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

DDA6010/CIE6010 · Assignment 3

Due: 23:59, October 20

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 Convex Sets (25pts).

(a) Verify whether the following sets are convex or not and explain your answer:

$$X_1 = \left\{(\boldsymbol{x}, t) \in \mathbb{R}^n \times \mathbb{R} : \boldsymbol{x}^\intercal \boldsymbol{x} \leq t^2\right\}$$
$$X_2 = \left\{\boldsymbol{x} \in \mathbb{R}^n : (\boldsymbol{a}^\intercal \boldsymbol{x})^2 \leq \alpha\right\},\ a \in \mathbb{R}^n, \alpha \geq 0$$

(b) Prove that

$$conv\{e_1, e_2, -e_1, -e_2\} = \{x \in \mathbb{R}^2 : |x_1| + |x_2| \leq 1\},$$

where $e_1 = (1, 0)^T, e_2 = (0, 1)^T$.

(c) Show that the conic hull of the set:

$$S = \{(x_1, x_2) : (x_1 - 1)^2 + x_2^2 = 1\}.$$

is the set

$$\{(x_1, x_2) : x_1 > 0\} \cup \{(0, 0)\}.$$

Remark: This is an example illustrating the fact that the conic hull of a closed set is not necessarily a closed set.

(d) Let S be a convex set. Prove that $x \in S$ is an extreme point of S if and only if $S \setminus \{x\}$ is convex.

Solution:

(a) The set X_1 is not convex. Counterexample:

$$n = 1, \begin{pmatrix} 1 \\ 1 \end{pmatrix}, \begin{pmatrix} 1 \\ -1 \end{pmatrix}.$$

The set X_2 is convex. We can rewrite X_2 as $X_2 = \{\boldsymbol{x} \in \mathbb{R}^n : g(\boldsymbol{x}) \leq 0\}$, where $g(\boldsymbol{x}) = (\boldsymbol{a}^\intercal \boldsymbol{x})^2 - \alpha$. Since $g(\boldsymbol{x})$ is a convex function, we can conclude that X_2 is a convex set.

(b) Let $x = (\lambda_1 - \lambda_3) e_1 + (\lambda_2 - \lambda_4) e_2$, when $\lambda_i > 0$ and $\lambda_1 + \lambda_2 + \lambda_3 + \lambda_4 = 1$.

$$\Rightarrow x = (\lambda_1 - \lambda_3, \lambda_2 - \lambda_4) \in \operatorname{conv}\{e_1, e_2, -e_1, -e_2\}$$
$$\Rightarrow |x_1| + |x_2| = |\lambda_1 - \lambda_3| + |\lambda_2 - \lambda_4|$$

Since $\lambda_i > 0 : |x_1| + |x_2| \leqslant |\lambda_1 + \lambda_3| + |\lambda_2 + \lambda_4| = 1$.
Then $\operatorname{conv}\{e_1, e_2, -e_1, -e_2\} \subseteq \left\{x \in \mathbb{R}^2 || x_1| + |x_2 | \leq 1\right\}$.
Let $x = (x_1, x_2)$ s.t. $|x_1| + |x_2| \leqslant 1$. Consider $x = \lambda_1 e_1 + \lambda_2 e_2 - \lambda_3 e_1 - \lambda_4 e_2$. A trivial representation:

$$\lambda_1 = \begin{cases} x_1, \text{ if } x_1 > 0 \\ 0, \text{ otherwise} \end{cases}, \lambda_2 = \begin{cases} x_2, \text{ if } x_2 > 0 \\ 0, \text{ otherwise} \end{cases}$$
$$\lambda_3 = \begin{cases} -x_1, \text{ if } x_1 < 0 \\ 0. \text{ otherwise} \end{cases}, \lambda_4 = \begin{cases} -x_2, \text{ if } x_2 < 0 \\ 0, \text{ otherwise} \end{cases}$$

Then $\lambda_1 + \lambda_2 + \lambda_3 + \lambda_4 \leq 1$ since $|x_1| + |x_2| \leq 1$.
If $\sum \lambda_i = 1$: done; if $\sum \lambda_i < 1$: W.L.O.G. assume $\lambda_1 = x_1, \lambda_2 = x_2$,

$$\text{let } \lambda_1 = x_1 + \frac{1 - x_1 - x_2}{2}, \lambda_2 = x_2, x_3 = \frac{1 - x_1 - x_2}{2}, \lambda_4 = 0$$

Then $\lambda_1 + \lambda_2 + \lambda_3 + \lambda_4 = 1$ and x is still (x_1, x_2). Then $x \in \operatorname{conv}\{e_1, e_2, -e_1, -e_2\}$.

$$\Rightarrow \left\{x \in \mathbb{R}^2 || x_1| + |x_2 | \leq 1\right\}.$$

Therefore, $\operatorname{con}\{e_1, e_2, -e_1, -e_2\} = \left\{x \in \mathbb{R}^2 || x_1| + |x_2 | \leq 1\right\}$.

