Math 484: Nonlinear Programming[1] Mikhail Lavrov

Chapter 5, Lecture 2: Separating and Supporting Inequalities

March 6, 2019 *University of Illinois at Urbana-Champaign*

1 The separation theorem

We apply the obtuse angle criterion to prove a useful property of closed convex sets: given any point outside the set, there is a linear inequality that distinguishes every element of the set from that point.

Theorem 1.1 (Separation theorem; Theorem 5.1.5). *If $C \subseteq \mathbb{R}^n$ is a closed convex set and $\mathbf{y} \notin C$, then there is a linear inequality separating $\mathbf{y}$ from C: we can choose $\mathbf{a} \in \mathbb{R}^n$, $b \in \mathbb{R}$ such that*

- $\mathbf{a} \cdot \mathbf{x} \leq b$ *for all* $\mathbf{x} \in C$, *but*
- $\mathbf{a} \cdot \mathbf{y} > b$.

In particular, $\mathbf{a} \cdot \mathbf{x} < \mathbf{a} \cdot \mathbf{y}$ for all $\mathbf{x} \in C$. (But the two-inequality version says something slightly stronger: it prevents $\mathbf{a} \cdot \mathbf{x}$ from even getting arbitrarily close to $\mathbf{a} \cdot \mathbf{y}$.)

Proof. This is just the obtuse angle criterion in another form.

Because C is closed, there is a point $\mathbf{x}^* \in C$ closest to $\mathbf{y}$. Then by the obtuse angle criterion,

$$(\mathbf{y} - \mathbf{x}^*) \cdot (\mathbf{x} - \mathbf{x}^*) \leq 0$$

for all $\mathbf{x} \in C$. We can rearrange this as

$$(\mathbf{y} - \mathbf{x}^*) \cdot \mathbf{x} \leq (\mathbf{y} - \mathbf{x}^*) \cdot \mathbf{x}^*,$$

which proves that $\mathbf{a} \cdot \mathbf{x} \leq b$ for all $\mathbf{x} \in C$¡ if we define $\mathbf{a} = \mathbf{y} - \mathbf{x}^*$ and $b = (\mathbf{y} - \mathbf{x}^*) \cdot \mathbf{x}^*$.

Now we have to prove $\mathbf{a} \cdot \mathbf{y} > b$ for the same $\mathbf{a}$ and b. Since $\mathbf{y} \notin C$, we know that $\mathbf{x}^* \neq \mathbf{y}$, so

$$(\mathbf{y} - \mathbf{x}^*) \cdot (\mathbf{y} - \mathbf{x}^*) = \|\mathbf{y} - \mathbf{x}^*\|^2 > 0.$$

We can rearrange this as

$$(\mathbf{y} - \mathbf{x}^*) \cdot \mathbf{y} > (\mathbf{y} - \mathbf{x}^*) \cdot \mathbf{x}^*,$$

or $\mathbf{a} \cdot \mathbf{y} > b$. □

Note that we could even get the inequality to have the form

$$\mathbf{a} \cdot \mathbf{x} < c < \mathbf{a} \cdot \mathbf{y}$$

for all $\mathbf{x} \in C$, by setting $c = \frac{b + \mathbf{a} \cdot \mathbf{y}}{2}$.

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

Can we still separate points outside C from C, if C is not closed? Well, not always: as you may suspect, trying to separate a point on the boundary of C from the rest of C causes a few problems.

To make this happen, we define a new set operation on top of int, ext, and bd. Given a set $S \subseteq \mathbb{R}^n$, we define the closure $\mathrm{cl}(S)$ to be the set $\mathrm{int}(S) \cup \mathrm{bd}(S)$.

The closure of S turns S into a closed set: informally, we just add in all the boundary points that were missing from S. (Technically, we ought to check that no new missing boundary points appear; this is too boring for us to spend time on it.)

Moreover, we can check that if $C \subseteq \mathbb{R}^n$ is a convex set, then so is $\mathrm{cl}(C)$. But now $\mathrm{cl}(C)$ is a closed set; so given a point $\mathbf{y} \notin \mathrm{cl}(C)$, we can use the separation theorem to separate $\mathbf{y}$ from $\mathrm{cl}(C)$ (and therefore from C).

Another way to say this is that the separation theorem applies to any convex set C and any point $\mathbf{y} \notin C$, as long as $\mathbf{y}$ is not a boundary point of C.

2 The Bolzano–Weierstrass theorem

We are getting into some pretty technical territory here. Before we proceed, we need a pretty technical theorem about closed and bounded sets.

