Optimization Methods **Fall 2019**

Homework 2

Instructor: Lijun Zhang *Name:* 高辰潇, *StudentId:* 181220014

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) : 充分性:

反证，假设有 $f\left(\frac{x_1+x_2}{2}\right) \leq \frac{1}{2}[f(x_1)+f(x_2)]$ 成立但 $f(x)$ 不为凸函数，则 $\exists a, b \in \mathbb{I}, \exists \theta \in [0,1]$，使得 $f(\theta a + (1-\theta)b) > \theta f(a) + (1-\theta)f(b)$.

令 $c = \theta a + (1-\theta)b$，$\Phi(x) = f(x) - \left[f(a) + \frac{f(b)-f(a)}{b-a}(x-a)\right]$.

易见 $\Phi(a) = \Phi(b) = 0$. 下面对 $\Phi(x)$ 进行讨论：

情况一，若 $\Phi(x)$ 在区间 (a,b) 上恒大于 0，则 $f(\frac{a+b}{2}) > \frac{1}{2}[f(a)+f(b)]$，矛盾，假设不成立。故 $f(x)$ 为凸函数。

情况二，若 $\Phi(x)$ 在区间 (a,b) 上不恒大于 0，由 $\Phi(c) > 0$ 结合零点存在定理知，存在区间 (a,b) 上的零点 $d_1, d_2, ..., d_n$。选取其中一段使得函数值恒大于 0 的区间 (d_i, d_{i+1})，由情况一知在这个区间上产生矛盾，假设不成立，$f(x)$ 为凸函数。

综上所述，$f(x)$ 为凸函数，充分性得证。

必要性:

若 f 为凸函数，取 $\theta = \frac{1}{2}$，可立即得到 $\forall x_1, x_2 \in \mathbf{I}$，都有 $f\left(\frac{x_1+x_2}{2}\right) \leq \frac{1}{2}[f(x_1)+f(x_2)]$。

b) : $\because f(x)=e^x$ 为凸函数

$\Leftrightarrow e^{\frac{x_1+x_2}{2}} \leq \frac{1}{2}(e^{x_1}+e^{x_2})$

$\Leftrightarrow e^{\frac{x_1-x_2}{2}} - \frac{1}{2}(e^{x_1-x_2}+1) \leq 0$

令 $t=e^{\frac{x_1-x_2}{2}}$, 由 $-\frac{1}{2}t^2+t-\frac{1}{2} \leq 0$ 知 $e^{\frac{x_1-x_2}{2}} - \frac{1}{2}(e^{x_1-x_2}+1) \leq 0$

因此 $f(x)=e^x$ 为凸函数

c) : 令 $g(x)=-\ln x$，取 $\theta=\frac{1}{p}$，则 $1-\theta=\frac{1}{q}$.

由凸函数定义知，$f(\theta m^p+(1-\theta)n^q)=-\ln(\frac{m^p}{p}+\frac{n^q}{q}) \leq \frac{1}{p}(-\ln m^p)+\frac{1}{q}(-\ln n^q)=-\ln(mn)$

即 $\ln(mn) \leq \ln(\frac{m^p}{p}+\frac{n^q}{q})$

对上式两边取对数即可得到 $mn \leq \frac{m^p}{p}+\frac{n^q}{q}$.

$\square$

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. 充要条件为 $\nabla^2 f(x)-\alpha\nabla f(x)\nabla f(x)^T \succcurlyeq 0$。证明如下:

证明:

$\nabla^2 f(x)-\alpha\nabla f(x)\nabla f(x)^T \succcurlyeq 0$

$\Leftrightarrow e^{-\alpha f(x)}(\nabla^2 f(x)-\alpha\nabla f(x)\nabla f(x)^T) \succcurlyeq 0$

$\Leftrightarrow -\alpha\left[e^{-\alpha f(x)}(\nabla^2 f(x)-\alpha\nabla f(x)\nabla f(x)^T)\right] \preccurlyeq 0$

$\Leftrightarrow \nabla^2(e^{-\alpha f(x)}) \preccurlyeq 0$

$\Leftrightarrow e^{-\alpha f(x)}$ 在 $\mathcal{D}$ 上为凹函数

$\Leftrightarrow f$ 在 D 上 α-exponentially concave.

$\square$

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) $f(x)=-(\det(A_0+x_1A_1+\cdots+x_nA_n))^{1/m}$, 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) : 对于 $\forall x_1,x_2 \in \mathbb{R}^n$ 和 $\forall\theta \in [0,1]$, 有

$$\begin{aligned} f(\theta x_1+(1-\theta)x_2) &= \|A[\theta x_1+(1-\theta)x_2]-b\| \\ &= \|\theta(Ax_1-b)+(1-\theta)(Ax_2-b)\| \\ &\leq \|\theta(Ax_1-b)\|+\|(1-\theta)(Ax_2-b)\| \\ &= \theta\|Ax_1-b\|+(1-\theta)\|Ax_2-b\| \\ &= \theta f(x_1)+(1-\theta)f(x_2) \end{aligned} \tag{2}$$

其中利用了范数的的三角不等式性质和范数的齐次性。由以上证明可知，f 为凸函数。

b)：首先证明函数 $h(X) = -\det(X)^{1/m}$ $X \in \mathbb{S}^m_{++}$ 关于 X 为凸函数。

将 $h(X)$ 的定义域限制在任意一条穿过定义域的直线 $A + tB$ 上，其中 $A, B \in \mathbb{S}^m, A \in \mathbb{S}^m_{++}$.

