Homework 4

Zhen

October 23, 2021

problem 1

$$H(x) = \sum_{i=1}^{n} -x_i \log x_i$$

Solve:

(a) Apparently, $\log x$ is concave. By the assumption that $\|\boldsymbol{x}\|_0 = k$ and the first k components of $\boldsymbol{x}$ are nonzero, we have

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

(b) $\forall i, x_i > 0$, we have

$$\frac{\partial^2 f}{\partial x_i \partial x_j} = \begin{cases} 0 & \text{if } i \neq j \\ -1/x_i & \text{if } i = j \end{cases}$$

which means that $H(x)$ is strictly concave (every eigenvalues of $\nabla^2 f$ are negative).

Let $C \overset{def}{=} \{\boldsymbol{x} \in \Delta_{n-1} : \boldsymbol{x} > \boldsymbol{0}\}$. Then we hold $\forall \boldsymbol{x} \in C$, $H(\boldsymbol{x}) \leq \log n = H([1/n, \ldots, 1/n]^{\mathcal{T}})$ and $[1/n, \ldots, 1/n]^{\mathcal{T}}$ is the **unique** maximum.

Also $\forall \boldsymbol{x} \in \Delta_{n-1}/C$, $H(\boldsymbol{x}) \leq \log \|\boldsymbol{x}\|_2 < \log n$.

Thus $[1/n, \ldots, 1/n]^{\mathcal{T}}$ is the **unique** maximum of $H(\boldsymbol{x})$.

problem 2

(a) Proof:

$$\begin{aligned} & \frac{f(\mu) - f(s)}{\mu - s} \leq \frac{f(u) - f(\mu)}{u - \mu} \\ \iff & uf(\mu) - (u - \mu)f(s) \leq (\mu - s)f(u) + sf(\mu) \\ \iff & f(\mu) \leq \frac{u - \mu}{u - s} f(s) + \frac{\mu - s}{u - s} f(u) \\ \iff & f(\frac{u - \mu}{u - s} s + \frac{\mu - s}{u - s} u) \leq \frac{u - \mu}{u - s} f(s) + \frac{\mu - s}{u - s} f(u) \\ \Longleftarrow & f \text{ is convex.} \end{aligned}$$

(b) $\exists \beta \in \mathbb{R}$ such that $f(x) \geq f(\mu) + \beta(x - \mu)$.

Proof: Let

$$\beta = \sup_{a < s < \mu} \frac{f(\mu) - f(s)}{\mu - s}$$

Case 1 $\mu < x < b$:

By (a) we have $\forall a < s < \mu$:

$$\frac{f(\mu) - f(s)}{\mu - s} < \frac{f(x) - f(\mu)}{x - \mu}$$

Thus $\beta = \sup_{a<s<\mu} \dfrac{f(\mu)-f(s)}{\mu - s} < \dfrac{f(x)-f(\mu)}{x-\mu}$, which means $\beta(x-\mu)+f(\mu) < f(x)$.

Case 2 $a < x < \mu$:

Obviously,

$$\frac{f(\mu)-f(x)}{\mu - x} \leq \sup_{a<s<\mu} \frac{f(\mu)-f(s)}{\mu - s} = \beta$$

Thus $\beta(x-\mu)+f(\mu) \leq f(x)$.

Case 3 $x = \mu$: The proof is trivial.

(c) By (b), $\exists \beta \in \mathbb{R}, \forall x \in (a,b) : f(x) \geq f(\mu) + \beta(x-\mu)$. Then $f(X) \geq f(\mu) + \beta(X-\mu)$.

Finally,

$$\begin{aligned}
\mathbb{E}[f(X)] &= \int_a^b \Pr(X=x) f(X=x)\mathrm{d}x \\
&\geq \int_a^b \Pr(X=x)[f(\mu)+\beta(x-\mu)]\mathrm{d}x \\
&= f(\mu) + \beta(\mu-\mu) \\
&= f(\mathbb{E}[X])
\end{aligned}$$

problem 3

$$S = \left\{ \boldsymbol{x} = (x_1, x_2) \in \mathbb{R}^2 : \max \left\{ \|\boldsymbol{A}\boldsymbol{x} + \boldsymbol{b}\|^3, \log\left(1 + e^{3x_1+2x_2}\right) \right\} \leq 2 \right\}$$

Solve: convex.

We have some true propositions as follows:

1. If $f : \mathbb{R}^n \to \mathbb{R}$ is convex, then $f(\boldsymbol{A}\boldsymbol{x}+\boldsymbol{b})$ is also convex ($\forall \boldsymbol{A} \in \mathbb{R}^{n\times m}, \boldsymbol{x} \in \mathbb{R}^m, \boldsymbol{b} \in \mathbb{R}^n$).
2. If $f : \mathbb{R}^n \to \mathbb{R}, g : \mathbb{R}^n \to \mathbb{R}$ are convex, then $h = \max\{f, g\}$ is convex.
3. If $f : \mathbb{R}^n \to \mathbb{R}$ is convex, then $S = \{\boldsymbol{x} : f(\boldsymbol{x}) \leq Const\}$ is convex.

First we have: $\|\boldsymbol{x}\|^3$ and $\log(1+e^x)$ are convex.
By proposition 1 we have $\|\boldsymbol{A}\boldsymbol{x}+\boldsymbol{b}\|^3$and $\log\left(1+e^{3x_1+2x_2}\right)$ are convex.
Then by proposition 2, $\max\left\{\|\boldsymbol{A}\boldsymbol{x}+\boldsymbol{b}\|^3, \log\left(1+e^{3x_1+2x_2}\right)\right\}$ is convex.
Thus, by proposition 4, S is convex.

problem 4

(a) Solve: It is a convex optimization problem.

The objective function $f(\boldsymbol{x}) = \boldsymbol{x}^{\mathcal{T}} \begin{bmatrix} 1 & -1 \\ -1 & 1 \end{bmatrix} \boldsymbol{x} + \begin{bmatrix} 1 & 1 \end{bmatrix} \boldsymbol{x}$. By the fact that $\begin{bmatrix} 1 & -1 \\ -1 & 1 \end{bmatrix} \succeq \boldsymbol{O}$, $f(\boldsymbol{x})$ is convex.

Inequality constraint function $g(\boldsymbol{x}) = (x_1 - x_2)^2 + 4x_1x_2 + e^{x_1+x_2} = (x_1+x_2)^2 + e^{x_1+x_2}$. Also, g is convex because $(x_1+x_2)^2$ and $e^{x_1+x_2}$ are convex.

Obviously equality constraint function $h(x) = x_1 - 3x_2$ is a affine function.

(b) Solve: It is **NOT** a convex optimization problem.

The equality constraint function $h(x) = 6x_1^2 - 7x_2$ is not a affine function.