Theorem 2.1 (Bolzano–Weierstrass)**.** *Let $S \subseteq \mathbb{R}^n$ be a closed and bounded set, and let $\mathbf{x}^{(1)}, \mathbf{x}^{(2)}, \dots$ be any sequence of points in S.*

Then we can pick out a subsequence of these points that has a limit $\mathbf{x}^* \in S$. Formally, we can find indices $i_1 < i_2 < i_3 < \dots$ such that the sequence $\mathbf{x}^{(i_1)}, \mathbf{x}^{(i_2)}, \mathbf{x}^{(i_3)}, \dots$ converges to $\mathbf{x} \in S$.

NOTE: In class, we didn't follow the notes as planned, but went on to discuss the statement of this theorem. So I've edited the notes to summarize that discussion, and to include only the topics not covered in the next lecture.

By convergence here, we mean that there's a point $\mathbf{x}^* \in S$ such that the distance

$$\left\|\mathbf{x}^{(i_k)} - \mathbf{x}^*\right\|$$

can be made arbitrarily small by taking k large enough. That is, for any $\epsilon > 0$ (any notion of "arbitrarily small") there is some integer $n > 0$ such that if $k \geq n$ (if k is "large enough") then $\left\|\mathbf{x}^{(i_k)} - \mathbf{x}^*\right\| < \epsilon$.

The two parts of the hypothesis (that S is closed, and that S is bounded) correspond to two parts of the conclusion:

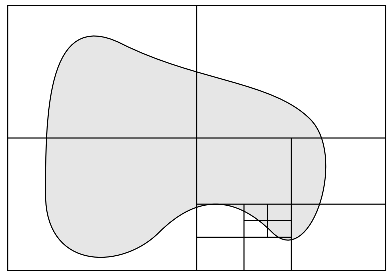
- If S is bounded, we can pick out a subsequence of $\mathbf{x}^{(1)}, \mathbf{x}^{(2)}, \dots$ that converges.

  Really, we can think of this part as a statement about sequences: given a sequence $\mathbf{x}^{(1)}, \mathbf{x}^{(2)}, \dots$, if there is some $M > 0$ such that $\|\mathbf{x}^{(k)}\| < M$ for all k, then the sequence has a convergent subsequence.

- If S is closed, the limit of this subsequence will also be an element of S.

Due to time, we won't prove the theorem in class, but I've left one proof of the theorem in the notes.

Proof. If S is bounded, we can draw a box around S. I'll give you a picture in $\mathbb{R}^2$, so the box will be a rectangle; in general, of course, it's n-dimensional.



Divide the box into 2^n parts. (In $\mathbb{R}^2$, we divide it in four parts.) Since the sequence of points we have is infinite, and there's finitely many parts, there's some part that has infinitely many points of the sequence.

Now zoom in on that part. It's another, smaller box; we divide it into 2^n parts again, and again we pick a part that has infinitely many points of the sequence. Keep doing this forever. We get a sequence of nested boxes shrinking to arbitrarily small size, each with infinitely many points of our sequence in them.

Finally, to get the subsequence we want: pick any i_1 such that $\mathbf{x}^{(i_1)}$ is in the first box. Then pick any $i_2 > i_1$ such that $\mathbf{x}^{(i_2)}$ is in the second box. Then pick any $i_3 > i_2$ such that $\mathbf{x}^{(i_2)}$ is in the third box. Keep going; at each step, there's a way to make the choice, because we have infinitely many options.

The size of the boxes we pick goes to 0, so the intersection of all these boxes is just a single point $\{\mathbf{x}^*\}$. We claim that $\mathbf{x}^*$ is actually the limit of the subsequence we chose.

To see this: for any $\epsilon > 0$, we can choose n such that the n^{th} box has diagonal shorter than ϵ. This means that any two points inside the k^{th} box are a distance of less than ϵ apart. Since $\mathbf{x}^*$ is in the n^{th} box (it's in all the boxes) and $\mathbf{x}^{(i_k)}$ is in the n^{th} box for all $k \geq n$, we have

$$\left\|\mathbf{x}^{(i_k)} - \mathbf{x}^*\right\| < \epsilon$$

for all $k \geq n$.

Finally, we show that $\mathbf{x}^* \in S$. Here is where we use the assumption that S is a closed set.

Suppose for the sake of contradiction that $\mathbf{x}^* \notin S$. Then $\mathbf{x}^*$ is definitely not an interior point of S (interior points are always in S), nor is it a boundary point of S (because S is closed, all boundary points of S are in S). So $\mathbf{x}^*$ is an exterior point of S. This means there's some positive radius $r > 0$ around $\mathbf{x}^*$ such that all points within that radius of $\mathbf{x}^*$ are still outside of S.

Take $\epsilon = r$. Then by the definition of convergence, there is some k such that $\|\mathbf{x}^{(k)} - \mathbf{x}^*\| < r$. This means that $\mathbf{x}^{(k)} \notin S$, because all points with in distance r of $\mathbf{x}^*$ are outside S. But we assumed that the sequence we started with was entirely contained in S. Contradiction! Therefore $\mathbf{x}^* \in S$. $\square$