

## 1 Review of convexity

**Definition 1.1.** A set  $C \subseteq \mathbb{R}^n$  is called *convex* if for any  $x, y \in C$  and  $\lambda \in [0, 1]$ ,  $\lambda x + (1 - \lambda)y \in C$ .

Some examples of convex sets:

- Halfspaces:  $\{x \in \mathbb{R}^n : \langle a, x \rangle \leq b\}$  where  $a \in \mathbb{R}^n \setminus \{0\}$  and  $b \in \mathbb{R}$ .
- The disk in  $\mathbb{R}^2$ :  $\{(x, y) : x^2 + y^2 \leq 1\}$ .
- The nonnegative orthant:  $\{x \in \mathbb{R}^n : x_1 \geq 0, \dots, x_n \geq 0\}$ .
- Nonnegative polynomials:  $\{(a_0, \dots, a_n) \in \mathbb{R}^{n+1} : a_0 + a_1x + \dots + a_nx^n \geq 0 \forall x \in \mathbb{R}\}$ .

**Proposition 1.1** (Operations that preserve convexity). *The following operations preserve convexity.*

- If  $C$  is a convex set in  $\mathbb{R}^n$  and  $A : \mathbb{R}^n \rightarrow \mathbb{R}^m$  is a linear map then  $A(C)$  is convex.
- If  $C_1, C_2$  are convex then  $C_1 \cap C_2$  are convex.
- If  $C \subseteq \mathbb{R}^n$  convex then  $\{(x, t) \in \mathbb{R}^n \times \mathbb{R} : t > 0 \text{ and } x/t \in C\} \subseteq \mathbb{R}^{n+1}$  is convex.

**Theorem 1.1** (Separating hyperplane theorem). *Assume  $C \subseteq \mathbb{R}^n$  is a convex subset of  $\mathbb{R}^n$ , and  $y \in \mathbb{R}^n$  with  $y \notin C$ . Then there exists  $a \in \mathbb{R}^n \setminus \{0\}$  and  $b \in \mathbb{R}$  such that*

$$\langle a, y \rangle \geq b \quad \text{and} \quad \langle a, x \rangle \leq b \quad \text{for all } x \in C. \quad (1)$$

Furthermore, if  $C$  is closed, the inequalities in (1) can be made strict.

*Proof.* We give the proof when  $C$  is closed. The general case is left as an exercise. If  $C$  is closed we can define the projection map on  $C$ , namely  $p_C(y) := \min\{\|y - x\| : x \in C\}$  is well defined and satisfies  $\langle y - p_C(y), x - p_C(y) \rangle \leq 0$  for any  $x \in C$ . Let  $a = y - p_C(y)$  and  $b = \langle a, p_C(y) \rangle + \frac{1}{2}\|a\|_2^2$ . Note that  $\langle a, y \rangle - b = \|a\|_2^2 - \frac{1}{2}\|a\|_2^2 > 0$ . Also for any  $x \in C$  we have  $\langle a, x \rangle - b = \langle y - p_C(y), x - p_C(y) \rangle - \frac{1}{2}\|a\|_2^2 < 0$  which is what we wanted.  $\square$

**Exercise 1.1.** 1. Prove Theorem 1.1 in the case where  $C$  is not necessarily closed.

2. Let  $C, D$  be two disjoint convex sets in  $\mathbb{R}^n$ . Use Theorem 1.1 to prove that there is  $a \in \mathbb{R}^n \setminus \{0\}$  and  $b \in \mathbb{R}$  such that  $\langle a, x \rangle \leq b$  for all  $x \in C$  and  $\langle a, y \rangle \geq b$  for all  $y \in D$  (Hint: consider the set  $C - D = \{x - y : x \in C, y \in D\}$ ). Give sufficient conditions on  $C$  and  $D$  that guarantee the existence of a strict separating hyperplane.

**Supporting hyperplane** An important application of the separating hyperplane theorem when  $C$  is not closed is to prove the existence of *supporting hyperplanes*: If  $C$  is a closed convex subset of  $\mathbb{R}^n$  and  $y$  is a point on the boundary  $C \setminus \text{int}(C)$  of  $C$ , a supporting hyperplane of  $C$  at  $y$  is a hyperplane  $H = \{x \in \mathbb{R}^n : \langle a, x \rangle = b\}$  such that  $y \in H$  and  $\langle a, x \rangle \leq b$  for all  $x \in C$ .

**Definition 1.2** (Convex hull). Assume  $S \subseteq \mathbb{R}^n$ . The *convex hull* of  $S$ , denoted  $\text{conv}(S)$ , is the smallest convex set containing  $S$ , i.e.,

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

**Exercise 1.2.** Let  $S \subseteq \mathbb{R}^n$ .

1. Show that the convex hull of  $S$  can also be expressed as

$$\begin{aligned} \text{conv}(S) = \left\{ x \in \mathbb{R}^n \quad : \quad \exists k \in \mathbb{N}_{\geq 1}, \lambda_1, \dots, \lambda_k \geq 0, s_1, \dots, s_k \in S \right. \\ \left. \text{s.t. } x = \sum_{i=1}^k \lambda_i s_i \text{ and } \sum_{i=1}^k \lambda_i = 1 \right\}. \end{aligned}$$

2. (Carathéodory theorem) Show that any point in  $\text{conv}(S)$  can be written as a convex combination of at most  $n + 1$  points of  $S$  (hint: any  $k$  points  $s_1, \dots, s_k$  with  $k \geq n + 2$  are affinely dependent i.e., there exist  $\mu_1, \dots, \mu_k$  such that  $\sum_{i=1}^k \mu_i s_i = 0$  and  $\sum_{i=1}^k \mu_i = 0$ ).

