mathbf{y}^t A = \mathbf{c}^t,\, \mathbf{y} \ge \mathbf{0}\}\,. \tag{13.2}$$

$A\mathbf{x} \le \mathbf{b}$ is TDI and $\mathbf{c}^t$ is integral, so the minimum is attained by an integer vector $\overline{\mathbf{y}}$. Suppose that $\mathbf{x} \in Q$. Then

$$\mathbf{c}^t\mathbf{x} \;=\; \overline{\mathbf{y}}^t A\mathbf{x} \;\le\; \overline{\mathbf{y}}^t\lfloor \mathbf{b} \rfloor \;\le\; \lfloor \overline{\mathbf{y}}^t\mathbf{b} \rfloor \;\le\; \lfloor \delta \rfloor,$$

which implies that $\mathbf{x} \in H_I$. So $Q \subseteq H_I$. □

Lemma 13.4. *Let P be a rational polyhedron and F a face of P. Then $F^{(t)} = P^{(t)} \cap F$. If $F^{(t)} \ne \varnothing$, then it is a non-empty proper face of $P^{(t)}$.*

Proof. It suffices to prove this in the case $t = 1$, with the general case then following by induction.
Let $P = \{\mathbf{x} \mid A\mathbf{x} \le \mathbf{b}\}$ for a TDI system $A\mathbf{x} \le \mathbf{b}$ with integral $A \in \mathbb{Z}^{m\times n}$. Let the face F be defined by

$$F := P \cap \{\mathbf{x} \mid \mathbf{c}^t\mathbf{x} = \delta\} = \{\mathbf{x} \mid A\mathbf{x} \le \mathbf{b},\, \mathbf{c}^t\mathbf{x} \le \delta,\, -\mathbf{c}^t\mathbf{x} \le -\delta\}\,. \tag{13.3}$$

for an integral vector $\mathbf{c}^t \in \mathbb{Z}^n$ and $\delta \in \mathbb{Z}$. The inequality $\mathbf{c}^t x \le \delta$ is valid for P, so by Proposition 12.2 the system $A\mathbf{x} \le \mathbf{b}$, $\mathbf{c}^t\mathbf{x} \le \delta$ is TDI. Theorem 12.7 implies that also $A\mathbf{x} \le \mathbf{b}$, $\mathbf{c}^t\mathbf{x} = \delta$ is TDI. δ is integral, so

$$P^{(1)} \cap F = \{\mathbf{x} \mid A\mathbf{x} \le \lfloor \mathbf{b} \rfloor,\, \mathbf{c}^t\mathbf{x} = \delta\} = \{\mathbf{x} \mid A\mathbf{x} \le \lfloor \mathbf{b} \rfloor,\, \mathbf{c}^t\mathbf{x} \le \lfloor \delta \rfloor, -\mathbf{c}^t\mathbf{x} \le \lfloor -\delta \rfloor\} = F^{(1)}\,,$$

where the last equation follows as the given system is the elementary closure of the right hand representation in (13.3). If $F^{(1)}$ is not empty, then we see from this representation that it is a face of $P^{(1)}$. □

Theorem 13.5. *Let $P \subseteq \mathbb{R}^n$ be a rational polyhedron. Then there is $t \in \mathbb{N}$ such that $P_I = P^{(t)}$.*

Proof. We prove this by induction on the dimension d of $P \subseteq \mathbb{R}^n$. If $d = -1$, then $P_I = P = P^{(0)} = \varnothing$. If $d = 0$, then P is a point and either $\varnothing = P_I = P^{(1)}$ or $P_I = P = P^{(0)}$. So in both cases, $t = 0$ or $t = 1$ suffices. So we can assume that $d > 0$ and the theorem is true for all rational polyhedra in dimensions strictly less than d. Let $H := \{\mathbf{x} \mid M\mathbf{x} = \mathbf{m}\}$ be the affine hull of P.

If $H \cap \mathbb{Z}^n = \varnothing$, then $P_I = \varnothing$. Using Theorem 8.17 (integral alternative theorem) we find some rational vector $\mathbf{y} \in \mathbb{Q}^m$ such that $\mathbf{c}^t := \mathbf{y}^t M$ is integral, but $\delta := \mathbf{y}^t \mathbf{m}$ is not. Let $\mathbf{x} \in H$. Then

$$\mathbf{c}^t \mathbf{x} = \mathbf{y}^t M \mathbf{x} = \mathbf{y}^t \mathbf{m} = \delta,$$

so $\{\mathbf{x} \mid \mathbf{c}^t \mathbf{x} = \delta\}$ is a supporting hyperplane of P and we obtain

$$P^{(1)} \subseteq \{\mathbf{x} \mid \mathbf{c}^t \mathbf{x} \le \lfloor \delta \rfloor\} \cap \{x \mid -\mathbf{c}^t \mathbf{x} \le \lfloor -\delta \rfloor\} = \varnothing,$$

as δ is not integer. Hence, $t = 1$ suffices in the case that $H \cap \mathbb{Z}^n = \varnothing$.

So we may assume that there is an integer point $\tilde{\mathbf{x}} \in H$. Translating everything by an integer vector does not affect the theorem, so we may assume that $H = \{\mathbf{x} \mid M\mathbf{x} = \mathbf{0}\}$. Further, we may assume that M has full row rank $n - d$. By Theorem 8.14 we can find a unimodular transformation U such that $M = [M_0 \;\; 0]$ is in Hermite normal form for a non-singular matrix M_0. Transforming P with U we obtain a polyhedron P' with affine hull

$$\{\mathbf{x} \mid [M_0 \;\; 0]\mathbf{x} = \mathbf{0}\} = \{\mathbf{0}\} \times \mathbb{R}^m,$$

and P' is full-dimensional in $\mathbb{R}^d$. U bijectively maps $\mathbb{Z}^n$ onto $\mathbb{Z}^n$ (by Proposition 8.9), and for any rational hyperplane H we have $(UH)_I = U(H_I)$. Hence, we may assume that P is full-dimensional.

