4. Theory of Linear Programming: First Steps

4.1 Equational Form

In the introductory chapter we explained how each linear program can be converted to the form

$$\text{maximize } \mathbf{c}^T\mathbf{x} \text{ subject to } A\mathbf{x} \leq \mathbf{b}.$$

But the simplex method requires a different form, which is usually called the *standard form* in the literature. In this book we introduce a less common, but more descriptive term *equational form*. It looks like this:

Equational form of a linear program:

$$\begin{array}{ll} \text{Maximize} & \mathbf{c}^T\mathbf{x} \\ \text{subject to} & A\mathbf{x} = \mathbf{b} \\ & \mathbf{x} \geq \mathbf{0}. \end{array}$$

As usual, $\mathbf{x}$ is the vector of variables, A is a given $m \times n$ matrix, $\mathbf{c} \in \mathbb{R}^n$, $\mathbf{b} \in \mathbb{R}^m$ are given vectors, and $\mathbf{0}$ is the zero vector, in this case with n components.

The constraints are thus partly equations, and partly inequalities of a very special form $x_j \geq 0$, $j = 1, 2, \ldots, n$, called **nonnegativity constraints**. (Warning: Although we call this form equational, it contains inequalities as well, and these must not be forgotten!)

Let us emphasize that *all* variables in the equational form have to satisfy the nonnegativity constraints.

In problems encountered in practice we often have nonnegativity constraints automatically, since many quantities, such as the amount of consumed cucumber, cannot be negative.

Transformation of an arbitrary linear program to equational form. We illustrate such a transformation for the linear program

$$\begin{array}{ll} \text{maximize} & 3x_1 - 2x_2 \\ \text{subject to} & 2x_1 - x_2 \leq 4 \\ & x_1 + 3x_2 \geq 5 \\ & x_2 \geq 0. \end{array}$$

We proceed as follows:

1. In order to convert the inequality $2x_1 - x_2 \leq 4$ to an equation, we introduce a new variable x_3, together with the nonnegativity constraint $x_3 \geq 0$, and we replace the considered inequality by the equation $2x_1 - x_2 + x_3 = 4$. The auxiliary variable x_3, which won't appear anywhere else in the transformed linear program, represents the difference between the right-hand side and the left-hand side of the inequality. Such an auxiliary variable is called a **slack variable**.
2. For the next inequality $x_1 + 3x_2 \geq 5$ we first multiply by -1, which reverses the direction of the inequality. Then we introduce another slack variable x_4 with the nonnegativity constraint $x_4 \geq 0$, and we replace the inequality by the equation $-x_1 - 3x_2 + x_4 = -5$.
3. We are not finished yet: The variable x_1 in the original linear program is allowed to attain both positive and negative values. We introduce two new, nonnegative, variables y_1 and z_1, $y_1 \geq 0$, $z_1 \geq 0$, and we substitute for x_1 the difference $y_1 - z_1$ everywhere. The variable x_1 itself disappears.

The resulting equational form of our linear program is

$$\begin{array}{ll} \text{maximize} & 3y_1 - 3z_1 - 2x_2 \\ \text{subject to} & 2y_1 - 2z_1 - x_2 + x_3 = 4 \\ & -y_1 + z_1 - 3x_2 + x_4 = -5 \\ & y_1 \geq 0,\ z_1 \geq 0,\ x_2 \geq 0,\ x_3 \geq 0,\ x_4 \geq 0. \end{array}$$

So as to comply with the conventions of the equational form in full, we should now rename the variables to $x_1, x_2, \ldots, x_5$.

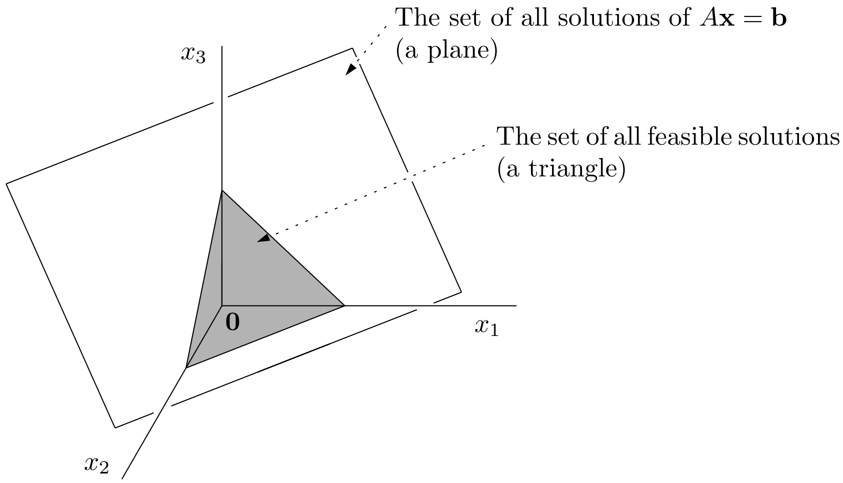
The presented procedure converts an arbitrary linear program with n variables and m constraints into a linear program in equational form with at most $m + 2n$ variables and m equations (and, of course, nonnegativity constraints for all variables).

