参考答案

2019 年 12 月 11 日

第四章习题

4.1

(1) 否，e^x 在 $x \in \mathbb{R}$ 上就没有局部最优解。

(2) 否，同 (1)，最优值为 0，有限，但不存在最优可行解。

(3) 否，考虑在 $(0,1]$ 上极小化 x，则最优值为 0 且在 $x^* = 0$ 处达到最优值，但 $x^* \notin (0,1]$，所以 x^* 不是最优解。

(4) 否，考虑在可行域 $(0,1]$ 上极小化 $\log(x)$，由于 $\log(x) \to -\infty$ 当 $x \to +0$，所以最优目标值无界，但可行解区域 $(0,1]$ 有界。

(5) 是，当可行解区域非空时，由于 $\mathcal{F}$ 有界，则 $cl(\mathcal{F})$ 为紧集。又因为，f 在全空间连续，故在 $cl(\mathcal{F})$ 上亦连续，于是 $f(x)$ 在 $cl(\mathcal{F})$ 上可以取到有限的最小值，不妨设为 $f(x^*)$，其中 $x^* \in cl(\mathcal{F})$。于是有 $f(x)$ 在 $\mathcal{F}$ 上的最小值不小于 $f(x^*)$，故有限。同理，如果是在 $\mathcal{F}$ 上极大化 $f(x)$，也可以类似得到最优值有限。

(6) 否，见书中例 4.4，$(0,0)$ 是最优解，但是不满足 KKT 条件。

4.2

(1) 令 $f(\boldsymbol{x}) = x_1 + x_2$，$g_1(\boldsymbol{x}) = x_1^2 + x_2^2 - 1$，则 $\nabla f(\boldsymbol{x}) = (1,\ 1)^\top$，$\nabla g_1(\boldsymbol{x}) = (2x_1,\ 2x_2)^\top$。于是此优化问题的 KKT 条件如下：

$$\begin{cases} \begin{pmatrix} 1 \\ 1 \end{pmatrix} + \lambda_1 \begin{pmatrix} 2x_1 \\ 2x_2 \end{pmatrix} = \begin{pmatrix} 2\lambda_1 x_1 + 1 \\ 2\lambda_1 x_2 + 1 \end{pmatrix} = \mathbf{0} \\ \lambda_1(x_1^2 + x_2^2 - 1) = 0 \\ x_1^2 + x_2^2 - 1 \le 0 \\ \lambda_1 \ge 0. \end{cases}$$

易知 $\lambda_1 \neq 0$，故 KKT 条件可写为

$$\begin{cases} x_1 = -\frac{1}{2\lambda_1} \\ x_2 = -\frac{1}{2\lambda_1} \\ x_1^2 + x_2^2 - 1 = 0 \\ \lambda_1 \ge 0. \end{cases}$$

可以得到 $\lambda_1 = \frac{\sqrt{2}}{2}$。于是满足 KKT 条件的点是 $\boldsymbol{x} = (-\sqrt{2}/2, -\sqrt{2}/2)^T$，该点既是局部最优解，亦是全局最优解。

(2) 令 $f(\boldsymbol{x}) = x_1^2 + x_2 - x_3$，$g_1(\boldsymbol{x}) = x_1^2 + 2x_2^2 + x_3^2 - 1$，$g_2(\boldsymbol{x}) = x_1 + x_2 + x_3 - 1$，则 $\nabla f(\boldsymbol{x}) = (2x_1\ 1\ -1)^T$，$\nabla g_1(\boldsymbol{x}) = (2x_1\ 4x_2\ 2x_3)^\top$，$\nabla g_2(\boldsymbol{x}) = (1\ 1\ 1)^T$。于是，此优化问题的 KKT 条件如下:

$$\begin{cases} \begin{pmatrix} 2x_1 \\ 1 \\ -1 \end{pmatrix} + \lambda_1 \begin{pmatrix} 2x_1 \\ 4x_2 \\ 2x_3 \end{pmatrix} + \lambda_2 \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} = \begin{pmatrix} 2x_1 + 2\lambda_1 x_1 + \lambda_2 \\ 1 + 4\lambda_1 x_2 + \lambda_2 \\ -1 + 2\lambda_1 x_3 + \lambda_2 \end{pmatrix} = \mathbf{0} \\ \lambda_1(x_1^2 + 2x_2^2 + x_3^2 - 1) = 0 \\ \lambda_2(x_1 + x_2 + x_3 - 1) = 0 \\ x_1^2 + 2x_2^2 + x_3^2 - 1 \le 0 \\ x_1 + x_2 + x_3 - 1 \le 0 \\ \lambda_1,\ \lambda_2 \ge 0 \end{cases}$$

