Linear and Convex Optimization Homework 04

Qiu Yihang, 2021/10/25-10/27

1.(a) ***Proof:***

$f(x) = \log x$ is obviously concave. (Since $\mathbf{dom} f = \mathbb{R}^+, f'(x) = \frac{1}{x}, f''(x) = -\frac{1}{x^2} < 0$.)

For any $\boldsymbol{x} = (x_1, x_2, \dots, x_n)^T \in \Delta_{n-1}$, we know $x_i \geq 0, \forall i \in \{1,2,\dots,n\}$ and $\sum_{i=1}^n x_i = 1$.

Let $\|\boldsymbol{x}\|_0 = k$. Obviously $k \leq n$.

Without loss of generality, assume the first k components of x are nonzero, **i.e.**

$$x_i \begin{cases} > 0, & \forall i \in \{1,2,\dots,k\} \\ = 0, & \forall i \in \{k+1,\dots,n\} \end{cases}.$$

Now we prove $H(\boldsymbol{x}) = -\sum_{i=1}^n x_i \log x_i \leq k$ by proving the proposition below inductively:

$$\frac{-\sum_{i=1}^p x_i \log x_i}{\sum_{i=1}^p x_i} = \frac{\sum_{i=1}^p x_i \log \frac{1}{x_i}}{\sum_{i=1}^p x_i} \leq \log\left(\frac{p}{\sum_{i=1}^p x_i}\right), 1 \leq p \leq k.$$

BASE STEP. When $p = 1$,

$$\frac{x_1}{x_1}\log\left(\frac{1}{x_1}\right) = \log\left(\frac{1}{x_1}\right).$$

INDUCTIVE STEP.

Suppose when $p = m$, $\frac{\sum_{i=1}^p x_i \log \frac{1}{x_i}}{\sum_{i=1}^p x_i} \leq \log\left(\frac{p}{\sum_{i=1}^p x_i}\right)$.

When $p = m+1 \leq k$, since $f(x) = \log x$ is concave, we have

$$\begin{aligned} \frac{\sum_{i=1}^{m+1} x_i \log \frac{1}{x_i}}{\sum_{i=1}^{m+1} x_i} &= \frac{\sum_{i=1}^m x_i}{\sum_{i=1}^{m+1} x_i} \frac{\sum_{i=1}^m x_i \log \frac{1}{x_i}}{\sum_{i=1}^{m+1} x_i} + \frac{x_{m+1} \log \frac{1}{x_{m+1}}}{\sum_{i=1}^{m+1} x_i} \\ &\leq \frac{\sum_{i=1}^m x_i}{\sum_{i=1}^{m+1} x_i} \log\left(\frac{m}{\sum_{i=1}^m x_i}\right) + \frac{x_{m+1}}{\sum_{i=1}^{m+1} x_i} \log \frac{1}{x_{m+1}} \\ &= \log\left(\frac{\sum_{i=1}^m x_i}{\sum_{i=1}^{m+1} x_i} \frac{m}{\sum_{i=1}^m x_i} + \frac{x_{m+1}}{\sum_{i=1}^{m+1} x_i} \frac{1}{x_{m+1}}\right) = \log\left(\frac{m+1}{\sum_{i=1}^{m+1} x_i}\right) \end{aligned}$$

i.e.

$$\frac{-\sum_{i=1}^p x_i \log x_i}{\sum_{i=1}^p x_i} = \frac{\sum_{i=1}^p x_i \log \frac{1}{x_i}}{\sum_{i=1}^p x_i} \leq \log\left(\frac{p}{\sum_{i=1}^p x_i}\right)$$

still holds when $p = m+1 \leq k$.

Let $p = k$. Since $\sum_{i=1}^n x_i = 1, 0\log 0 = 0$, we have

$$H(\boldsymbol{x}) = -\sum_{i=1}^n x_i \log x_i = -\sum_{i=1}^k x_i \log x_i \leq \log\left(\frac{k}{1}\right) = \log\|\boldsymbol{x}\|_0 \leq \log n.$$

Qed. ■

(b) ***Proof:***

Let $C = \{\boldsymbol{x} \in \Delta_{n-1} : \boldsymbol{x} > \boldsymbol{0}\}, \boldsymbol{x} = (x_1, x_2, \dots, x_n)^T$.

