DCS440 最优化理论 第一次作业: 凸集与凸函数

胡瑞康

22336087

问题 1

设 $C \subseteq \mathbb{R}^n$ 为一个凸集。证明: 对任意 k 个向量 $x_1, \cdots, x_k \in C$，以及 $\theta_1, \cdots, \theta_k \in \mathbb{R}$ 满足 $\theta_1 + \cdots + \theta_k = 1, \theta_i \geq 0$，都有 $\theta_1 x_1 + \cdots + \theta_k x_k \in C$。

解答:

目标：证明对于任意 $k \geq 2$，满足 $\theta_1 + \cdots + \theta_k = 1$ 且 $\theta_i \geq 0$ 的系数，线性组合 $\theta_1 x_1 + \cdots + \theta_k x_k$ 仍属于凸集 C。

证明方法：数学归纳法。

1. **基例** ($k = 2$)：

   凸集的定义即保证对于任何两个点 $x_1, x_2 \in C$ 和 $0 \leq \theta \leq 1$，点 $\theta x_1 + (1-\theta)x_2 \in C$。这与题目所述的 $k = 2$ 情况一致，基例成立。

2. **归纳假设**：

   假设对于某个 $k = m \geq 2$，任意满足 $\theta_1 + \cdots + \theta_m = 1$ 且 $\theta_i \geq 0$ 的系数，线性组合 $\sum_{i=1}^m \theta_i x_i \in C$。

3. **归纳步骤** ($k = m + 1$)：

   考虑 $k = m + 1$ 的情况，令 $\theta_1 + \cdots + \theta_{m+1} = 1$ 且 $\theta_i \geq 0$。

   定义 $\theta' = \sum_{i=1}^m \theta_i = 1 - \theta_{m+1}$。

   由于 $\theta_{m+1} \geq 0$，则 $\theta' \leq 1$ 且 $\theta' \geq 0$。

   线性组合可以表示为：

   $$\theta' \left(\frac{\theta_1}{\theta'} x_1 + \cdots + \frac{\theta_m}{\theta'} x_m \right) + \theta_{m+1} x_{m+1}$$

   其中，$\frac{\theta_i}{\theta'} \geq 0$ 且 $\sum_{i=1}^m \frac{\theta_i}{\theta'} = 1$，根据归纳假设，括号内的部分属于 C。
   此外，整体表达式是两个属于 C 的点的凸组合，因此根据凸集的定义，整个线性组合 $\sum_{i=1}^{m+1} \theta_i x_i \in C$。

4. **结论**：

   通过数学归纳法，命题对于所有 $k \geq 2$ 成立。

问题 2

设 $C \subseteq \mathbb{R}^n$ 为线性方程组的解集,即

$$C = \{x \in \mathbb{R}^n \mid Ax = b\}$$

其中 $A \in \mathbb{R}^{m\times n}, b \in \mathbb{R}^m$。证明: C 是凸集。

解答:

目标：证明线性方程组的解集 C 是凸集。

证明：

1. **定义**：

   $C = \{x \in \mathbb{R}^n \mid Ax = b\}$。
2. **凸性的定义**：

   若对于任意 $x, y \in C$ 和 $0 \leq \theta \leq 1$，有 $\theta x + (1-\theta)y \in C$，则 C 是凸集。
3. **验证凸性**：
   - 取任意 $x, y \in C$，即 $Ax = b$ 和 $Ay = b$。
   - 考虑任意 $0 \leq \theta \leq 1$，则:

   $$A(\theta x + (1-\theta)y) = \theta Ax + (1-\theta)Ay = \theta b + (1-\theta)b = b$$

   - 因此，$\theta x + (1-\theta)y \in C$。
4. **结论**：

   由于任意两个解的凸组合仍是解，C 是凸集。

问题 3

设 $C \subseteq \mathbb{R}^n$ 为二次不等式的解集,即

$$C = \left\{x \in \mathbb{R}^n \mid x^\top Ax + b^\top x + c \leq 0\right\},$$

其中 $A \in \mathbf{S}^n, b \in \mathbb{R}^n, c \in \mathbb{R}$。证明: 若 $A \succcurlyeq 0$ (即 A 是半正定矩阵),则 C 是凸集。

