Chapter 4

Convex Hull

There exists an incredible variety of point sets and polygons. Among them, some have certain properties that make them "nicer" than others in some respect. For instance, look at the two polygons shown below.

(a) A convex polygon. **(b)** A non-convex polygon.

Figure 4.1: *Examples of polygons: Which do you like better?*

As it is hard to argue about aesthetics, let us take a more algorithmic stance. When designing algorithms, the polygon shown on the left appears much easier to deal with than the visually and geometrically more complex polygon shown on the right. One particular property that makes the left polygon nice is that one can walk between any two vertices along a straight line without ever leaving the polygon. In fact, this statement holds true not only for vertices but for any two points within the polygon. A polygon or, more generally, a set with this property is called *convex*.

Definition 4.1 *A set* $P \subseteq \mathbb{R}^d$ *is* **convex** *if* $\overline{pq} \subseteq P$*, for any* $p, q \in P$.

An alternative, equivalent way to phrase convexity would be to demand that for every line $\ell \subset \mathbb{R}^d$ the intersection $\ell \cap P$ be connected. The polygon shown in Figure 4.1b is not convex because there are some pairs of points for which the connecting line segment is not completely contained within the polygon. An immediate consequence of the definition is the following

Observation 4.2 *For any family $(P_i)_{i\in I}$ of convex sets, the intersection $\bigcap_{i\in I} P_i$ is convex.*

Indeed there are many problems that are comparatively easy to solve for convex sets but very hard in general. We will encounter some particular instances of this phenomenon later in the course. However, not all polygons are convex and a discrete set of points is never convex, unless it consists of at most one point only. In such a case it is useful to make a given set P convex, that is, approximate P with or, rather, encompass P within a convex set $H \supseteq P$. Ideally, H differs from P as little as possible, that is, we want H to be a smallest convex set enclosing P.

At this point let us step back for a second and ask ourselves whether this wish makes sense at all: Does such a set H (always) exist? Fortunately, we are on the safe side because the whole space $\mathbb{R}^d$ is certainly convex. It is less obvious, but we will see below that H is actually unique. Therefore it is legitimate to refer to H as **the** smallest convex set enclosing P or—shortly—the *convex hull* of P.

4.1 Convexity

In this section we will derive an algebraic characterization of convexity. Such a characterization allows to investigate convexity using the machinery from linear algebra.

Consider $P \subset \mathbb{R}^d$. From linear algebra courses you should know that the *linear hull*

$$\operatorname{lin}(P) := \left\{ q \;\middle|\; q = \sum \lambda_i p_i \;\wedge\; \forall i : p_i \in P, \lambda_i \in \mathbb{R} \right\}$$

is the set of all *linear combinations* of P (smallest linear subspace containing P). For instance, if $P = \{p\} \subset \mathbb{R}^2 \setminus \{0\}$ then $\operatorname{lin}(P)$ is the line through p and the origin.

Similarly, the *affine hull*

$$\operatorname{aff}(P) := \left\{ q \;\middle|\; q = \sum \lambda_i p_i \;\wedge\; \sum \lambda_i = 1 \;\wedge\; \forall i : p_i \in P, \lambda_i \in \mathbb{R} \right\}$$

is the set of all *affine combinations* of P (smallest affine subspace containing P). For instance, if $P = \{p, q\} \subset \mathbb{R}^2$ and $p \neq q$ then $\operatorname{aff}(P)$ is the line through p and q.

It turns out that convexity can be described in a very similar way algebraically, which leads to the notion of *convex combinations*.

Proposition 4.3 *A set $P \subseteq \mathbb{R}^d$ is convex if and only if $\sum_{i=1}^n \lambda_i p_i \in P$, for all $n \in \mathbb{N}$, $p_1, \ldots, p_n \in P$, and $\lambda_1, \ldots, \lambda_n \geqslant 0$ with $\sum_{i=1}^n \lambda_i = 1$.*

Proof. "$\Leftarrow$": obvious with $n = 2$.

"$\Rightarrow$": Induction on n. For $n = 1$ the statement is trivial. For $n \geqslant 2$, let $p_i \in P$ and $\lambda_i \geqslant 0$, for $1 \leqslant i \leqslant n$, and assume $\sum_{i=1}^n \lambda_i = 1$. We may suppose that $\lambda_i > 0$, for all i. (Simply omit those points whose coefficient is zero.) We need to show that $\sum_{i=1}^n \lambda_i p_i \in P$.