P_I is a polyhedron, so there is a rational matrix A and rational vectors $\mathbf{b}$ and $\mathbf{b}'$ such that

$$P_I = \{\mathbf{x} \mid A\mathbf{x} \le \mathbf{b}\} \qquad \text{and} \qquad P = \{\mathbf{x} \mid A\mathbf{x} \le \mathbf{b}'\}$$

(take the union of sets of inequalities defining P and P', and choose the right hand sides of inequalities in $\mathbf{b}$ and $\mathbf{b}'$ not needed large enough so that they do not affect the polyhedron). By (13.1) we already know that

$$P_I \subseteq P^{(t)}$$

for all $t \in \mathbb{N}$. So we have to show that there is some $t \in \mathbb{N}$ such that the reverse inclusion is true. Let $\mathbf{a}^t x \le \beta$ be some inequality contained in $A\mathbf{x} \le \mathbf{b}$ and $\mathbf{a}^t x \le \beta'$ the corresponding inequality in $A\mathbf{x} \le \mathbf{b}'$. Let $H := \{\mathbf{x} \mid \mathbf{a}^t \mathbf{x} \le \beta\}$. Then it suffices to prove that $P^{(s)} \subseteq H$ for some finite $s \in \mathbb{N}$, as there are only a finite number of such half-spaces.

Assume by contradiction that there is no $s \in \mathbb{N}$ such that $P^{(s)} \subseteq H$. By definition,

$$P^{(1)} \subseteq \{\mathbf{x} \mid \mathbf{a}^t \mathbf{x} \le \lfloor \beta' \rfloor\}.$$

Hence, there is $\beta'' \in \mathbb{Z}$ and $r \in \mathbb{N}$ such that $\beta < \beta'' \le \lfloor \beta' \rfloor$ and for all $s \ge r$

$$P^{(s)} \subseteq \{\mathbf{x} \mid \mathbf{a}^t \mathbf{x} \le \beta''\} \qquad \text{but} \qquad P^{(s)} \not\subseteq \{\mathbf{x} \mid \mathbf{a}^t \mathbf{x} \le \beta'' - 1\}. \tag{13.4}$$

Let
$$F := P^{(r)} \cap \{\mathbf{x} \mid \mathbf{a}^t \mathbf{x} = \beta''\}.$$

Then F is a (possibly empty) proper face of $P^{(r)}$. Further, F does not contain any integral vectors, as $P_I \subseteq H = \{\mathbf{x} \mid \mathbf{a}^t \mathbf{x} \le \beta\}$ and $\beta < \beta''$. Hence, there is $u \in \mathbb{N}$ such that $F^{(u)} = \varnothing$ and we obtain

$$\varnothing = F^{(u)} = P^{(r+u)} \cap F = P^{(r+u)} \cap \{\mathbf{x} \mid \mathbf{a}^t \mathbf{x} = \beta''\}.$$

So
$$P^{(r+u)} \subseteq \{\mathbf{x} \mid \mathbf{a}^t \mathbf{x} < \beta''\},$$
and hence
$$P^{(r+u+1)} \subseteq \{\mathbf{x} \mid \mathbf{a}^t \mathbf{x} \le \beta'' - 1\},$$

as β'' is integral. However, this contradicts (13.4). □

Corollary 13.6. *Let P be a rational polyhedron. Then $P = P_I$ if and only if each rational supporting hyperplane contains an integral point.*

Proof. If $P = P_I$, then each supporting hyperplane contain a minimal face, which contains an integral point. If conversely each rational supporting hyperplane contains an integral point, then $P^{(t)} = P$ for all $t \geq 0$, and the claim follows from the previous Theorem 13.5. ☐

Corollary 13.7. *Let P be a (not necessarily rational) polytope. Then there is $t \in \mathbb{N}$ such that $P^{(t)} = P_I$.*

Proof. The only thing we need to show is that we can replace P by a rational polyhedron Q such that $Q_I = P_I$. Then the claim would follow from Theorem 13.5.
As P is a polytope, there is some $M \in \mathbb{N}$ such that

$$P \subseteq B := \{\mathbf{x} \mid \|\mathbf{x}\|_\infty \leq M\}.$$

For any $\mathbf{z} \in B \setminus P$ there is a rational hyperplane $H_\mathbf{z}$ such that P and $\mathbf{z}$ are on different sides of $H_\mathbf{z}$. B is bounded, so $|B \cap \mathbb{Z}^n|$ is finite. We can define Q to be the intersection of all half-spaces defined by the $H_\mathbf{z}$. ☐

Theorem 13.5 can be used to give an algorithm that computes the integer hull of a polyhedron. Observe however, that this is not a polynomial time algorithm, just because the size of a TDI system defining a polyhedron $P = \{\mathbf{x} \mid A\mathbf{x} \leq \mathbf{b}\}$ may be exponential in the size of A.

INTEGER HULL

INPUT: A rational polyhedron $P = \{\mathbf{x} \mid A\mathbf{x} \leq \mathbf{b}\}$.

OUTPUT: A description of P_I.

ALGORITHM: While P not integral do

```
(1) Replace $A\mathbf{x} \leq \mathbf{b}$ by a TDI system that describes $P$.
(2) Let $P' := \{\mathbf{x} \mid A\mathbf{x} \leq \lfloor \mathbf{b} \rfloor\}$.
(3) Set $P := P'$.
```

It is now natural to ask whether we can give a bound on the number of steps that this algorithm needs to compute the integer hull of a polyhedron. The next example shows that there cannot be a polynomial bound in the dimension only.