令 $g(t) = h(A + tB) = -\det(A + tB)^{1/m}$, 则

$$\begin{aligned} g(t) &= -\det(A + tB)^{1/m} \\ &= -\det(A^{\frac{1}{2}}(I + tA^{-\frac{1}{2}}BA^{-\frac{1}{2}})A^{\frac{1}{2}})^{1/m} \\ &= -\det(A)^{1/m} \cdot (\prod_{i=1}^{m}(1 + t\lambda_i))^{1/m} \end{aligned} \tag{3}$$

其中 λ_i 为 $A^{-\frac{1}{2}}BA^{-\frac{1}{2}})A^{\frac{1}{2}}$ 的特征值。由于 $(\prod_{i=1}^{m}(1 + t\lambda_i))^{1/m}$ 是关于 t 的几何均值，为 t 的凹函数，因此 $g(t)$ 为关于 t 的凸函数。

故 $h(X)$ 为关于矩阵 X 的凸函数。

又由于函数 $A_0 + x_1A_1 + \cdots + x_nA_n$ 为 x 的仿射，因此它与 h 的复合函数为凸函数。

c)：首先证明函数 $h(X) = \mathbf{tr}(X^{-1}), X \in \mathbb{S}^m_{++}$ 为关于 X 的凸函数。

采用和 b) 中相同的思路，令 $g(t) = h(A + tB) = \mathbf{tr}((A + tB)^{-1})$，则

$$\begin{aligned} g(t) &= \mathbf{tr}((A + tB)^{-1}) \\ &= \mathbf{tr}(A^{-\frac{1}{2}}(I + tA^{-\frac{1}{2}}BA^{-\frac{1}{2}})^{-1}A^{-\frac{1}{2}}) \\ &= \mathbf{tr}(A^{-1}(I + tA^{-\frac{1}{2}}BA^{-\frac{1}{2}})^{-1}) \\ &= \mathbf{tr}(A^{-1}(Q(1 + t\Lambda)Q^T)^{-1}) \\ &= \mathbf{tr}(Q^TA^{-1}Q(1 + t\Lambda)^{-1}) \\ &= \sum_{i=1}^{m}\frac{(Q^TA^{-1}Q)_{ii}}{1 + t\lambda_i} \end{aligned} \tag{4}$$

由于 $A \in \mathbb{S}^m_{++}$，故 $A^{-1} \in \mathbb{S}^m_{++}$. 又由于 Q 为正交矩阵，故 $Q^TA^{-1}Q \in \mathbb{S}^m_{++}$. 因此 $(Q^TA^{-1}Q)_{ii} > 0$.

由以上讨论知，$g(t)$ 可看作一系列凸函数 $\frac{1}{1+t\lambda_i}$ 的非负加权和，故函数 $g(t)$ 为凸函数，$h(X)$ 为关于 X 的凸函数。

由于 $f(x)$ 可看作 $h(X)$ 和仿射函数 $A_0 + x_1A_1 + \cdots + x_nA_n$ 的复合，因此 $f(x)$ 为凸函数。

$\square$

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) ：由共轭函数的定义知，$f^*(y) = \sup_{x\in\mathbb{R}}\{xy - \max\{0, 1-x\}\}$

当 $x < 1$ 时，$f^*(y) = \sup\{xy - 1 + x\}$, 且函数 $xy - 1 + x$ 在的导数为 $y + 1$。当 $y + 1 < 0$ 时函数无上界，$y + 1 \geq 0$ 时上界为 y.

当 $x \geq 1$ 时，$f^*(y) = \sup\{xy\}$, 若 $y \leq 0$，函数 xy 的上界为 y，若 $y > 0$，函数 xy 无上界。

综合以上讨论可知

$$f^*(y) = \begin{cases} y & (y \in [-1, 0]) \\ +\infty & (y \notin [-1, 0]) \end{cases}$$

b) ：$f^*(y) = \sup_{x\in\mathbb{R}}\{xy - \ln(1 + e^{-x})\}$

令 $g(x) = xy - \ln(1 + e^{-x})$，则 $g'(x) = y + \frac{e^{-x}}{1+e^{-x}}$，显然 $y \geq 0$ 时 $g(x)$ 单调增，无上界。

当 $y \leq -1$ 时，$g'(x) < 0$，因此 $g(x)$ 单调减。又由于 $\lim_{x\to-\infty} g(x) = (y + 1)x \to +\infty$，故此时 $g(x)$ 无上界。