Geometry of a linear program in equational form. Let us consider a linear program in equational form:

$$\text{Maximize } \mathbf{c}^T\mathbf{x} \text{ subject to } A\mathbf{x} = \mathbf{b},\ \mathbf{x} \geq \mathbf{0}.$$

As is derived in linear algebra, the set of all solutions of the system $A\mathbf{x} = \mathbf{b}$ is an affine subspace F of the space $\mathbb{R}^n$. Hence the set of all feasible solutions of the linear program is the intersection of F with the *nonnegative orthant*, which is the set of all points in $\mathbb{R}^n$ with all coordinates nonnegative.[1] The following picture illustrates the geometry of feasible solutions for a linear program with $n = 3$ variables and $m = 1$ equation, namely, the equation $x_1 + x_2 + x_3 = 1$:

[1] In the plane ($n = 2$) this set is called the nonnegative *quadrant*, in $\mathbb{R}^3$ it is the nonnegative *octant*, and the name *orthant* is used for an arbitrary dimension.



(In interesting cases we usually have more than 3 variables and no picture can be drawn.)

A preliminary cleanup. Now we will be talking about solutions of the system $A\mathbf{x} = \mathbf{b}$. By this we mean arbitrary real solutions, whose components may be positive, negative, or zero. So this is not the same as feasible solutions of the considered linear program, since a feasible solution has to satisfy $A\mathbf{x} = \mathbf{b}$ *and* have all components nonnegative.

If we change the system $A\mathbf{x} = \mathbf{b}$ by some transformation that preserves the set of solutions, such as a row operation in Gaussian elimination, it influences neither feasible solutions nor optimal solutions of the linear program. This will be amply used in the simplex method.

Assumption: We will consider only linear programs in equational form such that

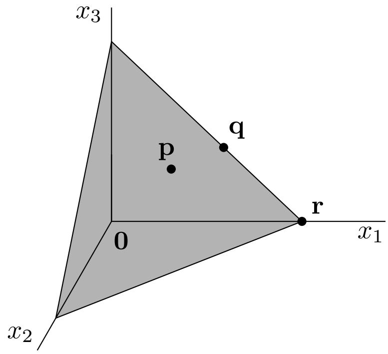
- the system of equations $A\mathbf{x} = \mathbf{b}$ has at least one solution, and
- the rows of the matrix A are linearly independent.

As an explanation of this assumption we need to recall a few facts from linear algebra. Checking whether the system $A\mathbf{x} = \mathbf{b}$ has a solution is easy by Gaussian elimination, and if there is no solution, the considered linear program has no feasible solution either, and we can thus disregard it.

If the system $A\mathbf{x} = \mathbf{b}$ has a solution and if some row of A is a linear combination of the other rows, then the corresponding equation is redundant and it can be deleted from the system without changing the set of solutions. We may thus assume that the matrix A has m linearly independent rows and (therefore) rank m.

4.2 Basic Feasible Solutions

Among all feasible solutions of a linear program, a privileged status is granted to so-called basic feasible solutions. In this section we will consider them only for linear programs in equational form. Let us look again at the picture of the set of feasible solutions for a linear program with $n = 3$, $m = 1$:



Among the feasible solutions $\mathbf{p}$, $\mathbf{q}$, and $\mathbf{r}$ only $\mathbf{r}$ is basic. Expressed geometrically and very informally, a basic feasible solution is a tip (corner, spike) of the set of feasible solutions. We will formulate this kind of geometric description of a basic feasible solution later (see Theorem 4.4.1).

The definition that we present next turns out to be equivalent, but it looks rather different. It requires that, very roughly speaking, a basic feasible solution have sufficiently many zero components. Before stating it we introduce a new piece of notation.

In this section A is always a matrix with m rows and n columns ($n \geq m$), of rank m. For a subset $B \subseteq \{1, 2, \ldots, n\}$ we let A_B denote the matrix consisting of the columns of A whose indices belong to B. For instance, for

$$A = \begin{pmatrix} 1 & 5 & 3 & 4 & 6 \\ 0 & 1 & 3 & 5 & 6 \end{pmatrix} \text{ and } B = \{2, 4\} \text{ we have } A_B = \begin{pmatrix} 5 & 4 \\ 1 & 5 \end{pmatrix}.$$

We will use a similar notation for vectors; e.g., for $\mathbf{x} = (3, 5, 7, 9, 11)$ and $B = \{2, 4\}$ we have $\mathbf{x}_B = (5, 9)$.

Now we are ready to state a formal definition.

A **basic feasible solution** of the linear program

$$\text{maximize } \mathbf{c}^T\mathbf{x} \text{ subject to } A\mathbf{x} = \mathbf{b} \text{ and } \mathbf{x} \geq \mathbf{0}$$

is a feasible solution $\mathbf{x} \in \mathbb{R}^n$ for which there exists an m-element set $B \subseteq \{1, 2, \ldots, n\}$ such that

- the (square) matrix A_B is nonsingular, i.e., the columns indexed by B are linearly independent, and
- $x_j = 0$ for all $j \notin B$.

For example, $\mathbf{x} = (0, 2, 0, 1, 0)$ is a basic feasible solution for

$$A = \begin{pmatrix} 1 & 5 & 3 & 4 & 6 \\ 0 & 1 & 3 & 5 & 6 \end{pmatrix}, \ \mathbf{b} = (14, 7)$$

with $B = \{2, 4\}$.

If such a B is fixed, we call the variables x_j with $j \in B$ the **basic variables**, while the remaining variables are called **nonbasic**. We can thus briefly say that all nonbasic variables are zero in a basic feasible solution.

Let us note that the definition doesn't consider the vector $\mathbf{c}$ at all, and so basic feasible solutions depend solely on A and $\mathbf{b}$.