Example 13.8. For any $k \in \mathbb{N}$ we define the matrix A_k, a vector $\mathbf{b}_k$, and the polygon P_k by

$$A_k := \begin{bmatrix} -1 & 0 \\ 1 & 2k \\ 1 & -2k \end{bmatrix} \qquad \mathbf{b}_k := \begin{bmatrix} 0 \\ 2k \\ 0 \end{bmatrix} \qquad P_k := \{\mathbf{x} \in \mathbb{R}^2 \mid A_k\mathbf{x} \leq \mathbf{b}_k\}\,.$$

P_k is the polygon with vertices

$$\mathbf{v}_1 := \begin{bmatrix} 0 \\ 0 \end{bmatrix} \qquad \mathbf{v}_2 := \begin{bmatrix} 0 \\ 1 \end{bmatrix} \qquad \mathbf{v}_3 := \begin{bmatrix} k \\ \frac{1}{2} \end{bmatrix},$$

while $P_I = \operatorname{conv}(\mathbf{v}_1, \mathbf{v}_2)$ is a line segment. To compute $P^{(1)}$ we need a TDI system defining P. For this we have to extend $A_k\mathbf{x} \leq \mathbf{b}_k$ by the inequalities

$$\begin{aligned} x_2 &\leq 1 \\ -x_2 &\leq 0 \\ x_1 + jx_2 &\leq k + \frac{j}{2} && \text{for } 0 \leq j \leq 2k-1 \\ x_1 - jx_2 &\leq k - \frac{j}{2} && \text{for } 1 \leq j \leq 2k-1\,. \end{aligned}$$

From this we can compute that $\begin{bmatrix} k-1 \\ \frac{1}{2} \end{bmatrix} \in P^{(1)}$, and by induction we get that

$$\begin{bmatrix} k-t \\ \frac{1}{2} \end{bmatrix} \in P^{(t)}$$

for $t < k$. So $P_I \neq P^{(t)}$ for $t < k$ and we need at least k steps in the algorithm. ◊

Example 13.9. Let $G = (V, E)$ be an undirected graph, and

$$Q_m(G) := \{\mathbf{x} \in \mathbb{R}^m \mid \mathbf{x}(\delta(v)) \leq 1,\ x_e \geq 0 \text{ for } v \in V, \text{ and } e \in E\}.$$

In Chapter 9 we have seen that if G is bipartite, then Q_m is the matching polytope of G. It can be shown that

$$\begin{aligned} x_e &\geq 0 && \text{for all } e \in E && (13.5a)\\ \mathbf{x}(\delta(v)) &\leq 1 && \text{for all } v \in V && (13.5b)\\ \mathbf{x}(E(U)) &\leq \frac{1}{2}|U| && \text{for all } U \subseteq V \text{ such that } |U| \text{ is odd.} && (13.5c) \end{aligned}$$

is a TDI system that defines $Q_m(G)$. So we can compute the elementary closure of Q_m:

$$Q_m^{(1)}(G) := \left\{ \mathbf{x} \in \mathbb{R}^m \;\middle|\; \begin{array}{c} \mathbf{x}(\delta(v)) \leq 1,\ x_e \geq 0,\ \mathbf{x}(\delta(U)) \leq \frac{1}{2}\lfloor |U| \rfloor \\ \text{for } v \in V,\ e \in E, \text{ and odd } U \subseteq V \end{array} \right\}.$$

We have seen in Theorem 9.8 that this defines the matching polytope $P_m(G)$ of G. So the general matching polytope is the first elementary closure of the bipartite matching inequalities.

For a rational polyhedron we can bound the number of facets and vertices by the dimension of A. This is not anymore true for the integer hull, as is shown by the following considerations.

A polyhedron $P := \{\mathbf{x} \mid A\mathbf{x} \leq \mathbf{b}\}$ for $A \in \mathbb{Q}^{m \times n}$ has at most m facets and $\binom{m}{n}$ vertices. Hence, the number of facets and vertices of a rational polyhedron is even independent of the size of A and $\mathbf{b}$. Now let in the above example $G = K_n$ be the complete graph on n vertices. We have seen that the matching polytope $P_m(G)$ of G is the elementary closure of $Q_m(G)$, and the corresponding matching polytope has $f := \binom{n}{2} + 2^{n-1}$ different facets. On the other hand, the size of (A, b) is $\langle (A, b) \rangle = n\binom{n}{2} + 3\binom{n}{2} = (n+3)\binom{n}{2} = \mathcal{O}(n^3)$. But 2^{n-1} cannot be bounded by a polynomial in n^3. ◊

We repeat the result of the previous example in the following proposition.

Proposition 13.10. *There is no polynomial φ such that for each rational polyhedron $P := \{\mathbf{x} \mid A\mathbf{x} \leq \mathbf{b}\}$ the integer hull P_I of P has at most $\varphi(\langle (A, b) \rangle)$ facets.* □

So we cannot say much about the number of facets of the integer hull of a rational polyhedron. However, it can be shown that at least the facet complexity of the integer hull can be polynomially bounded in the facet complexity of P.

Theorem 13.11. *Let $P \subset \mathbb{R}^m$ be a rational polyhedron with facet complexity φ. Then there is a constant $C \in \mathbb{R}_+$ such that the facet complexity φ_I of P_I is bounded by $\varphi_I \leq C\varphi^6$.* □

This theorem has two immediate consequences.

Corollary 13.12. *Let P be a rational polyhedron of facet complexity φ. If $P_I \neq \emptyset$, then P contains an integral point of size bounded by $Cn^3\varphi$ for some $C \in \mathbb{R}_+$.* □

Corollary 13.13. *The problem of deciding whether a rational inequality system $A\mathbf{x} \leq \mathbf{b}$ has an integral solution is in $\mathcal{NP}$.* □

The problem to decide whether $A\mathbf{x} \le \mathbf{b}$ has an integral solution is in general $\mathcal{NP}$-hard, but proving this is much more involved.

The above examples show that we cannot expect to find integer hulls of polyhedra within polynomial time. The situation is slightly better if we only want to know whether the integer hull is empty. Namely, if $P_I = \varnothing$, then we can check this in a finite number of steps that only depends on the dimension of the polytope. More precisely, we will prove that there is some $t \in \mathbb{N}$ only depending on the dimension, so that $P^{(t)} = \varnothing$ if $P_I = \varnothing$. We need a theorem from *geometry of numbers* for this. We will not prove this here, but see Barvinok's book [Bar02].

