Convexity

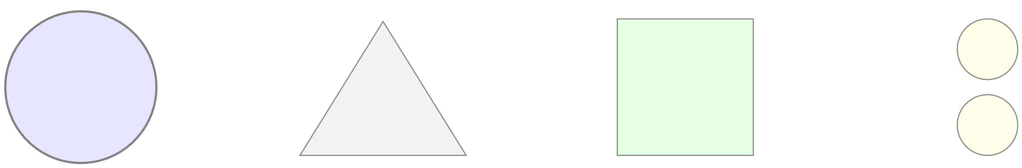
Geometric Methods in Computer Science

Instructor: Yury Makarychev

1 Convexity

Definition 1.1. *Let V be a linear (vector) space. A set $S \subseteq V$ is convex if for every two points $x, y \in S$, the segment $[x, y] \equiv \{\lambda x + (1-\lambda)y : \lambda \in [0,1]\}$ lies in S.*

Consider various examples: a circle, triangle, square, pair of circles. Are these sets convex?



Claim 1.2. *Let $\{S_\alpha\}_\alpha$ be a family of convex sets. Then their intersection $T = \bigcap_\alpha S_\alpha$ is also a convex set.*

Proof. Consider two points $x, y \in T$. We have $x, y \in S_\alpha$ for every index α. Since each set S_α is convex, $[x, y] \subseteq S_\alpha$ for every α. Therefore, $[x, y] \subseteq T$. We conclude that T is also convex. $\square$

Exercise 1. *Assume that S and T are convex. Can $S \cup T$ be convex? Is it necessarily true that $S \cup T$ is convex? Can the complement of S be convex? Is it necessarily true that the complement of S is convex?*

Exercise 2. *Assume that S is convex. Is it necessarily connected?*

2 Convex combinations

Definition 2.1. *Consider a set of points $v_1, \dots, v_n \in V$ and a set of non-negative weights $\lambda_1, \dots, \lambda_n$ that add up to 1: $\sum_{i=1}^n \lambda_i = 1$. Then $\sum_{i=1}^n \lambda_i$ is a convex combination of points v_i with weights λ_i.*

Note that we consider only finite convex combinations in Definition 2.1. The definition of convexity can be restated in terms of convex combinations: S is convex if and only if for every $x, y \in S$ every convex combination $\lambda_1 x + \lambda_2 y \in S$. In this definition, we consider only convex combinations involving two points. Can we consider arbitrary convex combinations instead? Obviously if every convex combination of points in S is in S, then so is every combination of two points, and thus S is convex. Now we show that if S is convex then all convex combinations of points from S are in S.

Claim 2.2. *Consider a convex set S. Let $u = \sum \lambda_i v_i$ be a convex combination of points $v_1, \dots, v_n$ in S. Then $u \in S$.*

Proof. We prove the claim by induction on n. For $n = 1$, the claim is trivial, as $u = v_1 \in S$. Assuming that the claim holds for $n - 1$ points, we prove it for n points.

Let $\mu_i = \lambda_i/(\lambda_1 + \dots + \lambda_{n-1}) = \lambda_i/(1 - \lambda_n)$ for $i \in [n-1] \equiv \{1, \dots, n-1\}$. Note that $\sum_{i=1}^{n-1} \mu_i = 1$ and all $\mu_i \geq 0$. Define

$$u' = \sum_{i=1}^{n-1} \mu_i v_i.$$

Point u' is a convex combination of $n-1$ points in S and thus belongs to S, by the induction hypothesis. Since both points u' and v_n are in S and S is convex, the entire segment $[u', v_n]$ lies in S. We conclude that $u = (1 - \lambda_n)u' + \lambda_n v_n \in [u', v_n] \subseteq S$, as required. $\square$

3 Convex hull

Now consider an arbitrary (not necessarily convex) subset S of V. We define the convex hull $\mathrm{conv}(S)$ of S as the "smallest" convex set that contains S.

Definition 3.1. *Consider a set $S \subseteq V$. Define its convex hull as*

$$\mathrm{conv}(S) = \bigcap_{\substack{T : S \subseteq T \\ T \text{ is convex}}} T.$$

Exercise 3. *Prove that*

1. $\mathrm{conv}(S)$ *is convex for every set* S
2. $\mathrm{conv}(S) \subseteq T$ *for every convex set* T *that contains* S
3. $\mathrm{conv}(S) = S$ *if* S *is convex*
4. $\mathrm{conv}(S') \subseteq \mathrm{conv}(S)$ *if* $S' \subseteq S$

The following claim provides an alternative characterization of $\mathrm{conv}(S)$.

Claim 3.2.

$$\operatorname{conv}(S) = \left\{ \sum_{i=1}^{n} \lambda_i v_i : v_1, \dots, v_n \in S \textit{ where } n \geq 1 \textit{ and } \sum_{i=1}^{n} \lambda_i = 1,\ \forall i:\ \lambda_i \geq 0 \right\}$$

Proof. Define $T = \operatorname{conv}(S)$ and

$$T' = \left\{ \sum_{i=1}^{n} \lambda_i v_i : v_1, \dots, v_n \in S \text{ where } n \geq 1 \text{ and } \sum_{i=1}^{n} \lambda_i = 1,\ \forall i:\ \lambda_i \geq 0 \right\}.$$

First, we show that $T' \subseteq T$. Indeed, consider a convex combination $u = \sum_{i=1}^{n} \lambda_i v_i$. We have, $v_i \in S \subseteq T$ for all i. Since T is convex, any convex combination of points in T is in T. In particular, $u = \sum_{i=1}^{n} \lambda_i v_i \in T$. We conclude that $T' \subseteq T$.