For some considerations it is convenient to reformulate the definition of a basic feasible solution a little.

4.2.1 Lemma. *A feasible solution $\mathbf{x}$ of a linear program in equational form is basic if and only if the columns of the matrix A_K are linearly independent, where $K = \{j \in \{1, 2, \ldots, n\} : x_j > 0\}$.*

Proof. One of the implications is obvious: If $\mathbf{x}$ is a basic feasible solution and B is the corresponding m-element set as in the definition, then $K \subseteq B$ and thus the columns of the matrix A_K are linearly independent.

Conversely, let $\mathbf{x}$ be feasible and such that the columns of A_K are linearly independent. If $|K| = m$, then we can simply take $B = K$. Otherwise, for $|K| < m$, we extend K to an m-element set B by adding $m - |K|$ more indices so that the columns of A_B are linearly independent. This is a standard fact of linear algebra, which can be verified using the algorithm described next.

We initially set the current B to K, and repeat the following step: If A has a column that is not in the linear span of the columns of A_B, we add the index of such a column to B. As soon as this step is no longer possible, that is, all columns of A are in the linear span of the columns of B, it is easily seen that the columns of A_B constitute a basis of the column space of A. Since A has rank m, we have $|B| = m$ as needed. $\square$

4.2.2 Proposition. *A basic feasible solution is uniquely determined by the set B. That is, for every m-element set $B \subseteq \{1, 2, \ldots, n\}$ with A_B nonsingular there exists at most one feasible solution $\mathbf{x} \in \mathbb{R}^n$ with $x_j = 0$ for all $j \notin B$.*

Let us stress right away that a single basic feasible solution may be obtained from many different sets B.

Proof of Proposition 4.2.2. For $\mathbf{x}$ to be feasible we must have $A\mathbf{x} = \mathbf{b}$. The left-hand side can be rewritten to $A\mathbf{x} = A_B\mathbf{x}_B + A_N\mathbf{x}_N$, where $N = \{1, 2, \ldots, n\} \setminus B$. For $\mathbf{x}$ to be a basic feasible solution, the vector $\mathbf{x}_N$ of nonbasic variables must equal $\mathbf{0}$, and thus the vector $\mathbf{x}_B$ of basic variables satisfies $A_B\mathbf{x}_B = \mathbf{b}$. And here we use the fact that A_B is a nonsingular square matrix: The system $A_B\mathbf{x}_B = \mathbf{b}$ has exactly one solution $\tilde{\mathbf{x}}_B$. If all components of $\tilde{\mathbf{x}}_B$ are nonnegative, then we have exactly one basic feasible solution for the considered B (we amend $\tilde{\mathbf{x}}_B$ by zeros), and otherwise, we have none. $\square$

We introduce the following terminology: We call an m-element set $B \subseteq \{1, 2, \ldots, n\}$ with A_B nonsingular a **basis**.[2] If, moreover, B determines a basic feasible solution, or in other words, if the unique solution of the system $A_B \mathbf{x}_B = \mathbf{b}$ is nonnegative, then we call B a **feasible basis**.

The following theorem deals with the existence of optimal solutions, and moreover, it shows that it suffices to look for them solely among basic feasible solutions.

4.2.3 Theorem. *Let us consider a linear program in equational form*

$$\text{maximize } \mathbf{c}^T \mathbf{x} \text{ subject to } A\mathbf{x} = \mathbf{b}, \mathbf{x} \geq \mathbf{0}.$$

(i) *("Optimal solutions may fail to exist only for obvious reasons.") If there is at least one feasible solution and the objective function is bounded from above on the set of all feasible solutions, then there exists an optimal solution.*

(ii) *If an optimal solution exists, then there is a basic feasible solution that is optimal.*

A proof is not necessary for further reading and we defer it to the end of this section. The theorem also follows from the correctness of the simplex method, which will be discussed in the next chapter.

The theorem just stated implies a finite, although entirely impractical, algorithm for solving linear programs in equational form. We consider all m-element subsets $B \subseteq \{1, 2, \ldots, n\}$ one by one, and for each of them we check whether it is a feasible basis, by solving a system of linear equations (we obtain at most one basic feasible solution for each B by Proposition 4.2.2). Then we calculate the maximum of the objective function over all basic feasible solutions found in this way.

Strictly speaking, this algorithm doesn't work if the objective function is unbounded. Formulating a variant of the algorithm that functions properly even in this case, i.e., it reports that the linear program is unbounded, we leave as an exercise. Soon we will discuss the considerably more efficient simplex method, and there we show in detail how to deal with unboundedness.

We have to consider $\binom{n}{m}$ sets B in the above algorithm.[3] For example, for $n = 2m$, the function $\binom{2m}{m}$ grows roughly like 4^m, i.e., exponentially, and this is too much even for moderately large m.

---

[2] This is a shortcut. The index set B itself is not a basis in the sense of linear algebra, of course. Rather the set of columns of the matrix A_B constitutes a basis of the column space of A.

[3] We recall that the binomial coefficient $\binom{n}{m} = \frac{n!}{m!(n-m)!}$ counts the number of m-element subsets of an n-element set.

As we will see in Chapter 5, the simplex method also goes through basic feasible solutions, but in a more clever way. It walks from one to another while improving the value of the objective function all the time, until it reaches an optimal solution.

Let us summarize the main findings of this section.