$\mathbb{Z}^n$-width

***Definition 13.14**. Let $K \subset \mathbb{R}^d$ be a closed and convex set. The* $\mathbb{Z}^n$**-width** *of K with respect to $\mathbf{c}^t \in \mathbb{Z}^n$ is*

$$\operatorname{width}(K, \mathbf{c}^t) := \max(\mathbf{c}^t\mathbf{x} \mid \mathbf{x} \in K) - \min(\mathbf{c}^t\mathbf{x} \mid \mathbf{x} \in K)\,.$$

***Theorem 13.15* (Flatness Theorem).** *There is a constant $w_n \in \mathbb{R}$ such that for any convex $K \subset \mathbb{R}^n$ with $K \cap \mathbb{Z}^n = \emptyset$ there is $\mathbf{c}^t \in \mathbb{Z}^n \setminus \{0\}$ such that*

$$\operatorname{width}(K, \mathbf{c}^t) \le w_n\,. \qquad \square$$

Remark 13.16. It has been shown that $w_n := n^{\frac{3}{2}}$ suffices in the flatness theorem. However, if one wants to explicitely compute a vector $\mathbf{c}$ realizing the bound in polynomial time, then one can only guarantee a bound $w_n = \mathcal{O}(2^n)$. Still, this only depends on the dimension of the polyhedron. ◊

***Theorem 13.17**. For each $d \in \mathbb{N}$ there is $t(d) \in \mathbb{N}$ such that if $P \subseteq \mathbb{R}^n$ is a rational d-dimensional polyhedron with $P_I = \varnothing$ then already $P^{(t(d))} = \varnothing$.*

Proof. We apply induction over the dimension d of P. We may assume that $t(d) \ge t(d-1)$ for all $d \ge 0$.
If $d = -1$ then $P = \varnothing$, so $t(-1) := 0$ suffices. If $d = 0$, then $P = \{\mathbf{x}\}$ for some $\mathbf{x} \in \mathbb{R}^n$. If $P_I = \varnothing$, then $\mathbf{x} \notin \mathbb{Z}^n$, so $P^{(1)} = \varnothing$ and $t(0) := 1$ suffices.
By the same argument as in the proof of Theorem 13.5, where we proved that $P_I = P^{(t)}$ for some t, we may assume that P is full-dimensional, so $d = n$. By Theorem 13.15 there is a primitive vector $\mathbf{c}^t \in \mathbb{Z}^d$ and a constant w_n that only depends on d such that

$$\max(\mathbf{c}^t\mathbf{x} \mid \mathbf{x} \in P) - \min(\mathbf{c}^t\mathbf{x} \mid \mathbf{x} \in P) \le w_d\,.$$

Let $\delta := \lfloor \max(\mathbf{c}^t\mathbf{x} \mid \mathbf{x} \in P) \rfloor$. We claim that

$$P^{(k+1+kt(d-1))} \subseteq \{\mathbf{x} \mid \mathbf{c}^t\mathbf{x} \le \delta - k\} \qquad (13.6)$$

for $0 \le k \le w_d + 1$. We prove this by induction over k. For $k = 0$ this specializes to

$$P^{(1)} \subseteq \{\mathbf{x} \mid \mathbf{c}^t\mathbf{x} \le \delta\}\,.$$

This follows immediately from the definition of $P^{(1)}$. Suppose that (13.6) is true for some $k \ge 0$. Let

$$F := P^{(k+1+kt(d-1))} \cap \{\mathbf{x} \mid \mathbf{c}^t\mathbf{x} = \delta - k\}\,, \qquad (13.7)$$

be a possibly empty proper face of $P^{(k+1+kt(d-1))}$. As the dimension of F is less than d, we know by induction that $F^{(t(d-1))} = \varnothing$. This implies

$$\left(P^{(k+1+kt(d-1))}\right)^{(t(d-1))} \cap \{x \mid \mathbf{c}^t\mathbf{x} = \delta - k\} = F^{(t(d-1))} = \varnothing\,.$$

This means that $P^{(k+1+(k+1)t(d-1))} \subseteq \{\mathbf{x} \mid \mathbf{c}^t\mathbf{x} < \delta - k\}$

and so $P^{(k+2+(k+1)t(d-1))} = \left(P^{(k+1+(k+1)t(d-1))}\right)^{(1)} \subseteq \{\mathbf{x} \mid \mathbf{c}^t\mathbf{x} \leq \delta - k - 1\}$.

So (13.6) holds for $k+1$. Take $k := w_d + 1$ in (13.6). Then

$$P^{(w_d+2+(w_d+1)t(d-1))} \subseteq \{\mathbf{x} \mid \mathbf{c}^t\mathbf{x} \leq \delta - w_d - 1\}.$$

On the other hand, $P \subseteq \{\mathbf{x} \mid \mathbf{c}^t\mathbf{x} > \delta - w_d - 1\}$, so that

$$t(d) := w_d + 2 + (w_d + 1)t(d-1)$$

satisfies the requirements of the theorem. □

Definition 13.18. *The* **Chvátal rank** *of a rational matrix $A \in \mathbb{Q}^{m\times n}$ is the smallest number $t \in \mathbb{N}$ such that*

$$\{\mathbf{x} \mid A\mathbf{x} \leq \mathbf{b}\}_I \;=\; \{\mathbf{x} \mid A\mathbf{x} \leq \mathbf{b}\}^{(t)}.$$

for each integral vector $\mathbf{b} \in \mathbb{Z}^m$.

We extend our definition of a unimodular matrix slightly and say that an integral matrix $A \in \mathbb{Z}^{m\times n}$ of rank r is unimodular if and only if for each sub-matrix B consisting of r linearly independent columns of A the gcd of all sub-determinants of order r of B is 1. If $\operatorname{rank}(A) = m$, then this clearly coincides with our earlier definition. The proof of the following characterization of integer polyhedra is completely analogous to that of Theorem 10.5.

Proposition 13.19. *Let $A \in \mathbb{Z}^{m\times n}$ be an integral matrix. Then the following are equivalent:*

(1) A is unimodular,
(2) $\{\mathbf{x} \mid A\mathbf{x} = \mathbf{b}, \mathbf{x} \geq \mathbf{0}\}$ is integral for any integral $\mathbf{b} \in \mathbb{Z}^m$,
(3) $\{\mathbf{y}^t \mid \mathbf{y}^t A \geq \mathbf{c}^t\}$ is integral for any integral $\mathbf{c}^t \in \mathbb{Z}^n$. □

This implies the following characterization of matrices with Chvátal rank 0.

