Optimization and Computational Linear Algebra – Brett Bernstein

Recitation 10

It may be helpful if after recitation you try to re-solve these problems by yourself, and use them as additional study problems for the class.

1. Which of the following sets are convex?

   (a) $\{x \in \mathbb{R}^2 : \|x\| = 1\}$
   (b) $\{x \in \mathbb{R}^2 : \|x\| \leq 1\}$
   (c) $\{x \in \mathbb{R}^2 : \|x\| \geq 1\}$
   (d) $\{x \in \mathbb{R}^2 : \|x\| < 1\}$
   (e) $\{x \in \mathbb{R}^2 : v^T x \geq a\}$ for fixed $v \in \mathbb{R}^2$ and $a \in \mathbb{R}$
   (f) $\{x \in \mathbb{R}^2 : v^T x = a\}$ for fixed $v \in \mathbb{R}^2$ and $a \in \mathbb{R}$
   (g) $\{x \in \mathbb{R}^2 : x_2 \geq x_1^2\}$
   (h) $\{x \in \mathbb{R}^2 : x_2 \leq x_1^2\}$

   Solution. b, d, e, f, g

2. For $f : \mathbb{R}^n \to \mathbb{R}$ define the epigraph $\text{epi}(f) \subset \mathbb{R}^{n+1}$ to be the set of all points above the graph of f:
$$\text{epi}(f) := \{(x, t) \in \mathbb{R}^{n+1} : t \geq f(x)\}.$$
   Prove that f is convex if and only if $\text{epi}(f)$ is convex.

   Solution.

   Proof. First assume that f is convex and let $(x_1, t_1), (x_2, t_2) \in \text{epi}(f)$ so that $t_1 \geq f(x_1)$ and $t_2 \geq f(x_2)$. For $\theta \in (0, 1)$, define $z \in \mathbb{R}^{n+1}$ by
$$z := \theta(x_1, t_1) + (1-\theta)(x_2, t_2) = (\theta x_1 + (1-\theta)x_2, \theta t_1 + (1-\theta)t_2).$$
   By convexity we know that
$$\theta t_1 + (1-\theta)t_2 \geq \theta f(x_1) + (1-\theta) f(x_2) \geq f(\theta x_1 + (1-\theta)x_2)$$
   proving $z \in \text{epi}(f)$.

   Conversely, suppose that $\text{epi}(f)$ is convex. Fix $x_1, x_2 \in \mathbb{R}^n$ and $\theta \in (0, 1)$. Since $(x_1, f(x_1)), (x_2, f(x_2)) \in \text{epi}(f)$ we have
$$(\theta x_1 + (1-\theta)x_2, \theta f(x_1) + (1-\theta) f(x_2)) \in \text{epi}(f)$$
   as well. But this means that
$$\theta f(x_1) + (1-\theta) f(x_2) \geq f(\theta x_1 + (1-\theta)x_2)$$
   completing the proof. □

3. Suppose $f, g : \mathbb{R}^n \to \mathbb{R}$ are convex functions. Prove that $h : \mathbb{R}^n \to \mathbb{R}$ defined by $h(x) = \max(f(x), g(x))$ is also convex.

   Solution. The epigraph of h is the intersection of the epigraphs of f and g. Since the intersection of convex sets is convex, the previous question proves h is convex.

4. Suppose $\lim_{t\to 0} f(t) < 0$. Prove there is an $\delta > 0$ such that $f(t) < 0$ for all $0 < |t| < \delta$.

   Solution.

   Proof. Let $\alpha > 0$ be such that $-\alpha = \lim_{t\to 0} f(t)$. By definition of the limit, there is an $\delta > 0$ such that
   $$|f(t) - (-\alpha)| < \alpha/2$$
   whenever
   $$0 < |t - 0| < \delta.$$
   Thus for $0 < |t| < \delta$ we have
   $$f(t) + \alpha \leq |f(t) + \alpha| < \alpha/2.$$
   Subtracting α from both sides shows
   $$f(t) < -\alpha/2$$
   completing the proof. □

5. Suppose you drive through a tunnel that is one mile long and has a 40 miles per hour speed limit. If driving through the tunnel takes one minute, did your velocity necessarily exceed the speed limit?

   Solution. Yes. Let $f(t)$ denote your distance traveled through the tunnel in miles, where t is measured in hours. We assume that f is continuous on $[0, 1/60]$ and differentiable on $(0, 1/60)$. By the mean value theorem we have
   $$f(1/60) - f(0) = f'(\xi)(1/60 - 0),$$
   for some $\xi \in (0, 1/60)$. Then we have
   $$f'(\xi) = 60(f(1/60) - f(0)) = 60.$$

   If we also assume f' is continuous then we can obtain another solution using the fundamental theorem of calculus: Suppose that $f'(t) < 60$ for $t \in (0, 1/60)$. Then we have
   $$1 = f(1/60) - f(0) = \int_0^{1/60} f'(t)\,dt < \int_0^{1/60} 60\,dt = 1,$$
   a contradiction.

6. Let $f : \mathbb{R} \to \mathbb{R}$ be differentiable and suppose we know the value of $f(x)$ and $f'(x)$ for some fixed $x \in \mathbb{R}$.

