SUN Yilin
520030910361

Homework 4

1

(a).

If $||\boldsymbol{x}||_0 = k$, without loss of generality, assume the first k components are nonzero, it is obvious that log$||\boldsymbol{x}||_0 \leq$log$n$. Now we show $H(\boldsymbol{x}) \leq$log$||\boldsymbol{x}||_0$.
$H(\boldsymbol{x}) = -\sum_{i=1}^{n} x_i \text{log} x_i = \sum_{i=1}^{k} x_i \text{log}\frac{1}{x_i}$ because the components of $\boldsymbol{x}$ from $k+1$ to n are zero and we treat 0log0=0. Then $H(\boldsymbol{x}) = \sum_{i=1}^{k} x_i \text{log}\frac{1}{x_i} + \sum_{i=k+1}^{n} x_i log 1$ by adding some terms which are equal to zero. Then by $\sum_{i=1}^{n} x_i = 1$ and the concavity of log$x$ we get $H(\boldsymbol{x}) \leq$log$(\sum_{i=1}^{k} \frac{x_i}{x_i} + 0)$ =log$k$. Thus $H(\boldsymbol{x}) \leq$log$||\boldsymbol{x}||_0 \leq$log$n$.

(b).

Firstly $\bar{\boldsymbol{x}}$ is the maximum of $H(\boldsymbol{x})$ because $H(\bar{\boldsymbol{x}})$ =log$n$ and $H(\boldsymbol{x}) \leq$log$n$ as we have shown in part (a). Now we prove the uniqueness.
Define $C = \{\boldsymbol{x} \in \Delta_{n-1} : \boldsymbol{x} > 0\}$. The Hessian Matrix of $H(\boldsymbol{x})$ is just $diag\{-\frac{1}{x_1}, -\frac{1}{x_2}, \ldots, -\frac{1}{x_n}\}$, which is negative definite on C. This means that $H(\boldsymbol{x})$ is strictly concave on C, the maximum of $H(\boldsymbol{x})$ on C must be unique. For $\boldsymbol{x} \in \Delta_{n-1} \backslash C$, at least one component of $\boldsymbol{x}$ is zero, which means that $||\boldsymbol{x}||_0$ defined in part(a) is smaller than n. So $H(\boldsymbol{x}) \leq$log$||\boldsymbol{x}||_0$ <log$n$, which means that $H(\boldsymbol{x})$ will not acquire its maximum on $\boldsymbol{x} \in \Delta_{n-1} \backslash C$. So $\bar{\boldsymbol{x}}$ is the unique maximum of $H(\boldsymbol{x})$ on Δ_{n-1}.

2

(a).

By convexity of f we have $f(\frac{s(u-\mu)+u(\mu-s)}{u-s}) \leq \frac{u-\mu}{u-s} f(s) + \frac{\mu-s}{u-s} f(u)$, which is equivalent to $f(\mu) \leq \frac{u-\mu}{u-s} f(s) + \frac{\mu-s}{u-s} f(u)$, then by transformation we can get $\frac{f(\mu)-f(s)}{\mu-s} \leq \frac{f(u)-f(\mu)}{u-\mu}$.

(b).

Take

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

and it is obvious that $\beta > -\infty$. As we have shown in part(a), $\frac{f(\mu)-f(s)}{\mu-s} \leq \frac{f(u)-f(\mu)}{u-\mu}$, so $\beta < +\infty$. Now we show that inequality $(\star)$ holds for $x \in (a, b)$.
When $x = \mu$, $(\star)$ is just $f(\mu) \geq f(\mu)$.

When $x \in (a,\mu)$, since $\beta = \sup_{a<s<\mu} \frac{f(\mu)-f(s)}{\mu-s}$, $\beta \geq \frac{f(\mu)-f(x)}{\mu-x}$, which is equivalent to $f(x) \geq f(\mu) + \beta(x-\mu)$.
When $x \in (\mu, b)$, from part(a) we know $\frac{f(\mu)-f(s)}{\mu-s} \leq \frac{f(x)-f(\mu)}{x-\mu}$. Since $\beta = \sup_{a<s<\mu} \frac{f(\mu)-f(s)}{\mu-s}$, $\frac{f(x)-f(\mu)}{x-\mu} \geq \beta$, which is equivalent to $f(x) \geq f(\mu) + \beta(x-\mu)$.
So ($\star$) holds for $x \in (a,b)$.

(c).

From part(b) we have shown $f(x) \geq f(\mu) + \beta(x-\mu), \forall x \in (a,b)$. Since X is a random variable taking values in (a,b), $f(X) \geq f(\mu) + \beta(X-\mu)$. Then we take expectations for this inequality and get $\mathbb{E}f(X) \geq f(\mu) + \beta(\mathbb{E}X - \mu)$, but $\mathbb{E}X$ is just μ, so we get $\mathbb{E}f(X) \geq f(\mathbb{E}X)$.

3

It is convex.
Let $f_1(x_1,x_2) = ||\boldsymbol{Ax}+\boldsymbol{b}||^3$, $f_2(x_1,x_2) = \log(1+e^{3x_1+2x_2})$. Then by letting $f(x_1,x_2) = \max(f_1,f_2)$, we know that S is the 2-sublevel set for f.
First we show that both f_1 and f_2 are convex functions. For f_1, notice that $||\boldsymbol{Ax}+\boldsymbol{b}||$ is just an affine composition for the norm function, since norm functions are convex, it must be convex. Then notice that $y = x^3$ is convex and increasing on $(0,+\infty)$, f_1 is a scalar composition of $y = x^3$ on$(0,+\infty)$ with $||\boldsymbol{Ax}+\boldsymbol{b}||$ on $\mathbb{R}^2$, it must be convex. For f_2, it is easy to show $\log(1+e^y)$ is convex because as a univariate function its second derivative is $\frac{e^y}{(1+e^y)^2}$, which is larger than 0. Then f_2 is just an affine composition of $\log(1+e^y)$ and $y = 3x_1+2x_2$, f_2 is also convex.
Now we have convex functions f_1 and f_2, notice that f is just a pointwise maximum of f_1 and f_2, it must be convex. Then S is a sublevel set of a convex function f, it is convex.

4

(a).

It is a convex optimization problem.
The objective function is just $(x_1-x_2)^2 + (x_1+x_2)$. It is an affine composition of $y_1^2 + y_2$ with $y_1 = x_1 - x_2$ and $y_2 = x_1 + x_2$, since $y_1^2 + y_2$ is convex, it is also convex.
The inequality constraint function is just $(x_1+x_2)^2 + e^{x_1+x_2}$, which is also an affine composition of $y^2 + e^y$ with $y = x_1 + x_2$. Since $y^2 + e^y$ is convex, it is also convex.

The equality constraint function is clearly an affine function.
The requirements for convex optimizations are all satisfied, thus it is a convex optimization problem.

(b).

It is not a convex optimization problem.
Clearly the equality constraint function $6x_1^2 - 7x_2 = 0$ is not an affine function, so it is not a convex optimization problem.