

凸优化

question1

证明如果 \mathcal{S}_1 和 \mathcal{S}_2 是 $\mathbb{R}^{m \times n}$ 中的凸集, 那么它们的部分和 $\mathcal{S} = \{(\mathbf{x}, \mathbf{y}_1 + \mathbf{y}_2) | \mathbf{x} \in \mathbb{R}^m, \mathbf{y}_1, \mathbf{y}_2 \in \mathbb{R}^n, (\mathbf{x}, \mathbf{y}_1) \in \mathcal{S}_1, (\mathbf{x}, \mathbf{y}_2) \in \mathcal{S}_2\}$ 也是凸的。

solve1

欲证明 \mathcal{S} 为凸, 只需证明对于任意 $(\mathbf{x}_1, \mathbf{z}_1), (\mathbf{x}_2, \mathbf{z}_2) \in \mathcal{S}$, $\theta \in [0, 1]$ 有 $\theta(\mathbf{x}_1, \mathbf{z}_1) + (1 - \theta)(\mathbf{x}_2, \mathbf{z}_2) \in \mathcal{S}$ 。(其中 $\mathbf{z}_1 = \mathbf{y}_1 + \mathbf{y}_2$, $\mathbf{z}_2 = \mathbf{y}_3 + \mathbf{y}_4$, $(\mathbf{x}_1, \mathbf{y}_1), (\mathbf{x}_2, \mathbf{y}_3) \in \mathcal{S}_1$, $(\mathbf{x}_1, \mathbf{y}_2), (\mathbf{x}_2, \mathbf{y}_4) \in \mathcal{S}_2$)。

$$\begin{aligned}\theta(\mathbf{x}_1, \mathbf{z}_1) + (1 - \theta)(\mathbf{x}_2, \mathbf{z}_2) &= (\theta\mathbf{x}_1 + (1 - \theta)\mathbf{x}_2, \theta\mathbf{z}_1 + (1 - \theta)\mathbf{z}_2) \\ &= (\theta\mathbf{x}_1 + (1 - \theta)\mathbf{x}_2, \theta\mathbf{y}_1 + (1 - \theta)\mathbf{y}_3 + \theta\mathbf{y}_2 + (1 - \theta)\mathbf{y}_4) \\ &= (\theta\mathbf{x}_1 + (1 - \theta)\mathbf{x}_2, \theta\mathbf{y}_1 + (1 - \theta)\mathbf{y}_3) \\ &\quad \oplus (\theta\mathbf{x}_1 + (1 - \theta)\mathbf{x}_2, \theta\mathbf{y}_2 + (1 - \theta)\mathbf{y}_4) \\ &= \mathbf{s}_1 \oplus \mathbf{s}_2\end{aligned}$$

其中 \oplus 代表部分和。

根据凸集的性质自然有 $(\theta\mathbf{x}_1 + (1 - \theta)\mathbf{x}_2, \theta\mathbf{y}_1 + (1 - \theta)\mathbf{y}_3) = \theta(\mathbf{x}_1, \mathbf{y}_1) + (1 - \theta)(\mathbf{x}_2, \mathbf{y}_3) = \mathbf{s}_1 \in \mathcal{S}_1$, $(\theta\mathbf{x}_1 + (1 - \theta)\mathbf{x}_2, \theta\mathbf{y}_2 + (1 - \theta)\mathbf{y}_4) = \theta(\mathbf{x}_1, \mathbf{y}_2) + (1 - \theta)(\mathbf{x}_2, \mathbf{y}_4) = \mathbf{s}_2 \in \mathcal{S}_2$ 。根据 \mathcal{S} 的定义, 自然有 $\mathbf{s}_1 \oplus \mathbf{s}_2 \in \mathcal{S}$, 这说明 \mathcal{S} 是一个凸集。

question2

对于任意 $\mathbf{x} \in \mathbb{R}^n$, 用 $\mathbf{x}_{[i]}$ 表示 \mathbf{x} 中第 i 大的分量, 即将 \mathbf{x} 的分量按照非升序进行排列得到下式 $\mathbf{x}_{[1]} \geq \mathbf{x}_{[2]} \geq \dots \geq \mathbf{x}_{[n]}$ 。证明对 \mathbf{x} 的最大 r 个分量进行求和所得到的函数 $f(\mathbf{x}) = \sum_{i=1}^r \mathbf{x}_{[i]}$ 是凸函数。

solve2

引理: 对于 $\mathbf{x} \in \mathbb{R}^n$, $\mathbf{x} = [x_1, \dots, x_n]^T$, $\max_i \{\mathbf{x}\} = \max_i \{x_1, \dots, x_n\}$ 是凸函数。证明如下:

任取 $\mathbf{x}, \mathbf{y} \in \mathbb{R}^n$, $\theta \in [0, 1]$, 有

$$\begin{aligned} & \theta \max_i \{\mathbf{x}\} + (1 - \theta) \max_i \{\mathbf{y}\} \\ &= \theta \max_i \{x_1, \dots, x_n\} + (1 - \theta) \max_i \{y_1, \dots, y_n\} \\ &= \theta x_{[1]} + (1 - \theta) y_{[1]} \\ &\geq \max_i \{\theta x_1 + (1 - \theta) y_1, \dots, \theta x_n + (1 - \theta) y_n\} = \max_i \{\theta \mathbf{x} + (1 - \theta) \mathbf{y}\} \end{aligned}$$