(c) For any (x_1, x_2) satisfying $(x_1 - 1)^2 + x_2^2 = 1$:
If $x_1 > 0 : (x_1, x_2) \in \{(x_1, x_2) : x_1 > 0) \cup \{(0, 0)\}$
If $x_1 = 0 : x_2 = 0$, then $(x_1, x_2) \in \{(x_1, x_2) : x_1 > 0) \cup \{(0, 0)\}$
Therefore $\left\{(x_1, x_2) \mid (x_1 - 1)^2 + x_2^2 = 1\right\} \subseteq \{(x_1, x_2) : x_1 > 0\} \cup \{(0, 0)\}$.
Since $\{(x_1, x_2) : x_1 > 0\} \cup \{(0, 0)\}$ is conic, $(x_1 > 0 \Rightarrow \lambda x_1 > 0, x_1 = x_2 = 0 \Rightarrow \lambda x_1 = \lambda x_2 = 0)$

$$LHS = \text{ cone } \left(\left\{(x_1, x_2) \mid (x_1 - 1)^2 + x_2^2 = 1\right\}\right) \subseteq \{(x_1, x_2) : x_1 > 0\} \cup \{(0, 0)\} = RHS$$

Now consider $(x_1, x_2) = (0, 0), (0, 0) \in LHS$ trivially;
for (x_1, x_2) with $x_1 > x_2 > 0$: let $\lambda_1 = \frac{x_1 - x_2}{2}, \lambda_2 = x_2$

$$\Rightarrow (x_1, x_2) = \lambda_1(2, 0) + \lambda_2(1, 1), \text{ and } (2 - 1)^2 + 0^2 = 1, (1 - 1)^2 + 1^2 = 1$$

for (x_1, x_2) with $x_1 > 0, x_2 \leqslant 0$: similarly $\lambda_1 = \frac{x_1 - x_2}{2}, \lambda_2 = -x_2,$

$$(x_1, x_2) = \lambda_1(2, 0) + \lambda_2(1, -1) \in LHS$$

for (x_1, x_2) with $x_2 > x_1 > 0$: suppose $x_2 = kx_1$, solve:

$$(x-1)^2 + k^2x^2 = 1 \Rightarrow x = 0 \text{ or } \frac{2}{k^2+1}$$

Then let $\lambda_1 = \frac{(k^2+1)x_1}{2}, (x_1, x_2) = \lambda_1 \left(\frac{2}{k^2+1}, \frac{2k}{k^2+1} \right) \in LHS.$ $\Rightarrow$ RHS $\subseteq$ LHS Therefore: $cone \left\{ (x_1, x_2) \mid (x_1 - 1)^2 + x_2^2 = 1 \right\} = \{(x_1, x_2) \mid x_1 > 0\} \cup \{(0, 0)\}.$

(d) $(\Rightarrow)$ If $x \in S$ is an extreme point:
for any $x_1, x_2 \in S$, if $x_1 \neq x_2$ and $\lambda \in (0, 1)$,

$$\lambda x_1 + (1 - \lambda)x_2 \neq x$$

this implies only when $x_1 = x_2 = x$.

$$\lambda x_1 + (1 - \lambda)x_2 = x$$

Then for any $x_1, x_2 \in S\backslash\{x\}$, since $x_1 \neq x, x_2 \neq x$:

$$\lambda x_1 + (1 - \lambda)x_2 \neq x$$

implying $\lambda x_1 + (1 - \lambda)x_2 \in S\backslash\{x\}$.
Then $S\backslash\{x\}$ is convex.
$(\Leftarrow)$ If $S\backslash\{x\}$ is convex, for any $x_1, x_2 \in S\backslash\{x\}, x_1 \neq x_2$, $\lambda \in (0, 1)$: $\quad \lambda x_1 + (1-\lambda)x_2 \in S\backslash\{x\}$, i.e. $\lambda x_1 + (1 - \lambda)x_2 \neq x$ Now consider replacing x_1 with x,

$$\lambda x + (1 - \lambda)x_2 = x$$

$\Leftrightarrow \lambda = 1$, which is not in (0.1).
Similarly, keep original x_1 and replacing $x_2 = x$:

$$\lambda x_1 + (1 - \lambda)x = x$$

$\Leftrightarrow \lambda = 0$, which is not in $(0, 1)$.
Therefore, for any $x_1, x_2 \in S, x_1 \neq x_2, \lambda \in (0, 1)$:

$$\lambda x_1 + (1 - \lambda)x_2 \neq x$$

Since $x \in S : x$ is an extreme point of S.