First we prove $H(\boldsymbol{x}) = -\sum_{i=1}^{n} x_i \log x_i$ is strictly concave on C.

$$\nabla H(\boldsymbol{x}) = (-\log x_1 - 1, -\log x_2 - 1, \dots, -\log x_n - 1)^T,$$

$$\nabla^2 H(\boldsymbol{x}) = \begin{pmatrix} -\frac{1}{x_1} & 0 & \cdots & 0 \\ 0 & -\frac{1}{x_2} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & -\frac{1}{x_n} \end{pmatrix} \prec \boldsymbol{O} \text{ (Since } x_i > 0, \forall i \in \{1,2,\dots,n\})$$

Thus, $H(\boldsymbol{x})$ is strictly concave.

The constrain is $g(\boldsymbol{x}) = \sum_{i=1}^{n} x_i - 1$. Use Lagrange Method to calculate the conditioned extreme value. Let $\mathcal{L}(\boldsymbol{x}, \lambda) = H(\boldsymbol{x}) + \lambda g(\boldsymbol{x})$.

When $\nabla \mathcal{L}(\boldsymbol{x}, \lambda) = \boldsymbol{O}$ **i.e.** $(-\log x_1 - 1 + \lambda, \dots, -\log x_n - 1 + \lambda, \sum_{i=1}^{n} x_i - 1)^T = \boldsymbol{O}$, we know

$$\begin{cases} x_1 = 1/n \\ \quad \vdots \\ x_n = 1/n \\ \lambda = 1 - \log n \end{cases}.$$

In this case, $H(\boldsymbol{x}) = \log n$. Thus, $H(\boldsymbol{x})$ reaches a local maximum at $\overline{\boldsymbol{x}} = \left(\frac{1}{n}, \frac{1}{n}, \dots, \frac{1}{n}\right)^T$ on C.

Since $H(\boldsymbol{x})$ is strictly concave, $H(\overline{\boldsymbol{x}})$ is the unique global maximum of $H(\boldsymbol{x})$ on C.

Now we consider $H(\boldsymbol{x})$ on $\Delta_{n-1} \backslash C$.

$\forall \boldsymbol{x} \in \Delta_{n-1} \backslash C$, $\|\boldsymbol{x}\|_0 < n$. From (a) we know $H(\boldsymbol{x}) \le \log \|\boldsymbol{x}\|_0 < \log n = H(\overline{\boldsymbol{x}})$.

Therefore, $H(\overline{\boldsymbol{x}})$ is the unique global maximum of $H(\boldsymbol{x})$ on Δ_{n-1}.

Qed. ■

2.(a) ***Proof:***

Considering $a < s < \mu < u < b$, let $\theta = \frac{\mu - s}{u - s} \in [0,1], \bar{\theta} = 1 - \theta = \frac{u - \mu}{u - s}$.

By convexity of f,

$$f(\mu) = f\left(\frac{\mu - s}{u - s} u + \frac{u - \mu}{u - s} s\right) = f(\theta u + \bar{\theta} s) \le \theta f(u) + \bar{\theta} f(s) = \frac{\mu - s}{u - s} f(u) + \frac{u - \mu}{u - s} f(s)$$

$$\Leftrightarrow (u - \mu + \mu - s) f(\mu) \le (\mu - s) f(u) + (u - \mu) f(s)$$

$$\Leftrightarrow (u - \mu)\big(f(\mu) - f(s)\big) \le (\mu - s)\big(f(u) - f(\mu)\big)$$

i.e.

$$\frac{f(\mu) - f(s)}{\mu - s} \le \frac{f(u) - f(\mu)}{u - \mu}.$$

Qed. ■

(b) ***Proof:***

Let $\beta = \sup\limits_{a<s<\mu} \frac{f(\mu)-f(s)}{\mu-s}$.

Obviously $\beta \geq \frac{f(\mu)-f(s)}{\mu-s} > -\infty$ (since $f: (a,b) \to \mathbb{R}$).

From (a) we know $\forall s \in (a,\mu), \frac{f(\mu)-f(s)}{\mu-s} \leq \frac{f(u)-f(\mu)}{u-\mu} < +\infty$. Thus, $-\infty < \beta < +\infty$.

CASE 01. $\mu < x < b$.

From (a) we know $\forall s \in (a,\mu), \frac{f(\mu)-f(s)}{\mu-s} \leq \frac{f(x)-f(\mu)}{x-\mu}$.

Thus, $\beta \leq \frac{f(x)-f(\mu)}{x-\mu}$, **i.e.** $f(\mu) + \beta(x-\mu) \leq f(\mu) + \frac{f(x)-f(\mu)}{x-\mu}(x-\mu) = f(x)$.