Corollary 13.20. *An integral matrix A has* Chvátal *rank* 0 *if and only if A^t is unimodular.* □

There is no such characterization known for matrices of higher rank in general. If A is totally unimodular, then it follows from Proposition 10.6, that A has Chvátal rank at most 1. A priori, it is not even clear that the Chvátal rank is finite for every matrix. In fact it is, and we will prove this with the following theorem.

Theorem 13.21. *Let $A \in \mathbb{Q}^{m\times n}$ be a rational matrix. Then there is $t \in \mathbb{N}$ such that for any rational $\mathbf{b} \in \mathbb{Q}^m$*

$$\{\mathbf{x} \mid A\mathbf{x} \leq \mathbf{b}\}_I \;=\; \{\mathbf{x} \mid A\mathbf{x} \leq \mathbf{b}\}^{(t)}.$$

We cannot prove this directly, but need some estimates on sizes of solutions first. This is interesting in its own right, so we formulate the results as separate theorems.

Lemma 13.22. *Let $A \in \mathbb{Z}^{m\times n}$ such that all minors of A have absolute value at most β. Let $C := \{\mathbf{x} \in \mathbb{R}^n \mid A\mathbf{x} \leq \mathbf{0}\}$. Let $\mathbf{y}_1, \ldots, \mathbf{y}_k$ be a set of primitive vectors spanning C and*

$$\Pi := \left\{\sum_{i=1}^{k} \lambda_i \mathbf{y}_i \mid 0 \leq \lambda_i \leq 1,\ 1 \leq i \leq k\right\}.$$

Then $\|\mathbf{x}\|_\infty \leq n\beta$ for any $\mathbf{x} \in \Pi$.

Proof. Each generator $\mathbf{y}$ of C is a solution of a subsystem $A_{*I}\mathbf{y} = 0$ of $A\mathbf{x} \leq \mathbf{0}$. Let $B := A_{*I}$ and assume that B has full row rank $\leq m-1$. We may reorder the rows of B such that $B = [B_1\, B_2]$ for a non-singular matrix B_1. Splitting $\mathbf{y} = (\mathbf{y}_1, \mathbf{y}_2)$ accordingly, we can construct a solution by setting $\mathbf{y}_2 = (1, 0, \dots, 0)$ and $y_1 := -B_2^{-1}(B_1)_{*1}$. By Cramer's rule, each entry of $\mathbf{y}_1$ is a quotient of minors of B. So, after clearing denominators, each generator has components of size at most β.

By Caratheodory's Theorem we know that for any $\mathbf{x} \in \Pi$ we need at most n generators in a conic combination of $\mathbf{x}$, so the claim follows. ☐

test set

The next theorem is the basis for a **test set** approach to integer programming. It shows that, if we have a feasible solution of an integer program that is not yet optimal, then there is another feasible solution not too far away from the original one that improves the objective value. In the test set method one wants to find a finite set of vectors $\mathscr{T}$ with the property that given some feasible, but not optimal solution $\mathbf{y}$, there is some $\mathbf{t} \in \mathscr{T}$ such that $\mathbf{y} + \mathbf{t}$ is feasible with better objective value. One iterates this step until an optimal solution is found (i.e. there is no $\mathbf{t} \in \mathscr{T}$ such that $\mathbf{y} + \mathbf{t}$ is feasible).

Theorem 13.23. *Let $P := \{\mathbf{x} \mid A\mathbf{x} \leq \mathbf{b}\}$ for $A \in \mathbb{Z}^{m \times n}$ such that the absolute values of all minors of A are at most β, $\mathbf{b} \in \mathbb{Z}^m$ and $\mathbf{c}^t \in \mathbb{R}^n$. Let $\mathbf{y}$ be a feasible but not optimal solution of the integer linear program*

$$\max(\mathbf{c}^t\mathbf{x} \mid A\mathbf{x} \leq \mathbf{b},\, \mathbf{x} \in Z^m)\,.$$

Then there exists a feasible solution $\tilde{\mathbf{y}}$ such that $\mathbf{c}^t\tilde{\mathbf{y}} > \mathbf{c}^t\mathbf{y}$ and $\|\mathbf{y} - \tilde{\mathbf{y}}\| \leq n\beta$.

Proof. If $\mathbf{y}$ is not optimal, then there is $\mathbf{y}' \in P \cap \mathbb{Z}^n$ such that $\mathbf{c}^t\mathbf{y}' > \mathbf{c}^t\mathbf{y}$. Split $[m]$ into subsets I and J such that

$$A_{I*}\mathbf{y}' \leq A_{I*}\mathbf{y} \qquad \text{and} \qquad A_{J*}\mathbf{y}' \leq A_{J*}\mathbf{y}\,.$$

Define the cone

$$C := \{\mathbf{x} \mid A_{I*}\mathbf{x} \leq \mathbf{0},\, -A_{J*}\mathbf{x} \leq \mathbf{0}\}$$

and let $\mathscr{H}_C := \{\mathbf{a}_1, \dots, \mathbf{a}_k\}$ be a Hilbert basis of C. Then $\mathbf{y}' - \mathbf{y} \in C \cap \mathbb{Z}^n$, so there are nonnegative integers $\lambda_1, \dots, \lambda_k$ such that

$$\mathbf{y}' - \mathbf{y} = \sum_{i=1}^{k} \lambda_i \mathbf{a}_i\,.$$

Then

$$0 < \mathbf{c}^t(\mathbf{y}' - \mathbf{y}) = \sum_{i=1}^{k} \lambda_i \mathbf{c}^t \mathbf{a}_i$$

implies that there is some index j such that $\lambda_j \geq 1$ and $\mathbf{c}^t\mathbf{a}_j > 0$. Set $\tilde{\mathbf{y}} := \mathbf{y} + \mathbf{a}_j$. We check that $\tilde{\mathbf{y}} \in P$:

$$\begin{aligned} A_{I*}\tilde{\mathbf{y}} &= A_{I*}(\mathbf{y} + \mathbf{a}_j) \leq \mathbf{b}_I \\ A_{J*}\tilde{\mathbf{y}} &= A_{J*}(\mathbf{y} + \mathbf{a}_j) = A_{J*}(\mathbf{y}' + \mathbf{a}_j - \sum_{i=1}^{k} \lambda_i \mathbf{a}_i) \\ &= A_{J*}\mathbf{y}' - (\lambda_j - 1)A_{J*} - \sum_{i \neq j} \lambda_i A_{J*}\mathbf{a}_i \leq \mathbf{b}_J\,, \end{aligned}$$

as $A_{I*}\mathbf{a}_i \leq \mathbf{0}$, $A_{J*}\mathbf{a}_i \geq \mathbf{0}$ for $1 \leq i \leq k$ and $\lambda_j - 1 \geq 0$. Further

$$\mathbf{c}^t\tilde{\mathbf{y}} = \mathbf{c}^t(\mathbf{y} + \mathbf{a}_j) < \mathbf{c}^t\mathbf{y}\,,$$

as $\mathbf{c}^t\mathbf{a}_j > 0$, and $\|\mathbf{y} - \tilde{\mathbf{y}}\| \leq n\beta$ by Lemma 13.22. Hence $\tilde{\mathbf{y}}$ is as required. ☐