其中 $x_{[1]}$ 、 $y_{[1]}$ 分别代表 $\{x_1, \dots, x_n\}$ 、 $\{y_1, \dots, y_n\}$ 中最大的元素。

接着证明比题目更强的命题: 对于任意 $\mathbf{x} \in \mathbb{R}^n$, 对 \mathbf{x} 的最大 r 个分量进行求和所得到的函数 $f(\mathbf{x}) = \sum_{i=1}^r \mathbf{x}_{[i]}$ 仍然是凸函数。证明如下:

构造矩阵 $\mathbf{A} \in \mathbb{R}^{C_n^r \times n}$, 其中 $C_n^r = \frac{n!}{r!(n-r)!}$ 。矩阵 \mathbf{A} 的每一行 \mathbf{a}_k 由 r 个 1 和 $n - r$ 个 0 组成, 容易证明 \mathbf{a}_k 有 C_n^r 种组合方式。令 $\mathbf{b} = \mathbf{A}\mathbf{x}$, 则 $f(\mathbf{x}) = \sum_{i=1}^r \mathbf{x}_{[i]} = \max_k \{\mathbf{b}\}$ 。根据引理可知, $\max_k \{\mathbf{b}\}$ 为凸函数。

特别的, 对于非升序排列的向量 \mathbf{x} , 令 $\mathbf{a}_1 = [1, \dots, 1, 0, \dots, 0]^T$, 则 $f(\mathbf{x}) \triangleq \max_k \{\mathbf{b}\} = b_1 = \mathbf{a}_1^T \mathbf{x}$ 同样是一个凸函数。

question3

判断下列函数是否是凸函数、凹函数、拟凸函数以及拟凹函数?

(a) 函数 $f(x) = e^x - 1$, 定义域为 \mathbb{R} 。

(b) 函数 $f(x_1, x_2) = x_1 x_2$, 定义域为 \mathbb{R}_{++}^2 。

(c) 函数 $f(x_1, x_2) = \frac{1}{x_1 x_2}$, 定义域为 \mathbb{R}_{++}^2 。

(d) 函数 $f(x_1, x_2) = \frac{x_1}{x_2}$, 定义域为 \mathbb{R}_{++}^2 。

(e) 函数 $f(x_1, x_2) = \frac{x_1^2}{x_2}$, 定义域为 $\mathbb{R} \times \mathbb{R}_{++}$ 。

(f) 函数 $f(x_1, x_2) = x_1^a x_2^{1-a}$, 其中 $0 \leq a \leq 1$, 定义域为 \mathbb{R}_{++}^2 。

solve3

容易验证, 以上所有函数的定义域是凸集。并且我们通过顺序主子式讨论 Hessian 矩阵的正定性, 在之后的证明过程中不再赘述。

(a) $f''(x) = e^x \geq 0$, 根据图像易得其凹凸性。

结论: 是凸函数, 非凹函数, 是拟凸函数, 是拟凹函数。

(b) $\nabla^2 f(x_1, x_2) = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}$ 是不定的, 所以非凸函数, 非凹函数, 是拟凸函数。而其 α 下水平集 $\mathcal{C}_\alpha = \{(x_1, x_2) \in \mathbb{R}_{++}^2 | x_1 x_2 \leq \alpha\}$ 不是凸集, 所以非拟凸函数; 其 α 上水平集 $\mathcal{S}_\alpha = \{(x_1, x_2) \in \mathbb{R}_{++}^2 | x_1 x_2 \geq \alpha\}$ 是凸集, 所以是拟凹函数。(这分别对应着正半轴上 $xy = k$ 双曲线的一支所分割成的两个平面)。

结论: 非凸函数, 非凹函数, 非拟凸函数, 是拟凹函数。

(c) $\nabla^2 f(x_1, x_2) = \begin{bmatrix} \frac{2}{x_1^3 x_2} & \frac{1}{x_1^2 x_2^2} \\ \frac{1}{x_1^2 x_2^2} & \frac{2}{x_1 x_2^3} \end{bmatrix} \succ 0$, 是凸函数, 非凹函数, 是拟凸函数。

而其 α 上水平集 $\mathcal{C}_\alpha = \{(x_1, x_2) \in \mathbb{R}_{++}^2 | x_1 x_2 \leq \alpha\}$ 不是凸集, 所以非拟凹函数。

结论: 是凸函数, 非凹函数, 是拟凸函数, 非拟凹函数。

(d) $\nabla^2 f(x_1, x_2) = \begin{bmatrix} 0 & -\frac{1}{x_2^2} \\ -\frac{1}{x_2^2} & \frac{2x_1}{x_2^3} \end{bmatrix}$ 是不定的。所以是非凸函数, 非凹函数。

