

## 凸优化 第二次作业

1. 用定义验证下列各集合是凸集:

$$(1) S = \{(x_1, x_2) \mid x_1 + 2x_2 \geq 1, x_1 - x_2 \geq 1\};$$

$$(2) S = \{(x_1, x_2) \mid x_2 \geq |x_1|\};$$

$$(3) S = \{(x_1, x_2) \mid x_1^2 + x_2^2 \leq 10\}.$$

**证明** 证 (1) 对集合  $S$  中任意两点  $\mathbf{x}^{(1)} = \begin{bmatrix} x_1^{(1)} \\ x_2^{(1)} \end{bmatrix}$ ,  $\mathbf{x}^{(2)} = \begin{bmatrix} x_1^{(2)} \\ x_2^{(2)} \end{bmatrix}$  及每个数  $\lambda \in [0, 1]$ , 有

$$\lambda \mathbf{x}^{(1)} + (1 - \lambda) \mathbf{x}^{(2)} = \begin{bmatrix} \lambda x_1^{(1)} + (1 - \lambda) x_1^{(2)} \\ \lambda x_2^{(1)} + (1 - \lambda) x_2^{(2)} \end{bmatrix}$$

由题设, 有

$$\begin{aligned} & \left[ \lambda x_1^{(1)} + (1 - \lambda) x_1^{(2)} \right] + 2 \left[ \lambda x_2^{(1)} + (1 - \lambda) x_2^{(2)} \right] \\ &= \lambda (x_1^{(1)} + 2x_2^{(1)}) + (1 - \lambda) (x_1^{(2)} + 2x_2^{(2)}) \geq \lambda + (1 - \lambda) = 1, \\ & \left[ \lambda x_1^{(1)} + (1 - \lambda) x_1^{(2)} \right] - \left[ \lambda x_2^{(1)} + (1 - \lambda) x_2^{(2)} \right] \\ &= \lambda (x_1^{(1)} - x_2^{(1)}) + (1 - \lambda) (x_1^{(2)} - x_2^{(2)}) \geq \lambda + (1 - \lambda) = 1, \end{aligned}$$

因此,  $\lambda \mathbf{x}^{(1)} + (1 - \lambda) \mathbf{x}^{(2)} \in S$ , 故  $S$  是凸集. (2) 对集合  $S$  中任意两点  $\mathbf{x}^{(1)} = \begin{bmatrix} x_1^{(1)} \\ x_2^{(1)} \end{bmatrix}$  和

$\mathbf{x}^{(2)} = \begin{bmatrix} x_1^{(2)} \\ x_2^{(2)} \end{bmatrix}$  及每个数  $\lambda \in [0, 1]$ , 有

$$\lambda \mathbf{x}^{(1)} + (1 - \lambda) \mathbf{x}^{(2)} = \begin{bmatrix} \lambda x_1^{(1)} + (1 - \lambda) x_1^{(2)} \\ \lambda x_2^{(1)} + (1 - \lambda) x_2^{(2)} \end{bmatrix}$$

由题设, 有

$$\lambda x_2^{(1)} + (1 - \lambda) x_2^{(2)} \geq \lambda |x_1^{(1)}| + (1 - \lambda) |x_1^{(2)}| \geq |\lambda x_1^{(1)} + (1 - \lambda) x_1^{(2)}|,$$

因此  $\lambda \mathbf{x}^{(1)} + (1 - \lambda) \mathbf{x}^{(2)} \in S$ , 故  $S$  是凸集. (3) 对集合  $S$  中任意两点  $\mathbf{x}^{(1)} = \begin{bmatrix} x_1^{(1)} \\ x_2^{(1)} \end{bmatrix}$  和

$\mathbf{x}^{(2)} = \begin{bmatrix} x_1^{(2)} \\ x_2^{(2)} \end{bmatrix}$  及每个数  $\lambda \in [0, 1]$ , 有

$$\lambda \mathbf{x}^{(1)} + (1 - \lambda) \mathbf{x}^{(2)} = \begin{bmatrix} \lambda x_1^{(1)} + (1 - \lambda) x_1^{(2)} \\ \lambda x_2^{(1)} + (1 - \lambda) x_2^{(2)} \end{bmatrix}$$

