Homework 2

Instructor: Lijun Zhang *Name:* Weikang Li, *StudentId:* 181220031

Notice

- The submission email is: **njuoptfall2019@163.com**.

- Please use the provided LaTeX file as a template. If you are not familiar with LaTeX, you can also use Word to generate a **PDF** file.

Problem 1: First-order Convexity Condition
If f is a continuous function on some interval **I**,

a) Prove that f is a convex function if and only if $\forall x_1, x_2 \in \mathbf{I}$,

$$f\left(\frac{x_1 + x_2}{2}\right) \leq \frac{1}{2}[f(x_1) + f(x_2)]. \tag{1}$$

b) Prove that $f(x) = e^x$ is a convex function.

c) If $m, n > 0, p > 1$ and $1/p + 1/q = 1$, prove that $mn \leq \frac{m^p}{p} + \frac{n^q}{q}$.

Solution.
   a) For this part, I referred to the Mathematical Analysis, Mei Jiaqiang.

Proof. The necessity is easy to prove by simply setting $\theta = \frac{1}{2}$.
Next, we are going to prove the sufficiency:
Let $a < b \in \mathbf{I}$, define $l(x)$ as a linear function that satisfies $l(a) = f(a), l(b) = l(b)$.
It's easy to see that $g(x) = f(x) - l(x)$ still fulfills the inequality above. According to the definition of convex function, we have to prove that $g(x) \leq 0, \forall x \in [a, b]$. Because of the continuity of $g(x)$, g max to M at $x_0 \in [a, b]$, so we just have to show that $M = 0$.
If $x_0 = a$ or $x_0 = b$, then $M = 0$.
If $x_0 \in (a, b)$, consider such set: $A = \{ y \in [a, b] \mid g(x) = M, \forall x \in [x_0, y] \}$, and $y_0 = sup(A)$. Because of the continuity of $g(x)$, we have $y_0 \in A$ and $g(x) = M, x \in [x_0, y_0]$.
We say that $y_0 = b$, in fact, if $y_0 < b, \exists \delta > 0, s.t. (y_0 - \delta, y_0 + \delta) \subset (a, b)$.
So, if $0 \leq h < \delta$:

$$M = g(y_0) = g\left(\frac{(y_0 - h) + (y_0 + h)}{2}\right) \leq \frac{1}{2}\left[g(y_0 - h) + g(y_0 + h)\right] \leq M$$

So we have $g(x) = M, \forall x \in (y_0 - \delta, y_0 + \delta)$, Specially, $[y_0, y_0 + \delta) \subset A$,which is contradicts with the definition of y_0, means that $y_0 = b$, and, $M = g(b) = 0$. $\square$

   b) According to the First-order Convexity Condition, consider $x, y \in \mathbf{dom}\ \mathbf{f}$, we have:

$$f(y) = e^y \ , \ f(x) + f'(x)(y - x) = e^x(1 + y - x)$$

It is easy to see that $e^{y-x} \geq 1 + y - x$, therefore we have:

$$f(y) \geq f(x) + f'(x)(y - x)$$

which means $f(x)$ is convex.
c) Consider $f(x) = log(x)$, it is easy to see that f is concave on $\mathbf{R}_{++}$, which means:

$$log(\theta x + (1 - \theta)y) \geq \theta log(x) + (1 - \theta)log(y), \ x, y \in \mathbf{dom}\ \mathbf{f}$$

Let $\theta = 1/p$, $x = m^p$, $y = n^q$, we have:

$$log(\frac{1}{p}m^p + \frac{1}{q}n^q) \geq \frac{1}{p}log(m^p) + \frac{1}{q}log(n^q),$$

Therefore:

$$mn \leq \frac{m^p}{p} + \frac{n^q}{q}$$

$\square$

Problem 2: Second-order Convexity Condition
Let $\mathcal{D} \subseteq \mathbf{R}^n$ be convex. For a function $f : \mathcal{D} \to \mathbf{R}$ and an $\alpha > 0$, we say that f is α-exponentially concave, if $\exp(-\alpha f(x))$ is concave on $\mathcal{D}$. Suppose $f : \mathcal{D} \to \mathbf{R}$ is twice differentiable, give the necessary and sufficient condition of that f is α-exponentially concave and the detailed proof.