> A linear program *in equational form* has finitely many basic feasible solutions, and if it is feasible and bounded, then at least one of the basic feasible solutions is optimal.
>
> Consequently, any linear program that is feasible and bounded has an optimal solution.

Proof of Theorem 4.2.3. We will use some steps that will reappear in the simplex method in a more elaborate form, and so the present proof is a kind of preparation. We prove the following statement:

> *If the objective function of a linear program in equational form is bounded above, then for every feasible solution* $\mathbf{x}_0$ *there exists a basic feasible solution* $\tilde{\mathbf{x}}$ *with the same or larger value of the objective function; i.e.,* $\mathbf{c}^T\tilde{\mathbf{x}} \geq \mathbf{c}^T\mathbf{x}_0$.

How does this imply the theorem? If the linear program is feasible and bounded, then according to the statement, for every feasible solution there is a basic feasible solution with the same or larger objective function. Since there are only finitely many basic feasible solutions, some of them have to give the maximum value of the objective function, which means that they are optimal. We thus get both (i) and (ii) at once.

In order to prove the statement, let us consider an arbitrary feasible solution $\mathbf{x}_0$. Among all feasible solutions $\mathbf{x}$ with $\mathbf{c}^T\mathbf{x} \geq \mathbf{c}^T\mathbf{x}_0$ we choose one that has the largest possible number of zero components, and we call it $\tilde{\mathbf{x}}$ (it need not be determined uniquely). We define an index set

$$K = \{j \in \{1, 2, \ldots, n\} : \tilde{x}_j > 0\}.$$

If the columns of the matrix A_K are linearly independent, then $\tilde{\mathbf{x}}$ is a basic feasible solution as in the statement, by Lemma 4.2.1, and we are done.

So let us suppose that the columns of A_K are linearly dependent, which means that there is a nonzero $|K|$-component vector $\mathbf{v}$ such that $A_K\mathbf{v} = \mathbf{0}$. We extend $\mathbf{v}$ by zeros in positions outside K to an n-component vector $\mathbf{w}$ (so $\mathbf{w}_K = \mathbf{v}$ and $A\mathbf{w} = A_K\mathbf{v} = \mathbf{0}$).

Let us assume for a moment that $\mathbf{w}$ satisfies the following two conditions (we will show later why we can assume this):

(i) $\mathbf{c}^T\mathbf{w} \geq 0$.
(ii) There exists $j \in K$ with $w_j < 0$.

For a real number $t \geq 0$ let us consider the vector $\mathbf{x}(t) = \tilde{\mathbf{x}} + t\mathbf{w}$. We show that for some suitable $t_1 > 0$ the vector $\mathbf{x}(t_1)$ is a feasible solution with more zero components than $\tilde{\mathbf{x}}$. At the same time, $\mathbf{c}^T\mathbf{x}(t_1) = \mathbf{c}^T\tilde{\mathbf{x}} + t_1\mathbf{c}^T\mathbf{w} \geq \mathbf{c}^T\mathbf{x}_0 + t_1\mathbf{c}^T\mathbf{w} \geq \mathbf{c}^T\mathbf{x}_0$, and so we get a contradiction to the assumption that $\tilde{\mathbf{x}}$ has the largest possible number of zero components.

We have $A\mathbf{x}(t) = \mathbf{b}$ for all t since $A\mathbf{x}(t) = A\tilde{\mathbf{x}} + tA\mathbf{w} = A\tilde{\mathbf{x}} = \mathbf{b}$, because $\tilde{\mathbf{x}}$ is feasible. Moreover, for $t = 0$ the vector $\mathbf{x}(0) = \tilde{\mathbf{x}}$ has all components from K strictly positive and all other components zero. For the jth component of $\mathbf{x}(t)$ we have $x(t)_j = \tilde{x}_j + tw_j$, and if $w_j < 0$ as in condition (ii), we get $x(t)_j < 0$ for all sufficiently large $t > 0$. If we begin with $t = 0$ and let t grow, then those $x(t)_j$ with $w_j < 0$ are decreasing, and at a certain moment $\tilde{t}$ the first of these decreasing components reaches 0. At this moment, obviously, $\mathbf{x}(\tilde{t})$ still has all components nonnegative, and thus it is feasible, but it has at least one extra zero component compared to $\tilde{\mathbf{x}}$. This, as we have already noted, is a contradiction.

Now what do we do if the vector $\mathbf{w}$ fails to satisfy condition (i) or (ii)? If $\mathbf{c}^T\mathbf{w} = 0$, then (i) holds and (ii) can be recovered by changing the sign of $\mathbf{w}$ (since $\mathbf{w} \neq \mathbf{0}$). So we assume $\mathbf{c}^T\mathbf{w} \neq 0$, and again after a possible sign change we can achieve $\mathbf{c}^T\mathbf{w} > 0$ and thus (i). Now if (ii) fails, we must have $\mathbf{w} \geq 0$. But this means that $\mathbf{x}(t) = \tilde{\mathbf{x}} + t\mathbf{w} \geq 0$ for all $t \geq 0$, and hence all such $\mathbf{x}(t)$ are feasible. The value of the objective function for $\mathbf{x}(t)$ is $\mathbf{c}^T\mathbf{x}(t) = \mathbf{c}^T\tilde{\mathbf{x}} + t\mathbf{c}^T\mathbf{w}$, and it tends to infinity as $t \to \infty$. Hence the linear program is unbounded. This concludes the proof. □

