第二次作业

2.(1) （1）**结论** 函数 $f(x_1,x_2)$ 是凸函数.

证明 函数 $f(x_1,x_2)$ 定义在 $\mathbf{R}^2$ 上，且二阶可微. 于是其黑塞矩阵为

$$\begin{pmatrix} \frac{\partial^2 f}{\partial x_1^2} & \frac{\partial^2 f}{\partial x_1 x_2} \\ \frac{\partial^2 f}{\partial x_2 x_1} & \frac{\partial^2 f}{\partial x_2^2} \end{pmatrix} = \begin{pmatrix} 2 & -2 \\ -2 & 2 \end{pmatrix}$$

为半正定矩阵. 且 dom f 为凸集. 因此，原函数为凸函数.

(2)

方法一:

(2) $f(x_1,x_2) = x_1 e^{-(x_1+x_2)}$

证明: $f(x_1,x_2)$ 的定义域 dom f 为凸集.

$$\frac{\partial f}{\partial x_1} = e^{-(x_1+x_2)} - x_1 e^{-(x_1+x_2)} \qquad \frac{\partial^2 f}{\partial x_1 \partial x_2} = x_1 e^{-(x_1+x_2)} - e^{-(x_1+x_2)}$$

$$\frac{\partial f}{\partial x_2} = -x_1 e^{-(x_1+x_2)} \qquad \frac{\partial^2 f}{\partial x_2 \partial x_1} = x_1 e^{-(x_1+x_2)} - e^{-(x_1+x_2)}$$

$$\frac{\partial^2 f}{\partial x_1^2} = x_1 e^{-(x_1+x_2)} - 2e^{-(x_1+x_2)}$$

$$\frac{\partial^2 f}{\partial x_2^2} = x_1 e^{-(x_1+x_2)}$$

则其 Hessian 矩阵为: $\nabla^2 f(x) = \begin{pmatrix} \frac{\partial^2 f}{\partial x_1^2} & \frac{\partial^2 f}{\partial x_1 \partial x_2} \\ \frac{\partial^2 f}{\partial x_2 \partial x_1} & \frac{\partial^2 f}{\partial x_2^2} \end{pmatrix} = \begin{pmatrix} x_1 e^{-(x_1+x_2)} - 2e^{-(x_1+x_2)} & x_1 e^{-(x_1+x_2)} - e^{-(x_1+x_2)} \\ x_1 e^{-(x_1+x_2)} - e^{-(x_1+x_2)} & x_1 e^{-(x_1+x_2)} \end{pmatrix}$

无法保证各阶主子式均大于等于0 $\Rightarrow$ Hessian 矩阵为不定矩阵.

$\Rightarrow$ 函数 $f(x_1,x_2)$ 不是凸函数.

方法二: （2）**结论** 函数 $f(x_1,x_2)$ 不是凸函数.

证明 用反证法. 如果 $f(x_1,x_2)$ 是凸函数，那么其在任意一条直线上都是凸的. 现在考虑函数 $f(x_1,x_2)$ 在直线 $x_2=0$ 上的情况，此时

$$f(x_1,x_2) = f(x_1,0) = x_1 e^{-x_1} \tag{2}$$

设函数

$$g(x) = \frac{x}{e^x}, \quad x>0, \tag{3}$$

则有

$$g'(x) = -\frac{x-1}{e^x}, \quad g''(x) = \frac{x-2}{e^x}. \tag{4}$$

但由于

$$g''(0) < 0, \tag{5}$$

$g(x)$ 不是凸函数，矛盾!

因此，$f(x_1,x_2)$ 不是凸函数.

3. **证明** 对于 $\mathbf{R}^n$ 中的任意两点 a,b，以及任意 $\theta\in\mathbf{R}:0\le\theta\le1$，都有

方法一：

$$f(a)=\sum_{i=1}^{n}c_ia_i,\quad f(b)=\sum_{i=1}^{n}c_ib_i.$$

于是

$$\theta f(a)=\theta\sum_{i=1}^{n}c_ia_i,\quad (1-\theta)f(b)=(1-\theta)\sum_{i=1}^{n}c_ib_i.$$

又因为

$$f(\theta a+(1-\theta)b)=\sum_{i=1}^{n}c_i(\theta a_i+(1-\theta)b_i)=\theta\sum_{i=1}^{n}c_ia_i+(1-\theta)\sum_{i=1}^{n}c_ib_i=\theta f(a)+\quad(1-\theta)f(b),$$

于是 $f(x)$ 是 $\mathbf{R}^n$ 上的凸函数. 命题得证.

