University of Illinois **Spring 2023**

ECE 490: Problem Set 2

Due: Tuesday February 14 through Gradescope by 11 AM
Reading: Lecture Notes 4-7; Secs 1.2 – 1.4 of text.

1. [**Convexity of Compositions**]
   (a) Suppose $f : \mathbb{R}^n \to \mathbb{R}$ is convex, and $g : \mathbb{R} \to \mathbb{R}$ is convex and non-decreasing. Prove that the composition of g and f is convex, i.e., show that for all $x, y \in \mathbb{R}^n$ and $\alpha \in [0, 1]$
   $$g(f(\alpha x + (1-\alpha)y)) \leq \alpha g(f(x)) + (1-\alpha)g(f(y)).$$
   (b) Give an example of $f : \mathbb{R} \to \mathbb{R}$ and $g : \mathbb{R} \to \mathbb{R}$, which are both convex, such that the composition $g(f(\cdot))$ is not convex.
   (c) (3 pts) Let Q be an $n \times n$ symmetric PSD matrix, and β be a positive scalar. Then prove that the function $h : \mathbb{R}^n \to \mathbb{R}$ defined by
   $$h(x) = e^{\beta x^\top Q x}$$
   is convex.

2. [**Armijo's Rule**]
   Consider the problem of minimizing the function $f : \mathbb{R}^2 \to \mathbb{R}$ defined as:
   $$f(x) = f(x_1, x_2) = 4x_1^2 + x_2^4$$
   using steepest descent with Armijo's Rule, with parameters $\tilde{\alpha} = 1$, $\sigma = 0.2$, and $\beta = 0.3$. Find α_k if $x_k = (1, 0) \equiv [1\ 0]^\top$.

3. [**Lipschitz Function**]
   Consider $f : \mathbb{R}^n \to \mathbb{R}$. Show that if
   $$\|\nabla f(x)\| \leq L, \quad \forall x \in \mathbb{R}^n$$
   then f is Lipschitz with Lipschitz constant L.

4. [**Convergence of Gradient Descent with Constant Step-Size**]
   Consider minimizing the function $f : \mathbb{R} \to \mathbb{R}$ defined by
   $$f(x) = |x|^3.$$
   Note that f is twice continuously differentiable, with $\nabla f(x) = 3x|x|$, and $\nabla^2 f(x) = 6|x|$.

   Suppose we run steepest descent on this function with constant step-size $\alpha = 1/3$, i.e.,
   $$x_{k+1} = x_k - \frac{\nabla f(x_k)}{3} = x_k(1 - |x_k|)$$
   with starting point x_0.
   (a) Establish that f has a unique minimum $x^* = 0$.
   (b) Show that if $|x_0| < 2$, then the sequence $\{|x_k|\}$ is *convergent* (as $k \to \infty$).
   (c) Now show that if $|x_0| < 2$, then x_k converges to 0, as $k \to \infty$, i.e., that
   $$\lim_{k\to\infty} x_k = \lim_{k\to\infty} |x_k| = 0.$$
   Hint: Suppose that $|x_k|$ converges to $c > 0$, and establish a contradiction.

5. **[Convergence with Varying Step Size]**
   Consider if $f : \mathbb{R} \to \mathbb{R}$ defined by:
$$f(x) = \frac{2}{3}|x|^3 + \frac{1}{2}x^2$$
   Suppose we use the steepest descent method with step size
$$\alpha_k = \frac{1}{k+1}.$$
   Show that if $|x_0| \geq 1$, then $|x_k| \geq k+1$ for all $k \geq 1$, i.e., the method diverges.

   Hint: Use mathematical induction.

6. **[Strong Convexity]**
   Suppose $f : \mathbb{R}^n \to \mathbb{R}$ is twice continuously differentiable and strongly convex, i.e.,
$$\nabla^2 f(x) \succeq mI, \ \forall x \in \mathbb{R}^n.$$
   Let x^* be the global minimum of f. Show that
$$\|x - x^*\| \leq \frac{\|\nabla f(x)\|}{m}, \ \forall x \in \mathbb{R}^n.$$
   Hint: You may want to start with the relation:
$$\nabla f(y) = \nabla f(x) + \int_0^1 \nabla^2 f(x + t(y-x))(y-x)dt$$
   and exploit the fact that $\nabla f(x^*) = 0$.

7. **[Condition Number]**
   Consider $f : \mathbb{R}^n \to \mathbb{R}$ defined as
$$f(x) = \frac{1}{2}x^\top Q(\epsilon)x + b(\epsilon)^\top x + c(\epsilon)$$
   where
$$Q(\epsilon) = \begin{bmatrix} 1 & \epsilon \\ \epsilon & 1 \end{bmatrix}, \ b(\epsilon) = \begin{bmatrix} 1+\epsilon \\ 1-\epsilon \end{bmatrix}, \ c(\epsilon) = \epsilon^2.$$
   (a) Find the condition number of the Hessian of f for $\epsilon \in (0,1)$.
   (b) What happens to the condition number as $\epsilon \to 0$ and $\epsilon \to 1$?