Math 484: Nonlinear Programming[1] | Mikhail Lavrov

Chapter 5, Lecture 1: The Obtuse Angle Criterion

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

1 The closest point problem

We can think of the material from Chapter 4 as a special case of the *closest point problem*: given a set $S \subseteq \mathbb{R}^n$, and a point $\mathbf{y} \notin S$, the goal is to find the point of S closest to $\mathbf{y}$. In other notation:

$$\begin{aligned} \underset{\mathbf{x}\in\mathbb{R}^n}{\text{minimize}} \quad & \|\mathbf{x}-\mathbf{y}\| \\ \text{subject to} \quad & \mathbf{x} \in S. \end{aligned}$$

In Chapter 4, the set S was a vector subspace of $\mathbb{R}^n$, or a translation of one. Now, we're going to consider more general sets.

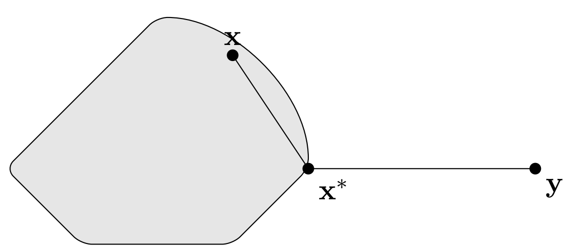
1.1 Uniqueness and the obtuse angle criterion

When solving this problem for vector spaces, the key was an orthogonality condition between $\mathbf{y}-\mathbf{x}$ and vectors of that subspace. A weaker form of that condition is the obtuse angle criterion, which holds for any convex set:

Theorem 1.1 (Obtuse angle criterion; Theorem 5.1.1)**.** *Let C be a convex set, and let $\mathbf{y}$ be a point outside of S. Then a point $\mathbf{x}^* \in C$ is the point of C closest to $\mathbf{y}$ if and only if*

$$(\mathbf{y}-\mathbf{x}^*)\cdot(\mathbf{x}-\mathbf{x}^*) \le 0 \qquad \textit{for all } \mathbf{x} \in C.$$

Why is this the obtuse angle criterion? Because it says that the angle between $\mathbf{y}-\mathbf{x}^*$ (the vector pointing from $\mathbf{x}^*$ to $\mathbf{y}$) and $\mathbf{x}-\mathbf{x}^*$ (the vector pointing from $\mathbf{x}^*$ to $\mathbf{x}$) is a right angle or an obtuse angle. See the diagram below:



Proof. Let $\mathbf{x}^*$ be a point of C that we think is the closest point, and let $\mathbf{x}$ be another point of C.

The proof mirrors our proof of a similar result for vector spaces. Once again, we are going to look at the restriction of $f(\mathbf{x}) = \|\mathbf{x}-\mathbf{y}\|^2$ to a line. In this case, we're going to consider the line through $\mathbf{x}^*$ and $\mathbf{x}$, and so we define the restriction

$$\phi(t) = f(\mathbf{x}^* + t(\mathbf{x}-\mathbf{x}^*)) = f(t\mathbf{x} + (1-t)\mathbf{x}^*).$$

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

Because C is convex, for all $t \in [0,1]$, the point $t\mathbf{x} + (1-t)\mathbf{x}^*$ is another point of C. We have:

- If $\mathbf{x}^*$ is the point of C closest to $\mathbf{y}$, then

$$\phi(t) = \|(t\mathbf{x} + (1-t)\mathbf{x}^*) - \mathbf{y}\|^2 \geq \|\mathbf{x}^* - \mathbf{y}\|^2 \geq \phi(0)$$

  for all $t \in [0,1]$, so 0 is the global minimizer of ϕ on $[0,1]$.

- If 0 is the global minimizer of ϕ on $[0,1]$ no matter which point $\mathbf{x} \in C$ we choose, then in particular $\phi(0) \leq \phi(1)$, or $\|\mathbf{x}^* - \mathbf{y}\| \leq \|\mathbf{x} - \mathbf{y}\|$, so $\mathbf{x}^*$ is a point of C closest to $\mathbf{y}$.

As a result, it's enough to figure out when 0 is the global minimizer of ϕ on $[0,1]$.

