Algorithms for Convex Optimization
Assignment 2

Note: All statements are based on the vectorial l_2-norm without special instructions.

1. Let $f : \mathbb{R}^n \to \overline{\mathbb{R}}$ be convex with $\bar{x} \in \mathrm{dom}(f)$, then we have

$$\partial f(\bar{x}) = \{g \in \mathbb{R}^n : (g, -1) \in N\left[(\bar{x}, f(\bar{x})); \mathrm{epi}(f)\right]\}.$$

2. Let $f : \mathbb{R}^n \to \overline{\mathbb{R}}$ be convex and lipschitz continuous around $\bar{x} \in \mathrm{dom}(f)$, then $\partial^\infty f(\bar{x}) = \{0\}$, where $\partial^\infty f(\bar{x})$ denote the singular subdifferential of f at $\bar{x}$:

$$\partial^\infty f(\bar{x}) = \left\{ g \in \mathbb{R}^n \middle| (g, 0) \in N\left[(\bar{x}, f(\bar{x})); \mathrm{epi}(f)\right] \right\}.$$

3. Let $f : \mathbb{R}^n \to \overline{\mathbb{R}}$ be not necessarily convex with $\bar{x} \in \mathrm{dom}(f)$ 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

$$\partial_\epsilon f(\bar{x}) = \left\{ g \in \mathbb{R}^n : g^T(x - \bar{x}) \leq f(x) - f(\bar{x}) + \epsilon \|x - \bar{x}\| \text{ for } \forall x \in \mathbb{R}^n \right\}.$$

4. Let φ is univariate convex defined on some open interval. Then φ is differential everywhere except a countable set.

5. Let $f : \mathbb{R}^n \to \overline{\mathbb{R}}$ be convex and dom(f) is open, consider the set

$$\Omega = \left\{ x \in \mathbb{R}^n : (g_1 - g_2)^T x \geq 1 \text{ for some } g_1, g_2 \in \partial f(x) \right\}.$$

   If f is continuous on its domain, then dom$(f) \backslash \Omega$ is open and dense in dom(f).

   Hint: Let $\varphi(t) = f$ (maybe some fixed point $+t$ some direction).

6. Show that for any $x = (x_1, \cdots, x_n)^T \in \mathbb{R}^n$,

$$\max\{x_1, \cdots, x_n\} = \max_{y \in \Delta_n} y^T x,$$

   where Δ_n is the unit simplex.

7. Let $f : \mathbb{R}^n \to \overline{\mathbb{R}}$ be proper but **not necessarily convex**, and let $x \in \operatorname{int}(\operatorname{dom}(f))$. Recall the directional derivative of f at x in direction $d \in \mathbb{R}^n$ is defined by

$$f'(x; d) = \lim_{\alpha \to 0^+} \frac{f(x + \alpha d) - f(x)}{\alpha}.$$

Next we consider a stronger concept of differentiability. Define the H-directional derivative of f at x in direction $d \in \mathbb{R}^n$ is:

$$f'_H(x; d) = \lim_{\substack{\alpha \to 0^+ \\ d' \to d}} \frac{f(x + \alpha d') - f(x)}{\alpha}.$$

Show

(a) For given $d \in \mathbb{R}^n$, the existence of $f'_H(x; d)$ implies that $f'(x; d)$ exists.

(b) If $f'_H(x; d)$ exists for any $d \in \mathbb{R}^n$. Then the mapping $d \mapsto f'_H(x; d)$ is continuous on $\mathbb{R}^n$.

(c) If $f'(x; d)$ exists for any $d \in \mathbb{R}^n$ and f is L_f-Lipschitz continuous on V (a small neighborhood of x), i.e., for all $y, z \in V$, it follows that

$$|f(y) - f(z)| \leq L_f \|y - z\|.$$

Then $f'_H(x; d)$ exists for any $d \in \mathbb{R}^n$ and the mapping $d \mapsto f'_H(x; d)$ is L_f-Lipschitz continuous on $\mathbb{R}^n$.