由题设, 有

$$\begin{aligned}
& \left[ \lambda x_1^{(1)} + (1-\lambda)x_1^{(2)} \right]^2 + \left[ \lambda x_2^{(1)} + (1-\lambda)x_2^{(2)} \right]^2 \\
&= \lambda^2 x_1^{(1)2} + 2\lambda(1-\lambda)x_1^{(1)}x_1^{(2)} + (1-\lambda)^2 x_1^{(2)2} + \lambda^2 x_2^{(1)2} + 2\lambda(1-\lambda)x_2^{(1)}x_2^{(2)} \\
&\quad + (1-\lambda)^2 x_2^{(2)2} = \lambda^2 \left[ x_1^{(1)2} + x_2^{(1)2} \right] + (1-\lambda)^2 \left[ x_1^{(2)2} + x_2^{(2)2} \right] + \lambda(1-\lambda) \left[ 2x_1^{(1)}x_1^{(2)} \right. \\
&\quad \left. + 2x_2^{(1)}x_2^{(2)} \right] \leq 10\lambda^2 + 10(1-\lambda)^2 + \lambda(1-\lambda) \left[ x_1^{(1)2} + x_1^{(2)2} + x_2^{(1)2} + x_2^{(2)2} \right] \\
&\leq 10\lambda^2 + 10(1-\lambda)^2 + 20\lambda(1-\lambda) = 10
\end{aligned}$$

因此  $\lambda x^{(1)} + (1-\lambda)x^{(2)} \in S$ , 故  $S$  是凸集. 2. 设  $C \subset \mathbb{R}^p$  是一个凸集,  $p$  是正整数. 证明下列集合  $S$  是  $\mathbb{R}^n$  中的凸集:

$$S = \{x \mid x \in \mathbb{R}^n, x = A\rho, \rho \in C\}$$

其中  $A$  是给定的  $n \times p$  实矩阵. 证对任意两点  $x^{(1)}, x^{(2)} \in S$  及每个数  $\lambda \in [0, 1]$ , 根据集合  $S$  的定义, 存在  $\rho_1, \rho_2 \in C$ , 使  $x^{(1)} = A\rho_1, x^{(2)} = A\rho_2$ , 因此必有  $\lambda x^{(1)} + (1-\lambda)x^{(2)} = \lambda A\rho_1 + (1-\lambda)A\rho_2 = A[\lambda\rho_1 + (1-\lambda)\rho_2]$ . 由于  $C$  是凸集, 必有  $\lambda\rho_1 + (1-\lambda)\rho_2 \in C$ , 因此  $\lambda x^{(1)} + (1-\lambda)x^{(2)} \in S$ , 故  $S$  是凸集.

2. 设  $C \subset \mathbb{R}^p$  是一个凸集,  $p$  是正整数. 证明下列集合  $S$  是  $\mathbb{R}^n$  中的凸集:

$$S = \{x \mid x \in \mathbb{R}^n, x = A\rho, \rho \in C\}$$

其中  $A$  是给定的  $n \times p$  实矩阵.

**证明** 证对任意两点  $x^{(1)}, x^{(2)} \in S$  及每个数  $\lambda \in [0, 1]$ , 根据集合  $S$  的定义, 存在  $\rho_1, \rho_2 \in C$ , 使  $x^{(1)} = A\rho_1, x^{(2)} = A\rho_2$ , 因此必有  $\lambda x^{(1)} + (1-\lambda)x^{(2)} = \lambda A\rho_1 + (1-\lambda)A\rho_2 = A[\lambda\rho_1 + (1-\lambda)\rho_2]$ . 由于  $C$  是凸集, 必有  $\lambda\rho_1 + (1-\lambda)\rho_2 \in C$ , 因此  $\lambda x^{(1)} + (1-\lambda)x^{(2)} \in S$ , 故  $S$  是凸集.

3. 判别下列函数是否为凸函数:

- (1)  $f(x_1, x_2) = x_1^2 - 2x_1x_2 + x_2^2 + x_1 + x_2$
- (2)  $f(x_1, x_2) = x_1^2 - 4x_1x_2 + x_2^2 + x_1 + x_2$ ;
- (3)  $f(x_1, x_2) = (x_1 - x_2)^2 + 4x_1x_2 + e^{x_1+x_2}$ ;
- (4)  $f(x_1, x_2) = x_1e^{-(x_1+x_2)}$ ;
- (5)  $f(x_1, x_2, x_3) = x_1x_2 + 2x_1^2 + x_2^2 + 2x_3^2 - 6x_1x_3$ .