如果 $\lambda_1 = 0$ 时，则由 $1 + 4\lambda_1 x_2 + \lambda_2 = 0$ 知 $\lambda_2 = -1$，矛盾。于是 $\lambda_1 > 0$，于是我们知道 $x_1^2 + 2x_2^2 + x_3^2 - 1 = 0$ 必然成立。

当 $\lambda_2 = 0$ 时，易知 $x_1 = 0$，KKT 条件可写为

$$\begin{cases} x_2 = -\frac{1}{4\lambda_1},\ x_3 = \frac{1}{2\lambda_1} \\ 2x_2^2 + x_3^2 - 1 = 0 \\ x_2 + x_3 - 1 \le 0 \\ \lambda_1 \ge 0 \end{cases}$$

解其方程，得 $\lambda_1 = \frac{\sqrt{6}}{4}, x_2 = -\frac{\sqrt{6}}{6}, x_3 = \frac{\sqrt{6}}{3}$。

当 $\lambda_2 > 0$ 时，上述的 KKT 条件可改写为

$$\begin{cases} \begin{pmatrix} 2x_1 + 2\lambda_1 x_1 + \lambda_2 \\ 1 + 4\lambda_1 x_2 + \lambda_2 \\ -1 + 2\lambda_1 x_3 + \lambda_2 \end{pmatrix} = \mathbf{0} \\ x_1^2 + 2x_2^2 + x_3^2 - 1 = 0 \\ x_1 + x_2 + x_3 - 1 = 0 \\ \lambda_1,\ \lambda_2 \ge 0 \end{cases}$$

无解。

故满足 KKT 条件的只有 $\boldsymbol{x} = (0, -\frac{\sqrt{6}}{6}, \frac{\sqrt{6}}{3})^T$。由于优化问题是凸优化问题，因此是全局最优解，也是局部最优解。

4.3 证明: $\mathcal{X}$ 是多面体，因此可以表示成有限个点的凸组合和有限个方向的非负线性组合。即对于任意的 $x \in \mathcal{X}$，均存在 $\sum_{i=1}^k \lambda_i = 1$，且 $\lambda_i \ge 0, i = 1, 2..., k, \mu_j \ge 0, j = 1, 2..., t$ 使得 $x = \sum_{i=1}^k \lambda_i v_i + \sum_{j=1}^t \mu_j d_j$，其中 $v_i, i = 1, \cdots, k$ 是给定的极点，$d_j, j = 1, \cdots, t$ 是给定的极方向。

于是 $\forall y \in A\mathcal{X}$，存在 $x \in \mathcal{X}$ 使得 $y = Ax$，而对此 x，存在 $\sum_{i=1}^k \lambda_i = 1$，且 $\lambda_i \ge 0, i = 1, 2..., k, \mu_j \ge 0, j = 1, 2..., t$ 使得 $x = \sum_{i=1}^k \lambda_i v_i + \sum_{j=1}^t \mu_j d_j$，于是 $y = \sum_{i=1}^k \lambda_i A v_i + \sum_{j=1}^t \mu_j A d_j$，于是我们可以取定新的一组极点 $Av_i, i = 1, \cdots, k$，新的一组极方向 $Ad_j, j = 1, \cdots, t$。反之，任意可以表示成这些极点的凸组合和极方向非负组合的点，也必然在 $A\mathcal{X}$ 中，所以由定义，$A\mathcal{X}$ 是多面体。 □

4.4

(1) 证明：令 $g(\boldsymbol{x}) \equiv f(\boldsymbol{x}) + \frac{1}{2\alpha}(\boldsymbol{x}-\boldsymbol{y})^\top(\boldsymbol{x}-\boldsymbol{y})$。我们证明对任意 $\boldsymbol{x} \in \mathbb{R}^n$，$g$ 的 Hesse 矩阵 $\nabla^2 g(\boldsymbol{x})$ 为正定矩阵即可。先注意，$\nabla g(\boldsymbol{x}) = \nabla f(\boldsymbol{x}) + \frac{1}{\alpha}(\boldsymbol{x}-\boldsymbol{y})$，$\nabla^2 g(\boldsymbol{x}) = \nabla^2 f(\boldsymbol{x}) + \frac{1}{a}I$。令 $\nabla^2 f(\boldsymbol{x})$ 的 n 个特征值为 $\lambda_1, \ldots, \lambda_n$（不妨设 $|\lambda_1| \le \cdots \le |\lambda_n|$），则 $(\nabla^2 f(\boldsymbol{x}))^2$ 的 n 个特征值为 $\lambda_1^2, \ldots, \lambda_n^2$。由

$$\|\nabla^2 f(\boldsymbol{x})\|_2 = \sqrt{\lambda_{\max}(\nabla^2 f(\boldsymbol{x})^T \nabla^2 f(\boldsymbol{x}))} = \sqrt{\lambda_{\max}((\nabla^2 f(\boldsymbol{x}))^2)} = |\lambda_n|,$$

