CSE 535: Theory of Optimization and Continuous Algorithms **Winter 2019**

Homework 1

Lecturer: Yin Tat Lee *Due Date: 14 Jan 2018*

You are allowed to discuss with others. If you discuss with others, please list your collaborators for each question. In any case, you must write your own solutions. The goal of the homework is to make sure you satisfies the prerequisite of the course.

1. Matrix Inequality (20 marks)

For any $n \times n$ symmetric matrices A, B, we define $A \succeq B$ if and only if $v^\top A v \geq v^\top B v$ for all $v \in \mathbb{R}^n$. We call a matrix A is PSD (positive semi-definite) if $A \succeq 0$.

- (5 marks) Show that $A^\top M_1 A \succeq A^\top M_2 A$ for any symmetric matrices $M_1 \succeq M_2$ and any matrix A.
- (5 marks) Show that $M(M+I)^{-1}M \succeq \frac{1}{2}I$ for any PSD matrix $M \succeq I$.
- (5 marks) Show that $\mathrm{tr}A^2 \geq \mathrm{tr}B^2$ for any PSD $A \succeq B$.
- (5 marks) Find a counter example for the statement "$A^2 \succeq B^2$ for any PSD $A \succeq B$". (Hint: It is probably easiest to do this by writing code)

2. Solving Linear Systems (20 marks)

Given a symmetric matrix $I \preceq A \preceq \kappa \cdot I$ for some $\kappa \geq 1$. Consider the algorithm $x^{(k+1)} = x^{(k)} + t \cdot (b - Ax^{(k)})$ where $t > 0$ is the step size to be decided.

- (5 marks) Define the residual $r^{(k)} \stackrel{\text{def}}{=} Ax^{(k)} - b$. Show that

$$\|r^{(k+1)}\|_2^2 = (r^{(k)})^\top (I - tA)^2 r^{(k)}.$$

- (8 marks) Pick $t = \frac{2}{\kappa+1}$, show that the eigenvalues of $(I - tA)^2$ lies in $[0, (\frac{\kappa-1}{\kappa+1})^2]$.
- (5 marks) Show that

$$\|r^{(k)}\|_2 \leq (1 - \frac{2}{\kappa+1})^k \|r^{(0)}\|_2.$$

- (2 marks) Explain why we picked $t = \frac{2}{\kappa+1}$.

3. Calculus Exercise (20 marks)

Given a twice differentiable function $f : \mathbb{R} \to \mathbb{R}$ and a $m \times n$ matrix A. Let $\Phi(x) = \sum_{i=1}^m f(\sum_{j=1}^n A_{ij} x_j)$. By extending the definition of f to $f : \mathbb{R}^n \to \mathbb{R}^n$ defined by $f(x)_i = f(x_i)$. We can write $\Phi(x)$ simply as $\Phi(x) = 1^\top f(Ax)$.

- (8 marks) Let $D\Phi(x)[h]$ as the directional derivative of Φ on direction h, namely $D\Phi(x)[h] = \lim_{t\to 0}(\Phi(x+th) - \Phi(x))/t$. Compute $D\Phi(x)[h]$. You can only express your answer using the following symbols A, x, f', $\top$ (transpose symbol) and h. Namely, the final answer cannot use $\sum$ and cannot use indices.
- (2 marks) Use previous part, show that $\nabla\Phi(x) = A^\top f'(Ax)$. (Again, we define f' by $f'(x)_i = f'(x_i)$.)
- (8 marks) Let $D^2\Phi(x)[h_1, h_2]$ as the second directional derivative of Φ on direction h_1, h_2, namely $D^2\Phi(x)[h_1, h_2] = \lim_{t\to 0}(D\Phi(x + th_2)[h_1] - D\Phi(x)[h_1])/t$. Compute $D^2\Phi(x)[h_1, h_2]$. Again you cannot use $\sum$ or indices in the final answer.
- (2 marks) Use previous part, show that $\nabla^2\Phi(x) = A^\top \mathrm{diag}(f''(Ax))A$ where $\mathrm{diag}(f''(Ax))$ is the diagonal matrix with diagonal given by $f''((Ax)_i)$. (Again, we define f'' by $f''(x)_i = f''(x_i)$.)

4. Your convex problem (20 marks)

State an optimization problem of the form $\min_{x \in K} f(x)$ where f is a convex function and K is a convex set. Explain some background about this minimization problem, such as, the meaning of each variable. Prove that the function f and the set K you give are indeed convex.
Remark: We prefer you write a convex problem that is close to your research/field.