Now we prove that $T \subseteq T'$. As T is a minimal convex set that contains S, it is sufficient to verify that T' contains S and is convex. By the definition of T', T' contains a trivial convex combination $1 \cdot u = u$ for every $u \in S$. Thus, $S \subseteq T'$. Now consider two convex combinations in T'. By introducing, zero coefficients if necessary, we may assume that both combinations use the same points $v_1, \dots, v_n$.

$$u_1 = \sum_{i=1}^{n} \mu_i v_i$$
$$u_2 = \sum_{i=1}^{n} \nu_i v_i.$$

We want to prove that $\lambda u_1 + (1-\lambda) u_2 \in T'$ for every $\lambda \in [0,1]$. We have,

$$\lambda u_1 + (1-\lambda) u_2 = \lambda \sum_{i=1}^{n} \mu_i v_i + (1-\lambda) \sum_{i=1}^{n} \nu_i v_i = \sum_{i=1}^{n} (\lambda \mu_i + (1-\lambda)\nu_i) v_i,$$

which is a convex combination of points $v_1, \dots, v_n$ with weights $\lambda \mu_i + (1-\lambda)\nu_i$.[1] $\square$

Example 3.1. *The convex hull of $k > 1$ points in $\mathbb{R}^2$ is a convex polygon with at most k vertices or a segment.*

Exercise 4. *Answer the questions below.*

1. *Is it true that the convex hull of a bounded set is necessarily bounded?*
2. *Is it true that the convex hull of a closed set necessarily closed?*
3. *Is it true that the convex hull of a compact set is necessarily compact?*
4. *Is it true that the convex hull of an open set is necessarily open?*

Do your answers on the questions above depend on whether the space is finite or infinite dimensional?

[1] Exercise: verify that $\sum_{i=1}^{n} \lambda \mu_i + (1-\lambda)\nu_i = 1$.

Figure 1: Illustration for Radon's theorem in $\mathbb{R}^2$: there exist two disjoint subsets A and B of S such that $\mathrm{conv}(A) \cap \mathrm{conv}(B) \neq \varnothing$, as long as $|S| \geq 4$.

4 Theorems about convex hulls

Theorem 4.1 (Radon's Theorem)**.** *Consider $S \subseteq \mathbb{R}^d$ with $|S| \geq d+2$. Then there exist disjoint subsets A and B of S with $\mathrm{conv}(A) \cap \mathrm{conv}(B) \neq \varnothing$ (see Figure 1).*

Proof. To simplify the notation, we prove the theorem when S is finite. If S is infinite, we can choose an arbitrary subset $S' \subseteq S$ of size $d+2$ and apply the theorem to it, obtaining desired sets A and B. Let $v_1, \dots, v_n$ be the points in S (where $n \geq d+2$). Define $v'_i = v_i \oplus 1 \in \mathbb{R}^{d+1}$. We have at least $d+2$ points $v'_1, \dots, v'_n$ in a $d+1$ dimensional space. The points must be linearly dependent. That is, we must have

$$\sum_{i=1}^{n} \lambda_i v'_i = 0$$

for some coefficients λ_i, some of which are non-zero.

! Note that coefficients λ_i are not necessarily positive. In general, $\sum_i \lambda_i v_i$ is not a convex combination.

Rewrite this equation in terms of the original vectors v_i.

$$\sum_{i=1}^{n} \lambda_i v_i = 0$$
$$\sum_{i=1}^{n} \lambda_i = 0$$

Let $A = \{v_i : \lambda_i > 0\}$ and $B = \{v_i : \lambda_i < 0\}$. Then

$$u \equiv \sum_{v_i \in A} \lambda_i v_i = \sum_{v_i \in B} (-\lambda_i) v_i$$
$$\Lambda \equiv \sum_{v_i \in A} \lambda_i = \sum_{v_i \in B} (-\lambda_i)$$

Note that in each of the two expressions for Λ all the terms are positive. In particular, $\Lambda > 0$. Let $\alpha_i = \lambda_i / \Lambda$ for $v_i \in A$ and $\beta_i = -\lambda_i / \Lambda$ for $v_i \in B$. We have, $\sum_{v_i \in A} \alpha_i = \sum_{v_i \in B} \beta_i = 1$

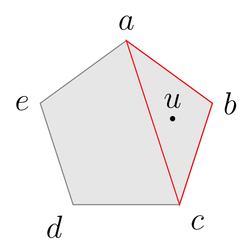


Figure 2: An illustration of Carathéodory's theorem in $\mathbb{R}^2$: The point u lies within the convex hull of five points, $a, b, c, d,$ and e. By Carathéodory's theorem, u can always be expressed as a convex combination of at most $d+1=3$ of these points. In this case, u is a convex combination of a, b, and c.

and all coefficients α_i and β_i are positive. Therefore,

$$\frac{u}{\Lambda} = \sum_{i:v_i \in A} \alpha_i v_i \in \operatorname{conv}(A) \qquad \text{and} \qquad u = \sum_{i:v_i \in B} \beta_i v_i \in \operatorname{conv}(B).$$

We conclude that $\operatorname{conv}(A) \cap \operatorname{conv}(B) \neq \varnothing$. □

Theorem 4.2 (Caratheódory's Theorem)**.** *Consider $S \subseteq \mathbb{R}^d$. Then every point $u \in \operatorname{conv}(S)$ can be expressed as a convex combination of at most $d+1$ points in S.*

Proof. Consider a convex combination for u with the smallest number of terms:

$$u = \sum_{i=1}^{n} \mu_i v_i$$

where all $v_i \in S$. If $n \leq d+1$, then we are done. So we assume that $n > d+1$ and then get a contradiction by providing another convex combination for u with a smaller number of terms.

