3
Properties of Convex Sets: A Glimpse

3.1 Convex Sets

Convex sets play a very important role in geometry. In this section we state and prove some of the "classics" of convex affine geometry: Carathéodory's theorem, Radon's theorem, and Helly's theorem. These theorems share the property that they are easy to state, but they are deep, and their proof, although rather short, requires a lot of creativity. We will return to convex sets when we study Euclidean geometry.

Given an affine space E, recall that a subset V of E is *convex* if for any two points $a, b \in V$, we have $c \in V$ for every point $c = (1 - \lambda)a + \lambda b$, with $0 \leq \lambda \leq 1$ ($\lambda \in \mathbb{R}$). Given any two points a, b, the notation $[a, b]$ is often used to denote the line segment between a and b, that is,

$$[a, b] = \{c \in E \mid c = (1 - \lambda)a + \lambda b,\ 0 \leq \lambda \leq 1\},$$

and thus a set V is convex if $[a, b] \subseteq V$ for any two points $a, b \in V$ ($a = b$ is allowed). The empty set is trivially convex, every one-point set $\{a\}$ is convex, and the entire affine space E is of course convex.

It is obvious that the intersection of any family (finite or infinite) of convex sets is convex. Then, given any (nonempty) subset S of E, there is a smallest convex set containing S denoted by $\mathcal{C}(S)$ and called the *convex hull of S* (namely, the intersection of all convex sets containing S).

A good understanding of what $\mathcal{C}(S)$ is, and good methods for computing it, are essential. First, we have the following simple but crucial lemma analogous to Lemma 2.5.3.

Lemma 3.1.1 *Given an affine space $\langle E, \overrightarrow{E}, +\rangle$, for any family $(a_i)_{i\in I}$ of points in E, the set V of convex combinations $\sum_{i\in I} \lambda_i a_i$ (where $\sum_{i\in I} \lambda_i = 1$ and $\lambda_i \geq 0$) is the convex hull of $(a_i)_{i\in I}$.*

Proof. If $(a_i)_{i\in I}$ is empty, then $V = \emptyset$, because of the condition $\sum_{i\in I} \lambda_i = 1$. As in the case of affine combinations, it is easily shown by induction that any convex combination can be obtained by computing convex combinations of two points at a time. As a consequence, if $(a_i)_{i\in I}$ is nonempty, then the smallest convex subspace containing $(a_i)_{i\in I}$ must contain the set V of all convex combinations $\sum_{i\in I} \lambda_i a_i$. Thus, it is enough to show that V is closed under convex combinations, which is immediately verified. $\square$

In view of Lemma 3.1.1, it is obvious that any affine subspace of E is convex. Convex sets also arise in terms of hyperplanes. Given a hyperplane H, if $f: E \to \mathbb{R}$ is any nonconstant affine form defining H (i.e., $H = \mathrm{Ker}\, f$), we can define the two subsets

$$H_+(f) = \{a \in E \mid f(a) \geq 0\} \quad \text{and} \quad H_-(f) = \{a \in E \mid f(a) \leq 0\},$$

called *(closed) half-spaces associated with f*.

Observe that if $\lambda > 0$, then $H_+(\lambda f) = H_+(f)$, but if $\lambda < 0$, then $H_+(\lambda f) = H_-(f)$, and similarly for $H_-(\lambda f)$. However, the set

$$\{H_+(f), H_-(f)\}$$

depends only on the hyperplane H, and the choice of a specific f defining H amounts to the choice of one of the two half-spaces. For this reason, we will also say that $H_+(f)$ and $H_-(f)$ are the *closed half-spaces associated with H*. Clearly, $H_+(f) \cup H_-(f) = E$ and $H_+(f) \cap H_-(f) = H$. It is immediately verified that $H_+(f)$ and $H_-(f)$ are convex. Bounded convex sets arising as the intersection of a finite family of half-spaces associated with hyperplanes play a major role in convex geometry and topology (they are called *convex polytopes*).