**解** 解 (1)  $\nabla^2 f(x) = \begin{bmatrix} 2 & -2 \\ -2 & 2 \end{bmatrix}$  为半正定矩阵, 故  $f(x_1, x_2)$  是凸函数. (2)  $\nabla^2 f(x) = \begin{bmatrix} 2 & -4 \\ -4 & 2 \end{bmatrix}$  为不定矩阵, 故  $f(x_1, x_2)$  不是凸函数. (3)  $\frac{\partial f}{\partial x_1} = 2(x_1 - x_2) + 4x_2 + e^{x_1+x_2}$ ,  $\frac{\partial f}{\partial x_2} = -2(x_1 - x_2) + 4x_1 + e^{x_1+x_2}$ ,

$$\frac{\partial^2 f}{\partial x_1^2} = 2 + e^{x_1+x_2}, \quad \frac{\partial^2 f}{\partial x_1 \partial x_2} = \frac{\partial^2 f}{\partial x_2 \partial x_1} = 2 + e^{x_1+x_2}, \quad \frac{\partial^2 f}{\partial x_2^2} = 2 + e^{x_1+x_2}$$

因此 Hesse 矩阵

$$\nabla^2 f(\mathbf{x}) = \begin{bmatrix} 2 + e^{x_1+x_2} & 2 + e^{x_1+x_2} \\ 2 + e^{x_1+x_2} & 2 + e^{x_1+x_2} \end{bmatrix} = (2 + e^{x_1+x_2}) \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix}$$

为半正定矩阵, 因此  $f(x)$  是凸函数. (4)  $\frac{\partial f}{\partial x_1} = e^{-(x_1+x_2)} - x_1 e^{-(x_1+x_2)} = (1-x_1)e^{-(x_1+x_2)}$ ,  $\frac{\partial f}{\partial x_2} = -x_1 e^{-(x_1+x_2)}$ ,

$$\frac{\partial^2 f}{\partial x_1^2} = (x_1 - 2)e^{-(x_1+x_2)}, \quad \frac{\partial^2 f}{\partial x_1 \partial x_2} = \frac{\partial^2 f}{\partial x_2 \partial x_1} = (x_1 - 1)e^{-(x_1+x_2)}, \quad \frac{\partial^2 f}{\partial x_2^2} = x_1 e^{-(x_1+x_2)}$$

于是 Hesse 矩阵

$$\nabla^2 f(x) = e^{-(x_1+x_2)} \begin{bmatrix} x_1 - 2 & x_1 - 1 \\ x_1 - 1 & x_1 \end{bmatrix}$$

为不定矩阵, 故  $f(\mathbf{x})$  不是凸函数. (5)  $f(x)$  的 Hesse 矩阵为

$$\nabla^2 f(x) = \begin{bmatrix} 4 & 1 & -6 \\ 1 & 2 & 0 \\ -6 & 0 & 4 \end{bmatrix}$$

做合同变换:

$$\begin{bmatrix} 4 & 1 & -6 \\ 1 & 2 & 0 \\ -6 & 0 & 4 \end{bmatrix} \longrightarrow \begin{bmatrix} 4 & 0 & 0 \\ 0 & \frac{7}{4} & \frac{3}{2} \\ 0 & \frac{3}{2} & -5 \end{bmatrix} \longrightarrow \begin{bmatrix} 4 & 0 & 0 \\ 0 & 7 & 0 \\ 0 & 0 & -\frac{44}{7} \end{bmatrix}$$

由此可得  $\nabla^2 f(\mathbf{x})$  为不定矩阵, 因此  $f(\mathbf{x})$  不是凸函数.

4. 设  $f(x_1, x_2) = 10 - 2(x_2 - x_1^2)^2$ ,  $S = \{(x_1, x_2) \mid -11 \leq x_1 \leq 1, -1 \leq x_2 \leq 1\}$ ,  $f(x_1, x_2)$  是否为  $S$  上的凸函数?