8. Let $\Phi : \mathbb{R}^n \rightrightarrows \mathbb{R}^m$ be a set-valued mapping, which maps $\mathbb{R}^n$ into the power set of $\mathbb{R}^m$ (the set of all subsets of $\mathbb{R}^m$), i.e., for any $x \in \mathbb{R}^n$, $\Phi(x) \subset \mathbb{R}^m$. We say Φ is convex if its graph

$$\operatorname{gph}(\Phi) := \left\{(x, y) \in \mathbb{R}^n \times \mathbb{R}^m \middle| y \in \Phi(x), x \in \mathbb{R}^n\right\}$$

is a convex set. Show Φ is convex if and only if for any $x, z \in \mathbb{R}^n$ and $0 \leq \lambda \leq 1$,

$$\lambda \Phi(x) + (1 - \lambda)\Phi(z) \subset \Phi(\lambda x + (1 - \lambda) z).$$

9. A set-valued mapping $\Phi : \mathbb{R}^n \rightrightarrows \mathbb{R}^n$ is called monotone if for any x, z with $\Phi(x) \neq \varnothing, \Phi(z) \neq \varnothing$, it follows that

$$\langle g_x - g_z, x - z\rangle \geq 0 \text{ whenever } g_x \in \Phi(x),\ g_z \in \Phi(z).$$

Show that the subdifferential mapping $\partial f : \mathbb{R}^n \rightrightarrows \mathbb{R}^n$ of a proper convex function $f : \mathbb{R}^n \to \overline{\mathbb{R}}$ is monotone.

10. For $f, g : \mathbb{R}^n \to \overline{\mathbb{R}}$ satisfying $f \leq g$, show that $f^* \geq g^*$.

11. Let $f, g : \mathbb{R}^n \to \overline{\mathbb{R}}$ be proper, consider the problem

$$\min_{x \in \mathbb{R}^n} \ f(x) + (g \circ \mathcal{A})(x),$$

where $\mathcal{A} : \mathbb{R}^n \to \mathbb{R}^n$ is a linear transformation. Show the Fenchel's dual problem is

$$\max_{y \in \mathbb{R}^n} \left\{ - \left(f \circ \mathcal{A}^T\right)^* (y) - g^*(-y) \right\}.$$

12. Let both $f : \mathbb{R}^n \to \overline{\mathbb{R}}$ and $\varphi : \mathbb{R}^n \times \mathbb{R}^m \to \overline{\mathbb{R}}$ with $\varphi(\cdot, 0) = f(\cdot)$ be not necessarily convex. Consider the following two unconstrained problems:

$$(P) \quad \min_{x \in \mathbb{R}^n} \ f(x),$$
$$(P_u) \quad \min_{x \in \mathbb{R}^n} \ \varphi(x, u).$$

We say (P_u) is the parametric problem of (P). Define the optimal value function associated with (P_u) is

$$v(u) := \inf_{x\in\mathbb{R}^n} \varphi(x,u).$$

Show that

(a) $\varphi(\cdot,u) : \mathbb{R}^n \to \overline{\mathbb{R}}$ is proper if and only if $v(u) < \infty$.

(b) $v^*(\cdot) = \varphi^*(0,\cdot)$. Recall that the conjugate φ^* is defined as

$$\varphi^*(y,w) = \sup_{(x,u)\in\mathbb{R}^n\times\mathbb{R}^m} \{\langle y,x\rangle + \langle w,u\rangle - \varphi(x,u)\}.$$

(c) Define the conjugate dual problem of (P_u):

$$(D_u) \quad \max_{w\in\mathbb{R}^m} \{\langle w,u\rangle - \varphi^*(0,w)\}.$$

Then $v(u)$ is no less than the optimal value of (D_u).

13. Show the 1-smoothness of

$$f(x) = \log\left(\sum_{i=1}^n \exp(x_i)\right)$$

w.r.t. l_2-norm and l_∞-norm.

14. If $f : \mathbb{R}^n \to \mathbb{R}$ is second order L_f-Lipschitz smooth, i.e., for any x, y we have

$$\left\|\nabla^2 f(x) - \nabla^2 f(y)\right\|_2 \le L_f \|x-y\|,$$

where the matrix norm $\|\cdot\|_2 : \mathbb{R}^{n\times n} \to \mathbb{R}_+$ is defined by

$$\|A\|_2 = \sup_{0\ne x\in\mathbb{R}^n} \frac{\|Ax\|}{\|x\|} \quad \text{for any } A\in\mathbb{R}^{n\times n}.$$

Then

$$f(y) \le f(x) + \langle\nabla f(x), y-x\rangle + \frac{1}{2}(y-x)^T\nabla^2 f(x)(y-x) + \frac{L_f}{6}\|y-x\|^3.$$

Hint: for $A\in\mathbb{R}^{n\times n}$ and $x\in\mathbb{R}^n$, we have $\|Ax\| \le \|A\|_2\|x\|$.