南方科技大学 SOUTHERN UNIVERSITY OF SCIENCE AND TECHNOLOGY

Course Name: Algorithms for Convex Optimization **Department:** Mathematics

Exam Duration: 120 minutes **Instructor:** Jin Zhang

Question No.	1	2	3	4	5	6
Score	20	10	15	25	15	15

This exam paper contains 6 exercises and the score is 100 in total. (Please hand in your answer sheet and your scrap paper to the proctor when the exam ends.)

Please note that all statements are based on the vectorial l_2-norm without special instructions.

1. True or False + Justification. (5+5+5+5=20 pts)

   (a) Let $f : \mathbb{R}^n \to \overline{\mathbb{R}}$ be proper closed and let $\mathrm{dom}(f)$ be bounded. Then f attains its minimal value over $\mathbb{R}^n$.

   (b) Let $f : \mathbb{R}^n \to \overline{\mathbb{R}}$ be proper coercive and bounded below **but not necessarily closed** . Then f attains its infimal value over $\mathbb{R}^n$.

   (c) Let $f : \mathbb{R}^n \times \mathbb{R}^m \to \overline{\mathbb{R}}$ be closed convex and let $\varnothing \neq \Omega \subset \mathbb{R}^m$ be an arbitrary set. We define $g : \mathbb{R}^n \to \overline{\mathbb{R}}$ with the values

   $$g(x) = \sup_{y \in \Omega} f(x, y).$$

   Then g is closed convex. (Hint: the epigraph of g)

   (d) Let $f_1 : \mathbb{R}^n \to \overline{\mathbb{R}}$ be proper convex and let $f_2 : \mathbb{R}^n \to \mathbb{R}$ be real-valued convex. If $x \in \mathrm{int}(\mathrm{dom}(f_1))$, then

   $$\partial(f_1 + f_2)(x) = \partial f_1(x) + \partial f_2(x).$$

2. **(affine function. 10 pts)** Let $f : \mathbb{R}^n \to \overline{\mathbb{R}}$ be convex and let $\Omega \subset \mathrm{dom}(\partial f)$ be a closed convex set. If $\cap_{x \in \Omega} \partial f(x) \neq \varnothing$. Then f is affine over Ω.

3. **(descent lemma. 15 pts)** Let $f : \mathbb{R}^n \to \overline{\mathbb{R}}$ be differentiable over an open convex set Ω. If

   $$\|\nabla f(x) - \nabla f(y)\| \leq L \|x - y\|^p \text{ for all } x, y \in \Omega,$$

   where $p \geq 0$ is a scalar. Then for any $x, y \in \Omega$,

   $$f(y) \leq f(x) + (y - x)^T \nabla f(x) + \frac{L}{p+1} \|y - x\|^{p+1}.$$

4. **(infimal convolution and conjugacy. 5+10+10=25 pts)** Let $f_1 : \mathbb{R}^n \to \overline{\mathbb{R}}$ be proper closed convex and let $f_2 : \mathbb{R}^n \to \mathbb{R}$ be real-valued, L-smooth convex. Assume that $f_1 \square f_2$ is real-valued. Then the following hold:

   (a)
   $$(f_1 \square f_2)^* = f_1^* + f_2^*,$$

   (b) $f_1 \square f_2$ is L-smooth.

   (c) Assume that $u(x)$ is a minimizer of
   $$\min_u \{f_1(u) + f_2(x-u)\}.$$
   Then $\nabla (f_1 \square f_2)(x) = \nabla f_2 (x - u(x))$.

5. **(subdifferential enlargement. 5+5+5=15 pts)** Let $f : \mathbb{R}^n \to \overline{\mathbb{R}}$ be not necessarily convex with $\bar{x} \in \operatorname{dom}(f)$ and let $\epsilon \geq 0$, we say
   $$\partial_\epsilon f(\bar{x}) = \left\{ g \in \mathbb{R}^n : \liminf_{x \to \bar{x}} \frac{f(x) - f(\bar{x}) - g^T(x - \bar{x})}{\|x - \bar{x}\|} \geq -\epsilon \right\}$$
   is the ϵ-subdifferential of f at $\bar{x}$. Show that

   (a) If $\partial_\epsilon f(\bar{x}) \neq \varnothing$, then it is convex.

   (b) $g \in \partial_\epsilon f(\bar{x})$ if and only if for every $\eta > 0$ the function
   $$f_{g,\eta}(x) = f(x) - f(\bar{x}) - g^T(x - \bar{x}) + (\epsilon + \eta)\|x - \bar{x}\|$$
   attains a local minimum at $\bar{x}$.

   (c) If f is convex, then
   $$\partial_\epsilon f(\bar{x}) = \left\{ g \in \mathbb{R}^n : g^T(x - \bar{x}) \leq f(x) - f(\bar{x}) + \epsilon \|x - \bar{x}\| \text{ for any } x \in \mathbb{R}^n \right\}.$$

6. **(KKT condition and constraint qualifications. 5+10=15 pts)** Consider the problem
   $$(P) \qquad \min f(x), \text{ s.t. } g_1(x) \leq 0, \cdots, g_m(x) \leq 0,$$
   where $f, g_1, \cdots, g_m : \mathbb{R}^n \to \mathbb{R}$ are real-valued convex functions. Suppose that x^* is a local optimal solution of (P).

   (a) Write down the Slater's condition and KKT condition.

   (b) Show that the Slater condition is equivalent to the Mangasarian-Fromovitz constraint qualification (MFCQ) at x^*. Recall the MFCQ holds at x^* means that $\exists d \in \mathbb{R}^n$ such that
   $$g_i'(x^*; d) < 0 \text{ for all } i \in \mathcal{I}(x^*) = \{i : g_i(x^*) = 0\}.$$
   Hint: use the directional derivative of the max-function $g(x) = \max\{g_1(x), \cdots, g_m(x)\}$.