Let us apply Radon's theorem to points $v_1, \ldots, v_n$. We get two disjoint sets $A \subseteq S$ and $B \subseteq S$ and positive weights α_i and β_i such that

$$w \equiv \sum_{v_i \in A} \alpha_i v_i = \sum_{v_i \in B} \beta_i v_i$$
$$\sum_{v_i \in A} \alpha_i = \sum_{v_i \in B} \beta_i = 1$$

Now let

$$\mu_i^{(t)} = \begin{cases} \mu_i - t\alpha_i, & \text{for } v_i \in A \\ \mu_i + t\beta_i, & \text{for } v_i \in B \\ \mu_i, & \text{otherwise} \end{cases}$$

We now verify that $\sum \mu_i^{(t)} v_i$ is also a convex combination for u, as long as t is small enough in absolute value to ensure that all the coefficients $\mu_i^{(t)}$ are non-negative.

$$\sum_{i=1}^{n} \mu_i^{(t)} v_i = \sum_{i=1}^{n} \mu_i v_i - t \sum_{v_i \in A} \alpha_i v_i + t \sum_{v_i \in B} \beta_i v_i = u - tw + tw = u.$$
$$\sum_{i=1}^{n} \mu_i^{(t)} = \sum_{i=1}^{n} \mu_i - t \sum_{v_i \in A} \alpha_i + t \sum_{v_i \in B} \beta_i = 1 - t + t = 1$$

We see that for every t, $\sum_{i=1}^{n} \mu_i^{(t)} v_i$ is indeed a convex combination for u as long as $\mu_i^{(t)} \geq 0$ for all i. Our goal now is to choose t so that this is a valid convex combination and at least one coefficient $\mu_i^{(t)}$ is 0.

? **Question:** What t should we use?
Answer: We can use $t = \min_{v_i \in A} \frac{\mu_i}{\alpha_i}$.

Then all $\mu_i^{(t)} \geq 0$ and at least one $\mu_i^{(t)} = 0$. We got a convex combination with fewer than n non-zero terms, as desired. □

Theorem 4.3 (Helly's Theorem). *Consider $n \geq d+1$ convex sets $S_1, \dots, S_n$ in $\mathbb{R}^d$. Assume that every $d+1$ of them have a non-empty intersection. Then $\bigcap_{i=1}^{n} S_i \neq \varnothing$.*

Proof. The proof is by induction on n. The claim is trivial when $n = d+1$. We prove it for $n > d+1$, assuming that it holds for $n' = n-1$.

We choose points $x_1, \dots, x_n$ as follows. Consider the intersection of all sets S_i other than S_j. It is non-empty by the induction hypothesis. Let x_j be an arbitrary point in this intersection $\bigcap_{i: i \neq j} S_i$. We obtain points $x_1, \dots, x_n$. By construction, $x_i \in S_j$ if $i \neq j$. Observe that if $x_i \in S_i$ for some i then we are done, since then x_i lies in all sets S_j, including $j = i$. So we assume below that $x_i \notin S_i$ for all i.

Now we apply Radon's theorem to the set of points $\{x_i : 1 \leq i \leq n\}$. We get two disjoint subsets A and B such that $\operatorname{conv}(A) \cap \operatorname{conv}(B) \neq \varnothing$. Choose $u \in \operatorname{conv}(A) \cap \operatorname{conv}(B)$. We prove that $u \in \bigcap_{i=1}^{n} S_i$ or, in other words, $u \in S_i$ for every i.

Fix some i. We know that x_i cannot belong to both A and B, as A and B are disjoint. Assume without loss of generality that $x_i \notin A$. All points $x_j \in A$ are in S_i. Therefore, $u \in \operatorname{conv}(A) \subseteq \operatorname{conv}(S_i) = S_i$. □

5 Extreme points

Consider a finite set of points in $\mathbb{R}^2$. Its convex hull is simply a convex polygon, which is uniquely determined by its vertices. Informally speaking, the vertices can be regarded as the most "important" points of the polygon. In higher dimensions, the convex hull is a polytope, which again is uniquely determined by its vertices. In this section, we generalize the notion of a vertex to arbitrary convex sets by introducing the concept of *extreme points*.

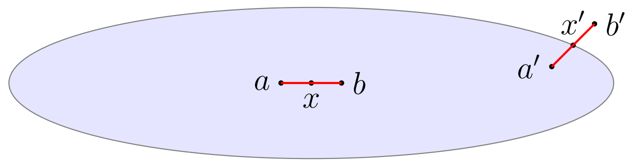


Figure 3: Point x is not an extreme points, since $x = \frac{a+b}{2}$ for some distinct $a, b \in S$. Point x' is an extreme point, since there are no distinct points $a', b' \in S$ such that $x' = \frac{a'+b'}{2}$.