而其 α 下水平集 $\mathcal{C}_\alpha = \{(x_1, x_2) \in \mathbb{R}_{++}^2 | \frac{x_1}{x_2} \leq \alpha\}$ 是凸集, 所以是拟凸函数; 而其 α 上水平集 $\mathcal{S}_\alpha = \{(x_1, x_2) \in \mathbb{R}_{++}^2 | \frac{x_1}{x_2} \geq \alpha\}$ 是凸集, 所以是拟凹函数(它们刚好都对应着半平面)。

结论: 非凸函数, 非凹函数, 是拟凸函数, 是拟凹函数。

(e) $\nabla^2 f(x_1, x_2) = \begin{bmatrix} \frac{2}{x_2} & -\frac{2x_1}{x_2^2} \\ -\frac{2x_1}{x_2^2} & \frac{2x_1^2}{x_2^3} \end{bmatrix} \succeq 0$, 是凸函数, 非凹函数, 是拟凸函数。而其 α 上水平集 $\mathcal{S}_\alpha = \{(x_1, x_2) \in \mathbb{R}_{++}^2 \mid \frac{x_1^2}{x_2} \geq \alpha\}$ 不是凸集, 所以非拟凹函数。

结论: 是凸函数, 非凹函数, 是拟凸函数, 非拟凹函数。

(f) $\nabla^2 f(x_1, x_2) = \begin{cases} \mathbf{O}_{2 \times 2}, & a = 1 \text{ or } 0 \\ \begin{bmatrix} a(a-1)x_1^{a-2}x_2^{1-a} & a(1-a)x_1^{a-1}x_2^{-a} \\ a(1-a)x_1^{a-1}x_2^{-a} & a(a-1)x_1^a x_2^{-a-1} \end{bmatrix} & 0 < a < 1 \end{cases}$
 $\nabla^2 f(x_1, x_2) \preceq 0$, 非凸函数, 是凹函数, 是拟凹函数。而其 α 下水平集 $\mathcal{C}_\alpha = \{(x_1, x_2) \in \mathbb{R}_{++}^2 \mid x_1^a x_2^{1-a} \leq \alpha\}$ 不是凸集, 所以非拟凸函数。

结论: 非凸函数, 是凹函数, 非拟凸函数, 是拟凹函数。

question4

考虑优化问题

$$\begin{aligned} \min_{x_1, x_2} \quad & f_0(x_1, x_2) \\ \text{s.t.} \quad & 2x_1 + x_2 \geq 1 \\ & x_1 + 3x_2 \geq 1 \\ & x_1 \geq 0, x_2 \geq 0 \end{aligned}$$

对其可行集进行概述。对下面每个目标函数, 给出最优解和最优值。

- (a) $f_0(x_1, x_2) = x_1 + x_2$
- (b) $f_0(x_1, x_2) = -x_1 - x_2$
- (c) $f_0(x_1, x_2) = x_1$
- (d) $f_0(x_1, x_2) = \max\{x_1, x_2\}$
- (e) $f_0(x_1, x_2) = x_1^2 + 9x_2^2$

solve4

构造拉格朗日函数 $L(x_1, x_2, \boldsymbol{\lambda}) = f_0(x_1, x_2) + \lambda_1(1 - 2x_1 - x_2) + \lambda_2(1 - x_1 - 3x_2) + \lambda_3(-x_1) + \lambda_4(-x_2)$, 其中 $\boldsymbol{\lambda} \geq 0$ 对于上述五个函数 $f_0(x_1, x_2)$, 容易验证它们都是关于 $\mathbf{x} = [x_1, x_2]^T$ 的凸函数。则 $\nabla_{\mathbf{x}} L(x_1, x_2, \boldsymbol{\lambda}) = \nabla f_0(x_1, x_2) + (-2\lambda_1 - \lambda_2 - \lambda_3, -\lambda_1 - 3\lambda_2 - \lambda_4)$ 。

并且给出满足以下凸问题的 KKT 条件:

$$\begin{cases} 2x_1 + x_2 \geq 1, x_1 + 3x_2 \geq 1, x_1 \geq 0, x_2 \geq 0 \\ \lambda_1, \lambda_2, \lambda_3, \lambda_4 \geq 0 \\ \lambda_1(1 - 2x_1 - x_2) = 0, \lambda_2(1 - x_1 - 3x_2) = 0, \lambda_3(-x_1) = 0, \lambda_4(-x_2) = 0 \\ \nabla_{\mathbf{x}} L(x_1, x_2, \boldsymbol{\lambda}) = 0 \end{cases}$$

先作简单的分析, 在二维平面内的仿射约束只会重合或两两交于一点。这说明了互补松弛条件中, $\lambda_1, \lambda_2, \lambda_3, \lambda_4$ 最多存在两个同时为 0。