   (a) Give an approximate value for $f(x+h)$.

   (b) Suppose we know that f is twice differentiable and $|f''(t)| < 3$ for all t between x and $x+h$. Give a bound on the error of your approximation in the previous previous part.

   Solution.

   (a) $f(x+h) \approx f(x) + hf'(x)$. By the definition of the derivative, we know the error in this approximation goes to zero quickly as $h \to 0$. More precisely, we can write

   $$f(x+h) = f(x) + hf'(x) + h\epsilon(h)$$

   where the error $\epsilon(h) \to 0$ as $h \to 0$. This statement is actually equivalent to being differentiable. We can also write this using "little-oh" notation:

   $$f(x+h) = f(x) + hf'(x) + o(h).$$

   (b) By Taylor's theorem we have

   $$f(x+h) = f(x) + hf'(x) + \frac{h^2}{2}f''(\xi)$$

   for some ξ between x and $x+h$. Thus we have

   $$|f(x+h) - f(x) - hf'(x)| \leq \frac{3h^2}{2}.$$

7. The directional derivative $f'(x; v)$ of $f : \mathbb{R}^n \to \mathbb{R}$ at $x \in \mathbb{R}^n$ in the direction $v \neq 0$ is defined by

   $$f'(x; v) := \lim_{t \downarrow 0} \frac{f(x+tv) - f(x)}{t},$$

   assuming the limit exists. If f is differentiable prove that

   $$f'(x; v) = \nabla f(x)^T v.$$

   Solution.

   Proof. By the definition of differentiability we have

   $$\lim_{h \to 0} \frac{f(x+h) - f(x) - \nabla f(x)^T h}{\|h\|} = 0,$$

where h takes values in $\mathbb{R}^n$. This implies that

$$\lim_{t\downarrow 0} \frac{f(x+tv) - f(x) - \nabla f(x)^T(tv)}{\|tv\|} = 0,$$

where we restrict h to approaching 0 along the direction v. Multiplying both sides by $\|v\|$ we obtain

$$\lim_{t\downarrow 0} \frac{f(x+tv) - f(x)}{t} - \nabla f(x)^T v = 0,$$

proving the result. □

8. Compute the gradient and Hessian of $f(x,y) = 3x^2 + 2xy + 5y^2$.

   Solution. We have

   $$\nabla f(x,y) = (6x + 2y, 2x + 10y)$$

   and

   $$\nabla^2 f(x,y) = \begin{bmatrix} 6 & 2 \\ 2 & 10 \end{bmatrix}.$$

9. Let $f : \mathbb{R}^n \to \mathbb{R}$ be given by $f(x) = x^T Ax$ for some symmetric $A \in \mathbb{R}^n$. Give conditions on A so that 0 is the global minimizer of f.

   Solution. If A is positive semidefinite then $f(0) = 0$ and $f(x) \geq 0$ for all $x \in \mathbb{R}^n$. If A has a negative eigenvalue λ with corresponding eigenvector v then

   $$v^T Av = \lambda\|v\|^2 < 0.$$

10. Assume $f : \mathbb{R}^n \to \mathbb{R}$ is differentiable and $\nabla f(x) \neq 0$. Compute the directions of steepest descent and steepest ascent:

    $$\arg\min_{\|v\|=1} f'(x;v) \quad \text{and} \quad \arg\max_{\|v\|=1} f'(x;v).$$

    Solution. By problem 7 above and homework 5.4 these are given by

    $$-\frac{\nabla f(x)}{\|\nabla f(x)\|} \quad \text{and} \quad \frac{\nabla f(x)}{\|\nabla f(x)\|},$$

    respectively.

11. Suppose $f : \mathbb{R} \to \mathbb{R}$ satisfies $f \in C^3$ (i.e., f is 3 times continuously differentiable). Show

    $$\lim_{h\to 0} \frac{f(x+h) + f(x-h) - 2f(x)}{h^2} = f''(x).$$

Solution. Applying Taylor's theorem twice we have

$$\begin{aligned}
f(x+h) &= f(x) + hf'(x) + \frac{h^2}{2} f''(x) + \frac{h^3}{3!} f'''(\xi_1) \\
f(x-h) &= f(x) - hf'(x) + \frac{h^2}{2} f''(x) - \frac{h^3}{3!} f'''(\xi_2),
\end{aligned}$$

for some ξ_1 between x and $x+h$, and some ξ_2 between x and $x-h$. Then we have

$$\begin{aligned}
\frac{f(x+h) + f(x-h) - 2f(x)}{h^2} &= \frac{h^2 f''(x) + h^3/6(f'''(\xi_1) - f'''(\xi_2))}{h^2} \\
&= f''(x) + \frac{h}{6}(f'''(\xi_1) - f'''(\xi_2)) \\
&\to f''(x),
\end{aligned}$$

as $h \to 0$.

This problem can be solved with less assumptions if we use the fact that if f is twice differentiable then

$$f(x+h) = f(x) + hf'(x) + \frac{h^2}{2} f''(x) + o(h^2).$$