4.3 ABC of Convexity and Convex Polyhedra

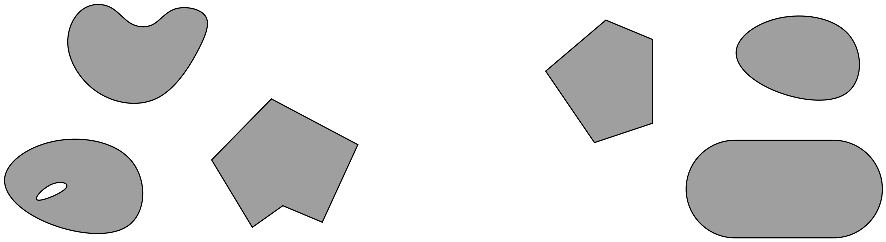
Convexity is one of the fundamental notions in all mathematics, and in the theory of linear programming it is encountered very naturally. Here we recall the definition and present some of the most basic notions and results, which, at the very least, help in gaining a better intuition about linear programming.

On the other hand, linear programming can be presented without these notions, and in concise courses there is usually no time for such material. Accordingly, this section and the next are meant as extending material, and the rest of the book should mostly be accessible without them.

> A set $X \subseteq \mathbb{R}^n$ is **convex** if for every two points $\mathbf{x}, \mathbf{y} \in X$ it also contains the segment $\mathbf{xy}$. Expressed differently, for every $\mathbf{x}, \mathbf{y} \in X$ and every $t \in [0, 1]$ we have $t\mathbf{x} + (1-t)\mathbf{y} \in X$.

A word of explanation might be in order: $t\mathbf{x} + (1-t)\mathbf{y}$ is the point on the segment $\mathbf{xy}$ at distance t from $\mathbf{y}$ and distance $1-t$ from $\mathbf{x}$, if we take the length of the segment as unit distance.

Here are a few examples of convex and nonconvex sets in the plane:



nonconvex convex

The convex set at the bottom right in this picture, a stadium, is worth remembering, since often it is a counterexample to statements about convex sets that may look obvious at first sight but are false.

In calculus one works mainly with convex *functions*. Both notions, convex sets and convex functions, are closely related: For instance, a real function $f\colon \mathbb{R} \to \mathbb{R}$ is convex if and only if its epigraph, i.e., the set $\{(x, y) \in \mathbb{R}^2 : y \geq f(x)\}$, is a convex set in the plane. In general, a function $f : X \to \mathbb{R}$ is called convex if for every $\mathbf{x}, \mathbf{y} \in X$ and every $t \in [0, 1]$ we have

$$f(t\mathbf{x} + (1-t)\mathbf{y}) \leq tf(\mathbf{x}) + (1-t)f(\mathbf{y}).$$

The function is called *strictly convex* if the inequality is strict for all $\mathbf{x} \neq \mathbf{y}$.

Convex hull and convex combinations. It is easily seen that the intersection of an arbitrary collection of convex sets is again a convex set. This allows us to define the convex hull.

Let $X \subset \mathbb{R}^n$ be a set. The **convex hull** of X is the intersection of all convex sets that contain X. Thus it is the smallest convex set containing X, in the sense that any convex set containing X also contains its convex hull.

the convex hull of X

This is not a very constructive definition. The convex hull can also be described using convex combinations, in a way similar to the description of the linear span of a set of vectors using linear combinations. Let $\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_m$ be points in $\mathbb{R}^n$. Every point of the form

$$t_1\mathbf{x}_1 + t_2\mathbf{x}_2 + \cdots + t_m\mathbf{x}_m, \text{ where } t_1, t_2, \ldots, t_m \geq 0 \text{ and } \sum_{i=1}^{m} t_i = 1,$$

is called a **convex combination** of $\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_m$. A convex combination is thus a particular kind of a linear combination, in which the coefficients are nonnegative and sum to 1.

Convex combinations of two points $\mathbf{x}$ and $\mathbf{y}$ are of the form $t\mathbf{x}+(1-t)\mathbf{y}$, $t \in [0,1]$, and as we said after the definition of a convex set, they fill exactly the segment $\mathbf{xy}$. It is easy but instructive to show that all convex combinations of three points $\mathbf{x}, \mathbf{y}, \mathbf{z}$ fill exactly the triangle $\mathbf{xyz}$ (unless the points are collinear, that is).

4.3.1 Lemma. *The convex hull C of a set $X \subseteq \mathbb{R}^n$ equals the set*

$$\tilde{C} = \left\{ \sum_{i=1}^{m} t_i \mathbf{x}_i : m \geq 1, \mathbf{x}_1, \ldots, \mathbf{x}_m \in X, t_1, \ldots, t_m \geq 0, \sum_{i=1}^{m} t_i = 1 \right\}$$

of all convex combinations of finitely many points of X.

Proof. First we prove by induction on m that each convex combination has to lie in the convex hull C. For $m = 1$ it is obvious and for $m = 2$ it follows directly from the convexity of C.

Let $m \geq 3$ and let $\mathbf{x} = t_1\mathbf{x}_1 + \cdots + t_m\mathbf{x}_m$ be a convex combination of points of X. If $t_m = 1$, then we have $\mathbf{x} = \mathbf{x}_m \in C$. For $t_m < 1$ let us put $t'_i = t_i/(1-t_m)$, $i = 1, 2, \ldots, m-1$. Then $\mathbf{x}' = t'_1\mathbf{x}_1 + \cdots + t'_{m-1}\mathbf{x}_{m-1}$ is a convex combination of the points $\mathbf{x}_1, \ldots, \mathbf{x}_{m-1}$ (the t'_i sum to 1), and by the inductive hypothesis $\mathbf{x}' \in C$. So $\mathbf{x} = (1-t_m)\mathbf{x}' + t_m\mathbf{x}_m$ is a convex combination of two points of the (convex) set C and as such it also lies in C.