Problem 2 Convex and Concave Functions (25pts).

(a) Let $C \subseteq \mathbb{R}^n$ be a nonempty set. The support function of C is the function $\sigma_C : \mathbb{R}^n \to (-\infty, \infty]$ given by

$$\sigma_C(\boldsymbol{y}) = \max_{\boldsymbol{x} \in C} \boldsymbol{y}^\intercal \boldsymbol{x}.$$

Show that σ_C is a closed and convex function.

(b) Let $f : \mathbb{R}^n \times \mathbb{R}^m \to \mathbb{R}$ and $C \subseteq \mathbb{R}^m$ ($C \neq \emptyset$) be given and assume that $f(\cdot, \boldsymbol{y}) : \mathbb{R}^n \to \mathbb{R}$ is convex for every fixed (but arbitrary) $\boldsymbol{y} \in C$. Show that the mapping $g : \mathbb{R}^n \to (-\infty, \infty]$, $g(\boldsymbol{x}) = \sup_{\boldsymbol{y} \in C} f(\boldsymbol{x}, \boldsymbol{y})$ is convex.

(c) Let $f : \mathbb{R}^n \times \mathbb{R}^m \to \mathbb{R}$ be L-smooth and $f(\boldsymbol{x}, \cdot)$ be μ-strongly concave (for an arbitrary but fixed $\boldsymbol{x}$). Suppose the constraint set $C \subseteq \mathbb{R}^m$ ($C \neq \emptyset$) is a convex and bounded set. Prove that the function $\mathbf{y}^*(\cdot) = \arg\max_{\boldsymbol{y} \in C} f(\cdot, \boldsymbol{y})$ is κ-Lipschitz, where $\kappa := L/\mu$. (Hint: You may refer to [1, Lemma 4.3].)

Solution:

(a) For a fixed $\boldsymbol{x}$, the linear function $\boldsymbol{y} \mapsto \boldsymbol{y}^\intercal \boldsymbol{x}$ is closed and convex. Since the support function is a maximum of a closed and convex function, we can conclude that it is closed and convex.

(b) Let $\boldsymbol{x}, \boldsymbol{z} \in \mathbb{R}^n$ and $\lambda \in [0, 1]$ be arbitrary. For every arbitrary but fixed $\boldsymbol{y} \in C$, we can now apply the convexity of the function $f(\cdot, \boldsymbol{y})$, which yields

$$f(\lambda \boldsymbol{x} + (1 - \lambda)\boldsymbol{z}, \boldsymbol{y}) \leq \lambda f(\boldsymbol{x}, \boldsymbol{y}) + (1 - \lambda) f(\boldsymbol{z}, \boldsymbol{y}).$$

Taking the supremum with respect to $\boldsymbol{y}$ over both sides, we then obtain

$$g(\lambda \boldsymbol{x} + (1 - \lambda)\boldsymbol{z}) \leq \sup_{\boldsymbol{y} \in C} [\lambda f(\boldsymbol{x}, \boldsymbol{y}) + (1 - \lambda) f(\boldsymbol{z}, \boldsymbol{y})] \leq \lambda g(\boldsymbol{x}) + (1 - \lambda) g(\boldsymbol{z}),$$

implying the convexity of g.

(c) The function $\mathbf{y}^*(\cdot)$ is well-defined since $f(\boldsymbol{x}, \cdot)$ is strongly concave for each $\boldsymbol{x} \in \mathbb{R}^n$. Let $\boldsymbol{x}_1, \boldsymbol{x}_2 \in \mathbb{R}^n$. The definition of $\mathbf{y}^*(\boldsymbol{x}_1)$ and $\mathbf{y}^*(\boldsymbol{x}_2)$ imply that

$$(\boldsymbol{y} - \mathbf{y}^*(\boldsymbol{x}_1))^\intercal \nabla_y f(\boldsymbol{x}_1, \mathbf{y}^*(\boldsymbol{x}_1)) \leq 0, \forall \boldsymbol{y} \in C \tag{1}$$
$$(\boldsymbol{y} - \mathbf{y}^*(\boldsymbol{x}_2))^\intercal \nabla_y f(\boldsymbol{x}_2, \mathbf{y}^*(\boldsymbol{x}_2)) \leq 0, \forall \boldsymbol{y} \in C. \tag{2}$$

