CS2601 Linear and Convex Optimization

Homework 4

Due: 2021.10.28

1. Let Δ_{n-1} be the probability simplex. The **entropy** of a probability distribution $\boldsymbol{x} \in \Delta_{n-1}$ is defined by

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

where we use the convention $0 \log 0 = 0$.

(a). Use the concavity of $\log x$ to show that $H(\boldsymbol{x}) \leq \log \|\boldsymbol{x}\|_0 \leq \log n$, where $\|\boldsymbol{x}\|_0 = \sum_{i=1}^{n} \mathbb{1}\{x_i \neq 0\}$ is the number of nonzero components of $\boldsymbol{x}$. Hint: If $\|\boldsymbol{x}\|_0 = k$, you can assume the first k components are nonzero without loss of generality.

(b). Show that the uniform distribution $\bar{\boldsymbol{x}}$ with $\bar{x}_i = \frac{1}{n}$ for $i = 1, 2, \ldots, n$ is the **unique** maximum of $H(\boldsymbol{x})$ on Δ_{n-1}. Hint: For uniqueness, show $H(\boldsymbol{x})$ is strictly convex on $C = \{\boldsymbol{x} \in \Delta_{n-1} : \boldsymbol{x} > \mathbf{0}\}$ and use part (a) for $\boldsymbol{x} \in \Delta_{n-1} \setminus C$.

2. Let $f : (a, b) \to \mathbb{R}$ be convex, where $-\infty \leq a < b \leq +\infty$. Let X be random variable taking values in (a, b). Suppose the expectations $\mathbb{E}X$ and $\mathbb{E}f(X)$ exist. Prove Jensen's inequality $f(\mathbb{E}X) \leq \mathbb{E}f(X)$ by completing the following steps.

(a). Let $\mu = \mathbb{E}X$. For $a < s < \mu < u < b$, show

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

(b). Show that there exists $\beta \in \mathbb{R}$ such that

$$f(x) \geq f(\mu) + \beta(x - \mu), \quad \forall x \in (a, b) \tag{$\star$}$$

Hint: You can take

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

Obviously $\beta > -\infty$. Use part (a) to show that $\beta < +\infty$ and satisfies ($\star$) (consider $a < x < \mu$ and $\mu < x < b$ separately).

(c). Show that

$$f(X) \geq f(\mu) + \beta(X - \mu).$$

and conclude $\mathbb{E}f(X) \geq f(\mathbb{E}X)$ by taking expectation.

Remark. If f is differentiable, we can take $\beta = f'(\mu)$ by the first-order condition. Part (a) shows that $(\star)$ holds without assuming differentiability. The number β used in the proof generalizes the concept of gradient (derivative) $f'(\mu)$. Any β satisfying $(\star)$ is called a **subgradient** of f at μ. For example, any $\beta \in [-1, 1]$ is a subgradient of $f(x) = |x|$ at 0.

3. Is the following set convex? Show your argument.

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

You can use any results we have proved in class.

4. Determine if the following optimization problems are convex optimization problems.

(a).

$$\begin{aligned} \min_{x_1, x_2} \quad & x_1^2 - 2x_1x_2 + x_2^2 + x_1 + x_2 \\ \text{s.t.} \quad & (x_1 - x_2)^2 + 4x_1x_2 + e^{x_1 + x_2} \leq 0 \\ & x_1 - 3x_2 = 0 \end{aligned}$$

(b).

$$\begin{aligned} \min_{x_1, x_2} \quad & x_1^2 + x_2^4 \\ \text{s.t.} \quad & x_1 e^{-(x_1 + x_2)} \leq 0 \\ & x_1^2 - 2x_1x_2 + x_2^2 + x_1 + x_2 \leq 0 \\ & 6x_1^2 - 7x_2 = 0 \end{aligned}$$