Optimization Hw1

Toby Harvey

October 30, 2019

Problem 1

Theorem 1. *$a^T x = \|a\| \, \|x\| \cos\theta$.*

Proof. Let $\vec{a} = [a_1, a_2]$ and $\vec{x} = [x_1, x_2] \in \mathbb{R}^2$. The vector $x - a$ creates a triangle in $\mathbb{R}^2$. Denote the angle opposite $\vec{x} - \vec{a}$ as θ. From the law of cosines we have:

$$\begin{gathered}
||\vec{x} - \vec{a}||^2 = ||\vec{x}||^2 + ||\vec{a}||^2 - 2\,\|\vec{x}\|\,\|\vec{a}\|\cos\theta \\
\implies 2\,\|\vec{x}\|\,\|\vec{a}\|\cos\theta = x_1^2 + x_2^2 + a_1^2 + a_2^2 - (x_1 - a_1)^2 + (x_2 - a_2)^2 \\
\implies 2\,\|\vec{x}\|\,\|\vec{a}\|\cos\theta = x_1^2 + x_2^2 + a_1^2 + a_2^2 - (x_1^2 - 2a_1x_1 + a_1^2 + x_2^2 - 2a_2x_2 + a_2^2) \\
\implies 2\,\|\vec{x}\|\,\|\vec{a}\|\cos\theta = 2(a_1x_1 + a_2x_2) \\
\implies \|\vec{x}\|\,\|\vec{a}\|\cos\theta = \vec{a} \cdot \vec{x}
\end{gathered}$$

□

Distance between two parallel hyperplanes:
Since $\vec{a}$ is normal to the the plane. The distance from orgin to the plane must be a multiple of $\vec{a}$ or $d\vec{a}$. Solving for d we get:

$$da_1^2 + da_2^2 = b_1 \implies d(a_1^2 + a_2^2) = b_1 \implies d = \frac{b_1}{\vec{a}}$$

We need the distance between the two planes or

$$|\frac{b_2}{\|\vec{a}\|} - \frac{b_1}{\|\vec{a}\|}| \implies \frac{|b_2 - b_1|}{\|a\|}$$

Problem 2

Proposition 1. *S is a polyhedron.*

Proof. As seen in Boyd and Vandenberghe if we can show that S is a half space we can show it is a poyhedron or is convex. To that end, we expand the inner products:

$$\begin{aligned}
\|x - x_0\| \leq \|x - x_i\| &\implies \|x - x_0\|^2 \leq \|x - x_i\|^2 \\
&\implies \langle x - x_0, x - x_0 \rangle \leq \langle x - x_i, x - x_i \rangle \\
&\implies \langle x, x \rangle + 2\langle x, x_0 \rangle + \langle x_0, x_0 \rangle \leq \langle x, x \rangle + 2\langle x, x_i \rangle + \langle x_i, x_i \rangle \\
&\implies x^T x + 2x^T x_0 + x_0^T x_0 \leq x^T x + 2x^T x_i + x_i^T x_i \\
&\implies 2(x_0 - x_i)x \leq x_i^T x_i - x_0^T x_0
\end{aligned}$$

This is by definition a halfspace. Where $A = x_0 - x_i$ for $i = 0, ..., k$, and $b = \frac{x_i^T x_i - x_0^T x_0}{2}$ for $i = 0, ..., k$.

□

Problem 3

Proposition 2. *f is convex if and only if $(\nabla f(x) - \nabla f(y))(x - y) \geq 0$.*

Proof. If f is convex then both $f(x) + \nabla f(x)(y - x) \leq f(y)$ and $f(y) + \nabla f(y)(x - y) \leq f(x)$ dependent on if x or y is larger. Adding these together we get:

$$\begin{aligned}
f(x) + \nabla f(x)(y - x) + f(y) + \nabla f(y)(x - y) \leq f(x) + f(y) \\
\implies \nabla f(x)(y - x) + \nabla f(y)(x - y) \leq 0 \implies -\nabla f(x)(y - x) - \nabla f(y)(x - y) \geq 0 \\
\implies (\nabla f(x) - \nabla f(y))^T (x - y) \geq 0
\end{aligned}$$

The other way:
If we say that $(\nabla f(x) - \nabla f(y))^T (x - y) \geq 0$ and without loss of generality $x > y$ then this can only be the case if $(\nabla f(x) - \nabla f(y))$ is positive, and by hypothesis $x - y > 0$ meaning we must also have $\nabla^2 \geq 0$ at the limit.

□

Problem 4

1.

Proof. Using the same technique as in class where we us the fact that $\max\{f_1(x), f_2(x), ..., f_m(x)\} = f_n(x)$ for some n in $1...m$. We must show $f(x_1t+x_2(1-t)) \leq f(x_1)t+f(x_2)(1-t)$.
To that end:

$$\begin{aligned} f(x_1t + x_2(1-t)) = f_n(x_1t + x_2(1-t)) \leq f_n(x_1) + f_n(x_2)(1-t) \\ \leq \max\{f_1(x_1), ..., f_m(x_1)\}t + \max\{f_1(x_2), ..., f_m(x_2)\}(1-t) \\ = f(x_1)t + f(x_2)(1-t) \end{aligned}$$

Where the second inequality is by the fact that f_n is in the set of functions that f maximizes over but is not necessarly the maximum at either x_1 or x_2. □

2.

Proof. We use the second order condition. Differentiating f with respect to x_{ij} we get:

$$\frac{\partial f}{\partial x_{ij}} = 1\ln(x_{ij}+1) + (x_{ij}+1)\frac{1}{x_{ij}+1} - 1 = \ln(x_{ij}+1)$$

and

$$\frac{\partial^2 f}{\partial x_{ij}^2} = \frac{1}{x_{ij}+1}$$

Where $\frac{\partial^2 f}{\partial x_{ij} x_{pq}} = 0$, $pq \neq ij$. Therefore the Hessian is $\frac{1}{x_{ij}+1}$ on the diagonal and 0 otherwise, and since $\frac{1}{x_{ij}+1}$ is positive, f is convex.

□

Problem 5
We can use all facts that we have already proved, plus proof that absolute value is convex.

Proposition 3. *absolute value is convex.*

Proof.

$$\begin{aligned} &|tx_1 + x_2(1-t)| \\ &\leq |tx_1| + |x_2(1-t)| \\ &= t|x_1| + |x_2|(1-t) \end{aligned}$$

□

We showed that $\ln \frac{1}{c^T x + d}$ is convex in class, and $|a^T x + b|$ is convex, and I just showed that the maximum of a set of convex functions is convex. Therefore $f(x)$ must also be convex.

Let x_a and x_b in S, so that $\|x_a - x_0\| \leq \|x_a - x_i\|$ and $\|x_b - x_0\| \leq \|x_b - x_i\|$. We must show that $\|(tx_a + (1-t)x_b) - x_0\| \leq \|(tx_a + (1-t)x_b) - x_i\|$. To that end we have:

$$\begin{aligned}\|tx_a + (1-t)x_b - x_0\| = \|tx_a + x_b - tx_b - x_0\| &\leq \|x_b - x_0\| + \|tx_a - tx_b\| \\ &\leq \|x_b - x_i\| + t\,\|x_a - x_b\|\end{aligned}$$