Letting $\boldsymbol{y} = \mathbf{y}^*(\boldsymbol{x}_1)$ in (1) and $\boldsymbol{y} = \mathbf{y}^*(\boldsymbol{x}_1)$ in (2) and adding them yields

$$[\mathbf{y}^*(\boldsymbol{x}_2) - \mathbf{y}^*(\boldsymbol{x}_1)]^\intercal [\nabla_y f(\boldsymbol{x}_1, \mathbf{y}^*(\boldsymbol{x}_1)) - \nabla_y f(\boldsymbol{x}_2, \mathbf{y}^*(\boldsymbol{x}_2))] \leq 0. \tag{3}$$

Recall that $f(\boldsymbol{x}_1, \cdot)$ is μ-strongly concave, we have

$$[\mathbf{y}^*(\boldsymbol{x}_2) - \mathbf{y}^*(\boldsymbol{x}_1)]^\intercal [\nabla_y f(\boldsymbol{x}_1, \mathbf{y}^*(\boldsymbol{x}_2)) - \nabla_y f(\boldsymbol{x}_1, \mathbf{y}^*(\boldsymbol{x}_1))] + \mu \|\mathbf{y}^*(\boldsymbol{x}_2) - \mathbf{y}^*(\boldsymbol{x}_1)\|^2 \leq 0. \tag{4}$$

Adding (3) and (4) with the L-smoothness of f, we have

$$\mu \|\mathbf{y}^*(\boldsymbol{x}_2) - \mathbf{y}^*(\boldsymbol{x}_1)\|^2 \leq L \|\mathbf{y}^*(\boldsymbol{x}_2) - \mathbf{y}^*(\boldsymbol{x}_1)\| \|\boldsymbol{x}_2 - \boldsymbol{x}_1\|,$$

indicating that $\mathbf{y}^*(\cdot)$ is κ-Lipschitz.

Problem 3 Log-Convexity (25pts).

Assume that $f : \mathbb{R}^n \to (0, \infty)$ is a positive-valued function. Show that $(a) \Rightarrow (b) \Rightarrow (c)$ and provide counterexamples to show that the converse implication is not true:

(a) $1/f$ is concave

(b) $log(f)$ is convex

(c) f is convex.

Hint: The generalized Arithmetic-Geometric inequality says that $\forall x, y, a, b \geq 0$ with $a + b = 1$ it holds that $ax + by \geq x^a y^b$.

Solution: $(a) \to (b)$: Since $\frac{1}{f}$ is concave:

$$\begin{aligned}
&\frac{1}{f\left(\lambda x_1+(1-\lambda) x_2\right))} \geqslant \frac{\lambda}{f\left(x_1\right)}+\frac{1-\lambda}{f\left(x_2\right)} \geqslant\left(\frac{1}{f\left(x_1\right)}\right)^\lambda\left(\frac{1}{f\left(x_2\right)}\right)^{1-\lambda} \\
&\Rightarrow -\log f\left(\lambda x_1+(1-\lambda) x_2\right) \geqslant -\lambda \log f\left(x_1\right)-(1-\lambda) \log f\left(x_2\right) \\
&\Rightarrow \log f\left(\lambda x_1+(1-\lambda) x_2\right) \leqslant \lambda \log f\left(x_1\right)+(1-\lambda) \log f\left(x_2\right) \\
&\Rightarrow \log f \text{ is convex}
\end{aligned}$$

$(b) \to (c)$: Since

$$\log f\left(\lambda x_1+(1-\lambda) x_2\right) \leq \lambda \log f\left(x_1\right)+(1-\lambda) \log f\left(x_2\right),$$

$$f\left(\lambda x_1+(1-\lambda) x_2\right) \leq f\left(x_1\right)^\lambda f\left(x_2\right)^{1-\lambda} \leq \lambda f\left(x_1\right)+(1-\lambda) f\left(x_2\right)$$

$$\Rightarrow f \text{ is convex}$$

Counterexamples: $f(x) = x$ is convex, but $\log f(x) = \log x$ is concave; $\log \frac{1}{x^2} = -2 \log x$ is convex, but $\frac{1}{f} = x^2$ is also convex.