Now we examine how far apart integer and linear optimal solutions can be. We have seen that we cannot give a bound in the dimension and the number of inequalities defining our polyhedron, see Figure 8.1. However, the following theorem shows that we at least can give a bound in the size of the constraint matrix A.

Theorem 13.24. *Let $A \in \mathbb{Z}^{m\times n}$ such that each minor has absolute value at most β, and let $\mathbf{b} \in Z^m$, $\mathbf{c} \in \mathbb{Z}^n$. Assume that the linear programs*

$$\max(\mathbf{c}^t\mathbf{x} \mid A\mathbf{x} \le \mathbf{b}) \tag{13.8}$$

$$\max(\mathbf{c}^t\mathbf{x} \mid A\mathbf{x} \le \mathbf{b},\, \mathbf{x} \in \mathbb{Z}^n) \tag{13.9}$$

are both finite. Then

1. *For each optimal solution $\overline{\mathbf{x}}$ of (13.8) there is an optimal solution $\overline{\mathbf{y}}$ of (13.9) such that $\|\mathbf{x}-\mathbf{y}\| \le n\beta$.*
2. *For each optimal solution $\overline{\mathbf{y}}$ of (13.9) there is an optimal solution $\overline{\mathbf{x}}$ of (13.8) such that $\|\mathbf{x}-\mathbf{y}\| \le n\beta$.*

Proof. Let $\overline{\mathbf{x}}$ be optimal for (13.8) and $\overline{\mathbf{y}}$ optimal for (13.9). Then there are $I \subset [m]$, $J := [m] - I$ such that

$$A_{I*}\overline{\mathbf{x}} < A_{I*}\overline{\mathbf{y}} \qquad\qquad A_{J*}\overline{\mathbf{x}} \ge A_{J*}\overline{\mathbf{y}}\,.$$

This implies $A_{I*}\overline{\mathbf{x}} < A_{I*}\overline{\mathbf{y}} \le \mathbf{b}_I$. We define the cone

$$C := \{\mathbf{x} \mid A_{I*}\mathbf{x} \le \mathbf{0},\, A_{J*}\mathbf{x} \ge \mathbf{0}\}\,.$$

By complementary slackness, we know that there is $\mathbf{u} \ge \mathbf{0}$ such that $\mathbf{u}^tA_{J*} = \mathbf{c}^t$. Hence, $\mathbf{c}^t\mathbf{x} \ge \mathbf{0}$ for all $\mathbf{x} \in C$. Further,

$$A_{I*}(\overline{\mathbf{x}}-\overline{\mathbf{y}}) = A_{I*}\overline{\mathbf{x}} - A_{I*}\overline{\mathbf{y}} \le \mathbf{0}$$
$$A_{J*}(\overline{\mathbf{x}}-\overline{\mathbf{y}}) = A_{J*}\overline{\mathbf{x}} - A_{J*}\overline{\mathbf{y}} \ge \mathbf{0}$$

so $\overline{\mathbf{x}} - \overline{\mathbf{y}} \in C$. This implies that there are $\mathbf{v}_1, \ldots, \mathbf{v}_k \in C$ and $\lambda_1, \ldots, \lambda_k \ge 0$ such that

$$\overline{\mathbf{x}} - \overline{\mathbf{y}} = \lambda_1\mathbf{v}_1 + \cdots + \lambda_k\mathbf{v}_k\,.$$

By Cramer's rule we can assume that $\|\mathbf{v}_i\|_\infty \le \beta$ for $1 \le i \le k$ (the proof of this is completely analogous to that of Corollary 5.8, except that we are only interested in the size of the largest entry, so we don't have to sum up). Now let $\mu_1, \ldots, \mu_k \ge 0$ be such that $0 \le \mu_i \le \lambda_i$ for all i. Then

$$\begin{aligned} & A_{I*}(\overline{\mathbf{y}} + \mu_1\mathbf{v}_1 + \cdots + \mu_k\mathbf{v}_k) \le A_{I*}\overline{\mathbf{y}} \le \mathbf{b}_I \\ \text{and}\quad & A_{J*}(\overline{\mathbf{y}} + \mu_1\mathbf{v}_1 + \cdots + \mu_k\mathbf{v}_k) \le A_{J*}\left(\overline{\mathbf{z}} - (\lambda_1-\mu_1)\mathbf{v}_1 - \cdots - (\lambda_k-\mu_k)\mathbf{v}_k\right) \\ & \qquad \le A_{J*}\overline{\mathbf{x}} \le \mathbf{b}_J\,, \end{aligned}$$

so $\overline{\mathbf{y}} + \mu_1\mathbf{v}_1 + \cdots + \mu_k\mathbf{v}_k \in C$. We define the vector

$$\tilde{\mathbf{y}} := \overline{\mathbf{y}} + \lfloor\lambda_1\rfloor\mathbf{v}_1 + \cdots + \lfloor\lambda_k\rfloor\mathbf{v}_k = \overline{\mathbf{x}} - \{\lambda_1\}\mathbf{v}_1 - \cdots - \{\lambda_k\}\mathbf{v}_k,.$$

