1 Convexity

Recall that if $x^*$ is a critical point of $f : \mathbb{R}^n \to \mathbb{R}$, and $Hf(x) \succeq 0$ for all $x \in \mathbb{R}^n$, then $x^*$ is a global minimizer.

This is a very convenient property for a function $f$ to have. We will call such functions convex functions. Almost. We’ll get to the definition of a convex function in the next lecture, and it will be more general; it won’t require having second derivatives.

Today we will begin heading in that direction by talking about convex sets.

First, some notation: given two points $x, y \in \mathbb{R}^n$, we write $[x, y]$ for the line segment whose endpoints are $x$ and $y$. (This generalizes the notation $[a, b]$ for the closed interval in $\mathbb{R}$ with endpoints $a$ and $b$.) The line segment $[x, y]$ has a convenient parametrization:

$$[x, y] = \{tx + (1 - t)y : 0 \leq t \leq 1\}.$$  

A set $S \subseteq \mathbb{R}^n$ is convex if, whenever, $x, y \in S$, we have $[x, y] \subseteq S$.

In the examples below, the set on the right is not convex: the endpoints of the dashed segment are in $S$, but some points in the interior are not. The set on the left is convex, though to check this, we would have to verify the definition for all possible segments.

2 Examples of convex sets

The empty set $\emptyset$, a single point $\{x\}$, and all of $\mathbb{R}^n$ are all convex sets.

For any $a \in \mathbb{R}^n$ and $b \in \mathbb{R}$, the half-spaces $\{x \in \mathbb{R}^n : a \cdot x \geq b\}$ and $\{x \in \mathbb{R}^n : a \cdot x > b\}$ are convex.

---

1This document comes from the Math 484 course webpage: https://faculty.math.illinois.edu/~mlavrov/courses/484-fall-2018.html
Proof. This is a good example of how we might prove that a set is convex.

Let $H$ be the closed half-space \( \{x \in \mathbb{R}^n : a \cdot x \geq b\} \). We pick two arbitrary points \( x, y \in H \). Our goal is to show that \( [x, y] \subseteq H \).

To do so, take an arbitrary \( t \in [0, 1] \). Since \( x \in H \), we have \( a \cdot x \geq b \), so \( a \cdot (tx) = t(a \cdot x) \geq tb \).

(Here, we use \( t \geq 0 \) so that the inequality doesn’t switch direction.)

Since \( y \in H \), we have \( a \cdot y \geq b \), so \( a \cdot ((1 - t)y) = (1 - t)(a \cdot y) \geq (1 - t)b \).

(Here, we use \( 1 - t \geq 0 \) so that the inequality doesn’t switch direction.)

Adding these two inequalities together, we get
\[
a \cdot (tx + (1 - t)y) = a \cdot (tx) + a \cdot ((1 - t)y) \geq tb + (1 - t)b = b.
\]

This shows that \( tx + (1 - t)y \in H \). This is true for any \( t \in [0, 1] \), so all of \( [x, y] \) is contained in \( H \). Therefore \( H \) is convex.

The ball \( B(x, r) = \{y \in \mathbb{R}^n : \|x - y\| < r\} \) is convex. This is also verified in the same way, though the proof is a bit more obnoxious. Take \( y, z \in B(x, r) \) and \( t \in [0, 1] \). Then
\[
\|x - (ty + (1 - t)z)\| = \|t(x - y) + (1 - t)(x - z)\|
\leq \|t(x - y)\| + \|(1 - t)(x - z)\|
= t\|x - y\| + (1 - t)\|x - z\|
\leq tr + (1 - t)r = r,
\]
so \( [x, y] \subseteq B(x, r) \).

If \( C_1 \) and \( C_2 \) are convex sets, so is their intersection \( C_1 \cap C_2 \); in fact, if \( C \) is any collection of convex sets, then \( \bigcap C \) (the intersection of all of them) is convex. The proof is short: if \( x, y \in \bigcap C \), then \( x, y \in C \) for each \( C \in C \). Therefore \( [x, y] \subseteq C \) for each \( C \in C \), which means \( [x, y] \subseteq \bigcap C \).

This gives us lots more examples, because we can take intersections of all of our previous examples. In particular, any set defined by a bunch of linear equations and inequalities is convex.