It is natural to wonder whether Lemma 3.1.1 can be sharpened in two directions: (1) Is it possible have a fixed bound on the number of points involved in the convex combinations? (2) Is it necessary to consider convex combinations of all points, or is it possible to consider only a subset with special properties?

The answer is yes in both cases. In case 1, assuming that the affine space E has dimension m, Carathéodory's theorem asserts that it is enough to consider convex combinations of $m + 1$ points. For example, in the plane $\mathbb{A}^2$, the convex hull of a set S of points is the union of all triangles (interior points included) with vertices in S. In case 2, the theorem of Krein and Milman asserts that a convex set that is also compact is the convex hull of its extremal points (given a convex set S, a point $a \in S$ is extremal if $S - \{a\}$ is also convex, see Berger [13] or Lang [109]). Next, we prove Carathéodory's theorem.

3.2 Carathéodory's Theorem

The proof of Carathéodory's theorem is really beautiful. It proceeds by contradiction and uses a minimality argument.

Theorem 3.2.1 *Given any affine space E of dimension m, for any (nonvoid) family $S = (a_i)_{i\in L}$ in E, the convex hull $\mathcal{C}(S)$ of S is equal to the set of convex combinations of families of $m+1$ points of S.*

Proof. By Lemma 3.1.1,

$$\mathcal{C}(S) = \left\{ \sum_{i\in I} \lambda_i a_i \mid a_i \in S,\ \sum_{i\in I} \lambda_i = 1,\ \lambda_i \geq 0,\ I \subseteq L,\ I \text{ finite} \right\}.$$

We would like to prove that

$$\mathcal{C}(S) = \left\{ \sum_{i\in I} \lambda_i a_i \mid a_i \in S,\ \sum_{i\in I} \lambda_i = 1,\ \lambda_i \geq 0,\ I \subseteq L,\ |I| = m+1 \right\}.$$

We proceed by contradiction. If the theorem is false, there is some point $b \in \mathcal{C}(S)$ such that b can be expressed as a convex combination $b = \sum_{i\in I} \lambda_i a_i$, where $I \subseteq L$ is a finite set of cardinality $|I| = q$ with $q \geq m+2$, and b cannot be expressed as any convex combination $b = \sum_{j\in J} \mu_j a_j$ of strictly fewer than q points in S, that is, where $|J| < q$. Such a point $b \in \mathcal{C}(S)$ is a convex combination

$$b = \lambda_1 a_1 + \cdots + \lambda_q a_q,$$

where $\lambda_1 + \cdots + \lambda_q = 1$ and $\lambda_i > 0$ $(1 \leq i \leq q)$. We shall prove that b can be written as a convex combination of $q-1$ of the a_i. Pick any origin O in E. Since there are $q > m+1$ points $a_1, \ldots, a_q$, these points are affinely dependent, and by Lemma 2.6.5, there is a family $(\mu_1, \ldots, \mu_q)$ all scalars not all null, such that $\mu_1 + \cdots + \mu_q = 0$ and

$$\sum_{i=1}^{q} \mu_i \mathbf{Oa_i} = 0.$$

Consider the set $T \subseteq \mathbb{R}$ defined by

$$T = \{t \in \mathbb{R} \mid \lambda_i + t\mu_i \geq 0,\ \mu_i \neq 0,\ 1 \leq i \leq q\}.$$

The set T is nonempty, since it contains 0. Since $\sum_{i=1}^{q} \mu_i = 0$ and the μ_i are not all null, there are some μ_h, μ_k such that $\mu_h < 0$ and $\mu_k > 0$, which implies that $T = [\alpha, \beta]$, where

$$\alpha = \max_{1\leq i\leq q} \{-\lambda_i/\mu_i \mid \mu_i > 0\} \quad \text{and} \quad \beta = \min_{1\leq i\leq q} \{-\lambda_i/\mu_i \mid \mu_i < 0\}$$

(T is the intersection of the closed half-spaces $\{t \in \mathbb{R} \mid \lambda_i + t\mu_i \geq 0,\ \mu_i \neq 0\}$). Observe that $\alpha < 0 < \beta$, since $\lambda_i > 0$ for all $i = 1, \ldots, q$.

