Optimization Methods **Fall 2019**

Homework 2

Instructor: Lijun Zhang *Name:* Renzhe Zhou, *StudentId:* 181220076

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 $\mathbf{I}$,

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) The necessity is obviously: take $\theta = 1/2$, we have if f is convex then $f(\frac{x_1+x_2}{2}) \leq \frac{1}{2}[f(x_1)+f(x_2)]$.
The sufficiency part is as following:
Firstly let's prove that any point in $[x,y]$ which is $\lambda x+(1-\lambda)y$ $(0 \leq \lambda \leq 1)$ can be expressed as $\frac{m}{2^k}x+(1-\frac{m}{2^k})y$. Using half-approximation, $\lambda x+(1-\lambda)y$ is in interval $\frac{x+y}{2}$ and $\{[\frac{x+y}{2}-\lambda x+(1-\lambda)y] > 0?x : y\}$, so continuously use half-approximation we can get $\lambda x+(1-\lambda)y+\epsilon = \frac{m}{2^k}x+(1-\frac{m}{2^k})y$ when $\epsilon \to 0$ and $k \to \infty$, in which $m/2^k = \frac{\lfloor 2^k\lambda \rfloor}{2^k}$. So when $k \to \infty$, $m/2^k \to \lambda$.
We have $f(\frac{3}{4}x+\frac{1}{4}y) = f(\frac{1}{2}(\frac{1}{2}(x+y)+x)) \leq \frac{1}{2}(f(\frac{1}{2}(x+y))+f(x)) \leq \frac{3}{4}f(x)+\frac{1}{4}f(y)$ because of $f(\frac{x+y}{2}) \leq \frac{1}{2}[f(x)+f(y)]$, repeating this process we can get $f(\frac{m}{2^k}x+(1-\frac{m}{2^k})y) \leq \frac{m}{2^k}f(x)+(1-\frac{m}{2^k})f(y)$, where $m \in \{2^i | i = 0, 1, 2, \cdots, k\}$. So because f is continuous, then $\lim_{\epsilon \to 0} f(x+\epsilon) = f(x)$ which means $f(\frac{m}{2^k}x+(1-\frac{m}{2^k})y) = f(\lambda x+(1-\lambda)y)$, $\frac{m}{2^k}f(x)+(1-\frac{m}{2^k})f(y) = \lambda f(x)+(1-\lambda)f(y)$, therefore, $f(\lambda x+(1-\lambda)y \leq \lambda f(x)+(1-\lambda)f(y)$, so f is convex.

b) Because of inequality of arithmetic and geometric means, $e^{\frac{x_1+x_2}{2}} = \sqrt{e^{x_1}e^{x_2}} \leq \frac{1}{2}(e^{x_1}+e^{x_2})$, which means $f(\frac{x_1+x_2}{2}) \leq \frac{1}{2}[f(x_1)+f(x_2)]$, through the conclusion from a), we can prove that $f(x) = e^x$ is a convex function.

c) Because $p > 1$,so $q = \frac{1}{1-\frac{1}{p}} > 1$ too. Take $g(x) = \ln x$, $g''(x) = -x^{-2} < 0$, so $g(x)$ is a concave function. Because $1/p+1/q = 1$, we have $g(\frac{1}{p}m^p+\frac{1}{q}n^q) \geq \frac{1}{p}g(m^p)+\frac{1}{q}g(n^q)$, which is $\ln(\frac{m^p}{p}+\frac{n^q}{q}) \geq \frac{1}{p}\ln(m^p)+\frac{1}{q}\ln(n^q) = \ln m+\ln n = \ln(mn)$. Because $g'(x) = 1/x > 0$, then we have $\frac{m^p}{p}+\frac{n^q}{q} \geq mn$.

done.

□

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.

- f is α-exponentially concave if and only if $\alpha\nabla f(x)\nabla f(x)^T - \nabla f^2(x) \preceq 0$.
- Prove: Let's take $g(x) = e^{-\alpha x}, h(x) = g(f(x))$, so f is α-exponentially concave if and only if $h(x)$ is a concave function if and only if $\nabla^2 h(x) \preceq 0$. From chain rules we have

$$\nabla h^2(x) = g'(f(x))\nabla^2 f(x) + g''(f(x))\nabla f(x)\nabla f(x)^T \tag{2}$$

and $g'(f(x)) = -\alpha e^{-\alpha f(x)}, g''(f(x)) = \alpha^2 e^{-\alpha f(x)}$, so

$$\nabla h^2(x) = \alpha e^{-\alpha f(x)}[\alpha\nabla f(x)\nabla f(x)^T - \nabla^2 f(x)] \tag{3}$$

because $\alpha > 0, e^{-\alpha f(x)} > 0$, so $\nabla h^2(x) \preceq 0$ if and only if $\alpha\nabla f(x)\nabla f^T(x) - \nabla f^2(x) \preceq 0$. Therefore, f is α-exponentially concave if and only if $\alpha\nabla f(x)\nabla f(x)^T - \nabla f^2(x) \preceq 0$.

done.