**证明** 解  $\frac{\partial f}{\partial x_1} = 8x_1(x_2 - x_1^2)$ ,  $\frac{\partial f}{\partial x_2} = -4(x_2 - x_1^2)$ ,

$$\frac{\partial^2 f}{\partial x_1^2} = 8(x_2 - 3x_1^2), \quad \frac{\partial^2 f}{\partial x_1 \partial x_2} = \frac{\partial^2 f}{\partial x_2 \partial x_1} = 8x_1, \quad \frac{\partial^2 f}{\partial x_2^2} = -4$$

函数  $f(x_1, x_2)$  的 Hesse 矩阵为

$$\nabla^2 f(x) = \begin{bmatrix} 8(x_2 - 3x_1^2) & 8x_1 \\ 8x_1 & -4 \end{bmatrix}$$

易知  $\nabla^2 f(\mathbf{x})$  在集合  $S$  上不是半正定矩阵, 如在点  $(0, 1)$  处的 Hesse 矩阵是  $\begin{bmatrix} 8 & 0 \\ 0 & -4 \end{bmatrix}$ , 是不定矩阵. 因此  $f(x_1, x_2)$  不是  $S$  上的凸函数.

5. 设  $f$  是定义在  $\mathbb{R}^n$  上的凸函数,  $\mathbf{x}^{(1)}, \mathbf{x}^{(2)}, \dots, \mathbf{x}^{(k)}$  是  $\mathbb{B}^n$  中的点,  $\lambda_1, \lambda_2, \dots, \lambda_k$  是非负数, 且满足  $\lambda_1 + \lambda_2 + \dots + \lambda_k = 1$ , 证明:

$$f(\lambda_1 x^{(1)} + \lambda_2 x^{(2)} + \dots + \lambda_k x^{(k)}) \leq \lambda_1 f(x^{(1)}) + \lambda_2 f(x^{(2)}) + \dots + \lambda_k f(x^{(k)}).$$

**证明** 证用数学归纳法. 当  $k = 2$  时, 根据凸函数的定义, 必有

$$f(\lambda_1 x^{(1)} + \lambda_2 x^{(2)}) \leq \lambda_1 f(x^{(1)}) + \lambda_2 f(x^{(2)})$$

设  $k = m$  时不等式成立. 当  $k = m + 1$  时, 有

$$\begin{aligned} & f(\lambda_1 x^{(1)} + \lambda_2 x^{(2)} + \cdots + \lambda_m x^{(m)} + \lambda_{m+1} x^{(m+1)}) \\ &= f\left(\sum_{i=1}^m \lambda_i \left(\frac{\lambda_1}{\sum_{i=1}^m \lambda_i} x^{(1)} + \frac{\lambda_2}{\sum_{i=1}^m \lambda_i} x^{(2)} + \cdots + \frac{\lambda_m}{\sum_{i=1}^m \lambda_i} x^{(m)}\right) + \lambda_{m+1} x^{(m+1)}\right) \end{aligned}$$

记

$$\hat{x} = \frac{\lambda_1}{\sum_{i=1}^m \lambda_i} x^{(1)} + \frac{\lambda_2}{\sum_{i=1}^m \lambda_i} x^{(2)} + \cdots + \frac{\lambda_m}{\sum_{i=1}^m \lambda_i} x^{(m)}$$

由于  $f(x)$  是凸函数,  $\sum_{i=1}^m \lambda_i + \lambda_{m+1} = 1, \lambda_i \geq 0$ , 根据凸函数定义, 有

$$f\left(\left(\sum_{i=1}^m \lambda_i\right) \hat{x} + \lambda_{m+1} x^{(m+1)}\right) \leq \left(\sum_{i=1}^m \lambda_i\right) f(\hat{x}) + \lambda_{m+1} f(x^{(m+1)})$$

根据归纳法假设, 有

$$f(\hat{x}) \leq \frac{\lambda_1}{\sum_{i=1}^m \lambda_i} f(x^{(1)}) + \frac{\lambda_2}{\sum_{i=1}^m \lambda_i} f(x^{(2)}) + \cdots + \frac{\lambda_m}{\sum_{i=1}^m \lambda_i} f(x^{(m)})$$

代入上式, 则有

$$f(\lambda_1 x^{(1)} + \lambda_2 x^{(2)} + \cdots + \lambda_{m+1} x^{(m+1)}) \leq \lambda_1 f(x^{(1)}) + \lambda_2 f(x^{(2)}) + \cdots + \lambda_{m+1} f(x^{(m+1)}),$$

即  $k = m + 1$  时, 不等式也成立. 从而得证.

《Convex Optimization》 3.1, 3.3, 3.4, 3.13, 3.17, 3.23, 3.32, 3.47, 3.57