3. 试证线性函数 $f(x)$ 是 R^n 上的凸函数.

方法二：

$f(x)=c^Tx=c_1x_1+c_2x_2+c_3x_3+\cdots+c_nx_n$

证明：$f(x)$ 的定义域 dom$f$ 是凸集.

$$\frac{\partial f}{\partial x_1}=c_1\quad \frac{\partial^2 f}{\partial x_1^2}=0\quad \frac{\partial^2 f}{\partial x_1\partial x_2}=0\ \cdots\ \frac{\partial^2 f}{\partial x_1\partial x_n}=0$$

$$\frac{\partial f}{\partial x_2}=c_2\quad \frac{\partial^2 f}{\partial x_2^2}=0\quad \frac{\partial^2 f}{\partial x_2\partial x_1}=0\ \cdots\ \frac{\partial^2 f}{\partial x_2\partial x_n}=0$$

$$\vdots$$

$$\frac{\partial f}{\partial x_n}=c_n\quad \frac{\partial^2 f}{\partial x_n\partial x_1}=0\quad \frac{\partial^2 f}{\partial x_n\partial x_2}=0\ \cdots\ \frac{\partial^2 f}{\partial x_n\partial x_{n-1}}=0$$

则其 Hessian 矩阵为：
$$\nabla^2 f(x)=\begin{pmatrix}\frac{\partial^2 f}{\partial x_1^2} & \frac{\partial^2 f}{\partial x_1\partial x_2} & \cdots & \frac{\partial^2 f}{\partial x_1\partial x_n}\\ \vdots & & & \vdots\\ \frac{\partial^2 f}{\partial x_n\partial x_1} & \frac{\partial^2 f}{\partial x_n\partial x_2} & \cdots & \frac{\partial^2 f}{\partial x_n^2}\end{pmatrix}=\begin{pmatrix}0 & 0 & 0 & \cdots & 0\\ \vdots & & & & \vdots\\ 0 & 0 & 0 & \cdots & 0\end{pmatrix}$$

$\Rightarrow$ 各阶主子式值均为0，也即非负.

$\Rightarrow f(x)$ Hessian 矩阵为半正定阵 $\Leftrightarrow$ 线性函数 $f(x)$ 是凸函数.

4、

4. $f(x)=(x-1)^2$，试证明 $f(x)$ 在 $(-\infty,\infty)$ 上是严格凸函数.

证明：$f(x)$ 的定义域 dom$f$：$(-\infty,\infty)$ 是凸集.

$$\frac{\partial f}{\partial x}=2(x-1)\qquad \frac{\partial^2 f}{\partial x^2}=2>0$$

则对于任意 $x\in$ dom$f$，有 $\nabla^2 f(x)>0\Rightarrow f(x)$ 在 $(-\infty,\infty)$ 上是严格凸函数.

5. **证明** 设函数

$$f(x)=x^n,\quad x>0,n\geq 2. \tag{10}$$

※ 需要证明 $f(x)$ 的凸性.

于是

$$f'(x)=nx^{n-1},\quad f''(x)=n(n-1)x^{n-2}>0. \tag{11}$$

因此函数 $f(s)$ 为 $(0,+\infty)$ 上的凸函数.

根据凸函数的性质，对于 $(0,+\infty)$ 上的任意 x,y 以及实数 $1/2$，都有

$$\frac{f(x)}{2}+\frac{f(y)}{2}\geq f\left(\frac{x+y}{2}\right), \tag{12}$$

即

$$\frac{x^n+y^n}{2}\geq\left(\frac{x+y}{2}\right)^n. \tag{13}$$

等号成立当且仅当 $x=y$. 而 $x\neq y$，故等号不成立. 原不等式得证.

6. **证明** 设函数

$$f(x)=x\ln x,\quad x>0.$$

于是

$$f'(x)=1+\ln x,\quad f''(x)=\frac{1}{x}>0,\quad x>0.$$

因此函数 $f(s)$ 为 $(0,+\infty)$ 上的凸函数.