CASE 02. $a < x < \mu$.

By the definition of β, we know $\beta \geq \frac{f(\mu)-f(x)}{\mu-x}$ (since $a < x < \mu$).

Thus, $f(\mu) + \beta(x-\mu) = f(\mu) - \beta(\mu - x) \leq f(\mu) - \frac{f(\mu)-f(x)}{\mu-x}(\mu-x) = f(x)$.

In conclusion, $f(x) \geq f(\mu) + \beta(x-\mu), \ \forall x \in (a,b)$.

Qed. ■

(c) ***Proof:***

From (b) we know $f(x) \geq f(\mu) + \beta(x-\mu), \ \forall x \in (a,b)$, where $\mu = \mathbb{E}X$.

Since X is a random variable taking values in (a,b), $f(X) \geq f(\mu) + \beta(X-\mu)$, □

By the isotonicity of expectation $\mathbb{E}$, we have

$$\mathbb{E}f(X) \geq \mathbb{E}\big(f(\mu) + \beta(X-\mu)\big).$$

By the linearity of expectation $\mathbb{E}$, we have

$$\mathbb{E}\big(f(\mu) + \beta(X-\mu)\big) = \mathbb{E}f(\mu) + \beta\mathbb{E}(X-\mu) = f(\mu) + \beta(\mathbb{E}X - \mathbb{E}\mu) = f(\mu) + \beta(\mathbb{E}X - \mu)$$
$$= f(\mu) + 0 = f(\mathbb{E}X).$$

Thus,

$$\mathbb{E}f(X) \geq f(\mathbb{E}X).$$

Qed. ■

3. ***Solution:***

S is a convex set. The proof is as follows. ■

Since $\|\cdot\|$ is a convex function, its affine composition $\|\boldsymbol{Ax}+\boldsymbol{b}\|$ is convex.

Considering x^3 is a convex function on $\mathbb{R}^+$, the composition of convex functions $\|\boldsymbol{Ax}+\boldsymbol{b}\|$ and x^3, **i.e.** $\|\boldsymbol{Ax}+\boldsymbol{b}\|^3$ is convex.

Since $\log(1+e^x)$ is convex and $3x_1+2x_2$ is affine, we know $\log(1+e^{3x_1+2x_2})$ is convex.

Therefore, $f(\boldsymbol{x}) = \max\{\|\boldsymbol{A}\boldsymbol{x}+\boldsymbol{b}\|^3, \log(1+e^{3x_1+2x_2})\}$ is a convex function.

By the definition of S, we know S is a **2**-sublevel set of f.

Since f is convex, its sublevel set S is convex. *Qed.* □

4.(a) ***Solution:***

The Problem (a) is a convex optimization problem. ■

(*Proof:*

Since $(x_1-x_2)^2$ is convex and x_1+x_2 is affine and thus convex, the objective function $f(\boldsymbol{x}) = (x_1-x_2)^2 + x_1 + x_2$ is convex.

Since $(x_1+x_2)^2$ and $e^{x_1+x_2}$ are both convex, inequality constraint function $g(\boldsymbol{x}) = (x_1+x_2)^2 + e^{x_1+x_2}$ is convex.

Regarding that objective function $f(\boldsymbol{x})$ and inequality constraint function $g(\boldsymbol{x})$ are both convex while equality constraint function $h(\boldsymbol{x}) = x_1 - 3x_2$ is affine, the problem (a) is a convex optimization problem.)

(b) ***Solution*:**

The Problem (b) is not a convex optimization problem. ■

(*Disproof:*

Consider the first inequality constraint function $g_1(\boldsymbol{x}) = x_1 e^{-(x_1+x_2)}$.

Let $\boldsymbol{x} = (3,-3), \boldsymbol{y} = (-1,3), \theta = \bar{\theta} = \frac{1}{2}$. Then $\theta\boldsymbol{x} + \bar{\theta}\boldsymbol{y} = (1,0)$.

$g_1(\theta\boldsymbol{x} + \bar{\theta}\boldsymbol{y}) = e > \theta g_1(\boldsymbol{x}) + \bar{\theta} g_1(\boldsymbol{y}) = \frac{3e^0 - e^2}{2} = \frac{3-e^2}{2}$. Thus, $g_1(\boldsymbol{x})$ is not convex.

Also, the equality constraint function $h(\boldsymbol{x}) = 6x_1^2 - 7x_2$ is not affine.

Therefore, the problem (b) is not a convex optimization problem.)