ENGG 5501: Foundations of Optimization	2018–19 First Term

Homework Set 2

Instructor: Anthony Man–Cho So — Due: October 3, 2018

INSTRUCTIONS: Problems 1 and 2 are compulsory. The remaining problems are for practice and will not be graded.

Problem 1 (20pts). Let $f : \mathbb{R}^n \to \mathbb{R}$ be a convex differentiable function.

(a) **(10pts).** Show that for any $x, y \in \mathbb{R}^n$, we have

$$(\nabla f(y) - \nabla f(x))^T (y - x) \geq 0.$$

(b) **(10pts).** Suppose *in addition* that f has Lipschitz continuous gradient; i.e.,

$$\|\nabla f(x) - \nabla f(y)\|_2 \leq L\|x - y\|_2 \quad \text{for all } x, y \in \mathbb{R}^n$$

for some constant $L > 0$. Show that for any $x, y \in \mathbb{R}^n$, we have

$$\left| f(y) - f(x) - \nabla f(x)^T (y - x) \right| \leq \frac{L}{2}\|x - y\|_2^2.$$

(*Hint: Fix $x, y \in \mathbb{R}^n$ and apply the Fundamental Theorem of Calculus to the function $t \mapsto f(x + t(y - x))$.*)

Problem 2 (10pts). Given a convex set $S \subseteq \mathbb{R}^n$ and a vector $c \in \mathbb{R}^n$, consider the optimization problem

$$\inf_{x \in S} c^T x.$$

Show that the minimum is attained at a point $\bar{x} \in \operatorname{rel\,int}(S)$ if and only if the function $x \mapsto c^T x$ is constant on S.

Problem 3. For any given $k \geq 1$, let $\lambda_1^k : \mathcal{S}^n \to \mathbb{R}$ be the function that returns the sum of the k largest eigenvalues of its argument.

(a) Show that

$$\begin{array}{rcll} \lambda_1^k(A) & = & \text{maximize} & \operatorname{tr}(AX) \\ & & \text{subject to} & \operatorname{tr}(X) = k, \\ & & & I \succeq X \succeq \mathbf{0}. \end{array}$$

(b) Using the result in (a), show that λ_1^k is convex for each $k \geq 1$.

Problem 4. Let $f : \mathbb{R}^n \to \mathbb{R} \cup \{+\infty\}$ be a convex function such that $\operatorname{epi}(f)$ is closed and f is not identically $+\infty$.

(a) Show that $f = f^{**}$, where $f^{**} = (f^*)^*$ is the conjugate of f^*.

(b) Using the result in (a), show that for any $x, y \in \mathbb{R}^n$, the following statements are equivalent:

(i) $y \in \partial f(x)$

(ii) $f(x) + f^*(y) = x^T y$

(iii) $x \in \partial f^*(y)$

REMARK: The subdifferential of f, ∂f, is a *set–valued mapping* in the sense that it assigns a set $\partial f(x) \subseteq \mathbb{R}^n$ to each $x \in \mathbb{R}^n$. The *inverse mapping* of ∂f, denoted by $(\partial f)^{-1}$, is simply defined as

$$(\partial f)^{-1}(y) = \{x \in \mathbb{R}^n : y \in \partial f(x)\}\,.$$

With these notations, the equivalence of (i) and (iii) above can be expressed as $(\partial f)^{-1} = \partial f^*$, which is another important relationship between f and f^*.