Define $\lambda = \sum_{i=1}^{n-1} \lambda_i$ and for $1 \leqslant i \leqslant n-1$ set $\mu_i = \lambda_i/\lambda$. Observe that $\mu_i \geqslant 0$ and $\sum_{i=1}^{n-1} \mu_i = 1$. By the inductive hypothesis, $q := \sum_{i=1}^{n-1} \mu_i p_i \in P$, and thus by convexity of P also $\lambda q + (1-\lambda)p_n \in P$. We conclude by noting that $\lambda q + (1-\lambda)p_n = \lambda \sum_{i=1}^{n-1} \mu_i p_i + \lambda_n p_n = \sum_{i=1}^{n} \lambda_i p_i$. □

Definition 4.4 *The* **convex hull** $\text{conv}(P)$ *of a set* $P \subseteq \mathbb{R}^d$ *is the intersection of all convex supersets of* P*.*

At first glance this definition is a bit scary: There may be a whole lot of supersets for any given P and it not clear that taking the intersection of all of them yields something sensible to work with. However, by Observation 4.2 we know that the resulting set is convex, at least. The missing bit is provided by the following proposition, which characterizes the convex hull in terms of exactly those convex combinations that appeared in Proposition 4.3 already.

Proposition 4.5 *For any* $P \subseteq \mathbb{R}^d$ *we have*

$$\text{conv}(P) = \left\{ \sum_{i=1}^{n} \lambda_i p_i \;\middle|\; n \in \mathbb{N} \wedge \sum_{i=1}^{n} \lambda_i = 1 \wedge \forall i \in \{1,\ldots,n\} : \lambda_i \geqslant 0 \wedge p_i \in P \right\}.$$

The elements of the set on the right hand side are referred to as *convex combinations* of P.

Proof. "$\supseteq$": Consider a convex set $C \supseteq P$. By Proposition 4.3 (only-if direction) the right hand side is contained in C. As C was arbitrary, the claim follows.

"$\subseteq$": Denote the set on the right hand side by R. Clearly $R \supseteq P$. We show that R forms a convex set. Let $p = \sum_{i=1}^{n} \lambda_i p_i$ and $q = \sum_{i=1}^{n} \mu_i p_i$ be two convex combinations. (We may suppose that both p and q are expressed over the same p_i by possibly adding some terms with a coefficient of zero.)

Then for $\lambda \in [0,1]$ we have $\lambda p + (1-\lambda)q = \sum_{i=1}^{n}(\lambda\lambda_i + (1-\lambda)\mu_i)p_i \in R$, as $\underbrace{\lambda\lambda_i}_{\geqslant 0} + \underbrace{(1-\lambda)}_{\geqslant 0}\underbrace{\mu_i}_{\geqslant 0} \geqslant 0$, for all $1 \leqslant i \leqslant n$, and $\sum_{i=1}^{n}(\lambda\lambda_i + (1-\lambda)\mu_i) = \lambda + (1-\lambda) = 1$. □

In linear algebra the notion of a basis in a vector space plays a fundamental role. In a similar way we want to describe convex sets using as few entities as possible, which leads to the notion of extremal points, as defined below.

Definition 4.6 *The convex hull of a finite point set* $P \subset \mathbb{R}^d$ *forms a* **convex polytope***. Each* $p \in P$ *for which* $p \notin \text{conv}(P \setminus \{p\})$ *is called a* **vertex** *of* $\text{conv}(P)$*. A vertex of* $\text{conv}(P)$ *is also called an* **extremal point** *of* P*. A convex polytope in* $\mathbb{R}^2$ *is called a* **convex polygon***.*

Essentially, the following proposition shows that the term vertex above is well defined.

Proposition 4.7 *A convex polytope in* $\mathbb{R}^d$ *is the convex hull of its vertices.*

Proof. Let $P = \{p_1, \ldots, p_n\}$, $n \in \mathbb{N}$, such that without loss of generality $p_1, \ldots, p_k$ are the vertices of $\mathcal{P} := \text{conv}(P)$. We prove by induction on n that $\text{conv}(p_1, \ldots, p_n) \subseteq \text{conv}(p_1, \ldots, p_k)$. For $n = k$ the statement is trivial.