Definition 5.1 (Minkowski's definition)**.** *We say that x is an extreme point of a convex set S if there are no distinct points $a, b \in S$ such that $x = \frac{a+b}{2}$.*

Exercise 5. *Check that in the definition of an extreme point, we can require that $x \notin (a, b)$ for all distinct points $a, b \in S$ (where (a, b) is the open interval between a and b).*

Theorem 5.2. *Let X be an arbitrary set. Then $x \in \operatorname{conv}(X)$ is an extreme point of $\operatorname{conv}(X)$ if and only if $x \notin \operatorname{conv}(X \setminus \{x\})$.*[2]

Proof. First, assume that $x \in \operatorname{conv}(X \setminus \{x\})$. We shall prove that x is not an extreme point of $\operatorname{conv}(X)$. That is, we show that there exist a and b such that $x \in (a, b)$. Since $x \in \operatorname{conv}(X \setminus \{x\})$, we have a convex combination $x = \sum_{i=1}^{n} \alpha_i x_n$ where all $x_i \in X \setminus \{x\}$ and all α_i are positive. Because all $x_i \neq x$, we must have $n > 1$. Let $a = \sum_{i=1}^{n-1} \frac{\alpha_i}{1-\alpha_n} x_n$ and $b = x_n$. Clearly, $a, b \in \operatorname{conv}(X \setminus \{x\})$. Then $x = (1 - \alpha_n)a + \alpha_n b \in (a, b)$, as desired.

Now, assume that x is not an extreme point of $\operatorname{conv}(X)$; that is, $x = \frac{a+b}{2}$ for some $a, b \in \operatorname{conv}(X)$. Since $a, b \in \operatorname{conv}(X)$, each of them is a convex combination of points in X. We may assume that the same points participate in both convex combinations (but possibly some coefficients are 0):

$$a = \sum_{i=1}^{n} \alpha_i x_i \qquad \text{and} \qquad b = \sum_{i=1}^{n} \beta_i x_i$$

If x is not among points $x_1, \ldots, x_n$ then

$$x = \frac{a+b}{2} = \sum_{i=1}^{n} \frac{\alpha_i + \beta_i}{2} x_i$$

is a convex combination of points in $X \setminus \{x\}$. Thus, $x \in \operatorname{conv}(X \setminus \{x\})$, as required. Now assume that one of the points x_i is x. Without loss of generality, $x_n = x$. Note that $\alpha_n < 1$ and $\beta_n < 1$, since $a \neq x$ and $b \neq x$, respectively. We have,

$$\sum_{i=1}^{n-1} \frac{\alpha_i + \beta_i}{2} x_i = \frac{a+b}{2} - \frac{\alpha_n + \beta_n}{2} x = \left(1 - \frac{\alpha_n + \beta_n}{2}\right) x.$$

[2]In particular, x must be in X, as otherwise $x \in \operatorname{conv}(X) = \operatorname{conv}(X \setminus \{x\})$.

Thus,

$$\sum_{i=1}^{n-1} \frac{\alpha_i + \beta_i}{2 - \alpha_n + \beta_n} x_i = x.$$

We conclude that x is a convex combination of points $x_1, \dots, x_{n-1}$ in $X \setminus \{x\}$. $\square$

Exercise 6. *Answer the following questions.*

1. *What is the set of extreme points of the closed unit disc* $\{x \in \mathbb{R}^2 : \|x\|_2 \leq 1\}$*?*
2. *What is the set of extreme points of the open unit disc* $\{x \in \mathbb{R}^2 : \|x\|_2 < 1\}$*?*
3. *What is the set of extreme points of a line in* $\mathbb{R}^2$.

Exercise 7. *Recall the definition of the boundary ∂X of a set X:*

$$\partial X = \{x \in X : B_\varepsilon(x) \setminus X \neq \varnothing \text{ for all } \varepsilon > 0\} \quad \text{where } B_\varepsilon(x) = \{y : \|x - y\|_2 < \varepsilon\}.$$

Prove that all extreme points of a convex set X lie on the boundary of X.

Exercise 8. *A polygon is uniquely determined by the set of its vertices. However, show that the extreme points of a convex set S do not determine S, in general.*

Theorem 5.3 (Minkowski, Krein–Milman)**.** *Assume that S is a compact*[3] *convex set in $\mathbb{R}^d$, then $S = \operatorname{conv}(X)$ where X is the set of extreme points of S.*

Before we proceed with the proof, we need some auxiliary definitions. For a point $x \in X$, let $L_x = \{v : x + \varepsilon v \in S \text{ and } x - \varepsilon v \in S \text{ for some } \varepsilon > 0\}$.

Lemma 5.4. *L_x is a linear subspace.*

Proof. It is clear from the definition that if $v \in L_x$ than so is $-v$. It is also clear that if $v \in L_x$ then $\alpha v \in L_x$ for every α. Now we verify that if $u, v \in L_x$ then $u + v \in L_x$.