□

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) $f(x) = -(\det(A_0 + x_1A_1 + \cdots + x_nA_n))^{1/m}$, on $\{x|A_0 + x_1A_1 + \cdots + x_nA_n \succ 0\}$ where $A_i \in \mathbf{S}^m$.

c) $f(x) = \mathbf{tr}((A_0 + x_1A_1 + \cdots + x_nA_n)^{-1})$, on $\{x|A_0 + x_1A_1 + \cdots + x_nA_n \succ 0\}$ where $A_i \in \mathbf{S}^m$.

Solution.

a) Make g(x)=$\|x\|$,then $f(x) = g(Ax - b)$, and $g(x)$(norm) is convex, so after affine mapping $f(x)$ is convex too.

b) Make $g(X) = -(det(X))^{1/m}$, then $f(x) = g(Ax^T + A_0)$ where $A = (A_1, A_2, \cdots, A_n), x = (x_1, x_2, \cdots, x_n)$, so f is a composition and an affine transformation of g. Let's prove g is convex so that f is convex too.
Let's transform g into a line $h(t) = -(det(Z + tV))^{1/m}$ and prove h(t) is convex.
$h(t) = -(detZ)^{1/m}(det(I + tZ^{-1/2}VZ^{-1/2}))^{1/m} = -(detZ)^{1/m}(\prod_{i=1}^m(1 + t\lambda_i))^{1/m}$
where λ_i is the eigenvalues of $Z^{-1/2}VZ^{-1/2}$.
So h is a convex function of t on $\{t|Z + tV \succ 0\}$ since $-(det(Z))^{1/m} < 0$ and the geometric mean $(\prod_{i=1}^m(1 + t\lambda_i))^{1/m}$ is concave.
Above all, $f(x)$ is a convex function.

c) Make $g(X) = tr(X)^{-1}$, then $f(x) = g(Ax^T + A_0)$ where $A = (A_1, A_2, \cdots, A_n), x = (x_1, x_2, \cdots, x_n)$, so f is a composition and an affine transformation of g. Let's prove g is convex so that f is convex too.
Let's transform g into a line $h(t) = tr(Z + tV)^{-1}$ and prove h(t) is convex.
$h(t) = tr(Z^{-1}(I + tZ^{-1/2}VZ^{-1/2})^{-1}) = tr(Z^{-1}Q(I + t\Lambda)^{-1}Q^T)$
$= tr(Q^TZ^{-1}Q(I + t\Lambda)^{-1}) = \sum_{i=1}^n(Q^TZ^{-1}Q)_{ii}(1 + t\Lambda_i)^{-1}$
where $Z^{-1/2}VZ^{-1/2} = Q\Lambda Q^T$ and actually h is a convex function because the sum of a positive value of $(Q^TZ^{-1}Q)_{ii}$ multiplies $(1 + t\Lambda_i)^{-1}$ which is convex.
Above all, $f(x)$ is a convex function.

done.

□

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)

$$f(x) = \begin{cases} 1 - x, x < 1 \\ 0, x \geq 1 \end{cases} \tag{4}$$

When $y > 0, sup(y^T x - f(x)) = +\infty$;
when $y = 0, sup(y^T x - f(x)) = sup(-f(x)) = 0$;
when $-1 < y < 0, sup(y^T x - f(x)) = +\infty$;
when $y = -1, sup(y^T x - f(x)) = -1$;
when $y < -1, sup(y^T x - f(x)) = sup_{x=1}(y^T x - f(x)) = y - f(1) = y$;
so the conjugate of f(x) is $f^*(y) = y, y \in (-\infty, -1) \bigcup \{0\}$.

