Math 484: Nonlinear Programming[1] Mikhail Lavrov

Chapter 2, Lecture 2: Convex functions

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

1 Definition of convex functions

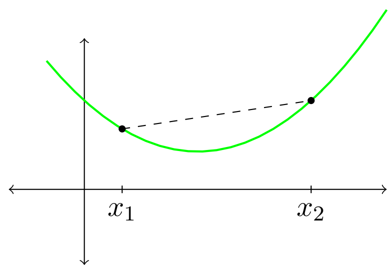
Convex functions $f : \mathbb{R}^n \to \mathbb{R}$ with $Hf(\mathbf{x}) \succeq 0$ for all $\mathbf{x}$, or functions $f : \mathbb{R} \to \mathbb{R}$ with $f''(x) \geq 0$ for all x, are going to be our model of what we want convex functions to be. But we actually work with a slightly more general definition that doesn't require us to say anything about derivatives.

Let $C \subseteq \mathbb{R}^n$ be a convex set. A function $f : C \to \mathbb{R}$ is convex on C if, for all $\mathbf{x}, \mathbf{y} \in C$, the inequality holds that

$$f(t\mathbf{x} + (1-t)\mathbf{y}) \leq tf(\mathbf{x}) + (1-t)f(\mathbf{y}).$$

(We ask for C to be convex so that $t\mathbf{x} + (1-t)\mathbf{y}$ is guaranteed to stay in the domain of f.)

This is easiest to visualize in one dimension:



The point $t\mathbf{x}+(1-t)\mathbf{y}$ is somewhere on the line segment $[\mathbf{x}, \mathbf{y}]$. The left-hand side of the definition, $f(t\mathbf{x} + (1-t)\mathbf{y})$, is just the value of the function at that point: the green curve in the diagram. The right-hand side of the definition, $tf(\mathbf{x}) + (1-t)f(\mathbf{y})$, is the dashed line segment: a straight line that meets f at $\mathbf{x}$ and $\mathbf{y}$.

So, geometrically, the definition says that secant lines of f always lie above the graph of f.

Although the picture we drew is for a function $\mathbb{R} \to \mathbb{R}$, nothing different happens in higher dimensions, because only points on the line segment $[\mathbf{x}, \mathbf{y}]$ (and f's values at those points) play a role in the inequality. In fact, we can check convexity of f by looking at f's restrictions onto lines:

Lemma 1.1. *Let $C \subseteq \mathbb{R}^n$ be a convex set. A function $f : C \to \mathbb{R}$ is convex if and only if, for all $\mathbf{x} \in C$ and $\mathbf{u} \in \mathbb{R}^n$, the function*

$$\phi(t) = f(\mathbf{x} + t\mathbf{u})$$

is a 1-variable convex function in t. (The domain of ϕ is the set of t for which $\mathbf{x} + t\mathbf{u} \in C$.)

Proof. The condition

$$f(\lambda\mathbf{x} + (1-\lambda)\mathbf{y}) \leq \lambda f(\mathbf{x}) + (1-\lambda)f(\mathbf{y})$$

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

can be checked for f by checking the condition

$$\phi(\lambda \cdot 0 + (1-\lambda)\cdot 1) \le \lambda\phi(0) + (1-\lambda)\phi(1)$$

for the restriction $\phi(t) = f(\mathbf{x} + t(\mathbf{y} - \mathbf{x}))$. □

The nice thing about the "secant line" definition of convex functions is that

- All the nice things we've said about functions with $Hf(\mathbf{x}) \succeq 0$, and more, still hold for convex functions in general.
- But we don't have to deal with derivatives to prove them.

For example, we can show the following result.

Theorem 1.1. *If $C \subseteq \mathbb{R}^n$ is a convex set, $f : C \to \mathbb{R}$ is a convex function, and $\mathbf{x}^* \in C$ is a local minimizer of f, then it is a global minimizer.*

Proof. Let's first prove this for convex functions $f : [a, b] \to \mathbb{R}$.

In this case, suppose $x^* \in [a, b]$ is a local minimizer of f: there's some interval $(x^* - r, x^* + r)$ where f stays above $f(x^*)$.

Here's the geometric argument. For any other point $y \in [a, b]$, draw the secant line through $(x^*, f(x^*))$ and $(y, f(y))$. As that secant line passes over either $x^* + \frac{r}{2}$ or $x^* - \frac{r}{2}$, it lies above the graph of f: above $f(x^* \pm \frac{r}{2})$, which is bigger than $f(x^*)$. So the secant line has gone up from $f(x^*)$, which means it has a nonnegative slope. This can only happen if $f(y) \ge f(x^*)$.

Algebraically: given $y \in [a, b]$, choose $t > 0$ small enough that

$$x^* + t(y - x^*) \in (x^* - r, x^* + r).$$

That is, choose t smaller than $\frac{|y - x^*|}{r}$. Also, make sure that $t \le 1$, as required by the definition of a convex function.