- (a) $f_0(x_1, x_2) = x_1 + x_2$, 最优解为 $(x_1, x_2) = (\frac{2}{5}, \frac{1}{5})$, 最优值为 $\frac{2}{5}$
- (b) $f_0(x_1, x_2) = -x_1 - x_2$, 不存在最优解与最优值 (最优解为 $(+\infty, +\infty)$ 最优值为 $-\infty$)
- (c) $f_0(x_1, x_2) = x_1$, 最优解为 $(x_1, x_2) = (0, x_2)$, 其中 $x_2 \geq 1$, 最优值为 0
- (d) $f_0(x_1, x_2) = \max\{x_1, x_2\}$, 最优解为 $(x_1, x_2) = (\frac{1}{3}, \frac{1}{3})$, 最优值为 $\frac{1}{3}$
- (e) $f_0(x_1, x_2) = x_1^2 + 9x_2^2$, 最优解为 $(x_1, x_2) = (\frac{1}{2}, \frac{1}{6})$, 最优值为 $\frac{1}{2}$

question5

给出下面每个线性规划 (LP) 的显式解。

- (a) 在仿射集合上极小化线性函数。

$$\begin{aligned} \min_{\mathbf{x}} \quad & \mathbf{c}^T \mathbf{x} \\ \text{s.t.} \quad & \mathbf{Ax} = \mathbf{b} \end{aligned}$$

(b) 在半空间上极小化线性函数。

$$\begin{aligned} \min_x \quad & \mathbf{c}^T \mathbf{x} \\ \text{s.t.} \quad & \mathbf{a}^T \mathbf{x} \leq b \end{aligned}$$

其中 $\mathbf{a} \neq \mathbf{0}$ 。

(c) 在矩阵上极小化线性函数。

$$\begin{aligned} \min_x \quad & \mathbf{c}^T \mathbf{x} \\ \text{s.t.} \quad & \mathbf{l} \preceq \mathbf{x} \preceq \mathbf{u} \end{aligned}$$

其中 \mathbf{l} 和 \mathbf{u} 满足 $\mathbf{l} \preceq \mathbf{u}$ 。

(d) 在概率单纯形上极小化线性函数。

$$\begin{aligned} \min_x \quad & \mathbf{c}^T \mathbf{x} \\ \text{s.t.} \quad & \mathbf{1}^T \mathbf{x} = 1 \\ & \mathbf{x} \succeq \mathbf{0} \end{aligned}$$

solve5

由于它们都是线性规划的问题，满足 weak-slater's condition。那么其对偶问题的最优解则为原问题的最优解。

(a)

$$\begin{aligned} \min_x \quad & \mathbf{c}^T \mathbf{x} \\ \text{s.t.} \quad & \mathbf{A}\mathbf{x} = \mathbf{b} \end{aligned}$$

构造拉格朗日函数 $L(\mathbf{x}, \boldsymbol{\nu}) = \mathbf{c}^T \mathbf{x} + \boldsymbol{\nu}^T (\mathbf{A}\mathbf{x} - \mathbf{b})$ ，其中 $\boldsymbol{\nu} \in \mathbb{R}^n$ ，对其求偏导可得 $\nabla_x L(\mathbf{x}, \boldsymbol{\nu}) = \mathbf{c} + \mathbf{A}^T \boldsymbol{\nu}$ 。

若 $\mathbf{A}\mathbf{x} = \mathbf{b}$ 无解，即 $\mathbf{b} \notin \mathcal{R}(\mathbf{A})$ ：

$$g(\boldsymbol{\nu}) = +\infty$$

若其有解：

$$g(\boldsymbol{\nu}) = \begin{cases} \mathbf{c}^T(\mathbf{A})^+\mathbf{b} & \mathbf{c} + \mathbf{A}^T\boldsymbol{\nu} = \mathbf{0} \\ -\infty & \text{otherwise} \end{cases}$$

其中 $(\cdot)^+$ 是广义逆运算。

综上，对偶问题为：

$$\max_{\boldsymbol{\nu}} g(\boldsymbol{\nu}) = \begin{cases} +\infty & \mathbf{b} \notin R(\mathbf{A}) \\ \mathbf{c}^T(\mathbf{A})^+\mathbf{b} & \mathbf{c} + \mathbf{A}^T\boldsymbol{\nu} = \mathbf{0} \\ -\infty & \text{otherwise} \end{cases}$$

其中 $\boldsymbol{\nu} \in \mathbb{R}^n$ 。

则显示解为：

$$\min_{\mathbf{x}} \mathbf{c}^T \mathbf{x} = \begin{cases} +\infty & \mathbf{b} \notin R(\mathbf{A}) \\ \mathbf{c}^T(\mathbf{A})^+\mathbf{b} & \mathbf{c} + \mathbf{A}^T\boldsymbol{\nu} = \mathbf{0} \\ -\infty & \text{otherwise} \end{cases}$$

(b)

$$\begin{aligned} \min_{\mathbf{x}} \mathbf{c}^T \mathbf{x} \\ \text{s.t. } \mathbf{a}^T \mathbf{x} \leq b \end{aligned}$$