We claim that there is some j $(1 \le j \le q)$ such that

$$\lambda_j + \alpha\mu_j = 0.$$

Indeed, since

$$\alpha = \max_{1\le i\le q} \{-\lambda_i/\mu_i \mid \mu_i > 0\},$$

as the set on the right hand side is finite, the maximum is achieved and there is some index j so that $\alpha = -\lambda_j/\mu_j$. If j is some index such that $\lambda_j + \alpha\mu_j = 0$, since $\sum_{i=1}^{q} \mu_i \mathbf{Oa_i} = 0$, we have

$$\begin{aligned} b = \sum_{i=1}^{q} \lambda_i a_i &= O + \sum_{i=1}^{q} \lambda_i \mathbf{Oa_i} + 0, \\ &= O + \sum_{i=1}^{q} \lambda_i \mathbf{Oa_i} + \alpha\left(\sum_{i=1}^{q} \mu_i \mathbf{Oa_i}\right), \\ &= O + \sum_{i=1}^{q} (\lambda_i + \alpha\mu_i)\mathbf{Oa_i}, \\ &= \sum_{i=1}^{q} (\lambda_i + \alpha\mu_i) a_i, \\ &= \sum_{i=1,\, i\neq j}^{q} (\lambda_i + \alpha\mu_i) a_i, \end{aligned}$$

since $\lambda_j + \alpha\mu_j = 0$. Since $\sum_{i=1}^{q} \mu_i = 0$, $\sum_{i=1}^{q} \lambda_i = 1$, and $\lambda_j + \alpha\mu_j = 0$, we have

$$\sum_{i=1,\, i\neq j}^{q} \lambda_i + \alpha\mu_i = 1,$$

and since $\lambda_i + \alpha\mu_i \ge 0$ for $i = 1, \ldots, q$, the above shows that b can be expressed as a convex combination of $q - 1$ points from S. However, this contradicts the assumption that b cannot be expressed as a convex combination of strictly fewer than q points from S, and the theorem is proved. $\square$

If S is a finite (of infinite) set of points in the affine plane $\mathbb{A}^2$, Theorem 3.2.1 confirms our intuition that $\mathcal{C}(S)$ is the union of triangles (including interior points) whose vertices belong to S. Similarly, the convex hull of a set S of points in $\mathbb{A}^3$ is the union of tetrahedra (including interior points) whose vertices belong to S. We get the feeling that triangulations play a crucial role, which is of course true!

We conclude this short chapter with two other classics of convex geometry.

3.3 Radon's and Helly's Theorems

We begin with *Radon's theorem.*

Theorem 3.3.1 *Given any affine space E of dimension m, for every subset X of E, if X has at least $m+2$ points, then there is a partition of X into two nonempty disjoint subsets X_1 and X_2 such that the convex hulls of X_1 and X_2 have a nonempty intersection.*

Proof. Pick some origin O in E. Write $X = (x_i)_{i\in L}$ for some index set L (we can let $L = X$). Since by assumption $|X| \geq m+2$ where $m = \dim(E)$, X is affinely dependent, and by Lemma 2.6.5, there is a family $(\mu_k)_{k\in L}$ (of finite support) of scalars, not all null, such that

$$\sum_{k\in L} \mu_k = 0 \quad \text{and} \quad \sum_{k\in L} \mu_k \mathbf{Ox_k} = 0.$$

Since $\sum_{k\in L} \mu_k = 0$, the μ_k are not all null, and $(\mu_k)_{k\in L}$ has finite support, the sets

$$I = \{i \in L \mid \mu_i > 0\} \quad \text{and} \quad J = \{j \in L \mid \mu_j < 0\}$$

are nonempty, finite, and obviously disjoint. Let

$$X_1 = \{x_i \in X \mid \mu_i > 0\} \quad \text{and} \quad X_2 = \{x_i \in X \mid \mu_i \leq 0\}.$$