We have thus proved $\tilde{C} \subseteq C$. For the reverse inclusion it suffices to prove that $\tilde{C}$ is convex, that is, to verify that whenever $\mathbf{x}, \mathbf{y} \in \tilde{C}$ are two convex combinations and $t \in (0,1)$, then $t\mathbf{x} + (1-t)\mathbf{y}$ is again a convex combination. This is straightforward and we take the liberty of omitting further details. $\square$

Convex sets encountered in the theory of linear programming are of a special type and they are called convex polyhedra.

Hyperplanes, half-spaces, polyhedra. We recall that a **hyperplane** in $\mathbb{R}^n$ is an affine subspace of dimension $n-1$. In other words, it is the set of all solutions of a single linear equation of the form

$$a_1x_1 + a_2x_2 + \cdots + a_nx_n = b,$$

where $a_1, a_2, \ldots, a_n$ are not all 0. Hyperplanes in $\mathbb{R}^2$ are lines and hyperplanes in $\mathbb{R}^3$ are ordinary planes.

A hyperplane divides $\mathbb{R}^n$ into two half-spaces and it constitutes their common boundary. For the hyperplane with equation $a_1x_1 + a_2x_2 + \cdots + a_nx_n = b$, the two half-spaces have the following analytic expression:

$$\left\{ \mathbf{x} \in \mathbb{R}^n : a_1x_1 + a_2x_2 + \cdots + a_nx_n \leq b \right\}$$

and

$$\left\{ \mathbf{x} \in \mathbb{R}^n : a_1x_1 + a_2x_2 + \cdots + a_nx_n \geq b \right\}.$$

More exactly, these are **closed half-spaces** that contain their boundary.

A **convex polyhedron** is an intersection of finitely many closed half-spaces in $\mathbb{R}^n$.

A half-space is obviously convex, and hence an intersection of half-spaces is convex as well. Thus convex polyhedra bear the attribute convex by right.

A disk in the plane is a convex set, but it is not a convex polyhedron (because, roughly speaking, a convex polyhedron has to be "edgy"... but try proving this formally).

A half-space is the set of all solutions of a single linear inequality (with at least one nonzero coefficient of some variable x_j). The set of all solutions of a system of finitely many linear inequalities, a.k.a. the set of all feasible solutions of a linear program, is geometrically the intersection of finitely many half-spaces, alias a convex polyhedron. (We should perhaps also mention that a hyperplane is the intersection of two half-spaces, and so the constraints can be both inequalities and equations.)

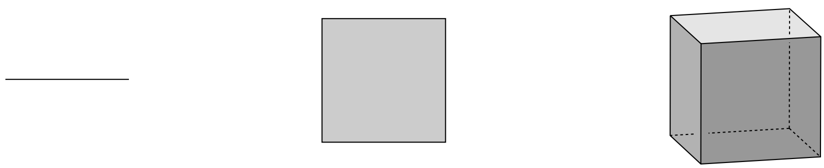
Let us note that a convex polyhedron can be unbounded, since, for example, a single half-space is also a convex polyhedron. A bounded convex polyhedron, i.e. one that can be placed inside some large enough ball, is called a *convex polytope*.

The **dimension** of a convex polyhedron $P \subseteq \mathbb{R}^n$ is the smallest dimension of an affine subspace containing P. Equivalently, it is the largest d for which P contains points $\mathbf{x}_0, \mathbf{x}_1, \ldots, \mathbf{x}_d$ such that the d-tuple of vectors $(\mathbf{x}_1 - \mathbf{x}_0, \mathbf{x}_2 - \mathbf{x}_0, \ldots, \mathbf{x}_d - \mathbf{x}_0)$ is linearly independent.

The empty set is also a convex polyhedron, and its dimension is usually defined as -1.

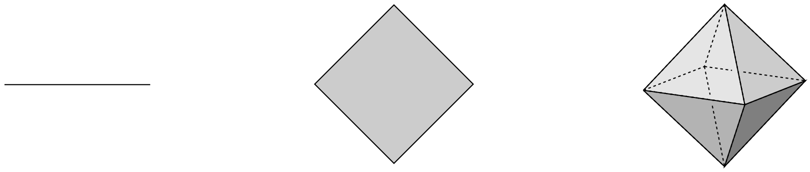
All convex polygons in the plane are two-dimensional convex polyhedra. Several types of three-dimensional convex polyhedra are taught at high schools and decorate mathematical cabinets, such as cubes, boxes, pyramids, or even regular dodecahedra, which can also be met as desktop calendars. Simple examples of convex polyhedra of an arbitrary dimension n are:

- The n-dimensional **cube** $[-1,1]^n$, which can be written as the intersection of $2n$ half-spaces (which ones?):



$n = 1$ $n = 2$ $n = 3$

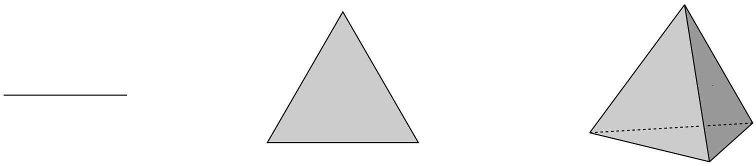
- the n-dimensional **crosspolytope** $\{\mathbf{x} \in \mathbb{R}^n : |x_1| + |x_2| + \cdots + |x_n| \leq 1\}$:



$n = 1$ $\qquad$ $n = 2$ $\qquad$ $n = 3$

For $n = 3$ we get the regular octahedron. For expressing the n-dimensional crosspolytope as an intersection of half-spaces we need 2^n half-spaces (can you find them?).