解答:

目标：证明当矩阵 A 半正定 $(A \succcurlyeq 0)$ 时，集合 C 是凸集。

证明：

1. **函数的凸性**：

   定义函数 $f(x) = x^\top Ax + b^\top x + c$。

- 该函数是二次函数。
- $f(x)$ 的海森矩阵为 $H = 2A$。
- 若 $A \succcurlyeq 0$，则 H 也是半正定的，因此 $f(x)$ 是凸函数。

2. **凸集的定义**：

   集合 C 可以表示为 $C = \{x \in \mathbb{R}^n \mid f(x) \leq 0\}$，即 C 是 f 的下水平集。

3. **凸函数的下水平集是凸集**：

   若 f 是凸函数，则对于任意 $x, y \in C$ 和 $0 \leq \theta \leq 1$，

   根据C的定义有 $f(x) \leq 0$，$f(y) \leq 0$

   结合凸函数定义得到

$$f(\theta x + (1-\theta)y) \leq \theta f(x) + (1-\theta)f(y) \leq \theta \cdot 0 + (1-\theta) \cdot 0 = 0$$

   因此，$\theta x + (1-\theta)y \in C$。

4. **结论**：

   因此，当 A 半正定时，C 是凸集。

问题 4

确定以下函数的凹凸性:

(a) $f(x) = \sum_{i=1}^{n} x_i \ln x_i, \quad x_i \in \mathbb{R}_{++},\ i = 1, \cdots, n$;

(b) $f(x_1, x_2) = x_1 x_2, \quad (x_1, x_2) \in \mathbb{R}_{++}^2$;

(c) $f(x_1, x_2) = x_1 / x_2, \quad (x_1, x_2) \in \mathbb{R}_{++}^2$;

解答:

(a) $f(x) = \sum_{i=1}^{n} x_i \ln x_i$, 其中 $x_i > 0$。

判断凹凸性：

1. **单变量分析**：

   对于每个 $x_i \ln x_i$，计算二阶导数：

$$\frac{d}{dx_i}(x_i \ln x_i) = \ln x_i + 1$$

$$\frac{d^2}{dx_i^2}(x_i \ln x_i) = \frac{1}{x_i} > 0 \quad （因为\ x_i > 0）$$

   因此，$x_i \ln x_i$ 在 $x_i > 0$ 上是凸函数。

2. **多变量情况**：

   函数 $f(x) = \sum_{i=1}^{n} x_i \ln x_i$ 是各变量的凸函数之和，因此 $f(x)$ 是凸函数。

结论：$f(x)$ 是**凸**函数。

(b) $f(x_1, x_2) = x_1 x_2$, 其中 $(x_1, x_2) \in \mathbb{R}^2_{++}$。

判断凹凸性：

1. **计算海森矩阵**：

$$f_{x_1} = x_2, \quad f_{x_2} = x_1$$

$$f_{x_1x_1} = 0, \quad f_{x_1x_2} = 1, \quad f_{x_2x_2} = 0$$

   海森矩阵为：

$$H = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}$$

2. **特征值分析**：
   计算特征值：

$$\det(H - \lambda I) = \det \begin{bmatrix} -\lambda & 1 \\ 1 & -\lambda \end{bmatrix} = \lambda^2 - 1 = 0 \implies \lambda = \pm 1$$

   因此，海森矩阵有正特征值和负特征值，表明 f 既不凸也不凹。

结论：$f(x_1, x_2)$ 既**不凸**也**不凹**。

(c) $f(x_1, x_2) = \frac{x_1}{x_2}$, 其中 $(x_1, x_2) \in \mathbb{R}^2_{++}$。

判断凹凸性：

1. **计算海森矩阵**：

$$f_{x_1} = \frac{1}{x_2}, \quad f_{x_2} = -\frac{x_1}{x_2^2}$$

$$f_{x_1x_1} = 0, \quad f_{x_1x_2} = -\frac{1}{x_2^2}, \quad f_{x_2x_2} = \frac{2x_1}{x_2^3}$$