构造拉格朗日函数 $L(\mathbf{x}, \lambda) = \mathbf{c}^T \mathbf{x} + \lambda(\mathbf{a}^T \mathbf{x} - b)$ ，其中 $\lambda \geq 0$ ，对其求偏导可得 $\nabla_{\mathbf{x}} L(\mathbf{x}, \lambda) = \mathbf{c} + \lambda \mathbf{a}$ ，则对偶问题为：

$$\begin{aligned} \max_{\lambda} g(\lambda) &= \begin{cases} -\lambda b & \mathbf{c} + \lambda \mathbf{a} = \mathbf{0} \\ -\infty & \text{otherwise} \end{cases} \\ \text{s.t. } \lambda &\geq 0 \end{aligned}$$

显示解为：

$$\min_{\mathbf{x}} \mathbf{c}^T \mathbf{x} = \begin{cases} -\lambda b & \mathbf{c} + \lambda \mathbf{a} = \mathbf{0} \\ -\infty & \text{otherwise} \end{cases}$$

(c)

$$\begin{aligned} \min_{\mathbf{x}} \quad & \mathbf{c}^T \mathbf{x} \\ \text{s.t.} \quad & \mathbf{l} \preceq \mathbf{x} \preceq \mathbf{u} \end{aligned}$$

其中 \mathbf{l} 和 \mathbf{u} 满足 $\mathbf{l} \preceq \mathbf{u}$ 。

构造拉格朗日函数 $L(\mathbf{x}, \boldsymbol{\lambda}) = \mathbf{c}^T \mathbf{x} + \boldsymbol{\lambda}_1^T (\mathbf{l} - \mathbf{x}) + \boldsymbol{\lambda}_2^T (\mathbf{x} - \mathbf{u})$, 其中 $\boldsymbol{\lambda}_1, \boldsymbol{\lambda}_2 \succeq \mathbf{0}$, 对其求偏导可得 $\nabla_{\mathbf{x}} L(\mathbf{x}, \boldsymbol{\lambda}) = \mathbf{c} - \boldsymbol{\lambda}_1 + \boldsymbol{\lambda}_2$, 则:

$$g(\boldsymbol{\lambda}) = \boldsymbol{\lambda}_1^T \mathbf{l} - \boldsymbol{\lambda}_2^T \mathbf{u}, \quad \mathbf{c} - \boldsymbol{\lambda}_1 + \boldsymbol{\lambda}_2 = \mathbf{0}$$

此时包含两个未知变量 $\boldsymbol{\lambda}_1$ 、 $\boldsymbol{\lambda}_2$, 无法得到显示解。

令 $\mathbf{c} = [c_1, \dots, c_n]^T$, 如果 $c_i < 0$, 则 $x_i = l_i$; 如果 $c_i > 0$, 则 $x_i = u_i$; 如果 $c_i = 0$, 则 x_i 为符合约束条件的任意值都可以。那么显示解为:

$$\min_{\mathbf{x}} \mathbf{c}^T \mathbf{x} = \sum_{c_i \geq 0, c_i \in \mathbf{c}} c_i l_i + \sum_{c_j < 0, c_j \in \mathbf{c}} c_j u_j$$

(d)

$$\begin{aligned} \min_{\mathbf{x}} \quad & \mathbf{c}^T \mathbf{x} \\ \text{s.t.} \quad & \mathbf{1}^T \mathbf{x} = 1 \\ & \mathbf{x} \succeq \mathbf{0} \end{aligned}$$

构造拉格朗日函数 $L(\mathbf{x}, \boldsymbol{\lambda}, \nu) = \mathbf{c}^T \mathbf{x} + \nu(\mathbf{1}^T \mathbf{x} - 1) + \boldsymbol{\lambda}^T (-\mathbf{x})$, 其中 $\nu \in \mathbb{R}, \boldsymbol{\lambda} \succeq \mathbf{0}$, 对其求偏导可得 $\nabla_{\mathbf{x}} L(\mathbf{x}, \boldsymbol{\lambda}, \nu) = \mathbf{c} + \nu \mathbf{1} - \boldsymbol{\lambda}$, 则:

$$g(\boldsymbol{\lambda}, \nu) = -\nu, \quad \mathbf{c} + \nu \mathbf{1} - \boldsymbol{\lambda} = \mathbf{0}$$

此时包含两个未知变量 ν 、 $\boldsymbol{\lambda}$, 无法得到显示解。

令 $\mathbf{c} = [c_1, \dots, c_n]^T$, 把 $\{c_1, \dots, c_n\}$ 按照升序排列变成 $\hat{\mathbf{c}} = [c_{[1]}, \dots, c_{[n]}]$ 。此时令 $\mathbf{x} = [1, 0, \dots, 0]^T$, 则 $\hat{\mathbf{c}}^T \mathbf{x}$ 则为目标函数的最小值, 其值为 $c_{[1]}$ 。那么:

$$\min_{\mathbf{x}} \mathbf{c}^T \mathbf{x} = \min_{c_i} \mathbf{c}$$

question6

给出如下优化函数式

$$\begin{aligned} \min_{\mathbf{x}} \quad & \mathbf{c}^T \mathbf{x} \\ \text{s.t.} \quad & \mathbf{Ax} = \mathbf{b} \\ & \mathbf{x} \succeq \mathbf{0} \end{aligned}$$