我们得到 $|\lambda_n| \le L$。于是对任意 $i = 1, \cdots, n$ 都有 $|\lambda_i| \le L$，即 $\lambda_i \ge -L$。由 $\nabla^2 g(\boldsymbol{x}) = \nabla^2 f(\boldsymbol{x}) + \frac{1}{\alpha}I$ 的 n 个特征值为 $\lambda_1 + \frac{1}{\alpha}, \ldots, \lambda_n + \frac{1}{\alpha}$ 且

$$\lambda_i + \frac{1}{\alpha} \ge -L + \frac{1}{\alpha} > 0 \quad (\forall i = 1, \cdots, n),$$

$\nabla^2 g(\boldsymbol{x})$ 为正定矩阵。 □

(2) 证明：由于 $\alpha > 0$，我们只需证明 $f(\boldsymbol{y}) - f(\boldsymbol{z}) \ge \frac{1}{2\alpha}(\boldsymbol{z}-\boldsymbol{y})^T(\boldsymbol{z}-\boldsymbol{y})$。令 $g(\boldsymbol{x}) \equiv f(\boldsymbol{x}) + \frac{1}{2\alpha}(\boldsymbol{x}-\boldsymbol{y})^\top(\boldsymbol{x}-\boldsymbol{y})$。由于 $\boldsymbol{z}$ 是 $g(\boldsymbol{x})$ 的最小值的点，我们有 $g(\boldsymbol{z}) = f(\boldsymbol{z}) + \frac{1}{2\alpha}(\boldsymbol{z}-\boldsymbol{y})^\top(\boldsymbol{z}-\boldsymbol{y}) \le g(\boldsymbol{y}) = f(\boldsymbol{y})$。于是 $f(\boldsymbol{y}) - f(\boldsymbol{z}) \ge \frac{1}{2\alpha}(\boldsymbol{z}-\boldsymbol{y})^\top(\boldsymbol{z}-\boldsymbol{y})$。 □

4.5

注意

$$\boldsymbol{x} \in \{0,1\}^n \Leftrightarrow x_i(x_i - 1) = 0, i = 1, \cdots, n$$

于是，我们得到二次约束二次规划模型：

$$\begin{aligned} \min \quad & \boldsymbol{x}^\top Q \boldsymbol{x} + \boldsymbol{q}^\top \boldsymbol{x} \\ \text{s.t.} \quad & x_i^2 - x_i \le 0 \quad (\forall i = 1, \cdots, n) \\ & -x_i^2 + x_i \le 0 \quad (\forall i = 1, \cdots, n) \end{aligned}$$

Lagrange 函数：

$$\begin{aligned} L(\boldsymbol{x}, \boldsymbol{\lambda}) &= \boldsymbol{x}^T Q \boldsymbol{x} + \boldsymbol{q}^T \boldsymbol{x} + \sum_{i=1}^n \lambda_i (x_i^2 - x_i) \\ &= \boldsymbol{x}^T Q \boldsymbol{x} + \boldsymbol{q}^\top \boldsymbol{x} + \boldsymbol{x}^T diag(\lambda_1, \cdots, \lambda_n) \boldsymbol{x} - \boldsymbol{\lambda}^T \boldsymbol{x} \ (\boldsymbol{\lambda} \in \mathbb{R}^n). \end{aligned}$$

Lagrange 对偶问题：

$$\begin{aligned} \max \quad & \sigma \\ \text{s.t.} \quad & L(\boldsymbol{x}, \boldsymbol{\lambda}) \ge \sigma \quad, \forall \boldsymbol{x} \in \mathbb{R}^n \\ & \sigma \in \mathbb{R}, \ \boldsymbol{\lambda} \in \mathbb{R}^n \end{aligned} \tag{0.1}$$