As we did before, we expand

$$\begin{aligned}\phi(t) &= \|\mathbf{x}^* + t(\mathbf{x} - \mathbf{x}^*) - \mathbf{y}\|^2 \\ &= \|\mathbf{x}^* - \mathbf{y}\|^2 + 2t(\mathbf{x} - \mathbf{x}^*) \cdot (\mathbf{x}^* - \mathbf{y}) + t^2\|\mathbf{x} - \mathbf{x}^*\|^2.\end{aligned}$$

So

$$\phi'(t) = 2\left[(\mathbf{x} - \mathbf{x}^*) \cdot (\mathbf{x}^* - \mathbf{y})\right] + 2t\|\mathbf{x} - \mathbf{x}^*\|^2 = \phi'(0) + 2t\|\mathbf{x} - \mathbf{x}^*\|^2.$$

When $\phi'(0) \geq 0$, we have $\phi'(t) \geq \phi'(0) \geq 0$ for all positive t. In this case, ϕ is increasing on $[0,1]$: so $t = 0$ is the global minimizer on $[0,1]$.

When $\phi'(0) < 0$, instead $\phi(t)$ is decreasing when t is small. This means that for sufficiently small t, $\phi(t) < \phi(0)$, and $t = 0$ is not the global minimizer on $[0,1]$.

Since $\phi'(0) = 2\left[(\mathbf{x} - \mathbf{x}^*) \cdot (\mathbf{x}^* - \mathbf{y})\right]$, we have $\phi'(0) \geq 0$ precisely when $(\mathbf{y} - \mathbf{x}^*) \cdot (\mathbf{x} - \mathbf{x}^*) \leq 0$: this is the obtuse angle criterion. □

The obtuse angle criterion is a generalization of our previous results for subspaces. In the textbook, Corollary 5.1.2 shows that from the obtuse angle criterion, we can deduce those results. (This is also a good exercise for you to try on your own: the key idea is that we can compare $\mathbf{x}^*$ to points $\mathbf{x}, \mathbf{x}'$ in the subspace that are "in opposite directions" from $\mathbf{x}^*$.)

In particular, we can show that the closest point, if it exists, must be unique.

Corollary 1.1 (Corollary 5.1.4)**.** *Let C be a convex set, and let $\mathbf{y}$ be a point outside of S. Then the minimum of $\|\mathbf{x} - \mathbf{y}\|$ for $\mathbf{x} \in C$ is achieved at most at one point of C.*

Proof. The textbook's approach is by using the obtuse angle criterion, which is good practice. Suppose that $\mathbf{x}^{(1)}, \mathbf{x}^{(2)}$ are both closest points in C to $\mathbf{y}$. Then we have

$$\begin{aligned}(\mathbf{y} - \mathbf{x}^{(1)}) \cdot (\mathbf{x}^{(2)} - \mathbf{x}^{(1)}) \leq 0, \\ (\mathbf{y} - \mathbf{x}^{(2)}) \cdot (\mathbf{x}^{(1)} - \mathbf{x}^{(2)}) \leq 0,\end{aligned}$$

by applying the obtuse angle criterion first to $\mathbf{x}^{(1)}$ (with $\mathbf{x}^{(2)}$ as the other point in C) and then to $\mathbf{x}^{(2)}$ (with $\mathbf{x}^{(1)}$ as the other point in C).

Rewrite the second inequality as $(\mathbf{x}^{(2)} - \mathbf{y}) \cdot (\mathbf{x}^{(2)} - \mathbf{x}^{(1)}) \leq 0$, to make their second factors the same, and then add them together. We get

$$\left((\mathbf{y} - \mathbf{x}^{(1)}) + (\mathbf{x}^{(2)} - \mathbf{y})\right) \cdot (\mathbf{x}^{(2)} - \mathbf{x}^{(1)}) \leq 0$$

which simplifies to $\|\mathbf{x}^{(1)} - \mathbf{x}^{(2)}\|^2 \le 0$. This can only happen if $\mathbf{x}^{(1)} = \mathbf{x}^{(2)}$.

There's another, shorter approach: since C is convex, we know that $\frac{1}{2}\mathbf{x}^{(1)} + \frac{1}{2}\mathbf{x}^{(2)} \in C$, and you can check that this point is closer to $\mathbf{y}$ than either $\mathbf{x}^{(1)}$ or $\mathbf{x}^{(2)}$ were, unless they are equal. □

It is a harder proof than I want to get into, but in fact convex sets are the most general sets for which this holds: if S is not convex, then we can find a point $\mathbf{y} \notin S$ for which two or more points of S "tie" for closeness.

1.2 Existence and boundary points

Now we know when the solution to the problem is unique. When can we guarantee existence?