- The regular n-dimensional **simplex**



$n = 1$ $\qquad$ $n = 2$ $\qquad$ $n = 3$

can be defined in a quite simple and nice way as a subset of $\mathbb{R}^{n+1}$:

$$\{\mathbf{x} \in \mathbb{R}^{n+1} : x_1, x_2, \ldots, x_{n+1} \geq 0,\ x_1 + x_2 + \cdots + x_{n+1} = 1\}.$$

We note that this is exactly the set of all feasible solutions of the linear program with the single equation $x_1 + x_2 + \cdots + x_{n+1} = 1$ and nonnegativity constraints;[4] see the picture in Section 4.1. In general, any n-dimensional convex polytope bounded by $n+1$ hyperplanes is called a simplex.

Many interesting examples of convex polyhedra are obtained as sets of feasible solutions of natural linear programs. For example, the LP relaxation of the problem of maximum-weight matching (Section 3.2) for a complete bipartite graph leads to the *Birkhoff polytope*. Geometric properties of such polyhedra

[4] On the other hand, the set of feasible solutions of a linear program in equational form certainly isn't always a simplex! The simplex method is so named for a rather complicated reason, related to an alternative geometric view of a linear program in equational form, different from the one discussed in this book. According to this view, the m-tuple of numbers in the jth *column* of the matrix A together with the number c_j is interpreted as a point in $\mathbb{R}^{m+1}$. Then the simplex method can be interpreted as a walk through certain simplices with vertices at these points. It was this view that gave Dantzig faith in the simplex method and convinced him that it made sense to study it.

are often related to properties of combinatorial objects and to solutions of combinatorial optimization problems in an interesting way. A nice book about convex polyhedra is

> G. M. Ziegler: *Lectures on Polytopes*, Springer-Verlag, Heidelberg, 1994 (corrected 2nd edition 1998).

The book

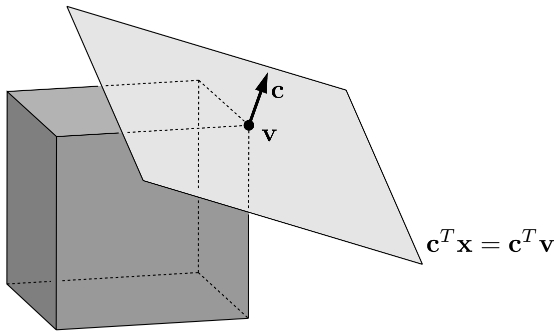
> B. Grünbaum: *Convex Polytopes*, second edition prepared by Volker Kaibel, Victor Klee, and Günter Ziegler, Springer-Verlag, Heidelberg, 2003

is a new edition of a 1967 classics, with extensive updates on the material covered in the original book.

4.4 Vertices and Basic Feasible Solutions

A vertex of a convex polyhedron can be thought of as a "tip" or "spike." For instance, a three-dimensional cube has 8 vertices, and a regular octahedron has 6 vertices.

Mathematically, a vertex is defined as a point where some linear function attains a unique maximum. Thus a point $\mathbf{v}$ is called a **vertex** of a convex polyhedron $P \subset \mathbb{R}^n$ if $\mathbf{v} \in P$ and there exists a nonzero vector $\mathbf{c} \in \mathbb{R}^n$ such that $\mathbf{c}^T\mathbf{v} > \mathbf{c}^T\mathbf{y}$ for all $\mathbf{y} \in P \setminus \{\mathbf{v}\}$. Geometrically it means that the hyperplane $\{\mathbf{x} \in \mathbb{R}^n : \mathbf{c}^T\mathbf{x} = \mathbf{c}^T\mathbf{v}\}$ touches the polyhedron P exactly at $\mathbf{v}$.



Three-dimensional polyhedra have not only vertices, but also edges and faces. A general polyhedron $P \subseteq \mathbb{R}^n$ of dimension n can have vertices, edges, 2-dimensional faces, 3-dimensional faces, up to $(n-1)$-dimensional faces. They are defined as follows: A subset $F \subseteq P$ is a ***k*-dimensional face** of a convex polyhedron P if F has dimension k and there exist a nonzero vector $\mathbf{c} \in \mathbb{R}^n$ and a number $z \in \mathbb{R}$ such that

$\mathbf{c}^T\mathbf{x} = z$ for all $\mathbf{x} \in F$ and $\mathbf{c}^T\mathbf{x} < z$ for all $\mathbf{x} \in P \setminus F$. In other words, there exists a hyperplane that touches P exactly at F. Since such an F is the intersection of a hyperplane with a convex polyhedron, it is a convex polyhedron itself, and its dimension is thus well defined. An **edge** is a 1-dimensional face and a vertex is a 0-dimensional face.

Now we prove that vertices of a convex polyhedron and basic feasible solutions of a linear program are the same concept.

