535526 Fall 2022: Optimization Algorithms **(Due: 2022/09/26, 11:00)**

Homework 0: Fundamentals

Submission Guidelines: Please compress all your write-ups into one PDF file (photos/scanned copies are acceptable; please make sure that the electronic files are of good quality and reader-friendly) and submit the file via E3.

Problem 1 (Second-Order Necessary Condition) (15 points)

Please prove the second-order necessary condition for optimality: Let $f : \mathbb{R}^n \to \mathbb{R}$ be twice continuously differentiable on an ϵ-neighborhood of $x^* \in \mathbb{R}^n$, and x^* is a local minimizer of f. Show that in addition to $\nabla f(x^*) = 0$, we must also have $\nabla^2 f(x^*) \succcurlyeq 0$. (Note: Please carefully justify every step of your proof)

Problem 2 (Optimality Conditions) (15+15=30 points)

In this problem, let's study a somewhat surprising example, where reckless intuition could be misleading. Let $f : \mathbb{R}^n \to \mathbb{R}$ be a differentiable function. Suppose that a point x^* is a local minimum of f along every line that passes through x^*, i.e., the function $h(t) := f(x^* + td)$ is minimized at $t = 0$ for all $d \in \mathbb{R}^n$.

(a) Show that x^* is a stationary point, i.e., $\nabla f(x^*) = 0$. (Note: Please carefully justify each step of your proof)

(b) Show by constructing a counterexample that x^* is not necessarily a local minimizer of f. (Hint: Consider the function of two variables $f(x_1, x_2) = (x_2 - px_1^2)(x_2 - qx_1^2)$, where $0 < p < q$. Show that $(0, 0)$ is indeed a local minimizer of f along every line that passes through $(0, 0)$. However, if we pick a scalar $m \in (p, q)$, then $f(x_1, mx_1^2) < 0$ if $x_1 \neq 0$ while $f(0, 0) = 0$).

Problem 3 (Convex Sets) (15+15=30 points)

In this problem, let's study the feasible set induced by a quadratic inequality, which will frequently show up in our subsequent lectures. Let $C \subseteq \mathbb{R}^n$ be the solution set of a quadratic inequality as follows:

$$C = \{x \in \mathbb{R}^n : x^\top Ax + b^\top x + c \leq 0\}, \tag{1}$$

where A is real symmetric square matrix, $b \in \mathbb{R}^n$, and $c \in \mathbb{R}$.

(a) Show that C is a convex set if $A \succcurlyeq 0$. Moreover, is the converse statement true?

(b) Show that the intersection of C and the hyperplane defined by $g^\top x + h = 0$ (where $g \neq 0$) is a convex set if $A + \lambda gg^\top \succcurlyeq 0$ for some $\lambda \in \mathbb{R}$.

Problem 4 (Subgradients) (15+15=30 points)

As will be discussed in Lecture 2, we would like to establish an optimality condition for non-differentiable functions by using "subgradients." Specifically, for a function $f : X \to \mathbb{R}$, a vector g is called a subgradient of f at some $x \in X$ if $f(z) \geq f(x) + g^\top(z - x)$, for all $z \in X$. The set of all subgradients at $x \in X$, denoted by $\partial f(x)$, is called the "subdifferential" of f at x. Suppose f is a convex function.

(a) Show that if f is differentiable at x, then $\partial f(x) = \{\nabla f(x)\}$. (Hint: Prove this by contradiction)

(b) Show that if $\partial f(x) = \{g_0\}$ (i.e., the subdifferential has only one element), then f is differentiable at x and $\nabla f(x) = g_0$. (Note: You may start from the 1-dimensional case to get some intuition)

Problem 5 (Convex Functions) (15 points)

One very useful property of convexity is that some compositions of functions could still preserve convexity. Consider the composition $f(x) = h(g(x))$, where $g : \mathbb{R}^n \to \mathbb{R}^k$, $h : \mathbb{R}^k \to \mathbb{R}$, and $f : \mathbb{R}^n \to \mathbb{R}$, and g, h are both twice differentiable. Without loss of generality, we can restrict ourselves to the case $n = 1$ since convexity is determined by the behavior of a function on arbitrary lines that intersect its domain. As a result, we can rewrite the composition as

$$f(x) = h(g(x)) = h(g_1(x), \cdots, g_k(x)), \tag{2}$$

where $x \in \mathbb{R}$. In this case, the second-order derivative of f can be obtained by

$$f''(x) = g'(x)^\top \nabla^2 h(g(x)) g'(x) + \nabla h(g(x))^\top g''(x). \tag{3}$$

Please verify the following two properties:

- If h is convex and <u>non-decreasing</u> in each input argument, and all g_i are <u>convex</u>, then f is also convex.
- If h is convex and <u>non-increasing</u> in each input argument, and all g_i are <u>concave</u>, then f is also convex.