By the property of being a local minimizer,

$$\begin{aligned} f(x^*) &\le f(x^* + t(y - x^*)) \\ &= f((1-t)x^* + ty) \\ &\le (1-t)f(x^*) + tf(y) \\ tf(x^*) &\le tf(y) \\ f(x^*) &\le f(y). \end{aligned}$$

The result holds in $\mathbb{R}^n$ by applying Lemma 1.1: if f is convex, so is the restriction of f to any line through $\mathbf{x}^*$, so $\mathbf{x}^*$ is the global minimizer along any such line, and this makes it the global minimizer on the entire domain of f. □

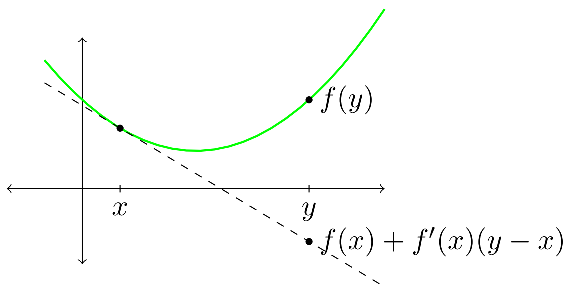
2 Derivatives

We can take the secant-line definition of convex functions and use it to say things about convex functions that *do* have first and second derivatives. We can get back our original definition this way. Along the way, we'll also get a first-derivative condition on convex functions.

We'll stick to the one-dimensional case here: as before, we can use Lemma 1.1 to get similar results for higher-dimensional functions.

2.1 First derivatives

Among functions with continuous first derivatives, convex functions are those which satisfy the tangent line condition. Geometrically, this condition says that f is convex if and only if tangent lines to f are always lower bounds on f:



Algebraically, the statement is as follows:

Theorem 2.1 (Theorem 2.3.5)**.** *Suppose that $f : (a, b) \to \mathbb{R}$ is a function such that $f'(x)$ is continuous on (a, b). Then f is convex if and only if*

$$f(y) \geq f(x) + f'(x)(y - x)$$

for all $x, y \in (a, b)$. For functions $f : C \to \mathbb{R}$ with $C \subseteq \mathbb{R}^n$, this becomes

$$f(\mathbf{y}) \geq f(\mathbf{x}) + \nabla f(\mathbf{x}) \cdot (\mathbf{y} - \mathbf{x}).$$

Proof. First, we will assume f is convex and try to prove the inequality. Take any $x, y \in (a, b)$, and assume $x \neq y$ because otherwise the inequality is already satisfied: it just says $f(x) \geq f(x)$. We have

$$f((1 - t)x + ty) \leq (1 - t)f(x) + tf(y)$$

whenever $0 \leq t \leq 1$, which we can rewrite as

$$\begin{aligned} f(x + t(y - x)) \leq (1 - t)f(x) + tf(y) &\implies f(x + t(y - x)) - f(x) \leq tf(y) - tf(x) \\ &\implies \frac{f(x + t(y - x)) - f(x)}{t(y - x)} \cdot (y - x) \leq f(y) - f(x). \end{aligned}$$

If we take the limit as $t \to 0$, then $t(y - x) \to 0$ as well, which means the left-hand side of this inequality approaches $f'(x) \cdot (y - x)$. The right-hand side does not depend on t, so it remains the same, and we get

$$f'(x) \cdot (y - x) \leq f(y) - f(x) \implies f(y) \geq f(x) + f'(x)(y - x).$$

Next, we will assume that the inequality holds, and try to prove that f is convex.

Let $u, v \in [a, b]$ and let $w = tu + (1-t)v$ with $0 \leq t \leq 1$. Then we have

$$f(u) \geq f(w) + f'(w)(u - w) \quad \text{and} \quad f(v) \geq f(w) + f'(w)(v - w)$$

so if we add t times the first inequality and $(1-t)$ times the second inequality, we get

$$\begin{aligned} tf(u) + (1-t)f(v) &\geq tf(w) + (1-t)f(w) + f'(w)(tu - tw + (1-t)v - (1-t)w) \\ &= f(w) + f'(w)(tu + (1-t)v - w) \\ &= f(w) + f'(w)(w - w) = f(w) \end{aligned}$$

and since $w = tu + (1-t)v$, this is exactly the inequality

$$tf(u) + (1-t)f(v) \geq f(tu + (1-t)v)$$

that proves that f is convex. □

2.2 Second derivatives

Theorem 2.2. *Suppose that $f : (a, b) \to \mathbb{R}$ has a continuous second derivative on (a, b). If $f''(x) \geq 0$ for all $x \in (a, b)$, then f is convex on (a, b).*

In higher dimensions, this turns into the condition $Hf(\mathbf{x}) \succeq 0$.

Proof. Here, we use the second-order Taylor series approximation to f: given $x, y \in C$, there is some $\xi \in [x, y]$ such that

$$f(y) = f(x) + f'(x)(y - x) + f''(\xi)\frac{(y - x)^2}{2}.$$

By using the previous theorem, we can take a shortcut here. If $f''(\xi) \geq 0$, then we get the inequality

$$f(y) \geq f(x) + f'(x)(y - x)$$

and we know that having this inequality for all x, y makes f convex. □