第一部分：基本理论

Apple Zhang

2021 年 9 月 16 日

题目 1. 证明 $f(\mathbf{x}) = (x_1 x_2 \cdots x_n)^{1/n}, \mathbf{x} \in \mathbb{R}^n_{++}$ 是凹函数.

证明. 显然定义域是一个凸锥，且为凸集. 因此其梯度

$$\frac{\partial f}{\partial x_i} = \frac{f(\mathbf{x})}{n x_i} \tag{1}$$

以及海森矩阵

$$\frac{\partial^2 f}{\partial x_i \partial x_j} = \frac{f(\mathbf{x})}{n^2 x_i x_j}, \quad \frac{\partial^2 f}{\partial x_i^2} = -(n-1)\frac{f(\mathbf{x})}{n^2 x_i^2}, \quad i \neq j. \tag{2}$$

或写成

$$\nabla^2 f(\mathbf{x}) = \frac{f(\mathbf{x})}{n^2}\left[\mathbf{z}\mathbf{z}^T - n\mathrm{diag}\left(\frac{1}{x_1^2}, \frac{1}{x_2^2}, \cdots, \frac{1}{x_n^2}\right)\right]. \tag{3}$$

其中，$\mathbf{z} = [1/x_1, 1/x_2, \cdots, 1/x_n]^T$. 则对于任意的非零向量 $\mathbf{w} \in \mathbb{R}^n$，有

$$\begin{aligned}
\mathbf{w}^T \nabla^2 f(\mathbf{x}) \mathbf{w} &= \frac{f(\mathbf{x})}{n^2}\left[\mathbf{w}^T \mathbf{z}\mathbf{z}^T \mathbf{w} - n\mathbf{w}^T \mathrm{diag}\left(\frac{1}{x_1^2}, \frac{1}{x_2^2}, \cdots, \frac{1}{x_n^2}\right)\mathbf{w}\right] \\
&= \frac{f(\mathbf{x})}{n^2}\left[\left(\sum_{i=1}^n w_i z_i\right)^2 - n\sum_{i=1}^n w_i^2 z_i^2\right] \\
&\leq 0.
\end{aligned}$$

其中最后一步用到了柯西不等式，即任意向量 $\mathbf{a}, \mathbf{b}$，都有

$$(\mathbf{a}^T \mathbf{b})^2 \leq (\mathbf{a}^T \mathbf{a})(\mathbf{b}^T \mathbf{b}). \tag{4}$$

此处取向量 $a_i = 1$ 和 $b_i = u_i z_i$ 即可. □

题目 2. 设 $f : \mathbb{R}^n \mapsto \mathbb{R}$，$g : \mathbb{R}^n \times \mathbb{R}_{++} \mapsto \mathbb{R}$，其中

$$g(\mathbf{x}, t) = tf\left(\frac{\mathbf{x}}{t}\right). \tag{5}$$

且 $\mathrm{dom}\, g = \{(\mathbf{x}, t) : t > 0,\ \mathbf{x}/t \in \mathrm{dom}\, f\}$. 则称 g 是关于函数 f 的透视. 证明如果 f 是凸函数，则 g 也是凸的. 同时，若 f 为凹函数，则 g 也是凹函数.

证明. 如果 f 是凸函数，则显然 $\mathrm{dom}\, f$ 是凸集，注意到 $\mathrm{dom}\, g = \{(\mathbf{x}, t) : t > 0, x/t\}$ 是 $\mathrm{dom}\, f$ 在透视函数下的原相，所以 $\mathrm{dom}\, g$ 也是凸集.