Since $u \in L_x$, the segment $[x - \varepsilon_1 u, x + \varepsilon_1 u]$ is in S for some $\varepsilon_1 > 0$. Since $v \in L_x$, the segment $[x - \varepsilon_2 v, x + \varepsilon_2 v]$ is in S for some $\varepsilon_2 > 0$. Since S is convex, the parallelogram Π (including its interior points) with vertices $x \pm \varepsilon_1 u$ and $x \pm \varepsilon_2 v$ lies in S. Let $\varepsilon_3 = \min(\varepsilon_1, \varepsilon_2)/2$. Then $x \pm \varepsilon_3(u + v) \in \Pi \subseteq S$. We conclude that $u + v \in L_x$. $\square$

We define $\operatorname{rank} x = \dim L_x$. Note that if x is not an extreme point then x belongs to some interval (a, b) with distinct endpoints $a, b \in S$. Thus, vector $a - b \in L_x$ and consequently $\operatorname{rank} x = \dim L_x \geq 1$. Thus, $\operatorname{rank} x = 0$ only if x is an extreme point of S.[4]

[3] Recall that $X \subseteq \mathbb{R}^d$ is compact if and only if it is closed and bounded.

[4] In fact, $\operatorname{rank} x = 0$ if and only if x is an extreme point.

Proof of Theorem 5.3. Clearly, $\operatorname{conv}(X) \subseteq \operatorname{conv}(S) = S$. So we need to prove that $S \subseteq \operatorname{conv}(X)$. That is, every point in S is a convex combination of extreme points. We are going to prove that by induction on $\operatorname{rank} y$. If $\operatorname{rank} y = 0$, then y is an extreme point and thus lies in $\operatorname{conv}(X)$.

Now assume that the induction hypothesis holds for points y with $\operatorname{rank} y \leq k-1$ and prove it for y with $\operatorname{rank} y = k$. Since y is not an extreme point, $y = \frac{a+b}{2}$ for some distinct $a, b \in S$. Consider the line ℓ that goes through a and b. Note that that $\ell \cap S$ is a closed (bounded) segment, since S is compact and convex. Denote the endpoints of this segment by y_1 and y_2. Then $y \in (a,b) \subseteq (y_1, y_2)$. We show that $\operatorname{rank} y_1 < k$ and similarly $\operatorname{rank} y_2 < k$.

Lemma 5.5. *We have,*

- $L_{y_1} \subseteq L_y$.
- $y_1 - y_2 \in L_{y_1} \setminus L_y$.

Proof. I. Consider $v \in L_{y_1}$. We have that $y_1 \pm \varepsilon v \in S$ for some small enough $\varepsilon > 0$. We also have that $y_2 \in S$. Since S is convex, the entire triangle Δ with vertices $y_1 + \varepsilon v, y_1 - \varepsilon v, y_2$ lies in S. Note that point y lies on the segment (cevian) $[y_1, y_2]$, which in turn is inside Δ. We get that

$$p_1 = \frac{\|y - y_2\|}{\|y_1 - y_2\|}(y_1 + \varepsilon v) + \frac{\|y - y_1\|}{\|y_1 - y_2\|} y_2 = y + \left(\frac{\|y - y_2\|}{\|y_1 - y_2\|}\varepsilon\right) v$$

is a convex combination of $y_1 + \varepsilon v$ and y_2 and thus lies inside Δ. Similarly,

$$p_2 = y - \left(\frac{\|y - y_2\|}{\|y_1 - y_2\|}\varepsilon\right) v$$

lies inside Δ. It follows that $p_1, p_2 \in S$ and hence $v \in L_y$.

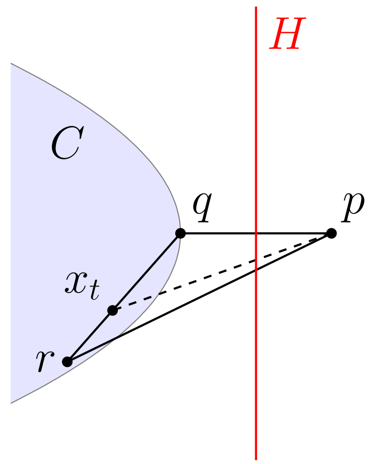
II. Recall that $a, b \in S$ and $y = \frac{a+b}{2}$. Therefore, $a - b \in L_y$. Now, $y_1 - y_2$ and $a - b$ are colinear so $y_1 - y_2 \in L_y$ as well. On the other hand, y_1 is an endpoint of the segment $S \cap \ell$. Therefore, $y_1 + \varepsilon(y_1 - y_2) \notin S$ for every $\varepsilon > 0$. We conclude that $y_1 - y_2 \notin L_{y_1}$. $\square$

We have proved that L_{y_1} is a proper subset of L_y. Thus, $\operatorname{rank} y_1 = \dim L_{y_1} < \dim L_y = \operatorname{rank} y$. Similarly, $\operatorname{rank} y_2 < \operatorname{rank} y$. By the induction hypothesis, $y_1, y_2 \in \operatorname{conv}(X)$. Since $\operatorname{conv}(X)$ is convex, $y \in [y_1, y_2] \subseteq \operatorname{conv}(X)$, as required. $\square$

6 Separating Hyperplanes

Definition 6.1. *Consider two sets A and B in a linear space. We say that an affine hyperplane H strictly separates A and B if A and B lie on opposite sides of H and $A \cap H = \varnothing$, $B \cap H = \varnothing$. We call such a hyperplane a strict separating hyperplane.*

If A and B lie on opposite sides of H but may share points with H, then we say that H weakly or non-strictly separates A and B.