### 3 Convex combinations

A convex combination of points \( x^{(1)}, x^{(2)}, \ldots, x^{(k)} \in \mathbb{R}^n \) is a “weighted average”: a linear combination
\[
\lambda_1 x^{(1)} + \lambda_2 x^{(2)} + \cdots + \lambda_k x^{(k)}
\]
where \( \lambda_1 + \lambda_2 + \cdots + \lambda_k = 1 \) and \( \lambda_1, \ldots, \lambda_k \geq 0 \).

The convex hull \( \text{conv}(S) \) of a set of points \( S \) is sometimes defined as the set of all convex combinations of points from \( S \). It’s also sometimes defined as the smallest convex set containing \( S \); we’ll prove those are equivalent in a bit.

In the plane, you can visualize \( \text{conv}(S) \) as the interior of a rubber band stretched around points in \( S \).
Here is an example of the convex hull of three points \( \text{conv}\{\mathbf{x}^{(1)}, \mathbf{x}^{(2)}, \mathbf{x}^{(3)}\} \):

![Convex Hull Example](image)

The definition of convex sets generalizes to the following result:

**Theorem 3.1.** If \( S \) is a convex set and \( \mathbf{x}^{(1)}, \mathbf{x}^{(2)}, \ldots, \mathbf{x}^{(k)} \in S \), then any convex combination \( \lambda_1 \mathbf{x}^{(1)} + \lambda_2 \mathbf{x}^{(2)} + \cdots + \lambda_k \mathbf{x}^{(k)} \) is also contained in \( S \).

**Proof.** The proof is by induction on \( k \): the number of terms in the convex combination.

When \( k = 1 \), this just says that each point of \( S \) is a point of \( S \). When \( k = 2 \), the statement of the theorem is the definition of a convex set: the set of convex combinations \( \lambda_1 \mathbf{x} + \lambda_2 \mathbf{y} \) is just the line segment \([\mathbf{x}, \mathbf{y}]\).

Now assume all length-\((k - 1)\) combinations are contained in \( S \), and take a length-\( k \) combination of points in \( S \):

\[
\lambda_1 \mathbf{x}^{(1)} + \lambda_2 \mathbf{x}^{(2)} + \cdots + \lambda_k \mathbf{x}^{(k)}.
\]

By the inductive hypothesis, we know that

\[
y = \frac{\lambda_1}{\lambda_1 + \cdots + \lambda_{k-1}} \mathbf{x}^{(1)} + \frac{\lambda_2}{\lambda_1 + \cdots + \lambda_{k-1}} \mathbf{x}^{(2)} + \cdots + \frac{\lambda_{k-1}}{\lambda_1 + \cdots + \lambda_{k-1}} \mathbf{x}^{(k-1)}
\]

is in \( S \). (This is only defined if \( \lambda_1 + \cdots + \lambda_{k-1} \neq 0 \); but if it’s 0, then \( \lambda_k \) is the only nonzero coefficient, so we effectively had a length-1 convex combination to begin with.) But now, the original convex combination can be written as

\[
\lambda_1 \mathbf{x}^{(1)} + \lambda_2 \mathbf{x}^{(2)} + \cdots + \lambda_k \mathbf{x}^{(k)} = (\lambda_1 + \cdots + \lambda_{k-1})y + \lambda_k \mathbf{x}^{(k)}
\]

which lies on the line segment \([y, \mathbf{x}^{(k)}]\), and therefore it is in \( S \) by the definition of a convex set.

By induction, convex combinations of all size must be contained in \( S \).

As a corollary, the other definition of \( \text{conv}(S) \) we saw is equivalent to the first:

**Corollary 3.1.** The convex hull \( \text{conv}(S) \) is the smallest convex set containing \( S \).

**Proof.** First of all, \( \text{conv}(S) \) contains \( S \): for every \( \mathbf{x} \in S \), \( \mathbf{1x} \) is a convex combination of size 1, so \( \mathbf{x} \in \text{conv}(S) \).

Second, \( \text{conv}(S) \) is a convex set: if we take \( \mathbf{x}, \mathbf{y} \in \text{conv}(S) \) which are the convex combinations of points in \( S \), then \( t \mathbf{x} + (1-t) \mathbf{y} \) can be expanded to get another convex combinations of points in \( S \).

All convex sets containing \( S \) must contain \( \text{conv}(S) \), and \( \text{conv}(S) \) is itself a convex set containing \( S \); therefore it’s the smallest such set.