Solution.
First, we give the conclusion: f is α-exponentially concave if and only if $\alpha \nabla f(x) \nabla f(x)^T - \nabla^2 f(x) \preccurlyeq 0$.
Proof. Let $g(x) = \exp(-\alpha x)$ and $h(x) = g(f(x))$.
According to the Second-order Convexity Condition, we have:

$$f \text{ is } \alpha\text{-exponentially concave} \iff h(x) \text{ is concave on } \mathcal{D} \iff \nabla^2 h(x) \preccurlyeq 0$$

From the chain rule of derivatives, we have:

$$\nabla^2 h(x) = g'(f(x))\nabla^2 f(x) + g''(f(x))\nabla f(x)\nabla f(x)^T$$
$$= \alpha e^{-\alpha f(x)} \left(\alpha \nabla f(x)\nabla f(x)^T - \nabla^2 f(x)\right) \preccurlyeq 0$$

$\because \alpha > 0$, $\alpha e^{-\alpha f(x)} > 0$. Therefore, we have $\alpha \nabla f(x)\nabla f(x)^T - \nabla^2 f(x) \preccurlyeq 0$

$\square$

Problem 3: Operations That Preserve Convexity
Show that the following functions $f : \mathbf{R}^n \to \mathbf{R}$ are convex.

a) $f(x) = \|Ax - b\|$, where $A \in \mathbf{R}^{m \times n}, b \in \mathbf{R}^m$ and $\| \cdot \|$ is a norm on $\mathbf{R}^m$.

b) $f(x) = -(\det(A_0 + x_1 A_1 + \cdots + x_n A_n))^{1/m}$, on $\{x | A_0 + x_1 A_1 + \cdots + x_n A_n \succ 0\}$ where $A_i \in \mathbf{S}^m$.

c) $f(x) = \mathbf{tr}((A_0 + x_1 A_1 + \cdots + x_n A_n)^{-1})$, on $\{x | A_0 + x_1 A_1 + \cdots + x_n A_n \succ 0\}$ where $A_i \in \mathbf{S}^m$.

Solution.
a) It is easy to see that $f(x) = \|Ax - b\|$ is composed with norm and an affine function. We know that norm is convex, so we have f is convex.
b) It is easy to see that $f(x)$ is composed with $g(X) = -(\det(X))^{1/m}$ and an affine transformation. Next, we will show that g is convex on its domain $\mathbf{S}^m_{++}$.
To show that, one of approaches is to convert the function to any line that intersects with its domain and verify its convexity by the single variable function that we just got.
Here, we define $h(t) = g(Z + tV)$, where $Z \succ 0$, and $V \in \mathbf{S}^n$.

$$
\begin{aligned}
h(t) &= -(\det(Z + tV))^{1/m} \\
&= -(\det(Z^{1/2}(I + tZ^{-1/2}VZ^{1/2})Z^{1/2}))^{1/m} \\
&= -(\det Z^{1/2} \det(I + tZ^{-1/2}VZ^{1/2}) \det Z^{1/2})^{1/m} \\
&= -(\det Z)^{1/m} \left(\prod_{i=1}^{n}(1 + t\lambda_i)\right)^{1/m}
\end{aligned}
$$

where $\lambda_1, ..., \lambda_i$ are the eigenvalues of $Z^{-1/2}VZ^{1/2}$. Because $\det Z > 0$ and the geometric mean $\left(\prod_{i=1}^{n}(1 + t\lambda_i)\right)^{1/m}$ is concave on $\mathbf{R}_{++}$, which can be proved directly by proving its Hessian matrix is negative semi-definite. Therefore, we have h is convex and f is convex.

c) It is easy to see that $f(x)$ is composed with $g(x) = \mathbf{tr}(X^{-1})$ and an affine transformation. Next, we will show that g is convex on its domain $\mathbf{S}_{++}^m$.

Here, we define $h(t) = g(Z + tV)$, where $Z \succ 0$, and $V \in \mathbf{S}^n$.

$$
\begin{aligned}
h(t) &= \mathbf{tr}((Z + tV)^{-1}) \\
&= \mathbf{tr}(Z^{-1}(I + tZ^{-1/2}VZ^{1/2})^{-1}) \\
&= \mathbf{tr}(Z^{-1}Q(I + t\mathbf{\Lambda})^{-1}Q^T) \qquad (Eigenvalue\ Decompopsition) \\
&= \mathbf{tr}(Q^T Z^{-1}Q(I + t\mathbf{\Lambda})^{-1}) \qquad (\mathbf{tr}(AB) = \mathbf{tr}(BA)) \\
&= \sum_{i=1}^{n}(Q^T Z^{-1}Q)_{ii}(1 + t\lambda_i)^{-1}
\end{aligned}
$$

It is easy to see that $(1 + t\lambda_i)^{-1}$ is convex. Therefore, h is convex and f is convex.

$\square$

Problem 4: Conjugate Function

Derive the conjugates of the following functions.

a) $f(x) = \max\{0, 1 - x\}$.

b) $f(x) = \ln(1 + e^{-x})$.

Solution.

a) $f(x) = \max\{0, 1 - x\} = \begin{cases} 1 - x & , x < 1 \\ 0 & , x \geq 1 \end{cases}$, $\quad f^*(y) = \sup_{x \in \mathbf{R}}(yx - f(x)) = \begin{cases} yx - 1 + x & , x < 1 \\ yx & , x \geq 1 \end{cases}$.

It is easy to see that:

If $y < -1$ and $y \geq 0$, $yx - f(x)$ has no upper bound;

If $-1 \leq y < 0$, $yx - f(x)$ reaches maximum value at $x = 1$.

