The Chinese University of Hong Kong, Shenzhen
School of Data Science

DDA6010/CIE6010 · Assignment 4

Due: 23:59, November 6

Instructions:

- Assignment problems must be carefully and clearly answered to receive full credit. Complete sentences that establish a clear logical progression are highly recommended.
- You must submit your assignment in Blackboard. Please upload a pdf file with codes. The file name should be in the format **last name-first name-student ID-hw1**, e.g. **Zhang-San-123456789-hw1**.
- Please make your solutions legible and write your solutions in English. You are strongly encouraged to type your solutions in LaTeX/Markdown or others.
- Late submission will **not** be graded.
- Each student **must not copy** assignment solutions from another student or from any other source.
- For those questions that ask you to write MATLAB/Python/other codes to solve the problem. Please attach your code in the **pdf file**. You also need to clearly state (write or type) the optimal solution and the optimal value you obtained. However, you do not need to attach the outputs in the command window of MATLAB/Python/others.

Problem 1 Epigraph (15pts).
Let $f : \mathbb{R}^n \to \mathbb{R} \cup \{\infty\}$ be an extended real-valued function. Show that f is convex if and only if $epi(f)$ is convex.

Problem 2 Optimality Condition (15pts).
Consider the problem

$$\text{(P)} \quad \begin{array}{ll} \min & f(x) \\ \text{s.t.} & g(x) \leq 0 \\ & x \in X, \end{array}$$

where f and g are convex functions over $\mathbb{R}^n$ and $X \subseteq \mathbb{R}^n$ is a convex set. Suppose that x^* is an optimal solution of (P) that satisfies $g(x^*) < 0$. Show that x^* is also an optimal solution to the problem

$$\begin{array}{ll} \min & f(x) \\ \text{s.t.} & x \in X. \end{array}$$

Problem 3 Pertubed Convex Optimization (40pts). Let $f : \mathbb{R}^d \to \mathbb{R}$ be a smooth and strongly convex function with smooth parameter L and strong convexity parameter μ, that is, for all $x, y \in \mathbb{R}^d$,

$$f(x) + \langle \nabla f(x), y - x \rangle + \frac{\mu}{2}\|y - x\|^2 \leq f(y) \leq f(x) + \langle \nabla f(x), y - x \rangle + \frac{L}{2}\|y - x\|^2 \tag{1}$$

Consider the problem

$$\text{(P)} \quad \min_{x \in \mathbb{R}^d} \quad f(x).$$

Let $\{x_k\}_{k \geq 0}$ be the sequence generated by the update:

$$x^{k+1} = x^k - \alpha_k \left(\nabla f(x^k) + e^k \right),$$

where e^k satisfies

$$\left\| e^k \right\|^2 \leq \sigma^2, \text{ and } \left\langle e^k, \nabla f(x^k) \right\rangle = \left\langle e^k, x^k - x^* \right\rangle = 0, \ k \geq 0.$$

Let x^* be the optimal solution of (P) and let f^* be the optimal value of (P). Prove

(a) It holds that

$$\|\nabla f(x)\|^2 \geq 2\mu \left[f(x) - f^* \right], \text{ for all } x \in \mathbb{R}^d.$$

(b) It holds for $\alpha_k \leq 1/L$ that

$$f(x^{k+1}) - f^* \leq (1 - \alpha_k \mu) \left[f(x_k) - f^* \right] + \frac{\alpha_k^2 L \sigma^2}{2}, \ k \geq 0.$$

(c) It holds that

$$\langle \nabla f(x) - \nabla f(y), x - y \rangle \geq \frac{L\mu}{L + \mu} \|x - y\|^2 + \frac{1}{L + \mu} \|\nabla f(x) - \nabla f(y)\|^2, \text{ for all } x, y \in \mathbb{R}^d.$$

(d) It holds for $\alpha_k \leq 1/(L + \mu)$ that

$$\left\| x^{k+1} - x^* \right\|^2 \leq \left(1 - \frac{2\alpha_k L \mu}{L + \mu} \right) \left\| x^k - x^* \right\|^2 - \frac{\alpha_k}{\mu + L} \left\| \nabla f(x^k) \right\|^2 + \alpha_k^2 \sigma^2, \ k \geq 0.$$

Problem 4 Stationary Point (15pts). Consider the minimization problem

$$\text{(P)} \quad \min\{f(x) : x \in \Delta_n\}$$

where f is a continuously differentiable function over Δ_n. Prove that $x^* \in \Delta_n$ is a stationary point of (P) if and only if there exists $\mu \in \mathbb{R}$ such that

$$\frac{\partial f}{\partial x_i}(x^*) \begin{cases} = \mu, & x_i^* > 0 \\ \geq \mu, & x_i^* = 0 \end{cases}$$

Problem 5 Convex Optimization (15pts). Consider the minimization problem

$$\text{(Q)} \quad \begin{array}{l} \min \ 2x_1^2 + 3x_2^2 + 4x_3^2 + 2x_1x_2 - 2x_1x_3 - 8x_1 - 4x_2 - 2x_3 \\ \text{s.t. } \ x_1, x_2, x_3 \geq 0 \end{array}$$

(i) Show that the vector $(\frac{17}{7}, 0, \frac{6}{7})^T$ is an optimal solution of (Q).

(ii) Employ the gradient projection method with constant stepsize $\frac{1}{L}$ (L being the Lipschitz constant of the gradient of the objective function). Plot the function values of the first 100 iterations and output the produced solution (Set the initial point as $(1, 1, 1)^\mathsf{T}$).