Problem 4 Optimality Conditions for Convex Problems (25pts).

We consider the convex optimization problem

$$\min_{\boldsymbol{x} \in X} f(\boldsymbol{x}) + \varphi(\boldsymbol{x}), \tag{5}$$

where $X \subseteq \mathbb{R}^n$ is a nonempty, convex set and $f : X \to \mathbb{R}$, $\varphi : X \to \mathbb{R}$ are given convex functions. Furthermore, let $\boldsymbol{x}^* \in X$ be a feasible point and suppose that the mapping f is continuously differentiable in an open neighborhood U containing the convex set X. Show that $\boldsymbol{x}^*$ is a global solution of the Problem (5) if and only if the following optimality condition is satisfied:

$$\boldsymbol{x}^* \in X, \text{ and } \nabla f(\boldsymbol{x}^*)^\intercal (\boldsymbol{x} - \boldsymbol{x}^*) + \varphi(\boldsymbol{x}) - \varphi(\boldsymbol{x}^*) \geq 0, \ \forall \boldsymbol{x} \in X.$$

Solution:

"$\Leftarrow$" Suppose we have $\boldsymbol{x}^* \in X$ and it holds that $\nabla f(\boldsymbol{x}^*)^\intercal (\boldsymbol{x} - \boldsymbol{x}^*) + \varphi(\boldsymbol{x}) - \varphi(\boldsymbol{x}^*) \geq 0, \ \forall \boldsymbol{x} \in X$. Utilizing the differentiability and convexity of f, it follows

$$f(\boldsymbol{x}) - f(\boldsymbol{x}^*) \geq \nabla f(\boldsymbol{x}^*)^\intercal (\boldsymbol{x} - \boldsymbol{x}^*), \ \forall x \in X.$$

Combining these two inequalities, we obtain

$$f(\boldsymbol{x}) + \varphi(\boldsymbol{x}) - [f(\boldsymbol{x}^*) + \varphi(\boldsymbol{x}^*)] \geq 0,$$

for all $\boldsymbol{x} \in X$. Hence, $\boldsymbol{x}^*$ is a solution of Problem (5).
"$\Rightarrow$" Let $\boldsymbol{x} \in X$ be arbitrary. For $t \in [0,1]$, we define the following line segment $\boldsymbol{x}(t) := \boldsymbol{x}^* + t(\boldsymbol{x} - \boldsymbol{x}^*)$. Since the set X is convex, we have $\boldsymbol{x}(t) \in X$ for all $t \in [0,1]$. Since $\boldsymbol{x}^*$ is a solution of Problem (5), we obtain $f(\boldsymbol{x}(t)) + \varphi(\boldsymbol{x}(t)) - [f(\boldsymbol{x}^*) + \varphi(\boldsymbol{x}^*)) \geq 0$ for all $t \in [0,1]$. Moreover, the convexity of φ implies

$$\frac{\varphi(\boldsymbol{x}(t)) - \varphi(\boldsymbol{x}^*)}{t} \leq \frac{t\varphi(\boldsymbol{x}) + (1-t)\varphi(\boldsymbol{x}^*) - \varphi(\boldsymbol{x}^*)}{t} = \varphi(\boldsymbol{x}) - \varphi(\boldsymbol{x}^*), \ \forall t \in [0,1].$$

Then,

$$\begin{aligned} 0 &\leq \lim_{t \to 0^+} \frac{f(\boldsymbol{x}(t)) - f(\boldsymbol{x}^*)}{t} + \frac{\varphi(\boldsymbol{x}(t)) - \varphi(\boldsymbol{x}^*)}{t} \\ &\leq \frac{f(\boldsymbol{x}(t)) - f(\boldsymbol{x}^*)}{t} + \varphi(\boldsymbol{x}) - \varphi(\boldsymbol{x}^*) \\ &= f(\boldsymbol{x}^*)^\intercal(\boldsymbol{x} - \boldsymbol{x}^*) + \varphi(\boldsymbol{x}) - \varphi(\boldsymbol{x}^*). \end{aligned}$$

References

[1] Tianyi Lin, Chi Jin, and Michael Jordan. On gradient descent ascent for nonconvex-concave minimax problems. In Hal Daumé III and Aarti Singh, editors, *Proceedings of the 37th International Conference on Machine Learning*, volume 119 of *Proceedings of Machine Learning Research*, pages 6083–6093. PMLR, 13–18 Jul 2020.