As $\mathbf{c}^t\mathbf{v}_i \ge 0$ for all i we know that $\tilde{\mathbf{y}}$ satisfies $\mathbf{c}^t\tilde{\mathbf{y}} \ge \mathbf{c}^t\overline{\mathbf{y}}$. So $\tilde{\mathbf{y}}$ is an integral optimal solution of (13.9). We compute

$$\|\overline{\mathbf{x}} - \tilde{\mathbf{y}}\|_\infty = \|\{\lambda_1\}\mathbf{v}_1 + \cdots + \{\lambda_k\}\mathbf{v}_k\|_\infty \le \|\mathbf{v}_1\|_\infty + \cdots + \|\mathbf{v}_k\|_\infty \le n\beta\,,$$

which implies the first claim. For the second we define

$$\tilde{\mathbf{x}} := \overline{\mathbf{x}} - \lfloor\lambda_1\rfloor\mathbf{v}_1 - \cdots - \lfloor\lambda_k\rfloor\mathbf{v}_k = \overline{\mathbf{y}} + \{\lambda_1\}\mathbf{v}_1 + \cdots + \{\lambda_k\}\mathbf{v}_k\,.$$

If for some $1 \le i \le k$ we have $\lfloor\lambda_i\rfloor > 0$ and $\mathbf{c}^t\mathbf{v}_i > 0$, then $\overline{\mathbf{y}} + \lfloor\lambda_i\rfloor\mathbf{v}_i$ is an integral solution of (13.9) with $\mathbf{c}^t(\overline{\mathbf{y}} + \lfloor\lambda_i\rfloor\mathbf{v}_i) > \mathbf{c}^t\overline{\mathbf{y}}$, which contradicts optimality of $\overline{\mathbf{y}}$. Hence,

$$\mathbf{c}^t\tilde{\mathbf{x}} \ge \mathbf{c}^t\overline{\mathbf{x}}\,,$$

and $\tilde{\mathbf{x}}$ is an optimal solution of (13.8). As above, the claim now follows from

$$\|\overline{\mathbf{y}} - \tilde{\mathbf{x}}\|_\infty = \|\{\lambda_1\}\mathbf{v}_1 + \cdots + \{\lambda_k\}\mathbf{v}_k\|_\infty \le \|\mathbf{v}_1\|_\infty + \cdots + \|\mathbf{v}_k\|_\infty \le n\beta\,. \qquad \square$$

Theorem 13.25. *For each $A \in \mathbb{Q}^{m\times n}$ there is some $M \in \mathbb{Z}^{l\times n}$ such that for each $\mathbf{b} \in \mathbb{Q}^m$ there is $\mathbf{d} \in \mathbb{Q}^l$ such that*

$$\{\mathbf{x} \mid A\mathbf{x} \le \mathbf{b}\}_I \;=\; \{\mathbf{x} \mid M\mathbf{x} \le \mathbf{d}\}\,.$$

If $A \in \mathbb{Z}^{m\times n}$ and all minors have an absolute value bounded by β, then we can find such an M where all entries of M have an absolute value at most $n^{2n}\beta^n$.

Proof. Let A be such that all minors have an absolute value bounded by β and

$$L := \{\mathbf{u} \mid \text{there is } \mathbf{y} \ge \mathbf{0} \text{ such that } \mathbf{y}^t A = \mathbf{u},\ \mathbf{u} \in \mathbb{Z}^n,\ \|\mathbf{u}\|_\infty \le n^{2n}\beta^n\}\,.$$

In other words, L is the set of all integral vectors in the cone cone(A) with infinity norm bounded by $n^{2n}\beta^n$. Let M be the matrix whose rows are the vectors in L. Observe that A is a sub-matrix of L. Let $I \subset \mathbb{N}$ be such that $A_I = M$.

By Proposition 3.4, any vector $\mathbf{c}^t \in \mathbb{R}^n$ is bounded over the polyhedron $P := \{\mathbf{x} \mid A\mathbf{x} \le \mathbf{b}\}$ if and only if $\mathbf{c}^t$ is in the polar cone cone(A) of the recession cone $\{\mathbf{x} \mid A\mathbf{x} \le \mathbf{0}\}$ of P. Hence, any linear functional $\mathbf{m}^t$ of M is bounded on $\{\mathbf{x} \mid A\mathbf{x} \le \mathbf{b}\}$, and $\mathbf{m}^t$ attains a finite maximum over P and the integer hull of P.

Now fix a right hand side $\mathbf{b}$. If $A\mathbf{x} \le \mathbf{b}$ has no solution, then by choosing $\mathbf{d}_I = \mathbf{b}$ we can achieve that also $M\mathbf{x} \le \mathbf{d}$ has no solution. If $A\mathbf{x} \le \mathbf{b}$ is feasible, but has no integral solution, then its recession cone $\{\mathbf{x} \mid A\mathbf{x} \le \mathbf{0}\}$ is not full dimensional, which implies that there is an implicit equality in the system $A\mathbf{x} \le 0$. Hence, both $\mathbf{a}$ and $-\mathbf{a}$ are in L, and we can choose $\mathbf{d}$ such that $M\mathbf{x} \le \mathbf{d}$ is infeasible.

So we may assume that $A\mathbf{x} \le \mathbf{b}$ has an integral solution. For each $\mathbf{c}^t \in \mathbb{R}^n$ we define

$$\delta_{\mathbf{c}} := \max(\mathbf{c}^t\mathbf{x} \mid A\mathbf{x} \le \mathbf{b},\ \mathbf{x} \in \mathbb{Z}^n)\,,$$

which is finite, by the above argument. It then suffices to show that

$$\{\mathbf{x} \mid A\mathbf{x} \le \mathbf{b}\}_I = \{\mathbf{x} \in \mathbb{R}^n \mid \mathbf{u}^t\mathbf{x} \le \delta_{\mathbf{u}} \text{ for all } \mathbf{u} \in L\}\,. \tag{13.10}$$

We show two inclusions. "$\subseteq$" just follows from the definition of $\delta_{\mathbf{u}}$ for any $\mathbf{u}$. So we need to show "$\supseteq$". Let $\mathbf{c}^t\mathbf{x} \le \delta$ be a valid inequality for the left hand side. We then have to show that $\mathbf{c}^t\mathbf{x} \le \delta$ is valid for the set on the right hand side. Let $\mathbf{z}$ be an optimal solution of

$$\max(\mathbf{c}^t\mathbf{x} \mid A\mathbf{x} \le \mathbf{b},\ \mathbf{x} \in \mathbb{Z}^n)\,,$$