写出其 Lagrange 方程，写出其对偶函数解。

solve6

构造拉格朗日函数 $L(\mathbf{x}, \boldsymbol{\lambda}, \boldsymbol{\nu}) = \mathbf{c}^T \mathbf{x} + \boldsymbol{\nu}^T (\mathbf{Ax} - \mathbf{b}) + \boldsymbol{\lambda}^T (-\mathbf{x})$ ，其中 $\boldsymbol{\nu} \in \mathbb{R}^n, \boldsymbol{\lambda} \succeq \mathbf{0}$ ，对其求偏导可得 $\nabla_{\mathbf{x}} L(\mathbf{x}, \boldsymbol{\lambda}) = \mathbf{c} + \mathbf{A}^T \boldsymbol{\nu} - \boldsymbol{\lambda}$ 。

若 $\mathbf{Ax} = \mathbf{b}$ 无解，即 $\mathbf{b} \notin \mathcal{R}(\mathbf{A})$ ：

$$g(\boldsymbol{\lambda}, \boldsymbol{\nu}) = +\infty$$

若其有解：

$$g(\boldsymbol{\lambda}, \boldsymbol{\nu}) = \begin{cases} -\boldsymbol{\nu}^T \mathbf{b} & \mathbf{c} + \mathbf{A}^T \boldsymbol{\nu} - \boldsymbol{\lambda} = \mathbf{0} \\ -\infty & \text{otherwise} \end{cases}$$

综上，对偶函数解析式为：

$$g(\boldsymbol{\lambda}, \boldsymbol{\nu}) = \begin{cases} +\infty & \mathbf{b} \notin \mathcal{R}(\mathbf{A}) \\ -\boldsymbol{\nu}^T \mathbf{b} & \mathbf{c} + \mathbf{A}^T \boldsymbol{\nu} - \boldsymbol{\lambda} = \mathbf{0} \\ -\infty & \text{otherwise} \end{cases}$$

question7

对下述问题，给出其对偶函数解析式，并给出其最优值的下界。

$$\begin{aligned} \min_{\mathbf{x}} \quad & \mathbf{x}^T \mathbf{W} \mathbf{x} \\ \text{s.t.} \quad & x_i^2 = 1, i = 1, \dots, n \end{aligned}$$

其中 $\mathbf{W} \in \mathbb{S}^n$ ， $\mathbf{x} = [x_1, \dots, x_n]^T$ 。

solve7

由于可行集 $\text{dom}f$ 不是凸集，并且目标函数 $f_0(\mathbf{x})$ 不是凸函数，所以这不是一个凸问题。

构造拉格朗日函数 $L(\mathbf{x}, \boldsymbol{\nu}) = \mathbf{x}^T \mathbf{W} \mathbf{x} + \sum_{i=1}^n \nu_i (x_i^2 - 1)$ ，其中 $\boldsymbol{\nu} = [\nu_1, \dots, \nu_n]^T$ 。对矩阵 \mathbf{W} 进行行分块 $[\mathbf{w}_1, \dots, \mathbf{w}_n]^T$ ，对 \mathbf{x} 的每个分量求偏导：

$$\begin{aligned} \frac{\partial L(\mathbf{x}, \boldsymbol{\nu})}{\partial x_i} &= 2\mathbf{w}_i^T \mathbf{x} + 2\nu_i x_i \\ &= 2w_{i1}x_1 + \dots + (2w_{ii}x_i + 2\nu_i)x_i + \dots + 2w_{in}x_n \end{aligned}$$

令 $\frac{\partial L(\mathbf{x}, \boldsymbol{\nu})}{\partial x_1} = \dots = \frac{\partial L(\mathbf{x}, \boldsymbol{\nu})}{\partial x_n} = 0$ ，等价于求解矩阵方程组 $\mathbf{S} \mathbf{x} = \mathbf{0}$ ，其中

$$\mathbf{S} = \begin{pmatrix} 2w_{11} + 2\nu_1 & 2w_{12} & \cdots & 2w_{1n} \\ 2w_{21} & 2w_{22} + 2\nu_2 & \cdots & 2w_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ 2w_{n1} & 2w_{n2} & \cdots & 2w_{nn} + 2\nu_n \end{pmatrix} = 2\mathbf{W} + 2\text{diag}\{\boldsymbol{\nu}\}。$$

若方程组有解，即 $\mathbf{W} \mathbf{x} = -\text{diag}\{\boldsymbol{\nu}\} \mathbf{x}$ 所以，对偶函数为：

$$g(\boldsymbol{\nu}) = \mathbf{x}^T \mathbf{W} \mathbf{x} = \sum_{i=1}^n -\nu_i x_i^2 = -\boldsymbol{\nu}^T \mathbf{1}$$

若方程组无解，根据 \mathbf{W} 的任意性，对偶函数为：

$$g(\boldsymbol{\nu}) = -\infty$$

综上，对偶函数为：

$$g(\boldsymbol{\nu}) = \begin{cases} -\boldsymbol{\nu}^T \mathbf{1} & \mathbf{S} \mathbf{x} = \mathbf{0} \\ -\infty & \text{otherwise} \end{cases}$$