Again, since the μ_k are not all null and $\sum_{k\in L} \mu_k = 0$, the sets X_1 and X_2 are nonempty, and obviously

$$X_1 \cap X_2 = \emptyset \quad \text{and} \quad X_1 \cup X_2 = X.$$

Furthermore, the definition of I and J implies that $(x_i)_{i\in I} \subseteq X_1$ and $(x_j)_{j\in J} \subseteq X_2$. It remains to prove that $\mathcal{C}(X_1) \cap \mathcal{C}(X_2) \neq \emptyset$. The definition of I and J implies that

$$\sum_{k\in L} \mu_k \mathbf{Ox_k} = 0$$

can be written as

$$\sum_{i\in I} \mu_i \mathbf{Ox_i} + \sum_{j\in J} \mu_j \mathbf{Ox_j} = 0,$$

that is, as

$$\sum_{i\in I} \mu_i \mathbf{Ox_i} = \sum_{j\in J} -\mu_j \mathbf{Ox_j},$$

where

$$\sum_{i\in I} \mu_i = \sum_{j\in J} -\mu_j = \mu,$$

with $\mu > 0$. Thus, we have

$$\sum_{i \in I} \frac{\mu_i}{\mu} \mathbf{Ox_i} = \sum_{j \in J} -\frac{\mu_j}{\mu} \mathbf{Ox_j},$$

with

$$\sum_{i \in I} \frac{\mu_i}{\mu} = \sum_{j \in J} -\frac{\mu_j}{\mu} = 1,$$

proving that $\sum_{i \in I} (\mu_i/\mu) x_i \in \mathcal{C}(X_1)$ and $\sum_{j \in J} -(\mu_j/\mu) x_j \in \mathcal{C}(X_2)$ are identical, and thus that $\mathcal{C}(X_1) \cap \mathcal{C}(X_2) \neq \emptyset$. $\square$

Finally, we prove a version of *Helly's theorem*.

Theorem 3.3.2 *Given any affine space E of dimension m, for every family $\{K_1, \ldots, K_n\}$ of n convex subsets of E, if $n \geq m+2$ and the intersection $\bigcap_{i \in I} K_i$ of any $m+1$ of the K_i is nonempty (where $I \subseteq \{1, \ldots, n\}$, $|I| = m+1$), then $\bigcap_{i=1}^{n} K_i$ is nonempty.*

Proof. The proof is by induction on $n \geq m+1$ and uses Radon's theorem in the induction step. For $n = m+1$, the assumption of the theorem is that the intersection of any family of $m+1$ of the K_i's is nonempty, and the theorem holds trivially. Next, let $L = \{1, 2, \ldots, n+1\}$, where $n+1 \geq m+2$. By the induction hypothesis, $C_i = \bigcap_{j \in (L - \{i\})} K_j$ is nonempty for every $i \in L$.

We claim that $C_i \cap C_j \neq \emptyset$ for some $i \neq j$. If so, as $C_i \cap C_j = \bigcap_{k=1}^{n+1} K_k$, we are done. So, let us assume that the C_i's are pairwise disjoint. Then, we can pick a set $X = \{a_1, \ldots, a_{n+1}\}$ such that $a_i \in C_i$, for every $i \in L$. By Radon's Theorem, there are two nonempty disjoint sets $X_1, X_2 \subseteq X$ such that $X = X_1 \cup X_2$ and $\mathcal{C}(X_1) \cap \mathcal{C}(X_2) \neq \emptyset$. However, $X_1 \subseteq K_j$ for every j with $a_j \notin X_1$. This is because $a_j \notin K_j$ for every j, and so, we get

$$X_1 \subseteq \bigcap_{a_j \notin X_1} K_j.$$

Symetrically, we also have

$$X_2 \subseteq \bigcap_{a_j \notin X_2} K_j.$$

Since the K_j's are convex and

$$\left(\bigcap_{a_j \notin X_1} K_j \right) \cap \left(\bigcap_{a_j \notin X_2} K_j \right) = \bigcap_{i=1}^{n+1} K_i,$$

it follows that $\mathcal{C}(X_1) \cap \mathcal{C}(X_2) \subseteq \bigcap_{i=1}^{n+1} K_i$, so that $\bigcap_{i=1}^{n+1} K_i$ is nonempty, contradicting the fact that $C_i \cap C_j = \emptyset$ for all $i \neq j$. $\square$

A more general version of Helly's theorem is proved in Berger [13]. An amusing corollary of Helly's theorem is the following result: Consider $n \geq 4$ parallel line segments in the affine plane $\mathbb{A}^2$. If every three of these line segments meet a line, then all of these line segments meet a common line.

