Homework 1a
Convex Sets and Convex Functions

CMU 10-725/36-725: Convex Optimization (Fall 2017)

OUT: Sep 1
DUE: Prob 1-3 Sep 11, 5:00 PM

START HERE: Instructions

- **Collaboration policy:** Collaboration on solving the homework is allowed, after you have thought about the problems on your own. It is also OK to get clarification (but not solutions) from books or online resources, again after you have thought about the problems on your own. There are two requirements: first, cite your collaborators fully and completely (e.g., "Jane explained to me what is asked in Question 3.4"). Second, write your solution independently: close the book and all of your notes, and send collaborators out of the room, so that the solution comes from you only.

- **Submitting your work:** Assignments should be submitted as PDFs using Gradescope unless explicitly stated otherwise. Each derivation/proof should be completed on a separate page. Submissions can be handwritten, but should be labeled and clearly legible. Else, submissions can be written in LaTeX. Upon submission, label each question using the template provided by Gradescope. Please refer to Piazza for detailed instruction for joining Gradescope and submitting your homework.

Problem 1: Convex Set (Hongyang - 20 pts)

1. **[5pts]** (Definition of convexity) Denote by $C \subseteq \mathbb{R}^n$ a convex set. Let $x_1, ..., x_k \in C$, and $\theta_1, ..., \theta_k \in \mathbb{R}$ satisfy $\theta_i \geq 0$ and $\sum_{i=1}^k \theta_i = 1$. Show that $\theta_1 x_1 + ... + \theta_k x_k \in C$ for arbitrary k (Recall that in the class, we define the convexity by the case of $k = 2$.)

2. (Example of convex set)
(a) **[5pts]** Show that if $a, b \geq 0$ and $0 \leq \theta \leq 1$, then $a^\theta b^{1-\theta} \leq \theta a + (1-\theta) b$. Hint: Use concavity of log functions.
(b) **[5pts]** Show that $S_n = \{x \in \mathbb{R}_+^n \mid \prod_{i=1}^n x_i \geq 1\}$ is convex.

3. **[5pts]** (Operations that preserve convexity) Suppose that S_1 and S_2 are convex sets in $\mathbb{R}^{m+n}$. Show that their partial sum

$$S = \{(x, y_1 + y_2) \mid x \in \mathbb{R}^m,\ y_1, y_2 \in \mathbb{R}^n,\ (x, y_1) \in S_1,\ (x, y_2) \in S_2\}$$

is a convex set.

Problem 2: Convex Functions (Devendra - 20 pts)

1. **[5pts]** Suppose that a non-convex function $f(x)$ is given as a sum of terms of the form $g : R^n_{++} \to R, g(x) = \beta x_1^{\alpha_1}..x_n^{\alpha_n}, \beta > 0, \alpha_i \in R$. Show that the substitution $y_i = \log x_i$ can transform $f(x)$, into a into a convex function in y.
   (Hint: Use a simple example such as $f(x_1, x_2) = x_1^{-2} + (x_1 x_2)^{\frac{1}{3}} + 2x_2^{-4}, x_1 > 0, x_2 > 0$ to prove the claim and then generalize it)

2. Show that the following functions are convex

   - **[5pts]** $f : R^n_{++} \to R$, $f(x) = \sum_{i=1}^n x_i \log x_i$, such that $\sum_{i=1}^n x_i = 1$. This is also called negative entropy function.
   - **[5pts]** $f : R^n_{++} \to R_-$, $f(x) = -(\sum_{i=1}^n x_i^a)^{\frac{1}{a}}$, given that $a < 1, a \neq 0$.
   - **[5pts]** $f : R_+ \to R$
   $$f(x) = \begin{cases} \frac{1}{x}\int_0^x g(y)dy & \text{if } x > 0 \\ 0 & \text{if } x = 0 \end{cases}$$
   given that $g : R_+ \to R$ is a convex, non-negative function and $g(0) = 0$.

Problem 3: Lipschitz gradients and strong convexity (Yifeng - 30 pts)

Let f be convex and twice differentiable.

1. **[6pts]** Prove the *monotonicity of gradient* ∇f, i.e.,

   f is convex if and only if $(\nabla f(x) - \nabla f(y))^T(x - y) \geq 0$, for all x, y.

   (Note: Feel free to use this property in the proof of problem 2 or 3 if necessary.)

2. **[12pts]** (Smoothness of f) Show that the following statements are equivalent.

   i. ∇f is Lipschitz with constant L;

   ii. $(\nabla f(x) - \nabla f(y))^T(x - y) \leq L\|x - y\|_2^2$ for all x, y;

   iii. $\nabla^2 f(x) \preceq LI$ for all x;

   iv. $f(y) \leq f(x) + \nabla f(x)^T(y - x) + \frac{L}{2}\|y - x\|_2^2$ for all x, y.

3. **[12pts]**(Curvature of f) Show that the following statements are equivalent.

   i. f is strongly convex with constant m;

   ii. $(\nabla f(x) - \nabla f(y))^T(x - y) \geq m\|x - y\|_2^2$ for all x, y;

   iii. $\nabla^2 f(x) \succeq mI$ for all x;

   iv. $f(y) \geq f(x) + \nabla f(x)^T(y - x) + \frac{m}{2}\|y - x\|_2^2$ for all x, y.