我们可以得到

$$\begin{aligned} L(\boldsymbol{x}, \boldsymbol{\lambda}) \ge \sigma, \forall \boldsymbol{x} \in \mathbb{R}^n &\Leftrightarrow \boldsymbol{x}^\top Q \boldsymbol{x} + \boldsymbol{x}^\top diag(\lambda_1, \ldots, \lambda_n) \boldsymbol{x} + \boldsymbol{q}^\top \boldsymbol{x} - \boldsymbol{\lambda}^\top \boldsymbol{x} - \sigma \ge 0, \forall \boldsymbol{x} \in \mathbb{R}^n \\ &\Leftrightarrow (1 \quad \boldsymbol{x}) \begin{pmatrix} -\sigma & \frac{1}{2}(\boldsymbol{q}-\boldsymbol{\lambda})^\top \\ \frac{1}{2}(\boldsymbol{q}-\boldsymbol{\lambda}) & Q + diag(\lambda_1, \ldots, \lambda_n) \end{pmatrix} \begin{pmatrix} 1 \\ \boldsymbol{x} \end{pmatrix} \ge 0, \forall \boldsymbol{x} \in \mathbb{R}^n \\ &\Leftrightarrow \begin{pmatrix} -\sigma & \frac{1}{2}(\boldsymbol{q}-\boldsymbol{\lambda})^\top \\ \frac{1}{2}(\boldsymbol{q}-\boldsymbol{\lambda}) & Q + diag(\lambda_1, \ldots, \lambda_n) \end{pmatrix} \in \mathcal{S}_+^{n+1} \end{aligned}$$

最后一个等价可参考教材 p129-p130，则我们可以将对偶问题写称下述半正定规划问题

$$
\begin{aligned}
\max\quad & \sigma \\
\text{s.t.}\quad & \begin{pmatrix} -\sigma & \frac{1}{2}(\boldsymbol{q}-\boldsymbol{\lambda})^\top \\ \frac{1}{2}(\boldsymbol{q}-\boldsymbol{\lambda}) & Q+diag(\lambda_1,\dots,\lambda_n) \end{pmatrix} \in \mathcal{S}_+^{n+1}, \sigma\in\mathbb{R},\ \boldsymbol{\lambda}\in\mathbb{R}^n
\end{aligned}
$$

4.6

一般的 QCQP 模型如下：

$$
\begin{aligned}
\min\quad & \frac{1}{2}x^TQ_0x+q_0^Tx+c_0 \\
s.t.\quad & \frac{1}{2}x^TQ_ix+q_i^Tx+c_i\le 0, \forall i=0,1,\cdots,n. \\
& x\in\mathbb{R}^n
\end{aligned}
$$

其 Lagrange 对偶模型如下

$$
\begin{aligned}
\max\quad & \sigma \\
s.t.\quad & \begin{pmatrix}1\\x\end{pmatrix}^T U \begin{pmatrix}1\\x\end{pmatrix}\ge 0, \forall x\in\mathbb{R}^n. \\
& \sigma\in\mathbb{R}, \lambda\in\mathbb{R}_+^m
\end{aligned}
$$

其中

$$
U=\begin{pmatrix} -2(\sigma-c_0-\sum_{i=1}^m\lambda_ic_i) & (q_0+\sum_{i=1}^m\lambda_iq_i)^T \\ q_0+\sum_{i=1}^m\lambda_iq_i & Q_0+\sum_{i=1}^m\lambda_iQ_i \end{pmatrix}
$$

进一步等价为

$$
\begin{aligned}
\max\quad & \sigma \\
s.t.\quad & \begin{pmatrix} -2(\sigma-c_0-\sum_{i=1}^m\lambda_ic_i) & (q_0+\sum_{i=1}^m\lambda_iq_i)^T \\ q_0+\sum_{i=1}^m\lambda_iq_i & Q_0+\sum_{i=1}^m\lambda_iQ_i \end{pmatrix}\in\mathcal{S}_+^{n+1} \\
& \sigma\in\mathbb{R}, \lambda\in\mathbb{R}_+^m
\end{aligned}
$$

4.7 由于 $\mathcal{X}$ 是多面体，则存在矩阵 $A\in\mathbb{R}^{m\times n}$，向量 $b\mathbb{R}^m$，使得 $\mathcal{X}=\{x\in\mathbb{R}^n|Ax\le b\}$，另外 $f(x)$ 为线性函数，所以不妨设其为 $f(x)=a^Tx$。于是 $f^*(y)=max_{Ax\le b}(y-a)^Tx$，由线性规划的对偶理论可知，$min_{y\in\mathcal{Y}}f^*(y)$ 模型如下：

$$
\begin{aligned}
min\ b^Ty \\
s.t.\ A^Tz-y=-a \\
z\ge 0
\end{aligned}
$$

另解：

由于 $\mathcal{X}$ 是多面体，因此可以表示成有限个点的凸组合和有限个方向的非负线性组合。即 $x=\sum_{i=1}^k\lambda_iv_i+\sum_{j=1}^t\mu_jd_j$，其中 $\sum_{i=1}^k\lambda_i=1$，且 $\lambda_i\ge 0, i=1,2...,k, \mu_j\ge 0, j=1,2...,t$。$f(x)$ 为线性函数，所以不妨设其为 $f(x)=a^Tx$，则共轭函数为 $f^*(y)=max_{x\in\mathcal{X}}\{(y-a)^Tx\}$，不难知道 $\mathcal{Y}=\{y|(y-a)^Td_j\le 0, j=1,2,...,t\}$，且此时 $f^*(y)=max_{i\in 1,2...k}\{(y-a)^Tv_i\}$。于是 $min_{y\in\mathcal{Y}}f^*(y)$ 模型如下：

$$
\begin{aligned}
min\ \sigma \\
s.t.\ (y-a)^Tv_i\le\sigma, i=1,2,...,k \\
(y-a)^Td_j\le 0, j=1,2,...,t
\end{aligned}
$$