3.4 Problems

Problem 3.1 Let a, b, c, be any distinct points in $\mathbb{A}^3$, and assume that they are not collinear. Let H be the plane of the equation

$$\alpha x + \beta y + \gamma z + \delta = 0.$$

(i) What is the intersection of the plane H and of the solid triangle determined by a, b, c (the convex hull of a, b, c)?

(ii) Give an algorithm to find the intersection of the plane H and the triangle determined by a, b, c.

(iii) (**extra credit**) Implement the above algorithm so that the intersection can be visualized (you may use *Maple*, *Mathematica*, *Matlab*, etc.).

Problem 3.2 Given any two affine spaces E and F, for any affine map $f\colon E \to F$, any convex set U in E, and any convex set V in F, prove that $f(U)$ is convex and that $f^{-1}(V)$ is convex. Recall that

$$f(U) = \{b \in F \mid \exists a \in U,\, b = f(a)\}$$

is the *direct image of* U *under* f, and that

$$f^{-1}(V) = \{a \in E \mid \exists b \in V,\, b = f(a)\}$$

is the *inverse image of* V *under* f.

Problem 3.3 Consider the subset S of $\mathbb{A}^2$ consisting the points belonging to the right branch of the hyperbola of the equation $x^2 - y^2 = 1$, i.e.,

$$S = \{(x, y) \in \mathbb{R}^2 \mid x^2 - y^2 \geq 1,\, x \geq 0\}.$$

Prove that S is convex. What is the convex hull of $S \cup \{(0, 0)\}$? Is the convex hull of a closed subset of $\mathbb{A}^m$ necessarily a closed set?

Problem 3.4 Use the theorem of Carathéodory to prove that if S is a compact subset of $\mathbb{A}^m$, then its convex hull $\mathcal{C}(S)$ is also compact.

Problem 3.5 Let S be any nonempty subset of an affine space E. Given some point $a \in S$, we say that S is *star-shaped with respect to* a iff the line segment $[a, x]$ is contained in S for every $x \in S$, i.e., $(1 - \lambda)a + \lambda x \in S$ for all λ such that $0 \leq \lambda \leq 1$. We say that S is *star-shaped* iff it is star-shaped w.r.t. to some point $a \in S$.

(1) Prove that every nonempty convex set is star-shaped.

(2) Show that there are star-shaped subsets that are not convex. Show that there are nonempty subsets that are not star-shaped (give an example in $\mathbb{A}^n$, $n = 1, 2, 3$).

(3) Given a star-shaped subset S of E, let $N(S)$ be the set of all points $a \in S$ such that S is star-shaped with respect to a. Prove that $N(S)$ is convex.

Problem 3.6 Consider $n \geq 4$ parallel line segments in the affine plane $\mathbb{A}^2$. If every three of these line segments meet a line, then all of these line segments meet a common line.
Hint. Choose a coordinate system such that the y axis is parallel to the common direction of the line segments. For any line segment S, let

$$CS = \{(\alpha, \beta) \in \mathbb{R}^2, \text{ the line } y = \alpha x + \beta \text{ meets } S\}.$$

Show that CS is convex and apply Helly's theorem.

Problem 3.7 Given any two convex sets S and T in the affine space $\mathbb{A}^m$, and given $\lambda, \mu \in \mathbb{R}$ such that $\lambda + \mu = 1$, the *Minkowski sum* $\lambda S + \mu T$ is the set

$$\lambda S + \mu T = \{\lambda p + \mu q \mid p \in S, q \in T\}.$$

(i) Prove that $\lambda S + \mu T$ is convex. Draw some Minkowski sums, in particular when S and T are tetrahedra (with T upside down).

(ii) Show that the Minkowski sum does not preserve the center of gravity.