Theorem 6.2. *Let $p \in \mathbb{R}^d$ be a point and $C \subseteq \mathbb{R}^d$ be a non-empty closed convex set. Assume that $p \notin C$. Then there is a strict separating hyperplane H between p and C.*

Proof. First, we find the point q in C that is closest to p. Why does such a point exist? Consider the function $f(x) = \|x - p\|_2$ on C. Note that f is continuous.

If C is compact, then f attains its minimum on C, so we simply define $q = \operatorname{argmin}_x f(x)$.

If C is not compact, let $\Delta = \inf_{x \in C} \|x - p\|_2$ and define

$$C' = C \cap \{x : \|x - p\| \leq \Delta + 1\}.$$

Since C' is the intersection of two closed sets—C and a closed ball of radius $\Delta + 1$—it is also closed. Moreover, because the ball is bounded, so is C'. Thus, C' is compact. Applying the argument above to C', we obtain the desired point q at distance Δ from p.

Note that $\|p - q\|_2 > 0$ because $p \notin C$. Now let H be the perpendicular bisector hyperplane of the segment $[p, q]$; in other words, $H = \{x : \|x - p\|_2 = \|x - q\|_2\}$. Clearly, the distance from p to H is $\|p - q\|_2/2 > 0$. Thus, $p \notin H$. We claim that H does not intersect C. Assume to the contrary that there exists $r \in C \cap H$. Consider the triangle with vertices p, q, and r. Since $r \in H$, $\|p - r\| = \|q - r\|$. Therefore, the triangle is isosceles and thus $\angle pqr < \pi/2$. Since $q, r \in C$, we have $[q, r] \subset C$ and thus $x_t \equiv q + t(r - q) \in C$ for $t \in [0, 1]$. Now

$$\begin{aligned} \|p - x_t\|^2 &= \|p - q\|^2 + t^2\|r - q\|^2 - 2t \cdot \|p - q\| \cdot \|r - q\| \cdot \cos\angle pqr \\ &= \|p - q\|^2 - 2t \cdot \underbrace{\|p - q\| \cdot \|r - q\| \cdot \cos\angle pqr}_{>0} + O(t^2) \end{aligned}$$

For sufficiently small $t > 0$, we have $\|p - x_t\|_2 < \|p - q\|_2$, contradicting the assumption that q is the closest to p point in C.

We conclude that $p \notin H$ and C_2 lies entirely on one side of H. Since the segment $[p, q]$ intersects H, point p and set C lie on opposite sides of H. $\square$

Theorem 6.3. *Let $C_1 \subseteq \mathbb{R}^d$ be a compact convex set and $C_2 \subseteq \mathbb{R}^d$ be a closed convex set. Assume that $C_1 \cap C_2 = \varnothing$ and both sets are not empty. Then there is a strict separating hyperplane H between C_1 and C_2.*

Proof sketch. Let $f(x) = \inf_{y \in C_2} \|x - y\|$ be the distance from $x \in C_1$ to C_2. Function $f(x)$ is continuous (and, in fact, 1-Lipschitz) and thus attains its minimum on compact set C_1. Let p be the point where it attains its minimum. We use Theorem 6.2 to find a separating hyperplane H between p and C_2. Now the same argument as in Theorem 6.2 shows that C_1 does not intersect H. $\square$

Exercise 9. *Is Theorem 6.3 true if we only require that C_1 and C_2 be closed convex sets (that is, we no longer require that C_1 be compact).*

If we do not require that the hyperplane strictly separate A abd B, then we only need that both sets are convex and non-empty.

Theorem 6.4. *Let $A, B \subseteq \mathbb{R}^d$ be non-empty disjoint convex sets. Then there is a hyperplane H that non-strictly separates A and B.*

The proof constructs appropriately defined sequences of compact convex sets A_i and B_i, applying Theorem 6.2 to each pair (A_i, B_i) to obtain separating hyperplanes H_i, and then taking the limit of these hyperplanes. We omit the details here.

7 Polar Sets

The Krein–Milman theorem states that a compact convex body is determined by its extreme points. This is analogous to defining a polygon or polyhedron by specifying its vertices. However, we can also define a polygon or polyhedron by specifying its facets instead of its vertices. In fact, this approach is used to define the feasible polytope in a linear program. Let us generalize this idea to arbitrary convex sets. Consider all closed affine half-spaces H that contain a given convex set S and their intersection

$$\bigcap_{H : S \subseteq H} H.$$

Q: Is this intersection equal to S?
A: The intersection of closed affine half-spaces is a closed set. So if S is not closed, then the intersection is not equal to S.

Claim 7.1. *If S is a closed convex set, then $S = \bigcap_{H : S \subseteq H} H$.*

Proof. Since all H in the intersection contain S, so does their intersection. On the other hand, if $p \notin S$, then by Theorem 6.2, there is a separating hyperplane P that separates p and C. Hyperplane P defines a half-space that contains C but not p. We conclude that $p \notin \bigcap_{H : S \subseteq H} H$. $\square$

Note that a half-space H can be written as $\{x : \langle c, x \rangle \leq b\}$ for some vector c and scalar b. Assume for a moment that S contains the origin. Then if H contains S, it also contains 0, and thus $b \geq \langle c, 0 \rangle = 0$. Further, it is easy to see that Claim 7.1 holds for S even if we

Figure 4: A set S and its polar set S° (shown separately for clarity).

exclude half-spaces with $b = 0$, since all hyperplanes from Theorem 6.2 strictly separate p and S and thus do not go through the origin. The formula for a half-space H with $b > 0$ can be simplified: $H = H_y = \{x : \langle y, x\rangle \leq 1\}$ where $y = c/b$. That is,

$$S = \bigcap_{y: S \subseteq H_y} H_y \tag{1}$$

here we may assume that $H_0 = \mathbb{R}^d$ also participates in the intersection, even though H_0 is not a half-space.