对偶问题为：

$$\max_{\boldsymbol{\nu}} g(\boldsymbol{\nu}) = \begin{cases} -\boldsymbol{\nu}^T \mathbf{1} & \mathbf{S} \mathbf{x} = \mathbf{0} \\ -\infty & \text{otherwise} \end{cases}$$

其中 $\mathbf{x} = [x_1, \dots, x_n]^T$ ，并且 $x_i^2 = 1, i = 1, \dots, n$ 。 $\mathbf{S} = 2\mathbf{W} + \text{diag}\{\boldsymbol{\nu}\}$ 。

由于最小值为 $-\infty$ 没有任何意义，因此根据对偶问题给出的一个较优的下界为 $-\boldsymbol{\nu}^T \mathbf{1}$ 。

question8

考虑优化问题

$$\begin{aligned} \min_x \quad & x^2 + 1 \\ \text{s.t.} \quad & (x - 2)(x - 4) \leq 0 \end{aligned}$$

其中变量 $x \in \mathbb{R}$ 。

- (a) 分析原问题。求解可行集，最优值以及最优解。
- (b) Lagrange 函数以及对偶函数。绘制目标函数根据 x 变化的图像。在同一幅图中，标出可行集，最优点及最优值，选择一些正的 Lagrange 乘子 λ ，绘出 Lagrange 函数 $L(x, \lambda)$ 关于 x 的变化曲线。利用图像，证明下界性质（即对任意 $\lambda \geq 0$ ， $p^* \geq \inf_x L(x, \lambda)$ ）。推导 Lagrange 对偶函数 g 并大致描绘其图像。
- (c) Lagrange 对偶问题。描述对偶问题，证明它是一个凹极大化问题。求解对偶最优值以及对偶最优解 λ^* 。此时强对偶性是否成立？