根据凸函数的性质，对于 $(0,+\infty)$ 上的任意 x,y 以及实数 $1/2$，都有

$$\frac{f(x)}{2}+\frac{f(y)}{2}\geq f\left(\frac{x+y}{2}\right),$$

即

$$x\ln x+y\ln y\geq(x+y)\ln\left(\frac{x+y}{2}\right).$$

等号成立当且仅当 $x=y$. 而 $x\neq y$，故等号不成立. 原不等式得证.

7.

先证明. R^m上的范数 $\|\cdot\|$ 是凸函数.

设 $x_1, x_2 \in \mathbb{R}^m$, $\theta \in [0,1]$. 则有

$$\theta\cdot\|x_1\| + (1-\theta)\cdot\|x_2\| = \|\theta\cdot x_1\| + \|(1-\theta)\cdot x_2\|$$
$$\geq \|\theta x_1 + (1-\theta)\cdot x_2\|.$$

故 $\|\cdot\|$ 为凸函数.

由于复合仿射映射是保凸运算. 故 $\|A^{(i)}x+b^{(i)}\|$ 也是凸函数. $i=1,2,\cdots,k$.

由于逐点最大是保凸运算. 故 $\max\limits_{1\le i\le k}\|A^{(i)}x+b^{(i)}\|$ 是凸函数. 证毕.

8.（P_{111}. 3.22(c)） x, u, v 都是变量.

方法一：

证明函数 $f(x,u,v) = -\log(uv - x^T x)$ 为凸函数，其定义域为 $\mathrm{dom} f = \{(x,u,v) \mid uv > x^T x,\ u,v > 0\}$

证明：设 $h(y) = -\log y$，$h'(y) = -\frac{1}{y}$，$h''(y) = \frac{1}{y^2}$，定义域 $Y = \begin{vmatrix} u & x^T \\ x & v \end{vmatrix}$

Y：一阶 $u,v>0$，二阶：$uv - x^T x > 0$，故 Y 为半正定矩阵，~~Y为凸集.~~ X

因此 $h(y)$ 为凸函数，且单调非增 $\Rightarrow h''(y) > 0$，$h'(y) < 0$

$g(x,u,v) = uv - x^T x$

$H = \begin{vmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & -2 \end{vmatrix}$，一阶：0, 0, −2；二阶：−1, 0, 0；三阶：−2

$\Rightarrow g(x,u,v)$ 为凹函数，$g''(x) \le 0$

由 $f = h \circ g$，$f''(x) = \underbrace{h''(g(x))}_{>0} g'(x)^2 + \underbrace{h'(g(x))}_{<0} \underbrace{g''(x)}_{<0} > 0$

故 $f(x,u,v)$ 为凸函数，得证

$u \quad x^T$

$\begin{bmatrix} & \end{bmatrix}$

$[\ ]$

方法二：也可用 $\frac{x^T x}{t}$ 的凸性. P84. 例3.18

$f(x,u,v) = -\log(uv - x^T x) = -\log u - \log(v - x^T x/u)$

因为 $y(x,w) = \frac{x^T x}{t}$ 为凸函数（$y(x,w)$ 是 $f(x) = x^T x$ 的透视函数），$-\frac{x^T x}{t}$ 凹，v 凹

则 $v - \frac{x^T x}{u}$ 为凹函数，$-\log(v - \frac{x^T x}{u})$ 为凸函数

$-\log(y)$ 凸、非增

易知 $f(x,u,v)$ 为凸函数

加凸集的证明.

9. (1)

函数： $f(x) = c \cdot x$，其中 $x \in [a, b]$ 且 c 为常数。

共轭函数求解过程：

- $f^*(y) = \sup_{x \in [a,b]} \{yx - cx\}$。
- 结果取决于 y 和 c 的相对大小，以及区间 $[a, b]$ 的范围。

共轭函数结果： $f^*(y) = \begin{cases} ay - ca & 当\ y \leq c \\ by - cb & 当\ y > c \end{cases}$。

(2)

函数： $f(x) = \frac{1}{2}(x - b)^2$

共轭函数求解过程：

- 定义 $f^*(y) = \sup_x \{yx - \frac{1}{2}(x - b)^2\}$。
- 通过配方法或求导法求解 $f^*(y)$ 的表达式。
- 结果表明，共轭函数取决于参数 b 的位置。

共轭函数结果： $f^*(y) = yb + \frac{1}{2}y^2$。