We conclude that the set $S^\circ \equiv \{y : S \subseteq H_y\}$ uniquely defines S if S is a closed convex set and $0 \in S$. We call S° the *polar set* of S. In the following definition of the polar set, we use that $S \subseteq H_y$ if and only if $\langle x, y\rangle \leq 1$ for all $x \in S$.

Definition 7.2. *Consider an arbitrary set S in Euclidean space $\mathbb{R}^d$. The polar set of S is*

$$S^\circ = \{y : S \subseteq H_y\} = \{y : \langle x, y\rangle \leq 1 \textit{ for all } x \in S\}.$$

Note that we defined S° for all sets S. However, the definition is mostly useful when S is a closed convex set containing the origin.

Exercise 10. *Find the polar sets of the following sets.*

- *B_R, the closed Euclidean ball of radius R centered at the origin*
- *$\{x\}$ where $x \in \mathbb{R}^d$*
- *a half-space H_y*
- *a regular polygon P centered at the origin*
- *a cube centered at the origin*

Exercise 11. *Prove that $0 \in S^\circ$ for every set S.*

Exercise 12. *Let S be an arbitrary set and T be a non-degenerate linear transform of $\mathbb{R}^d$. Prove that $(TS)^\circ = (T^\top)^{-1} S^\circ$, where $TS \equiv \{Tx : x \in S\}$.*

Now observe that (1) can be written as follows for closed convex sets containing 0:

$$S = \bigcap_{y \in S^\circ} H_y.$$

On the other hand (for every S),

$$S^\circ = \{y : \langle x, y \rangle \le 1 \text{ for all } x \in S\} = \bigcap_{x \in S} \{y : \langle x, y \rangle \le 1\} = \bigcap_{x \in S} H_x.$$

We see the duality between S and S°. Thus, we have proved the following theorem.

Theorem 7.3. *If S is a closed convex set containing 0, then $S^{\circ\circ} = S$.*

Let us now prove some other basic properties of S°.

Claim 7.4. *The following properties hold.*

1. *Set S° is a convex closed set for every S.*
2. *If $S \subseteq T$ then $S^\circ \supseteq T^\circ$.*
3. *$(S \cup T)^\circ = S^\circ \cap T^\circ$.*
4. *More generally, let $\{S_\alpha\}_\alpha$ be a family of sets in $\mathbb{R}^d$. Then $(\bigcup_\alpha S_\alpha)^\circ = \bigcap_\alpha S_\alpha^\circ$.*

Proof. **1.** We have, $S^\circ = \bigcap_{x \in S} H_x$ is an intersection of closed convex sets and thus is a closed convex set itself.

2. We need to prove that $\bigcap_{x \in S} H_x \supseteq \bigcap_{x \in T} H_x$. This inclusion holds since each half-space that participates in the intersection on the left also participates in one on the right.

3.

$$(S \cup T)^\circ = \bigcap_{x \in S \cup T} H_x = \Big(\bigcap_{x \in S} H_x\Big) \cap \Big(\bigcap_{x \in T} H_x\Big) = S^\circ \cap T^\circ.$$

4. The proof is essentially identical to that of item 3. □

Claim 7.5. *Assume that S and T are closed convex sets containing the origin. Then*

$$(S \cap T)^\circ = \overline{\mathrm{conv}(S^\circ \cup T^\circ)}$$

Here $\overline{A}$ denotes the closure of set A. Note that $S^\circ \cup T^\circ$ is generally speaking a non-convex set. We will study polar sets of non-convex sets in the next section and then prove Claim 7.5.

8 Polar Sets of Arbitrary Sets

As we discussed above, polar sets are particularly useful when S is a closed convex set containing 0. Many properties hold only for such sets (e.g. $S = S^{\circ\circ}$ only for such sets). In this section, we give some properties of polar sets of arbitrary sets.

Claim 8.1. *Consider a set $S \subseteq \mathbb{R}^d$. Then*

- $S^\circ = (S \cup \{0\})^\circ$
- $S^\circ = \mathrm{conv}(S)^\circ$
- $S^\circ = (\overline{S})^\circ$

In particular, $S^\circ = \left(\overline{\mathrm{conv}(S \cup \{0\})}\right)^\circ$.

Proof. Since $S \subseteq S \cup \{0\}$, $S \subseteq \mathrm{conv}(S)$, and $S \subseteq \overline{S}$, from Claim 7.4, we get $S^\circ \supseteq (S \cup \{0\})^\circ$, $S^\circ \supseteq \mathrm{conv}(S)^\circ$, and $S^\circ \supseteq \overline{S}^\circ$. So we need to prove that $S^\circ \subseteq (S \cup \{0\})^\circ$, $S^\circ \subseteq \mathrm{conv}(S)^\circ$, and $S^\circ \subseteq \overline{S}^\circ$.