凸函数 $L(\cdot,\bar{\lambda})$ 达最小值 $\Longleftrightarrow \nabla_x L(\bar{x},\bar{\lambda})=0$

$\therefore f(\bar{x}) = L(\bar{x},\bar{\lambda}) \le L(x,\bar{\lambda}) \le f(x)$

4.8

证明：对偶问题为

$$\max_{\boldsymbol{\lambda}\geq \mathbf{0}}[\min_{\boldsymbol{x}\in\mathbb{R}^n} L(\boldsymbol{x},\boldsymbol{\lambda})]$$

$\max \quad v(\lambda)$

$\text{s.t.} \quad L(x,\lambda) \geq v(\lambda) \quad \forall x$ 成立

其中，$L(\boldsymbol{x},\boldsymbol{\lambda}) = f(\boldsymbol{x}) + \sum_{i=1}^m \lambda_i g_i(\boldsymbol{x})$。令

$$v(\boldsymbol{\lambda}) \equiv \min_{\boldsymbol{x}\in\mathbb{R}^n} L(\boldsymbol{x},\boldsymbol{\lambda})$$

由 $(\bar{\boldsymbol{x}},\bar{\boldsymbol{\lambda}})$ 满足 KKT 条件，因此

$$\nabla f(\bar{\boldsymbol{x}}) + \sum_{i=1}^m \bar{\lambda}_i \nabla g(\bar{\boldsymbol{x}}) = \mathbf{0}.$$

由 f 和 g_i $(i=1\cdots,m)$ 均为凸函数，$L(\cdot,\bar{\boldsymbol{\lambda}})$ 亦为 $\mathbb{R}^n$ 上的凸函数，故 $\bar{\boldsymbol{x}}$ 为 $L(\cdot,\bar{\boldsymbol{\lambda}})$ 的全局最小解。于是，记 v_p 为原问题的最优值，v_d 为对偶问题的最优值，则

$$v_d \geq v(\bar{\boldsymbol{\lambda}}) = L(\bar{\boldsymbol{x}},\bar{\boldsymbol{\lambda}}) = f(\bar{\boldsymbol{x}}) + \sum_{i=1}^m \bar{\lambda}_i g_i(\bar{\boldsymbol{x}}) = f(\bar{\boldsymbol{x}}) \geq v_p$$

另一方面，由弱对偶定理，$v_p \geq v_d$，因此 $v_p = v_d$。因此 $f(\bar{\boldsymbol{x}}) = v_p, v(\bar{\boldsymbol{\lambda}}) = v_d$，即 $\bar{\boldsymbol{x}}$ 是原始问题的最优解，$\bar{\boldsymbol{\lambda}}$ 为对偶问题的最优解。 □

4.9

$$\mathcal{X} = \{u\in\mathbb{R}^{m+1} | \quad u_i = a^i\cdot x - b_i \quad i=1,\cdots,m \quad u_{m+1} = c\cdot x \quad x\in\mathcal{K}\},$$

$$\mathcal{K}_0 = \{u\in\mathbb{R}^{m+1} | u_i \geq 0, \quad i=1,\cdots,m, \quad u_{m+1}\in\mathbb{R}\}.$$

$$\begin{aligned}
f^*(v_1,\cdots,v_{m+1}) &= \max_{x\in\mathcal{K}}\{\sum_{i=1}^{m+1} u_i v_i - u_{m+1}\} \\
&= \max_{x\in\mathcal{K}}\{\sum_{i=1}^{m}(a^i\cdot x - b_i)v_i + (v_{m+1}-1)c\cdot x\} \\
&= \max_{x\in\mathcal{K}}\{[(v_{m+1}-1)c + \sum_{i=1}^{m} v_i a^i]\cdot x - \sum_{i=1}^{m} b_i v_i\} \\
&= \begin{cases} -\sum_{i=1}^m b_i v_i, & -(v_{m+1}-1)c - \sum_{i=1}^m v_i a^i \in \mathcal{K}^* \\ +\infty, & \text{otherwise} \end{cases}
\end{aligned}$$

$$\mathcal{Y} = \{(v_1,\cdots,v_{m+1})^T\in\mathbb{R}^{m+1} | \quad -(v_{m+1}-1)c - \sum_{i=1}^m v_i a^i \in \mathcal{K}^*\},$$

$$\mathcal{K}^* = \{(v_1,\cdots,v_{m+1})^T\in\mathbb{R}^{m+1} | \quad v\in\mathbb{R}^m_+, \quad v_{m+1} = 0\}.$$