当 $-1 < y < 0$ 时，令 $g'(x) = 0$，得到 $x = \ln(-\frac{y+1}{y})$ 时 $g(x)$ 取得最大值 $(y + 1)\ln(y + 1) - y\ln(-y)$

综合以上讨论可知

$$f^*(y) = \begin{cases} (y + 1)\ln(y + 1) - y\ln(-y) & (y \in (-1, 0)) \\ +\infty & (y \notin (-1, 0)) \end{cases}$$

$\square$

Problem 5: Optimality Condition

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

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

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. 显然 $x^\star$ 在可行域内。令 $f(x) = (1/2)x^T Px + q^T x + r$，则 $\nabla f(x) = P^T x + q$。

则 $\nabla f(x^\star) = (-1, 0, 5)^T$. 任取可行域内一点 y，则 $\nabla f(x)^T(y - x) = 5y_3 - y_1 + 6$.

由于 $-1 \leq y_1 \leq 1$,$-1 \leq y_3 \leq 1$

故 $\nabla f(x)^T(y - x) \geq -6 + 6 = 0$

由最优解的一阶判定准则知，点 $x^\star = (1, 1, -1)$ 是原优化问题的最优解。

$\square$

Problem 6: Equivalent Problems

Consider a problem of the form

$$\begin{array}{ll} \text{minimize} & f_0(x)/\left(c^T x + d\right) \\ \text{subject to} & f_i(x) \le 0, \quad i = 1, \ldots, m \\ & Ax = b \end{array} \tag{5}$$

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 (**??**) is a quasiconvex optimization problem.

b) Show that the problem (**??**) is equivalent to

$$\begin{array}{ll} \text{minimize} & g_0(y, t) \\ \text{subject to} & g_i(y, t) \le 0, \quad i = 1, \ldots, m \\ & Ay = bt \\ & c^T y + dt = 1 \end{array} \tag{6}$$

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 (**??**) is convex.

Solution.

a)：为了证明该问题是一个拟凸优化问题，只需要证明目标函数 $g_0(x) = \frac{f_0(x)}{c^T x + d}$ 是拟凸的，也就只需要证明目标函数的任意非空下水平集为凸集。（空集本身就可以看作是凸的，不做讨论）

对于 $\forall t \in \mathbb{R}$，令 $g_0(x) \le t$，得到满足该条件的下水平集 $S_t = \{x | f_0(x) \le t(c^T x + d)\}$

假设该下水平集非空，任取 $x_1, x_2 \in S_t$，有 $f_0(x_i) \le t(c^T x_i + d)$，$i \in \{1, 2\}$

故对于任意的 $\theta \in [0, 1]$，有

$$\begin{aligned} & f_0(\theta x_1 + (1-\theta) x_2) \\ & \le \theta f_0(x_1) + (1-\theta) f_0(x_2) \\ & \le \theta t(c^T x_1 + d) + (1-\theta) t(c^T x_2 + d) \\ & = t(c^T(\theta x_1 + (1-\theta) x_2) + d) \end{aligned} \tag{7}$$

即 $\theta x_1 + (1-\theta) x_2 \in S_t$.

故目标函数的任意非空下水平集均为凸集。

因此原问题是拟凸优化问题。

b)：对于原问题中的任意一点 x，令 $t = \frac{1}{c^T x + d}, y = tx$，下面证明该从 x 到 (y, t) 的映射为双射。

对于原问题可行域中的任意一点 x，对应的 $t = \frac{1}{c^T x + d} > 0$，且 $y/t = x \in \mathbf{dom} f_i$。同时代入 $x = y/t$ 可发现 (y, t) 满足问题 (6) 中的所有约束，因此 x 的像 (y, t) 在问题 (6) 的可行域中。

对于 (y, t)，若在原问题中存在两个不同的 x_1, x_2 与之对应，则有 $tx_1 = tx_2$，与 x_1, x_2 互异矛盾。因此，上述映射为单射。

另一方面，对于问题 (6) 可行域中的任意一点 (y, t)，其原像为 $x = y/t \in \mathbf{dom} f_i$. 同时将问题 (6) 中的 (y, t) 替换为 x 可发现 x 满足原问题的所有约束，因此对于问题 (6) 可行域中的每一个点 (y, t)，在原问题的可行域中都存在一个点 x 与 (y, t) 对应。因此，上述映射为满射。

最后，在上述映射下可以验证二者目标函数值相同。由于 x 到 (y, t) 的映射为双射，由变量替换规则可知原问题和问题 (6) 等价。

c)：对于问题 (6)，易见其中的等式约束均为仿射的。

由函数 g_i 定义的形式可知，$g_i(x)$ 为 $f_i(x)$ 的视角函数。由于视角函数与原函数的凹凸性相同，因此 $g_i(x)$ 也是凸函数

故问题 (6) 的目标函数、约束函数均为凸函数，且等式约束均为仿射。因此问题 (6) 为凸优化问题。

$\square$