   海森矩阵为：

$$H = \begin{bmatrix} 0 & -\frac{1}{x_2^2} \\ -\frac{1}{x_2^2} & \frac{2x_1}{x_2^3} \end{bmatrix}$$

2. **判定正定性**：
   计算主子式：

$$\text{Det}(H) = 0 \cdot \frac{2x_1}{x_2^3} - \left(-\frac{1}{x_2^2}\right)^2 = -\frac{1}{x_2^4} < 0$$

   因为行列式为负，海森矩阵不正定也不负定。

3. **结论**：
   由于海森矩阵既不正定也不负定，$f(x_1, x_2)$ 既**不凸**也**不凹**。

问题 5

设 $h : \mathbb{R}^m \to \mathbb{R}, g : \mathbb{R}^n \to \mathbb{R}^m$**，则复合函数** $f := h \circ g : \mathbb{R}^n \to \mathbb{R}$ **定义为**

$$f(x) := h(g(x)) = h(g_1(x), \cdots, g_m(x)),$$

$$\operatorname{dom} f := \{x \in \operatorname{dom} g \mid g(x) \in \operatorname{dom} h\}.$$

证明: 若 g_i 是凹函数, h 是凸函数,且 h 关于其每个分量是非增的,则复合函数 $f := h \circ g$ 是凸函数。

解答:

目标: 在满足 g_i 为凹函数，h 为凸且关于每个分量非增的条件下，证明复合函数 $f = h \circ g$ 是凸函数。

证明:

1. **凸函数的定义**:

   函数 f 是凸的，当且仅当对于任意 $x, y \in \operatorname{dom} f$ 和 $0 \leq \theta \leq 1$，有:

   $$f(\theta x + (1-\theta)y) \leq \theta f(x) + (1-\theta)f(y)$$

2. **利用 h 的凸性和非增性**:

   - h 是凸函数，满足:

     $$h(\theta z + (1-\theta)w) \leq \theta h(z) + (1-\theta)h(w)$$

     对于任意 $z, w \in \mathbb{R}^m$。
   - h 关于每个分量非增，意味着如果 $z \leq w$ （逐分量），则 $h(z) \geq h(w)$。

3. **利用 g_i 的凹性**:

   - 每个 g_i 是凹函数，满足:

     $$g_i(\theta x + (1-\theta)y) \geq \theta g_i(x) + (1-\theta)g_i(y)$$

     对于任意 $x, y \in \mathbb{R}^n$。

4. **综合应用**:

   - 设 $z = g(x)$ 和 $w = g(y)$，则:

     $$g(\theta x + (1-\theta)y) = (\theta g_1(x) + (1-\theta)g_1(y), \ldots, \theta g_m(x) + (1-\theta)g_m(y))$$

     由于 g_i 是凹函数，故:

     $$g(\theta x + (1-\theta)y) \geq \theta g(x) + (1-\theta)g(y)$$

     （逐分量不等式成立）。
   - 因为 h 关于每个分量非增，且 $g(\theta x + (1-\theta)y) \geq \theta g(x) + (1-\theta)g(y)$，则:

     $$h(g(\theta x + (1-\theta)y)) \leq h(\theta g(x) + (1-\theta)g(y))$$

   - 利用h的凸性:

     $$h(\theta g(x) + (1-\theta)g(y)) \leq \theta h(g(x)) + (1-\theta)h(g(y)) = \theta f(x) + (1-\theta)f(y)$$

   - 综合上述不等式:

     $$f(\theta x + (1-\theta)y) = h(g(\theta x + (1-\theta)y)) \leq \theta f(x) + (1-\theta)f(y)$$

5. **结论**:

   因此，复合函数 $f = h \circ g$ 是**凸**函数。