因此对偶模型为：

$$\begin{aligned}
-\min \quad & -\sum_{i=1}^m b_i v_i \\
s.t. \quad & -(v_{m+1}-1)c - \sum_{i=1}^m v_i a^i \in \mathcal{K}^* \\
& v_{m+1} = 0 \\
& v\in\mathbb{R}^m_+
\end{aligned}$$

令 $y=(v_1,\cdots,v_m)^T$ 则模型等价于:

$$\begin{aligned}\max\quad & b^Ty\\ s.t.\quad & \sum_{i=1}^m y_ia^i+s=c\\ & s\in\mathcal{K}^*\\ & y\in\mathbb{R}_+^m\end{aligned}$$

4.10

(1) 根据原问题，我们得到 $\mathcal{X}\equiv\{(x_1\ x_2\ x_3)^\top\mid x_2-x_3=0\}$。于是 f 的共轭函数 f^* 为

$$\begin{aligned}f^*(\boldsymbol{y})&=\max_{\boldsymbol{x}\in\mathcal{X}}\{\boldsymbol{x}^\top\boldsymbol{y}-f(\boldsymbol{x})\}=\max_{\boldsymbol{x}\in\mathbb{R}^3,\ x_2-x_3=0}\{x_1y_1+x_2y_2+x_3y_3+2x_1\}\\&=\max_{x_1,x_2\in\mathbb{R}}\{(y_1+2)x_1+(y_2+y_3)x_2\}\\&=\begin{cases}0 & y_1=-2,y_2+y_3=0)\\+\infty & \text{其它}\end{cases}\end{aligned}$$

于是共轭对偶问题是

$$\begin{aligned}\text{max}\quad & 0\\ \text{s.t.}\quad & y_1=-2\\ & y_2+y_3=0\\ & \boldsymbol{y}\in\mathcal{L}^3\end{aligned}\tag{0.2}$$

易知对偶问题不可行，故最优值为 $-\infty$。对原问题，必须有有 $x_1=0$，故目标函数值恒为 0，即原始最优值为 0，原问题和对偶问题不具有强对偶性。

(2) 由 $x_1+x_3=0,\ x_2+x_4=-1,x_1+x_5=0$ 且 $(x_3\ x_4\ x_5)^\top\in\mathcal{L}^3$，可知 $x_4=0,x_2=-1$，因此目标函数值恒为 1，于是原问题的最优值为 1。另外此时

$$\mathcal{X}\equiv\left\{\boldsymbol{x}\in\mathbb{R}^7\ \middle|\ \begin{array}{l}x_1+x_3=0,\ x_2+x_4=-1,\ x_1+x_5=0,\\ -x_1+x_2+x_6=0,\ x_1+x_7=0\end{array}\right\},$$

于是 f 的共轭函数 f^* 为

$$\begin{aligned}f^*(\boldsymbol{y})&=\max_{\boldsymbol{x}\in\mathcal{X}}\{\boldsymbol{x}^\top\boldsymbol{y}-f(\boldsymbol{x})\}\\&=\max_{x_1,x_2\in\mathbb{R}}\{(y_1-y_3-y_5+y_6-y_7)x_1+(y_2-y_4-y_6+1)x_2-y_4\}\\&=\begin{cases}-y_4 & y_1-y_3-y_5+y_6-y_7=0,y_2-y_4-y_6=-1\\+\infty & \text{其它}\end{cases}\end{aligned}$$

共轭对偶问题是

$$\begin{aligned}\text{max}\quad & y_4\\ \text{s.t.}\quad & -y_3-y_5+y_6-y_7=0\\ & -y_4-y_6=-1\\ & (y_3\ y_4\ y_5)^T\in\mathcal{L}^3\\ & (y_6\ y_7)^T\in\mathcal{L}^2\end{aligned}$$

对偶问题显然可行，$(y_3,y_4,y_5,y_6,y_7)=(0,0,0,1,1)$ 是一个可行解。由弱对偶定理，对任意的对偶可行解，均有 $y_4\le 1$，故 $y_6=-y_4+1\ge 0$，再由此和 $(y_6\ y_7)^\top\in\mathcal{L}^2$ 可知 $y_6=-y_4+1\le y_7$。由 $-y_3-y_5+y_6-y_7=-y_3-y_5-y_4+1-y_7=0$ 可知 $y_3+y_5\le 0$，又由 $(y_3,y_4,y_5)^\top\in\mathcal{L}^3$ 可知 $y_4=0$，因此对偶问题的最优值为 0。所以原始对偶问题不具有强对偶性。