对于任意 $t_1, t_2 > 0$, $\mathbf{x}_1, \mathbf{x}_2 \in \operatorname{dom} f$，$\theta \in [0,1]$，并且 $\mathbf{x}_1/t_1, \mathbf{x}_2/t_2 \in \operatorname{dom} f$，有

$$
\begin{aligned}
& g(\theta(\mathbf{x}_1, t_1) + (1-\theta)(\mathbf{x}_2, t_2)) \\
=& g(\theta\mathbf{x}_1 + (1-\theta)\mathbf{x}_2, \theta t_1 + (1-\theta)t_2) \\
=& [\theta t_1 + (1-\theta)t_2] f\left(\frac{\theta\mathbf{x}_1 + (1-\theta)\mathbf{x}_2}{\theta t_1 + (1-\theta)t_2}\right) \\
=& [\theta t_1 + (1-\theta)t_2] f\left(\frac{\theta t_1\mathbf{x}_1/t_1 + (1-\theta)t_2\mathbf{x}_2/t_2}{\theta t_1 + (1-\theta)t_2}\right) \\
\leq& [\theta t_1 + (1-\theta)t_2]\left[\frac{\theta t_1}{\theta t_1 + (1-\theta)t_2} f\left(\frac{\mathbf{x}_1}{t_1}\right) + \frac{(1-\theta)t_2}{\theta t_1 + (1-\theta)t_2} f\left(\frac{\mathbf{x}_2}{t_2}\right)\right] \\
=& \theta t_1 f\left(\frac{\mathbf{x}_1}{t_1}\right) + (1-\theta)t_2 f\left(\frac{\mathbf{x}_2}{t_2}\right) \\
=& \theta g(\mathbf{x}_1, t_1) + (1-\theta) g(\mathbf{x}_2, t_2).
\end{aligned}
$$

其中第 5 行用到了 f 为凸函数. 根据以上及凸函数的定义，可知 g 是凸的. 凹函数的情况完全可以类似证明. □

题目 3 (Boyd 2.1)**.** Let $C \subseteq \mathbb{R}^n$ be a convex set, with $\mathbf{x}_1, \cdots, \mathbf{x}_k \in C$, and let $\theta_1, \cdots, \theta_k \in \mathbb{R}$ satisfy $\theta_i \geq 0$, $\theta_1 + \cdots + \theta_k = 1$. Show that $\theta_1\mathbf{x}_1 + \cdots + \theta_k\mathbf{x}_k \in C$. (The definition of convexity is that this holds for $k = 2$; you must show it for arbitrary k.) *Hint*: Use induction on k.

证明. 当 $k = 2$ 时，根据凸集的性质，结论显然成立. 假设当 $k = l$ 时结论成立，即对于任意 $\mathbf{x}_1, \cdots, \mathbf{x}_l \in C$ 和 $\theta_i \geq 0, \theta_1 + \cdots + \theta_l = 1$，有

$$\theta_1\mathbf{x}_1 + \theta_2\mathbf{x}_2 + \cdots + \theta_l\mathbf{x}_l \in C. \tag{6}$$

则当 $k = l + 1$ 时，任意取 $\mathbf{x}_1, \cdots, \mathbf{x}_{l+1} \in C$ 和 $\theta_i \geq 0, \theta_1 + \cdots + \theta_{l+1} = 1$，由于我们知道

$$\sum_{j=1}^{l} \frac{\theta_j\mathbf{x}_j}{\sum_{i=1}^{l}\theta_i} \in C. \tag{7}$$

因此，我们有

$$\begin{aligned}
&\theta_1 \mathbf{x}_1 + \theta_2 \mathbf{x}_2 + \cdots + \theta_{l+1} \mathbf{x}_{l+1} \\
= &\theta_1 \mathbf{x}_1 + \theta_2 \mathbf{x}_2 + \cdots + \left(1 - \sum_{i=1}^{l} \theta_i\right) \mathbf{x}_{l+1} \\
= &\left(\sum_{i=1}^{l} \theta_i\right) \left(\sum_{j=1}^{l} \frac{\theta_j \mathbf{x}_j}{\sum_{i=1}^{l} \theta_i}\right) + \left(1 - \sum_{i=1}^{l} \theta_i\right) \mathbf{x}_{l+1} \\
\in &C.
\end{aligned}$$

根据数学归纳法，结论得证. $\square$