b) We have $f'(x) = \frac{-e^{-x}}{1+e^{-x}}$. Because $f'(x) < 0$, so $f(x)$ is a decreasing function whose limit is 0 when x tends to be positive infinity, so when $y > 0, sup(y^T x - f(x)) = +\infty$;
when $y = 0$, $sup(y^T x - f(x)) = sup(-f(x)) = 0$;
when $y < 0$, make $g(x) = yx - \ln(1 + e^{-x}), g'(x) = y + \frac{e^{-x}}{1+e^{-x}}$, $\frac{e^{-x}}{1+e^{-x}} \in (0, 1)$, so when $y \leq -1, f'(x) < 0$ which means $g(x)$ has no upper bound; when $-1 < y < 0$, make $f'(x) = 0$, we can get $x_0 = -\ln \frac{-y}{1+y}$, so $sup(g(x)) = g(x_0) = -y \ln \frac{-y}{1+y} + \ln(1 + y)$. Above all,

$$f^*(y) = \begin{cases} 0, y = 0 \\ -y \ln \dfrac{-y}{1+y} + \ln(1 + y), -1 < y < 0 \end{cases} \tag{5}$$

done.

□

Problem 5: Optimality Condition
Prove that $x^\star = (1, 1, -1)$ is optimal for the optimization problem

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

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.

· Make $f(x) = \frac{1}{2}x^T P x + q^T x + r$, the gradient of f(x) is $\nabla f(x) = Px + q$, and $\nabla f_0(x^*) = (-1, 0, 5)$, because for all y subject to $-1 \leq y_i \leq 1$, $\nabla f_0(x^*)^T(y - x) = -1 * (y_1 - 1) + 0 * (y_2 - 1) + 5 * (y_3 + 1) = (1 - y_1) + 5(y_3 + 1)$. And $-1 \leq y_i \leq 1$, so $(1 - y_1) + 5(y_3 + 1) \geq 0$, so x^* is optimal.

done.

□

Problem 6: Equivalent Problems
Consider a problem of the form

$$\begin{array}{ll} \text{minimize} & f_0(x)/\left(c^T x + d\right) \\ \text{subject to} & f_i(x) \leq 0, \quad i = 1, \dots, m \\ & Ax = b \end{array} \tag{6}$$

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

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

a) Show that the problem (6) is a quasiconvex optimization problem.

b) Show that the problem (6) is equivalent to

$$\begin{array}{ll} \text{minimize} & g_0(y,t) \\ \text{subject to} & g_i(y,t) \le 0, \quad i = 1,\dots,m \\ & Ay = bt \\ & c^T y + dt = 1 \end{array} \tag{7}$$

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

c) Show that the problem (7) is convex.

Solution.

a) Make $g(x) = f_0(x)/(c^T x + d)$, the domain of $g(x)$ is convex because f_0 is convex. Let's see the sublevel sets $S_\alpha = \{x \in \textbf{dom}\ g | g(x) \le \alpha\}$. From $\frac{f_0(x)}{c^T x+d} \le \alpha$ and $c^T x + d > 0$ we can get $f_0(x) \le \alpha(c^T x + d)$, so S_α is convex, therefore, the problem(6) is a quasiconvex optimization problem.

b) · Assume x is feasible for problem(6). Take $t = \frac{1}{c^T x+d}, y = \frac{x}{c^T x+d}$, let's prove y, t is feasible for problem(7). So $g_0(y,t) = tf_0(y/t) = f_0(x)/(c^T x + d)$, $g_i(y,t) = tf_i(y/t) = f_i(x)/(c^T x + d), (i = 1, \cdots, m)$, We have $g_i(y,t) \le 0$ is equivalent to $f_i(x) \le 0$ because $c^T x + d > 0$;
$Ay = bt$ is equivalent to $A\frac{x}{c^T x+d} = b\frac{1}{c^T x+d}$ which is $Ax = b$;
And $c^T y + dt = c^T \frac{x}{c^T x+d} + d\frac{1}{c^T x+d} = 1$.
Therefore, y, t is feasible for problem(7).

· On the contrary, assume y, t are feasible for problem(7), let's prove x is feasible for problem(6).
make $x = \frac{y}{t}$, because g_i is the perspective function of f_i then we must have $t > 0$.
$c^T x + d = (c^T(y/t) + d) = \frac{c^T y+dt}{t} = \frac{1}{t}$, so $f_0(x)/(c^T x + d) = tf_0(y/t) = g_0(y,t)$;
And $f_i(x) = g_i(y,t)/t \le 0$ because $g_i(y,t) \le 0$;
And $Ax = b$ is equivalent to $A(y/t) = b$ which is $Ay = bt$.
So x is feasible for problem(6).

Above all, problem(6) is equivalent to problem(7).

c) Because f_i is convex, so their perspective function g_i is convex too. And $Ay - bt = 0$,$c^T y + dt - 1 = 0$ are affine, so problem(7) is convex.

done.

□