For $n > k$, p_n is not a vertex of $\mathcal{P}$ and hence p_n can be expressed as a convex combination $p_n = \sum_{i=1}^{n-1} \lambda_i p_i$. Thus for any $x \in \mathcal{P}$ we can write $x = \sum_{i=1}^{n} \mu_i p_i = \sum_{i=1}^{n-1} \mu_i p_i + \mu_n \sum_{i=1}^{n-1} \lambda_i p_i = \sum_{i=1}^{n-1} (\mu_i + \mu_n \lambda_i) p_i$. As $\sum_{i=1}^{n-1} (\mu_i + \mu_n \lambda_i) = 1$, we conclude inductively that $x \in \text{conv}(p_1, \ldots, p_{n-1}) \subseteq \text{conv}(p_1, \ldots, p_k)$. □

4.2 Classical Theorems for Convex Sets

Next we will discuss a few fundamental theorems about convex sets in $\mathbb{R}^d$. The proofs typically use the algebraic characterization of convexity and then employ some techniques from linear algebra.

Theorem 4.8 (Radon [8]) *Any set $P \subset \mathbb{R}^d$ of $d + 2$ points can be partitioned into two disjoint subsets P_1 and P_2 such that $\text{conv}(P_1) \cap \text{conv}(P_2) \neq \emptyset$.*

Proof. Let $P = \{p_1, \ldots, p_{d+2}\}$. No more than $d + 1$ points can be affinely independent in $\mathbb{R}^d$. Hence suppose without loss of generality that p_{d+2} can be expressed as an affine combination of $p_1, \ldots, p_{d+1}$, that is, there exist $\lambda_1, \ldots, \lambda_{d+1} \in \mathbb{R}$ with $\sum_{i=1}^{d+1} \lambda_i = 1$ and $\sum_{i=1}^{d+1} \lambda_i p_i = p_{d+2}$. Let P_1 be the set of all points p_i for which λ_i is positive and let $P_2 = P \setminus P_1$. Then setting $\lambda_{d+2} = -1$ we can write $\sum_{p_i \in P_1} \lambda_i p_i = \sum_{p_i \in P_2} -\lambda_i p_i$, where all coefficients on both sides are non-negative. Renormalizing by $\mu_i = \lambda_i/\mu$ and $\nu_i = \lambda_i/\nu$, where $\mu = \sum_{p_i \in P_1} \lambda_i$ and $\nu = -\sum_{p_i \in P_2} \lambda_i$, yields convex combinations $\sum_{p_i \in P_1} \mu_i p_i = \sum_{p_i \in P_2} \nu_i p_i$ that describe a common point of $\text{conv}(P_1)$ and $\text{conv}(P_2)$. □

Theorem 4.9 (Helly) *Consider a collection $\mathcal{C} = \{C_1, \ldots, C_n\}$ of $n \geqslant d+1$ convex subsets of $\mathbb{R}^d$, such that any $d+1$ pairwise distinct sets from $\mathcal{C}$ have non-empty intersection. Then also the intersection $\bigcap_{i=1}^{n} C_i$ of all sets from $\mathcal{C}$ is non-empty.*

Proof. Induction on n. The base case $n = d + 1$ holds by assumption. Hence suppose that $n \geqslant d + 2$. Consider the sets $D_i = \bigcap_{j \neq i} C_j$, for $i \in \{1, \ldots, n\}$. As D_i is an intersection of $n - 1$ sets from $\mathcal{C}$, by the inductive hypothesis we know that $D_i \neq \emptyset$. Therefore we can find some point $p_i \in D_i$, for each $i \in \{1, \ldots, n\}$. Now by Theorem 4.8 the set $P = \{p_1, \ldots, p_n\}$ can be partitioned into two disjoint subsets P_1 and P_2 such that $\text{conv}(P_1) \cap \text{conv}(P_2) \neq \emptyset$. We claim that any point $p \in \text{conv}(P_1) \cap \text{conv}(P_2)$ also lies in $\bigcap_{i=1}^{n} C_i$, which completes the proof.

Consider some C_i, for $i \in \{1, \ldots, n\}$. By construction $D_j \subseteq C_i$, for $j \neq i$. Thus p_i is the only point from P that may not be in C_i. As p_i is part of only one of P_1 or P_2, say, of P_1, we have $P_2 \subseteq C_i$. The convexity of C_i implies $\text{conv}(P_2) \subseteq C_i$ and, therefore, $p \in C_i$. □