By Theorem 13.23

$$\begin{aligned} C &:= \operatorname{cone}(\mathbf{x} - \mathbf{z} \mid A\mathbf{x} \le \mathbf{b},\ \mathbf{x} \in \mathbb{Z}^n) \\ &= \operatorname{cone}(\mathbf{x} - \mathbf{z} \mid A\mathbf{x} \le \mathbf{b},\ \mathbf{x} \in \mathbb{Z}^n, \|\mathbf{x} - \mathbf{z}\|_\infty \le n\beta) \end{aligned}$$

C is the cone spanned by the integral vectors in the cone spanned by all directions pointing into the polyhedron $A\mathbf{x} \le \mathbf{b}$ from $\mathbf{z}$. It follows that there is a subset L_1 of L such that

$$C = \{\mathbf{y} \mid \mathbf{u}^t\mathbf{y} \le 0 \text{ for all } \mathbf{u} \in L_i\}.$$

Here Cramer's rule ensures that the vectors of bounded length in L suffice (by the same argument already used in the proof of Theorem 13.24). For each $\mathbf{u} \in L_1$ we have that $\delta_{\mathbf{u}} = \mathbf{u}^t\mathbf{z}$. By construction, we also know that

$$\mathbf{c}^t\mathbf{y} \le 0$$

for all $\mathbf{y} \in C$, so $\mathbf{c}^t \in C^*$. This dual cone is spanned by the vectors in L_1, hence, there are $\mathbf{u}_1^t, \ldots, \mathbf{u}_k^t \in L_1$ and $\lambda_1, \ldots, \lambda_k \ge 0$ such that

$$\mathbf{c}^t = \lambda_1\mathbf{u}_1^t + \cdots + \lambda_k\mathbf{u}_k^t$$

and

$$\delta_{\mathbf{c}} = \mathbf{c}^t\mathbf{z} = \lambda_1\mathbf{u}_1^t\mathbf{z} + \cdots + \lambda_k\mathbf{u}_k^t\mathbf{z} = \lambda_1\delta_{\mathbf{u}_1} + \cdots + \lambda_k\delta_{\mathbf{u}_k}.$$

Hence, the inequality $\mathbf{c}^t\mathbf{x} \le \delta_{\mathbf{c}}$ is valid for the right hand side of (13.10). □

Now we can finally prove that the CHVÁTAL RANK of any integral matrix A is finite.

Proof (of Theorem 13.21). We may assume that $A \in \mathbb{Z}^{m\times n}$ and $\mathbf{b} \in \mathbb{Z}^m$. Let β be the maximum absolute value of a minor of A. We claim that

$$t := \max\big(t(n), n^{2n+2}\beta^{n+1}(1+t(n-1))+1\big)$$

suffices, where $t(n)$ is the function obtained in Theorem 13.17. Fix a vector $\mathbf{b}$ and define

$$P^{\mathbf{b}} := \{\mathbf{x} \mid A\mathbf{x} \le \mathbf{b}\}\,.$$

If $P_I^{\mathbf{b}} = \emptyset$, then t suffices by Theorem 13.17. So we assume that $P_I^{\mathbf{b}} \neq \emptyset$. By the previous Theorem 13.25, there is a matrix M and a vector $\mathbf{d}$ such that all entries of M have absolute value at most $n^{2n}\beta$ such that

$$P_I^{\mathbf{b}} = \{\mathbf{x} \mid M\mathbf{x} \le \mathbf{d}\}\,.$$

Let $\mathbf{m}^t\mathbf{x} \le \delta$ be a valid inequality of $M\mathbf{x} \le \mathbf{d}$. We can clearly assume that $\delta = \max(\mathbf{m}^t\mathbf{x} \mid \mathbf{x} \in P_I^{\mathbf{b}})$. Let

$$\delta' := \lfloor \max(\mathbf{m}^t\mathbf{x} \mid \mathbf{x} \in P) \rfloor\,.$$

By Theorem 13.24, $\delta' - \delta \le \|\mathbf{m}\|_1 n\beta \le n^{2n+2}\beta^{n+1}$. Now one shows in the same way as in Theorem 13.17 that

$$(P^{\mathbf{b}})^{(k+1+k\cdot t(n-1))} \subseteq \{\mathbf{x} \mid \mathbf{m}^t\mathbf{x} \le \delta' - k\}\,.$$

Hence, for $k := \delta' - \delta$ we have that $P_I^{\mathbf{b}} \subseteq \{\mathbf{x} \mid \mathbf{m}^t\mathbf{x} \le \delta\}$. $\mathbf{m}^t\mathbf{x} \le \delta$ was an arbitrary inequality of $M\mathbf{x} \le \mathbf{d}$, so this implies the theorem. $\square$

This finally completes the proof that the CHVÁTAL rank of an integral matrix is finite. However, there is no characterization known of matrices with a given CHVÁTAL rank r, except for $r = 0$.

Bibliography

[Bar02] Alexander Barvinok. *A Course in Convexity*. Oxford University Press, 2002.

[Chv83] Vašek Chvátal. *Linear programming.* A series of books in the mathematical sciences. New York - San Francisco: W. H. Freeman and Company., 1983.

[KV08] Bernhard Korte and Jens Vygen. *Combinatorial optimization. Theory and algorithms. 4th ed.* Algorithms and Combinatorics 21. Berlin: Springer., 2008.

[MG07] Jiří Matoušek and Bernd Gärtner. *Understanding and using linear programming.* Universitext. Berlin: Springer., 2007.

[Sch86] Alexander Schrijver. *Theory of linear and integer programming.* Wiley-Interscience Series in Discrete Mathematics. A Wiley-Interscience Publication. Chichester: John Wiley & Sons Ltd., 1986.

[Sch03] Alexander Schrijver. *Combinatorial optimization. Polyhedra and efficiency (3 volumes)*. Algorithms and Combinatorics 24. Berlin: Springer., 2003.

[Zie95] Günter M. Ziegler. *Lectures on polytopes.* Graduate Texts in Mathematics. 152. Berlin: Springer., 1995.