Math 484: Nonlinear Programming[1] | Mikhail Lavrov

Chapter 2, Lecture 1: Convex sets

February 4, 2019 | *University of Illinois at Urbana-Champaign*

1 Convexity

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

Our proof worked by being able to compare $\mathbf{x}^*$ to any other point $\mathbf{x} \in \mathbb{R}^n$ along the line through $\mathbf{x}$ and $\mathbf{x}^*$, and applying the one-dimensional second derivative test. So when the domain of f is not all of $\mathbb{R}^n$, we may run into trouble: if the line through $\mathbf{x}$ and $\mathbf{x}^*$ leaves the domain of f, then the second derivative test doesn't work.

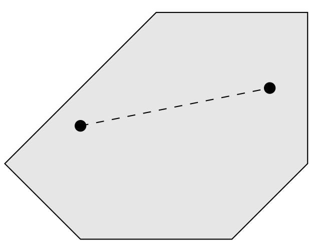
A convex set is, informally, the kind of set where this problem doesn't occur, and we can still make conclusions about global minimizers in the same way as for $\mathbb{R}^n$. Here's how we define this carefully.

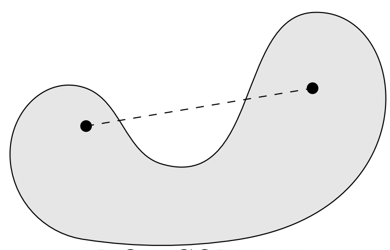
Given two points $\mathbf{x}, \mathbf{y} \in \mathbb{R}^n$, we write $[\mathbf{x}, \mathbf{y}]$ for the line segment whose endpoints are $\mathbf{x}$ and $\mathbf{y}$. (This generalizes the notation $[a, b]$ for the closed interval in $\mathbb{R}$ with endpoints a and b.) The line segment $[\mathbf{x},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

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

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

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 $\varnothing$, a single point $\{\mathbf{x}\}$, and all of $\mathbb{R}^n$ are all convex sets.

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

[1] This document comes from the Math 484 course webpage: https://faculty.math.illinois.edu/~mlavrov/courses/484-spring-2019.html

Proof. This is a good example of how we might prove that a set is convex.

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

To do so, take an arbitrary $t \in [0, 1]$. Since $\mathbf{x} \in H$, we have $\mathbf{a} \cdot \mathbf{x} \geq b$, so $\mathbf{a} \cdot (t\mathbf{x}) = t(\mathbf{a} \cdot \mathbf{x}) \geq tb$. (Here, we use $t \geq 0$ so that the inequality doesn't switch direction.)

Since $\mathbf{y} \in H$, we have $\mathbf{a} \cdot \mathbf{y} \geq b$, so $\mathbf{a} \cdot ((1-t)\mathbf{y}) = (1-t)(\mathbf{a} \cdot \mathbf{y}) \geq (1-t)b$. (Here, we use $1 - t \geq 0$ so that the inequality doesn't switch direction.)

Adding these two inequalities together, we get

$$\mathbf{a} \cdot (t\mathbf{x} + (1-t)\mathbf{y}) = \mathbf{a} \cdot (t\mathbf{x}) + \mathbf{a} \cdot ((1-t)\mathbf{y}) \geq tb + (1-t)b = b.$$

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

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

$$\begin{aligned}
\|\mathbf{x} - (t\mathbf{y} + (1-t)\mathbf{z})\| &= \|t(\mathbf{x} - \mathbf{y}) + (1-t)(\mathbf{x} - \mathbf{z})\| \\
&\leq \|t(\mathbf{x} - \mathbf{y})\| + \|(1-t)(\mathbf{x} - \mathbf{z})\| \\
&= t\|\mathbf{x} - \mathbf{y}\| + (1-t)\|\mathbf{x} - \mathbf{z}\| \\
&\leq tr + (1-t)r = r,
\end{aligned}$$

so $[\mathbf{x}, \mathbf{y}] \subseteq B(\mathbf{x}, r)$.

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

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 $\mathbf{x}^{(1)}, \mathbf{x}^{(2)}, \ldots, \mathbf{x}^{(k)} \in \mathbb{R}^n$ is a "weighted average": a linear combination

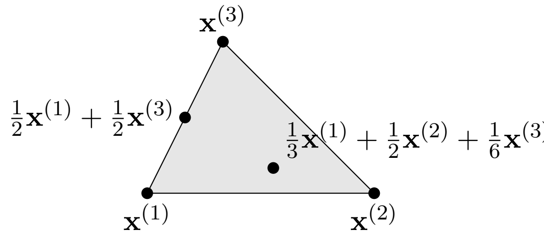
$$\lambda_1 \mathbf{x}^{(1)} + \lambda_2 \mathbf{x}^{(2)} + \cdots + \lambda_k \mathbf{x}^{(k)}$$

where $\lambda_1 + \lambda_2 + \cdots + \lambda_k = 1$ and $\lambda_1, \ldots, \lambda_k \geq 0$.

The *convex hull* $\mathrm{conv}(S)$ 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 $\mathrm{conv}(S)$ as the interior of a rubber band stretched around points in S.

Here is an example of the convex hull of three points $\operatorname{conv}\{\mathbf{x}^{(1)}, \mathbf{x}^{(2)}, \mathbf{x}^{(3)}\}$:



The definition of convex sets generalizes to the following result:

Theorem 3.1. *If S is a convex set and $\mathbf{x}^{(1)}, \mathbf{x}^{(2)}, \dots, \mathbf{x}^{(k)} \in S$, 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

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

is in S. (This is only defined if $\lambda_1 + \cdots + \lambda_{k-1} \neq 0$; but if it's 0, then λ_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

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

which lies on the line segment $[\mathbf{y}, \mathbf{x}^{(k)}]$, 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 $\operatorname{conv}(S)$ we saw is equivalent to the first:

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

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

Second, $\operatorname{conv}(S)$ is a convex set: if we take $\mathbf{x}, \mathbf{y} \in \operatorname{conv}(S)$ 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 $\operatorname{conv}(S)$, and $\operatorname{conv}(S)$ is itself a convex set containing S; therefore it's the smallest such set. ☐