Operations Research Assignment 2

1. Consider the unconstrained optimization problem

$$\min f(x) \quad \text{s.t.} \quad x \in \mathbb{R}^n, \tag{P}$$

where $f : \mathbb{R}^n \to \mathbb{R}$ is convex and L_f-smooth. Assume $X^* \subset \mathbb{R}^n$, the optimal set of (P), is nonempty. Let f^* be the optimal value. Recall that the following useful details on GD method (convex case).

   (a) the iterative process:

$$x^{k+1} = x^k - \frac{1}{L_k} \nabla f(x^k),$$

   (b) $1/L_k$ is the k-th stepsize chosen by the constant-stepsize or the backtracking procedure.

   (c) L_k satisfies for any $x, y \in \mathbb{R}^n$,

$$f(x) - f\left(T_{L_k}(y)\right) \geq \frac{L_k}{2} \left\|x - T_{L_k}(y)\right\|^2 - \frac{L_k}{2} \|x - y\|^2 + f(x) - f(y) - \langle \nabla f(y), x - y \rangle.$$

   (d) the bounds on L_k (independent with k) is

$$\beta L_f \leq L_k \leq \alpha L_f.$$

   (e) the theorem on sequence convergence under Fejér monotonicity.

Show that

   (a) $\left\{f\left(x^k\right)\right\}_{k\geq 0}$ is nonincreasing.

   (b) $\left\{\left\|x^k - x^*\right\|\right\}_{k\geq 0}$ is nonincreasing for any $x^* \in X^*$.

   (c) $f(x^k) - f^* \leq \frac{\alpha L_f \left\|x^0 - x^*\right\|^2}{2k}$ for any $k \geq 1$ and $x^* \in X^*$.

   (d) $\left\{x^k\right\}_{k\geq 0}$ converges to some optimal solution as $k \to \infty$.

   (e) $\min\limits_{n=0,1,\cdots,k} \left\|\nabla f\left(x^n\right)\right\| \leq \frac{2\alpha^{1.5} L_f \left\|x^0 - x^*\right\|}{\sqrt{\beta} k}$ for any $k \geq 1$ and $x^* \in X^*$.

   (f) Under the constant stepsize rule in which $L_k \equiv L_f$ and $\alpha = \beta = 1$,

$$\left\|\nabla f\left(x^k\right)\right\| \leq \frac{2L_f \left\|x^0 - x^*\right\|}{k}$$

for any $k \geq 1$ and $x^* \in X^*$.

Hint: prove the norm of the gradient $\left\{\left\|\nabla f(x^k)\right\|\right\}_{k\geq 0}$ is nonincreasing.

2. Let $f : \mathbb{R}^n \to \mathbb{R}$ be L-smooth and f^* be the minimum of f over $\mathbb{R}^n$. Apply the iteration $x^{k+1} = x^k - t_k \nabla f\left(x^k\right)$. Next is the update scheme of $\{t_k\}$:

> Define $\varphi_k(t) = f\left(x^k - t\nabla f\left(x^k\right)\right)$ and $\hat{\varphi}_k(t) = f\left(x^k\right) - \frac{t}{3}\left\|\nabla f\left(x^k\right)\right\|^2$. Then $t_k = \frac{1}{2^{i_k}}$, where i_k is the smallest nonnegative integer i for which
>
> $$\varphi_k\left(\frac{1}{2^i}\right) \leq \hat{\varphi}_k\left(\frac{1}{2^i}\right).$$

Show that $\left\{f\left(x^k\right)\right\}_{k\geq 0}$ is nonincreasing and $\nabla f\left(x^k\right) \to 0$ as $k \to \infty$.

3. (a) Consider

$$f(x) = \begin{cases} -cx, & \text{if } x < 0 \\ x(\frac{x^2}{3} - c), & \text{if } x \geq 0, \end{cases}$$

where c is a positive parameter and $x \in \mathbb{R}$. If we use the Newton's method to find the global minimizer and assume that the initial point $x^0 \neq \sqrt{c}$, please write the iterative formula for the generated sequence $\{x^k\}_{k\geq 0}$ and derive the convergence by the formula.

(b) Consider the following logistic regression model,

$$\min_{x\in\mathbb{R}^n} l(x) = \frac{1}{m}\sum_{i=1}^{m} \ln\left(1 + \exp\left(-b_i a_i^T x\right)\right) + \lambda \|x\|^2,$$

where $\{(a_i, b_i)\}_{i=1}^m \subseteq \mathbb{R}^n \times \mathbb{R}$ is a given data set, $\lambda > 0$ is a parameter. Write down the iterative process for the Newton's method.

4. Let Ω be a nonempty, convex and bounded subset of $\mathbb{R}^n$. Define

$$\mu_\Omega(x) := \sup\{\|x - \omega\| \mid \omega \in \Omega\},\ x \in \mathbb{R}^n.$$

Show that

(a) μ_Ω is a well-defined real-valued function on $\mathbb{R}^n$. i.e $\mu_\Omega(x) < \infty$ for any $x \in \mathbb{R}^n$.

(b) Use the definition of convex functions to show that μ_Ω is convex.