Therefore, $\mathbf{dom}\ f^* = \{y \,|\, -1 \leq y < 0\}$, $f^*(y) = y$.

b) $f(x) = \ln(1 + e^{-x})$, $f^*(y) = \sup_{x \in \mathbf{R}}(yx - f(x)) = f^*(y) = \sup_{x \in \mathbf{R}}(yx - \ln(1 + e^{-x}))$.

Observed that $f(x)$ is monotonically decreasing on $\mathbf{dom}\ f$, it is easy to see that $yx - f(x)$ has no upper bound for all $y > 0$.

Thus when $y \leq 0$, $yx - f(x)$ reaches maximum value if $f'(x) = y$, which is $x = ln(1 + y) - ln(-y)$, $-1 < y < 0$.

Therefore $f^*(y) = (1 + y)ln(1 + y) - yln(-y)$, $\mathbf{dom}\ f^* = \{y \,|\, -1 < y < 0\}$.

$\square$

Problem 5: Optimality Condition

Prove that $x^\star = (1, 1, -1)$ is optimal for the optimization problem

$$
\begin{aligned}
\text{minimize} \quad & (1/2)x^T Px + q^T x + r \\
\text{subject to} \quad & -1 \leq x_i \leq 1, \quad i = 1, 2, 3
\end{aligned}
$$

where

$$
P = \begin{bmatrix} 13 & 12 & -2 \\ 12 & 17 & 6 \\ -2 & 6 & 12 \end{bmatrix}, \qquad q = \begin{bmatrix} -28.0 \\ -23.0 \\ 13.0 \end{bmatrix}, \qquad r = 1.
$$

Solution.

According to the optimality condition of differentiable function:

$$
x \text{ is optimal} \iff x \in \mathbf{X} \text{ and } \nabla f_0(x)^T(y - x) \geq 0, \forall y \in \mathbf{X}
$$

Let $f_0(x) = (1/2)x^T Px + q^T x + r$, we have:

$$
\nabla f_0(x^*) = x^* P + q^T = (-1, 0, 5).
$$

Let $y = (y_1, y_2, y_3)$, where $-1 \leq y_i \leq 1$, $i = 1, 2, 3$. Thus, we have:

$$
\nabla f_0(x^*)^T(y - x^*) = -y_1 + 5y_3 + 6 \geq 0.
$$

So, x^* is optimal for this problem.

$\square$

Problem 6: Equivalent Problems

Consider a problem of the form

$$
\begin{array}{ll}
\text{minimize} & f_0(x)/\left(c^T x + d\right) \\
\text{subject to} & f_i(x) \le 0, \quad i = 1, \ldots, m \\
& Ax = b
\end{array}
\tag{2}
$$

where $f_0, f_1, \ldots, f_m$ are convex, and the domain of the objective function is defined as

$$
\{x \in \mathbf{dom} \ f_0 \mid c^T x + d > 0\}.
$$

a) Show that the problem (2) is a quasiconvex optimization problem.

b) Show that the problem (2) is equivalent to

$$
\begin{array}{ll}
\text{minimize} & g_0(y, t) \\
\text{subject to} & g_i(y, t) \le 0, \quad i = 1, \ldots, m \\
& Ay = bt \\
& c^T y + dt = 1
\end{array}
\tag{3}
$$

where $g_i(y, t) = t f_i(y/t)$ and $\mathbf{dom} \ g_i = \{(y, t) \mid y/t \in \mathbf{dom} \ f_i, t > 0\}$, for $i = 0, 1, \ldots, m$. The variables are $y \in \mathbf{R}^n$ and $t \in \mathbf{R}$.

c) Show that the problem (3) is convex.

Solution.

a) To strat with, we need to verify the problem is quasiconvex. We have known that f_0 is convex, so $\mathbf{dom} \ f_0$ is convex. Also, $f_0(x)/c^T x + d \le \alpha$ is equivalent to $f_0(x) - \alpha(c^T x + d) \le 0$, which is a convex constraint, so the sublevel sets are convex.

b) Assuming that x is feasible in problem (2)

Here, we define $t = \frac{1}{c^T x + d}$, $y = \frac{x}{c^T x + d}$. Then we have $t > 0$, and it's easy to see that t, y is also feasible for problem (3) when $g_0(y, t) = f_0(x)/(c^T x + d)$.

On the contrary, assuming that y, t are feasible for the problem (3), which must leads to $t > 0$. Define that $x = y/t$, we have that $x \in \mathbf{dom} f_i$.

Then, $f_i(x) = g_i(y, t)/t \le 0$, and $A(y/t) = b$, which means $c^T x + d = (c^T y + dt)/t = 1/t$.

Therefore, we have $t = 1/(c^T x + d)$, and $f_0(x)/(c^T x + d) = t f_0(x) = g_0(y, t)$. So x is feasible in problem (2).

Thus, problem (2) is equivalent to problem (3).

$\square$