4.4.1 Theorem. *Let P be the set of all feasible solutions of a linear program in equational form (so P is a convex polyhedron). Then the following two conditions for a point $\mathbf{v} \in P$ are equivalent:*

(i) $\mathbf{v}$ *is a vertex of the polyhedron P.*
(ii) $\mathbf{v}$ *is a basic feasible solution of the linear program.*

Proof. The implication (i)$\Rightarrow$(ii) follows immediately from Theorem 4.2.3 (with $\mathbf{c}$ being the vector defining $\mathbf{v}$). It remains to prove (ii)$\Rightarrow$(i).

Let us consider a basic feasible solution $\mathbf{v}$ with a feasible basis B, and let us define a vector $\tilde{\mathbf{c}} \in \mathbb{R}^n$ by $\tilde{c}_j = 0$ for $j \in B$ and $\tilde{c}_j = -1$ otherwise. We have $\tilde{\mathbf{c}}^T\mathbf{v} = 0$, and $\tilde{\mathbf{c}}^T\mathbf{x} \leq 0$ for any $\mathbf{x} \geq \mathbf{0}$, and hence $\mathbf{v}$ maximizes the objective function $\tilde{\mathbf{c}}^T\mathbf{x}$. Moreover, $\tilde{\mathbf{c}}^T\mathbf{x} < 0$ whenever $\mathbf{x}$ has a nonzero component outside B. But by Proposition 4.2.2, $\mathbf{v}$ is the *only* feasible solution with all nonzero components in B, and therefore $\mathbf{v}$ is the only point of P maximizing $\tilde{\mathbf{c}}^T\mathbf{x}$. $\square$

Basic feasible solutions for arbitrary linear programs. A similar theorem is valid for an arbitrary linear program, not only for one in equational form. We will not prove it here, but we at least say what a basic feasible solution is for a general linear program:

4.4.2 Definition. *A* **basic feasible solution** *of a linear program with n variables is a feasible solution for which some n linearly independent constraints hold with equality.*

A constraint that is an equation always has to be satisfied with equality, while an inequality constraint may be satisfied either with equality or with a strict inequality. The nonnegativity constraints satisfied with equality are also counted. The linear independence of constraints means that the vectors of the coefficients of the variables are linearly independent. For example, for $n = 4$, the constraint $3x_1 + 5x_3 - 7x_4 \leq 10$ has the corresponding vector $(3, 0, 5, -7)$.

As is known from linear algebra, a system of n linearly independent linear equations in n variables has exactly one solution. Hence, if $\mathbf{x}$ is a basic feasible solution and it satisfies some n linearly independent constraints with equality, then it is the only point in $\mathbb{R}^n$ that satisfies these n constraints with equality. Geometrically speaking, the constraints satisfied with equality determine hyperplanes, $\mathbf{x}$ lies on some n of them, and these n hyperplanes meet in a single point.

The definition of a basic feasible solution for the equational form looks quite different, but in fact, it is a special case of the new definition, as we now indicate. For a linear program in equational form we have m linearly independent equations always satisfied with equality, and so it remains to satisfy with equality some $n-m$ of the nonnegativity constraints, and these must be linearly independent with the equations. The coefficient vector of the nonnegativity constraint $x_j \geq 0$ is $\mathbf{e}_j$, with 1 at position j and with zeros elsewhere. If $\mathbf{x}$ is a basic feasible solution according to the new definition, then there exists a set $N \subseteq \{1, 2, \ldots, n\}$ of size $n-m$ such that $x_j = 0$ for all $j \in N$ and the rows of the matrix A together with the vectors $(\mathbf{e}_j : j \in N)$ constitute a linearly independent collection. This happens exactly if the matrix A_B has linearly independent rows, where $B = \{1, 2, \ldots, n\} \setminus N$, and we are back at the definition of a basic feasible solution for the equational form.

For a general linear program none of the optimal solutions have to be basic, as is illustrated by the linear program

$$\text{maximize } x_1 + x_2 \text{ subject to } x_1 + x_2 \leq 1.$$

This contrasts with the situation for the equational form (cf. Theorem 4.2.3) and it is one of the advantages of the equational form.

Vertices and extremal points. The intuitive notion of a "tip" of a convex set can be viewed mathematically in at least two ways. One of them is captured by the above definition of a vertex of a convex polyhedron: A tip is a point for which some linear function attains a unique maximum. The other one leads to a definition talking about points that cannot be "generated by segments." These are called **extremal points**; thus a point $\mathbf{x}$ is an extremal point of a convex set $C \subseteq \mathbb{R}^n$ if $\mathbf{x} \in C$ and there are no two points $\mathbf{y}, \mathbf{z} \in C$ different from $\mathbf{x}$ such that $\mathbf{x}$ lies on the segment $\mathbf{yz}$.

For a convex polyhedron it is not difficult to show that the extremal points are exactly the vertices. Hence we have yet another equivalent description of a basic feasible solution.

A convex polytope is the convex hull of its vertices. A general convex polyhedron need not have any vertices at all—consider a half-space. However, a convex *polytope* P, i.e., a bounded convex polyhedron, always has vertices, and even more is true: P equals the convex hull of the set of its vertices. This may look intuitively obvious from examples in dimensions 2 and 3, but a proof is nontrivial (Ziegler's book cited in the previous section calls this the "Main Theorem" of polytope theory). Consequently, every convex polytope can be represented either as the intersection of finitely many half-spaces or as the convex hull of finitely many points.