(3) 对任意 $k\ge 0$，$(x_1,x_2,x_3,t_1,x_4,x_5,t_2)=(1,0,-2,0,0,0,k)$ 是一个原问题的可行解并且此时的目标函数值为 $-5k$。因 k 的任意性，$-5k\to-\infty$ $(k\to+\infty)$，故原问题的最优值为 $-\infty$。原问题等价为:

$$\begin{aligned}
\min\quad & 2t_1-5t_2\\
\text{s.t.}\quad & x_1'-x_1=-1\\
& x_3'-x_3=2\\
& \boldsymbol{x}=(x_1',x_2,x_3',t_1,x_4,x_5,t_2,x_1,x_3)\in\mathcal{L}^4\times\mathcal{L}^3\times\mathbb{R}^2
\end{aligned}\tag{0.3}$$

于是，令 $\mathcal{X}=\{\boldsymbol{x}\in\mathbb{R}\mid x_1'-x_1=-1,x_3'-x_3=2\}$，则

$$\begin{aligned}
& f^*(\boldsymbol{y})=\max_{\boldsymbol{x}\in\mathcal{X}}\{\boldsymbol{x}^\top\boldsymbol{y}-f(\boldsymbol{x})\}\\
=& \max_{(x_1',x_2,x_3',t_1,x_4,x_5,t_2)\in\mathbb{R}^7}\{x_1'y_1+x_2y_2+x_3'y_3+t_1y_4+x_4y_5+x_5y_6+t_2y_7+(x_1'+1)y_8+(x_3'-2)y_9-2t_1+5t_2\}\\
=& \max_{(x_1',x_2,x_3',t_1,x_4,x_5,t_2)\in\mathbb{R}^7}\{(y_1+y_8)x_1'+y_2x_2+(y_3+y_9)x_3'+(y_4-2)t_1+y_5x_4+y_6x_5+(y_7+5)t_2+y_8-2y_9\}\\
=& \begin{cases} y_8-2y_9 & y_1+y_8=0,\ y_2=0,\ y_3+y_9=0,\ y_4=2,\ y_5=0,\ y_6=0,\ y_7=-5\\ +\infty & \text{其它}\end{cases}
\end{aligned}$$

于是，此优化问题的共轭对偶问题是

$$\begin{aligned}
\max\quad & -y_8+2y_9\\
\text{s.t.}\quad & y_1+y_8=0\\
& y_2=0\\
& y_3+y_9=0\\
& y_4=2\\
& y_5=0\\
& y_6=0\\
& y_7=-5\\
& \boldsymbol{y}\in\mathcal{L}^4\times\mathcal{L}^3\times\{0\}^2
\end{aligned}$$

可以进一步化为

$$\begin{aligned}
\max\quad & 0\\
\text{s.t.}\quad & y_1=0\\
& y_2=0\\
& y_3=0\\
& y_4=2\\
& y_5=0\\
& y_6=0\\
& y_7=-5\\
& \boldsymbol{y}\in\mathcal{L}^4\times\mathcal{L}^3
\end{aligned}$$

由于 $y_7=-5$，所以对偶问题不可行，对偶问题最优值亦为 $-\infty$。原始对偶问题的目标值相等，但是由于对偶不可行，故强对偶不成立。

(4) 由

$$S=\begin{pmatrix}0 & 1-x\\ 1-x & 0\end{pmatrix}\in\mathcal{S}^2_+,$$

可知必有 $x=1$，S 是全零矩阵，此时的目标函数值为 -1。因此原问题的最优值为 -1 且可达。令

$$\mathcal{X}\equiv\left\{(x,S)\middle| x\begin{pmatrix}0 & 1\\ 1 & 0\end{pmatrix}+S=\begin{pmatrix}0 & 1\\ 1 & 0\end{pmatrix}\right\},$$

于是共轭函数

$$\begin{aligned}
f^*(y,T)&=\max_{(x,S)\in\mathcal{X}}\{xy+S\bullet T+x\}\\
&=\max_{x\in\mathbb{R}}\left\{xy+\begin{pmatrix}0 & 1-x\\ 1-x & 0\end{pmatrix}\bullet\begin{pmatrix}t_{11} & t_{12}\\ t_{21} & t_{22}\end{pmatrix}+x\right\}\\
&=\max_{x\in\mathbb{R}}\{xy+t_{12}(1-x)+t_{21}(1-x)+x\}\\
&=\max_{x\in\mathbb{R}}\{x(y-t_{12}-t_{21}+1)+t_{12}+t_{21}\}\\
&=\begin{cases}t_{12}+t_{21} & y-t_{12}-t_{21}+1=0\\ +\infty & \text{其它}\end{cases}
\end{aligned}$$

max $-f^*(y)$

于是，此优化问题的共轭对偶问题是

$$\begin{aligned}
\max\quad & -t_{12}-t_{21}\\
\text{s.t.}\quad & y-t_{12}-t_{21}=-1\\
& (y,T)\in\{0\}\times\mathcal{S}^2_+
\end{aligned}$$

