EN.520.618: Modern Convex Optimization, Fall 2022, Assignment 1
Due: Sep. 8th 5:00 p.m.

Submission formats:

1. LaTeX: you will be provided with a template and we will consider bonus points for LaTeX submissions.
2. Digitally written on a notepad, iPad, etc.
3. Handwritten on paper.

Hard copies of handwritten homework must be turned in the class or at office hours (if applicable). Otherwise, you can submit your homework on Canvas. You are allowed to discuss the problems with your classmates but the final write down must be yours. Provide the name(s) of your collaborators.

Problem 1 (Quadratics of a Matrix) Suppose $A \in \mathbb{S}_{+}^{n\times n}$, i.e., A is a positive semidefinite matrix of size n. Show that for any real vector x with $\| x \|_2 = 1$ we have

$$\lambda_{\min}(A) \leq x^\top A x \leq \lambda_{\max}(A)$$

Problem 2 (PSD properties) Let A and B be $n \times n$ symmetric real matrices. Prove or disprove each of the following statements:

1. If $A \succeq 0$ and $\mathrm{trace}(A) = 0$ then A $= 0$.
2. If $A \succeq 0, B \succeq 0$, and $AB = 0$ than $A = 0$ or $B = 0$.
3. If $A \succeq 0$, then the largest entry in absolute value must be on the diagonal.
4. If $A \succeq 0, B \succeq 0$, and $A + B = 0$, then $A = B = 0$.

Problem 3 (Chain Rule) Suppose $x(\alpha)$ is a twice differentiable curve and consider $f(x(\alpha))$ as a function of α. Define

$$s = \frac{dx}{d\alpha}(\alpha_0), \quad t = \frac{d^2x}{d\alpha^2}(\alpha_0)$$

Compute $\dfrac{df}{d\alpha}(\alpha_0)$ and $\dfrac{d^2f}{d\alpha^2}(\alpha_0)$ in terms of s and t.

Problem 4 (Pseudoinverse) Suppose $A \in \mathbb{R}^{m\times n}$ has singular value decomposition $A = U\Sigma V^\top$ and has rank $k < \min(m, n)$. The columns of U and V are vectors $u_i, i = 1, \cdots, m$ and $v_i, i = 1, \cdots, n$, respectively. Using the singular values or the (right/left) singular vectors

1. Provide a basis for the range space of A.
2. Provide a basis for the null space of A.
3. Provide an expression for $\| A \|_{\mathrm{F}}^2$

Problem 5 (Taylor Approximation) Suppose $f\colon \mathbb{R}^n \mapsto \mathbb{R}$ and $x, y \in \mathbb{R}^n$.

1. Given that f is differentiable, show that

$$f(y) = f(x) + \nabla f(x + t(y - x))^\top (y - x), \quad \text{for some } t \in (0, 1)$$

2. Given that f is twice differentiable, show that

$$f(y) = f(x) + \nabla f(x)^\top (y - x) + \frac{1}{2}(y - x)^\top \nabla^2 f(x + t(y - x))(y - x), \quad \text{for some } t \in (0, 1)$$

3. Given that $mI \preceq \nabla^2 f(x) \preceq LI$ for some $0 < m < L < \infty$, show that

$$f(x) + \nabla f(x)^\top (y - x) + \frac{m}{2}\|y - x\|_2^2 \leq f(y) \leq f(x) + \nabla f(x)^\top (y - x) + \frac{L}{2}\|y - x\|_2^2.$$

Problem 6 (First/Second-Order Necessary Conditions) Suppose $f\colon \mathbb{R}^n \mapsto \mathbb{R}$. Suppose that $x^\star$ is a local minimizer of f.

1. Given that ∇f is Lipschitz continuous, prove that $\nabla f(x^\star) = 0$.
2. Given that ∇f is Lipschitz continuous, and $\nabla^2 f$ exists and is continuous in an open neighborhood around $x^\star$, show that $\nabla f(x^\star) = 0$ and $\nabla^2 f(x^\star) \succeq 0$.

Hint: Problem 5 can be useful in the course of the solution.

Problem 7 (Backtracking Line Search) Suppose $f\colon \mathbb{R}^n \mapsto \mathbb{R}$ is continuously differentiable with Lipschitz gradient, i.e., there exists an $L > 0$ such that $\| \nabla f(y) - \nabla f(x) \|_2 \leq L \| y - x \|_2$ for all $x, y \in \mathrm{dom} f$. Consider the gradient descent method $x^{k+1} = x^k - t^k \nabla f(x^k)$. The Armijo step size selection criterion states that we need to choose a step size that satisfies the conditions

$$f(x^{k+1}) \leq f(x^k) - \alpha t^k \| \nabla f(x^k) \|_2^2$$

If we use backtracking line search with parameter $\beta \in (0, 1)$, prove the following:

1. The step size chosen at each step satisfies $t^k \geq \dfrac{2}{L}\beta(1 - \alpha)$.
2. $f(x^{k+1}) - f(x^k) \leq -\dfrac{2}{L}\alpha(1 - \alpha)\beta \| \nabla f(x^k) \|_2^2$.
3. Given $f^\star = \inf_x f(x)$, the global minimum of f,

$$\min_{0 \leq k \leq N} \| \nabla f(x^k) \|_2 \leq \sqrt{\frac{L(f(x^0) - f^\star)}{2\alpha\beta(1 - \alpha)(N + 1)}}.$$