solve8

(a) 这是一个二次约束二次规划的问题。可行集为 $x \in [2, 4]$ ，最优值为 5，最优解为 $x = 2$ 。

(b) 1. 目标函数 $f_0(x)$

图 1给出了函数 $f_0(x) = x^2 + 1$ 的图像。

2. 拉格朗日函数 $L(x, \lambda)$

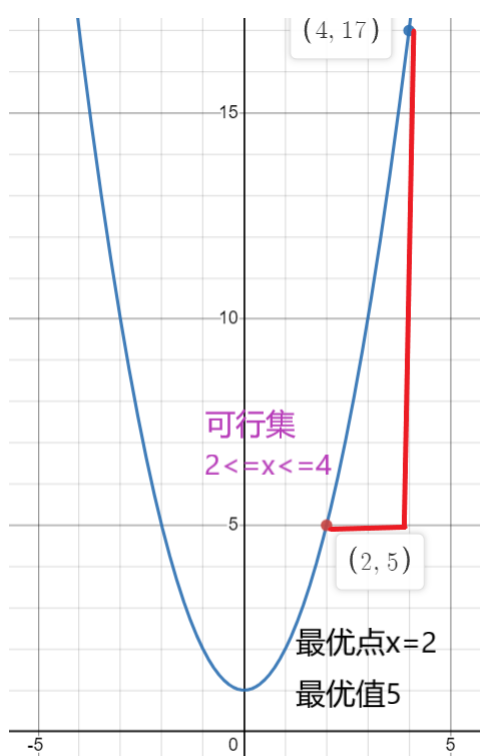


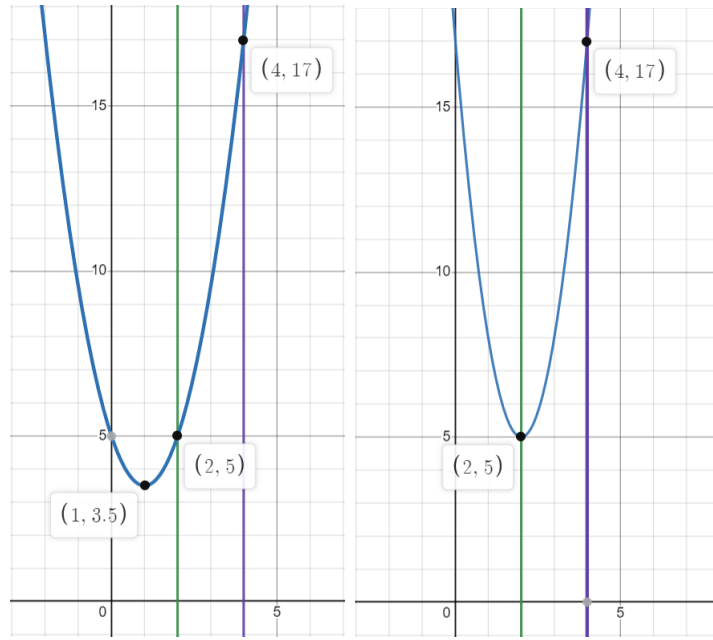
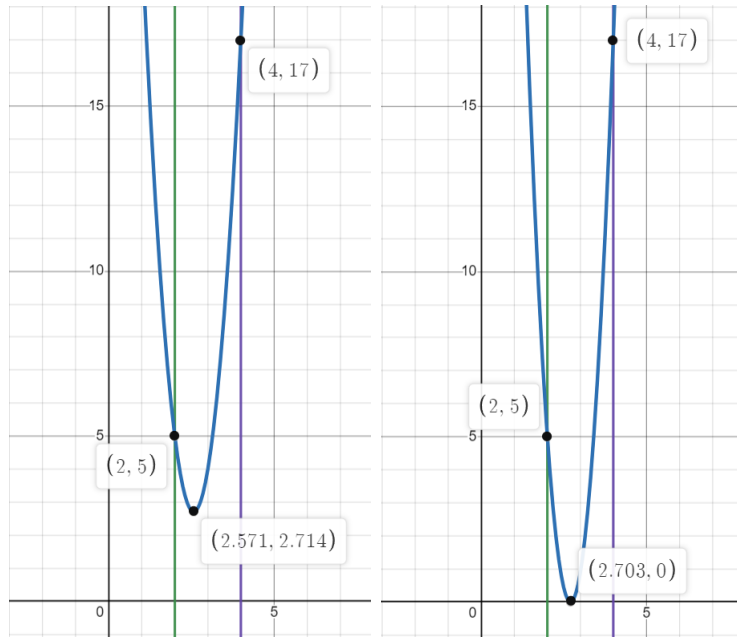
图 1: $f_0(x) = x^2 + 1$ 函数图像

构造拉格朗日函数 $L(x, \lambda) = x^2 + 1 + \lambda(x - 2)(x - 4)$ ，其中 $\lambda \geq 0$ 。对 x 求偏导可得：

$$\frac{\partial L(x, \lambda)}{\partial x} = 2x + \lambda(2x - 6) \quad (1)$$

图 2给出了拉格朗日函数 $L(x, \lambda)$ 的图像。若 $\lambda \leq 2$ ， $\inf_x L(x, \lambda) \leq 5 = p^*$ ；若 $\lambda > 2$ ， $\inf_x L(x, \lambda) < 5 = p^*$ 。这便说明了 $\inf_x L(x, \lambda) \leq p^*$ 。

3. 对偶函数 $g(\lambda)$

(a) $L(x, 0.5)$ (b) $L(x, 2)$ (c) $L(x, 6)$ (d) $L(x, \frac{\sqrt{85}+9}{2})$ 图 2: $L(x, \lambda)$ 函数图像

令式(1)=0, 可得 $x = \frac{3\lambda}{1+\lambda}$, 则对偶函数 $g(\lambda)$ 为:

$$\begin{aligned} g(\lambda) &= \left(\frac{3\lambda}{1+\lambda}\right)^2 + 1 + \lambda\left(\frac{3\lambda}{1+\lambda} - 2\right)\left(\frac{3\lambda}{1+\lambda} - 4\right) \\ &= \frac{-\lambda^3 + 8\lambda^2 + 10\lambda + 1}{(\lambda + 1)^2} \end{aligned} \quad (2)$$

式(2)给出了对偶函数 g 的表达式, 图 3给出了对偶函数 g 的大致图像。

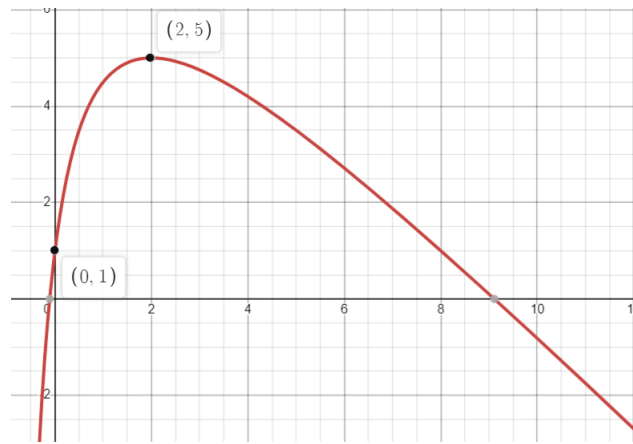


图 3: g 函数图像

(c) 对偶问题是:

$$\begin{aligned} \max \quad & \frac{-\lambda^3 + 8\lambda^2 + 10\lambda + 1}{(\lambda + 1)^2} \\ \text{s.t.} \quad & \lambda \geq 0 \end{aligned}$$

对 $g(\lambda)$ 求导, 并令其等于 0, 可得 $\lambda = 2$ 。并且 $g''(\lambda) \leq 0$ 恒成立, 当且仅当 $\lambda = 2$ 时, $g''(\lambda) = 0$, 这说明了它是一个凹极大化问题。此时对偶问题取到最大值 $d^* = g(\lambda)_{\max} = g(2) = 5$, 与原问题的最优值相同, 这表明了 $p^* = d^*$, 则强对偶成立。