5. Recall that function $f : \mathbb{R}^n \to \mathbb{R}$ is $\sigma-$strongly convex if the following inequality holds for any $x, y \in \mathbb{R}^n$ and $\lambda \in [0, 1]$:

$$f\left(\lambda x + (1-\lambda)y\right) \leq \lambda f(x) + (1-\lambda)f(y) - \frac{\sigma}{2}\lambda(1-\lambda)\|x - y\|^2,$$

Show that

(a) $g_1(\cdot) := f(\cdot) - \frac{\sigma}{2}\|\cdot\|^2$ is convex.

(b) assume that x^* is the minimizer of f, then $g_2(\cdot) := f(\cdot) - \frac{\sigma}{2} \|\cdot - x^*\|^2$ is convex.

(c) assume that $h : \mathbb{R}^n \to \mathbb{R}$ is convex, then $f + h$ is also $\sigma-$strongly convex.

6. Let $\Omega \subseteq \mathbb{R}^n$ be nonempty, closed and convex, define the distance of $x \in \mathbb{R}^n$ to Ω by

$$d_\Omega(x) := \inf\{\|x - z\| \mid z \in \Omega\},$$

and the projection of $x \in \mathbb{R}^n$ to Ω by

$$P_\Omega(x) := \{\omega \in \Omega \mid \|x - w\| = d_\Omega(x)\}.$$

Show that $P_\Omega(x)$ is a singleton of $\mathbb{R}^n$ for any $x \in \mathbb{R}^n$.

7. (a) Let $f : \mathbb{R}^n \to \mathbb{R}$ be a real-valued convex function, we denote the collection of all subgradients of f at x by

$$\partial f(x) := \left\{ v \in \mathbb{R}^n \,\middle|\, f(y) \geq f(x) + \langle v, y - x\rangle \text{ for all } y \in \mathbb{R}^n \right\}.$$

Then let $\Omega \subseteq \mathbb{R}^n$ be a convex set with $\bar{x} \in \Omega$. Show $0 \in \partial f(\bar{x}) + N_\Omega(\bar{x})$ implies that $\bar{x}$ is a minimizer of f over Ω.

(b) Let $\Omega \subset \mathbb{R}^n$ be a convex set and $f : \mathbb{R}^n \to \mathbb{R}$ be of class C^2 with $\bar{x} \in \Omega$. If for any $\bar{x} \neq x \in \Omega$ we have

i. $(x - \bar{x})^T \nabla f(\bar{x}) = 0$;

ii. $(x - \bar{x})^T \nabla^2 f(\bar{x})(x - \bar{x}) > 0$.

Then, $\bar{x}$ is a strict local minimizer of f over Ω.

8. Let $\Omega \subset \mathbb{R}^n$ be closed and $\bar{x} \in \Omega$. Recall the tangent directions of Ω at $\bar{x}$ is given by

$$T(\bar{x}) = \left\{ d \in \mathbb{R}^n \,\middle|\, \begin{array}{l} \exists \{t_k, d^k\} \subset \mathbb{R}_{++} \times \mathbb{R}^n \text{ with } ① t_k \downarrow 0, d^k \to d \text{ as } k \to \infty; \\ \hfill ② \bar{x} + t_k d^k \in \Omega \text{ for all } k \geq 0. \end{array} \right\}$$

If in addition Ω is convex, then

$$T(\bar{x}) = \mathrm{cl}\left\{\alpha(x - \bar{x}) \,\middle|\, \forall \alpha \geq 0 \text{ and } x \in \Omega\right\}.$$

Hint: WTS $A = \mathrm{cl}(B)$. Firstly, show that $B \subset A$, then derive $A \subset \mathrm{cl}(B)$. Finally we shall check the closedness of A.

9. Let $\Omega \subset \mathbb{R}^n$ be closed and $\bar{x} \in \Omega$. Define the normal directions of Ω at $\bar{x}$ is given by

$$N(\bar{x}) = \left\{ d \in \mathbb{R}^n \,\middle|\, \limsup_{x \xrightarrow{\Omega} \bar{x}} \frac{\langle d, x - \bar{x}\rangle}{\|x - \bar{x}\|} \leq 0 \right\},$$

where $x \overset{\Omega}{\to} \bar{x}$ means $x \to \bar{x}$ with $x \in \Omega$. If in addition Ω is convex, then

$$N(\bar{x}) = \left\{ d \in \mathbb{R}^n \middle| \langle d, x - \bar{x} \rangle \leq 0 \ \forall x \in \Omega \right\}.$$

10. Suppose that $\Omega \subset \mathbb{R}^n$ is closed convex, and $D \subset \mathbb{R}^n$ is compact convex.

    (a) Prove that the set

    $$D - \Omega := \left\{ d - w \middle| \forall d \in D \text{ and } w \in \Omega \right\}$$

    is convex.

    (b) If in addition $D \cap \Omega = \varnothing$, then $\exists p \in \mathbb{R}^n$ with

    $$\inf_{x \in D} \langle p, x \rangle > \sup_{y \in \Omega} \langle p, y \rangle$$

    (c) Given an example to show part (b) fails if D is closed convex but not necessarily bounded.

    Hint: consider special convex function and its asymptote.