First, $(S \cup \{0\})^\circ = S^\circ \cap \{0\}^\circ = S^\circ \cap \mathbb{R}^d = S^\circ$. Then, since H_y is convex, if $S \subseteq H_y$ then $\mathrm{conv}(S) \subseteq H_y$. Thus,

$$S^\circ = \{y : S \subseteq H_y\} \subseteq \{y : \mathrm{conv}(S) \subseteq H_y\} = \mathrm{conv}(S)^\circ.$$

Finally, since H_y is closed, if $S \subseteq H_y$ then $\bar{S} \subseteq H_y$, as above we get

$$S^\circ = \{y : S \subseteq H_y\} \subseteq \{y : \bar{S} \subseteq H_y\} = \bar{S}^\circ.$$

□

Claim 8.2. *Let S be an arbitrary set in $\mathbb{R}^d$. Then $S^{\circ\circ} = \overline{\mathrm{conv}(S \cup \{0\})}$.*

Proof. Define $\hat{S} = \overline{\mathrm{conv}(S) \cup \{0\}}$. By Claim 8.1, $S^\circ = \hat{S}^\circ$. Now $\hat{S}$ is a closed convex set containing 0. Thus, $\hat{S}^{\circ\circ} = \hat{S}$. We get,

$$S^{\circ\circ} = (\hat{S}^\circ)^\circ = \hat{S}^{\circ\circ} = \hat{S},$$

as required. □

Proof of Claim 7.5. We apply Claim 7.4, item 3, to sets S° and T°. We get

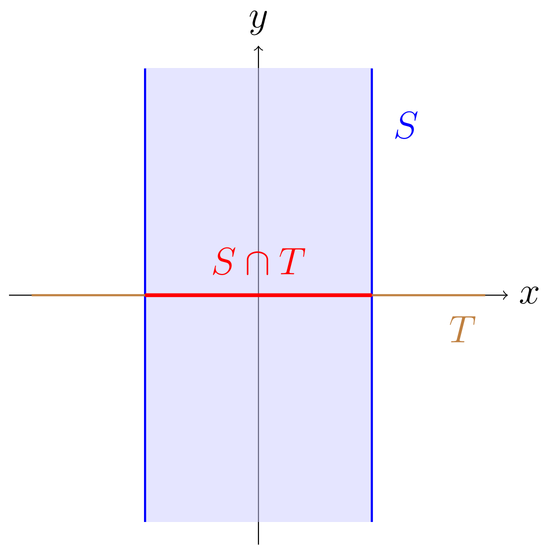
$$(S^\circ \cup T^\circ)^\circ = S^{\circ\circ} \cap T^{\circ\circ} = S \cap T.$$

Thus, $(S \cap T)^\circ = (S^\circ \cup T^\circ)^{\circ\circ} = \overline{\mathrm{conv}(S^\circ \cup T^\circ)}$. Here we used that $S^\circ \cup T^\circ$ contains the origin. □

Exercise 13. *Prove that*

$$(S \cap T)^\circ \neq \overline{\mathrm{conv}(S^\circ \cup T^\circ)}$$

for the following sets S and T:



- $S = \{(x, y) : x > 0,\ y > 0\}$ *and* $T = \{(x, y) : x < 0,\ y > 0\}$
- $S = \{(1, y) : y \in \mathbb{R}\}$ *and* $T = \{(x, 1) : x \in \mathbb{R}\}$

Exercise 14. *Prove that*

$$(S \cap T)^\circ \neq \mathrm{conv}(S^\circ \cup T^\circ)$$

for the following closed convex sets S and T in $\mathbb{R}^2$ containing the origin:

$$S = [-1, 1] \times \mathbb{R} \qquad \textit{and} \qquad T = \mathbb{R} \times \{0\}.$$

Claim 8.3. *Let P be a linear subspace of $\mathbb{R}^d$ and π be the orthogonal projection on P. Let $S \subset \mathbb{R}^d$. Then*

$$(\pi S)^\circ \cap P = S^\circ \cap P$$

if S is a closed convex set containing 0 then

$$(S \cap P)^\circ \cap P = \pi(S^\circ)$$

Proof. It is straightforward to verify these identities directly using the definition of the polar set. However, we will prove them using polar set properties we established above. Consider $P^\perp$, the orthogonal complement to P. Note that $P^\circ = P^\perp$. Observe that for every set A

$$\overline{\mathrm{conv}(\pi A \cup P^\perp)} = \overline{\mathrm{conv}(A \cup P^\perp)} = \pi A + P^\perp \equiv \{x' + x'' : x' \in \pi A,\ x'' \in P^\perp\} \tag{2}$$

$$= A + P^\perp \equiv \{x' + x'' : x' \in A,\ x'' \in P^\perp\}. \tag{3}$$

Therefore, $(\pi A \cup P^\perp)^\circ = (A \cup P^\perp)^\circ = \pi A + P^\perp$. We start with proving the first identity. We apply the statement we just proved with $A = S$.

$$(\pi S)^\circ \cap P = (\pi S)^\circ \cap (P^\perp)^\circ = (\pi S \cup P^\perp)^\circ \stackrel{(2)}{=} (S \cup P^\perp)^\circ = S^\circ \cap (P^\perp)^\circ = S^\circ \cap P.$$

Now we prove the second identity. Here, we let $A = S^\circ$.

$$(S \cap P)^\circ \cap P = \overline{\operatorname{conv}(S^\circ \cup P^\circ)} \cap P \stackrel{(2)}{=} (\pi(S^\circ) + P^\perp) \cap P = \pi(S^\circ).$$

$\square$