**Definition 1.3** (Face). Let  $C \subseteq \mathbb{R}^n$  be a convex set. A subset  $F$  of  $C$  is called a *face* of  $C$  if the following two conditions hold:

1.  $F$  is convex
2. For any  $x \in F$ , if  $a, b \in C$  and  $0 < \lambda < 1$  are such that  $x = \lambda a + (1 - \lambda)b$ , then  $a, b \in F$ .

**Definition 1.4** (Extreme point). Let  $C \subseteq \mathbb{R}^n$  be a convex set. A point  $x \in C$  is called *extreme* if the singleton  $\{x\}$  is a face of  $C$ .

**Definition 1.5** (Dimension). Let  $C \subseteq \mathbb{R}^n$  be a convex set. We define the *dimension* of  $C$  to be the dimension of the smallest affine space that contains  $C$ . We say that  $C$  is *full-dimensional* if it has dimension  $n$ .

Recall:

- An *affine subspace* of  $\mathbb{R}^n$  is, by definition, a translate of a linear subspace, i.e., a set of the form  $x + V$  where  $x \in \mathbb{R}^n$  and  $V$  is a linear subspace of  $\mathbb{R}^n$ . The dimension of an affine subspace is the dimension of the corresponding linear subspace  $V$ .
- An *affine combination* of two points  $x, y \in \mathbb{R}^n$  is a combination of the form  $\lambda x + (1 - \lambda)y$  where  $\lambda$  is an arbitrary real number (not necessarily nonnegative). Affine subspaces are closed under affine combinations.

**Proposition 1.2.** Let  $C \subset \mathbb{R}^n$  be a convex set with nonempty interior.

- (i) If  $F \subseteq G \subseteq C$  where  $F$  is a face of  $G$  and  $G$  a face of  $C$ , then  $F$  is a face of  $C$ .
- (ii) Assume  $C$  is closed. Then any point  $x \in C \setminus \text{int}(C)$  lies on a face  $F$  of  $C$  of dimension strictly smaller than  $n$ .

*Proof.* Item (i) is easy to verify. For item (ii) we use the separating hyperplane theorem. Since  $x \notin \text{int}(C)$  we can find a hyperplane that separates  $x$  from  $\text{int}(C)$ , i.e.,  $\langle a, x \rangle = b$  and  $\langle a, x \rangle \leq b$  for any  $x \in \text{int}(C)$ . Define  $F = C \cap \{z \in \mathbb{R}^n : \langle a, z \rangle = b\}$ . It is easy to verify that  $F$  satisfies the conditions that we want: namely  $F$  is a face of dimension at most  $n - 1$  that contains  $x$ . This completes the proof.  $\square$

**Theorem 1.2** (Minkowski theorem). *Let  $C$  be a closed and bounded convex subset of  $\mathbb{R}^n$ . Let  $\text{ext}(C)$  be the set of extreme points of  $C$ . Then  $C = \text{conv}(\text{ext}(C))$ .*

*Proof.* The inclusion  $C \supseteq \text{conv}(\text{ext}(C))$  is clearly true. We have to show that  $C \subseteq \text{conv}(\text{ext}(C))$ , namely that any point in  $C$  can be written as a convex combination of elements in  $\text{ext}(C)$ . We proceed by induction on the dimension of  $C$ . The claim is clearly true if  $C$  is a point (zero-dimensional). Assume  $C$  is a convex subset of  $\mathbb{R}^n$  of dimension  $k$ . By considering the affine space of dimension  $k$  that contains  $C$ , we can think of  $C$  as a full-dimensional convex set in  $\mathbb{R}^k$ . Let  $v$  be an arbitrary vector in  $\mathbb{R}^k$  and consider the line  $L = \{x + \alpha v, \alpha \in \mathbb{R}\}$ . Since  $C$  is closed and bounded we know that  $C \cap L$  is a segment; let  $x_1, x_2$  be its two extreme points and note that  $x \in \text{conv}(\{x_1, x_2\})$ . Observe that  $x_1, x_2 \in C \setminus \text{int}(C)$ . Thus by Proposition 1.2(ii) they lie on low-dimensional faces  $F_1$  and  $F_2$  of  $C$ . By using the induction hypothesis on  $x_i \in F_i$  (for  $i = 1, 2$ ) we know that  $x_i$  is a convex combination of the extreme points of  $F_i$ . By Proposition 1.2(i) we know that the extreme points of  $F_i$  are extreme points of  $C$ . Thus since  $x_1$  and  $x_2$  are convex combinations of extreme points of  $C$ , and  $x$  is a convex combination of  $\{x_1, x_2\}$  the claim follows.  $\square$