To do this, we have to go back to some details of open and closed sets we introduced earlier in the semester. Recall that the open ball $B(\mathbf{x}, r)$, where $\mathbf{x} \in \mathbb{R}^n$, is the set of all points at distance less than r to $\mathbf{x}$:

$$B(\mathbf{x}, r) = \{\mathbf{y} \in \mathbb{R}^n : \|\mathbf{x} - \mathbf{y}\| < r\}.$$

Given a set $S \subseteq \mathbb{R}^n$, points of $\mathbb{R}^n$ are divided into three types, relative to S:

- *interior points of* S: points $\mathbf{x} \in S$ for which some open ball $B(\mathbf{x}, r)$ is entirely contained in S.

  We write $\text{int}(S)$ for the set of all interior points of S.

- *exterior points of* S: points $\mathbf{x} \notin S$ for which some open ball $B(\mathbf{x}, r)$ contains no points of S.

  We write $\text{ext}(S)$ for the set of all exterior points of S.

- *boundary points of* S: all other points. Here, for any open ball $B(\mathbf{x}, r)$ we take, both points in S and points outside S will be included.

  We write $\text{bd}(S)$ for the set of all boundary points of S.

In general, we know that $\text{int}(S)$ is a subset of S, while $\text{ext}(S)$ is disjoint to S. For $\text{bd}(S)$, we don't know in advance; boundary points could go either way.

We say that a set S is *open* if $S = \text{int}(S)$ (that is, if no boundary point is included). A set S is *closed* if $S = \text{int}(S) \cup \text{bd}(S)$ (that is, if all boundary points are included).

Theorem 1.2 (Theorem 5.1.3)**.** *If S is a nonempty closed set, then the closest point problem always has an optimal solution.*

Proof. Recall a fact from Chapter 1: a continuous function always has a global minimizer when we're working over a nonempty, closed and bounded set. Here, S is closed and nonempty; it is not necessarily bounded. But we can fix that.

Let $\mathbf{x}^{(0)}$ be any point in S. If $\mathbf{x}^{(0)}$ is the closest point to $\mathbf{y}$, we are done. Otherwise, we are looking for a point closer to $\mathbf{y}$ than $\mathbf{x}^{(0)}$ is; so it's enough to minimize, not over S, but over

$$S' = \{\mathbf{x} \in S : \|\mathbf{x} - \mathbf{y}\| \le \|\mathbf{x}^{(0)} - \mathbf{y}\|\}.$$

The set S' is still closed.[2] It's still nonempty, because $\mathbf{x}^{(0)} \in S'$. But now it's bounded, because it's contained in a closed ball around $\mathbf{y}$.

Therefore the function $f(\mathbf{x}) = \|\mathbf{x} - \mathbf{y}\|$ has a global minimizer on S', which is the closest point that we wanted. $\square$

So when S is closed, we're happy. What if S is not closed?

Well, in general, solutions to the closest point problem ought to be boundary points of S. If $\mathbf{x} \in \text{int}(S)$, then there's an open ball $B(\mathbf{x}, r)$ contained in S; in particular, for some small ϵ, $\mathbf{x} + \epsilon(\mathbf{y} - \mathbf{x})$ is still in S, and it's closer to $\mathbf{y}$ than $\mathbf{x}$ is.

This is fine if S is closed. But if S is not closed, then there are some points in $\text{bd}(S)$ which are not in S, and that's a problem. Let $\mathbf{y}$ be an element of $\text{bd}(S)$ which is not in S.

Now there is no closest point to $\mathbf{y}$: for any $\mathbf{x} \in S$, we can find an $\mathbf{x}'$ closer to $\mathbf{y}$ than $\mathbf{x}$ is. To do this, take the open ball $B(\mathbf{y}, \|\mathbf{x} - \mathbf{y}\|)$: since $\mathbf{y}$ is a boundary point of S, this open ball contains some point $\mathbf{x}' \in S$, and $\|\mathbf{x}' - \mathbf{y}\| < \|\mathbf{x} - \mathbf{y}\|$ by definition of the open ball.

Altogether, we conclude that if we want to guarantee that the closest point problem has a unique solution, we should be looking at closed convex sets.

[2] If we're being careful: S' is the intersection of S with the closed ball $\{\mathbf{x} \in \mathbb{R}^n : \|\mathbf{x} - \mathbf{y}\| \le \|\mathbf{x}^{(0)} - \mathbf{y}\|\}$, and the intersection of two closed sets is closed.