等价为

$$\begin{aligned}
\max\quad & -t_{12}-t_{21}\\
\text{s.t.}\quad & t_{12}+t_{21}=1\\
& T\in\mathcal{S}^2_+
\end{aligned}$$

该问题显然可行，且目标函数值恒为 -1。于是对偶问题的最优值为 -1 且可达。因此，原始对偶问题具有强对偶性。易知对偶问题存在严格可行内点，且目标值有限，强对偶成立。

(5) 由定义

$$x_{11}\geq 0,\ \ x_{22}\geq 0, x_{11}x_{22}-\frac{1}{4}\geq 0. \tag{0.4}$$

于是

$$x_{11}+x_{22}\geq 2\sqrt{x_{11}x_{22}}=1,$$

故原问题的最优值大于等于 1。显然

$$\begin{pmatrix}1/2 & 1/2\\ 1/2 & 1/2\end{pmatrix},$$

可行，且此时的目标值为 1，故原问题的最优值为 1。

若开始取 $X = \{ X \in M(2,2) \mid 2X_{12} = 1 \}$ 则有 $f^*(Y) = \begin{cases} \frac{1}{2}Y_{12}, & Y_{21} = 0 \\ +\infty & Y_{11} = Y_{22} = 1 \end{cases}$

令 $\mathcal{X} = \{X \in \mathcal{S}^2 \mid 2x_{12} = 1\}$，则共轭函数为

$$
\begin{aligned}
f^*(Y) &= \max_{X \in \mathcal{X}} \{X \bullet Y - f(X)\} \\
&= \max_{X \in \mathcal{S}^2, 2x_{12}=1} \{x_{11}y_{11} + x_{12}y_{12} + x_{21}y_{21} + x_{22}y_{22} - x_{11} - x_{22}\} \\
&= \max_{x_{11}, x_{22} \in \mathbb{R}} \{(y_{11} - 1)x_{11} + (y_{22} - 1)x_{22} + \frac{1}{2}y_{12} + \frac{1}{2}y_{21}\} \\
&= \begin{cases} \frac{1}{2}y_{12} + \frac{1}{2}y_{21} & y_{11} = 1, y_{22} = 1 \\ +\infty & \text{其它} \end{cases}
\end{aligned}
$$

$f^*(Y)$

$\max -\frac{1}{2}Y_{12}$

于是共轭对偶问题是

$$
\begin{aligned}
\max \quad & -\frac{1}{2}y_{12} - \frac{1}{2}y_{21} \\
\text{s.t.} \quad & y_{11} = 1 \\
& y_{22} = 1 \\
& Y \in \mathcal{S}_+^2
\end{aligned}
$$

取

$$
Y^* = \begin{pmatrix} 1 & -1 \\ -1 & 1 \end{pmatrix},
$$

则 Y^* 是对偶问题的可行解并且此时的目标函数值为 1。由弱对偶定理和原问题的最优值为 1，可知 Y^* 为对偶问题的最优解且对偶问题的最优值亦为 1，故原始对偶问题具有强对偶性。我们可以看出对偶问题的可行域与 $\mathcal{S}_{++}^2$ 相交非空，且目标值有限，强对偶成立。

(6) 由原问题的约束可知

$$
S = \begin{pmatrix} -x_2 & 0 & 0 \\ 0 & -x_1 & 1 + x_2 \\ 0 & 1 + x_2 & 0 \end{pmatrix} \in \mathcal{S}_+^3 \quad (x_1, x_2 \in \mathbb{R}),
$$

$-x_2 \geq 0$

$-x_1 \geq 0$

因此必有 $x_2 = -1$，此时的目标函数值为 1，原问题显然可行，我们可以取

$x_2 = -1$

$$
x_1^* = 0, \; x_2^* = -1, \; S = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix},
$$

则 (x_1^*, x_2^*, S) 是原问题的可行解，故原问题的最优值为 1。

令

$$
\mathcal{X} \equiv \left\{ (x_1, x_2, S) \in \mathbb{R}^2 \times \mathcal{S}^3 \;\middle|\; x_1 \begin{pmatrix} 0 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{pmatrix} + x_2 \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & -1 \\ 0 & -1 & 0 \end{pmatrix} + S = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{pmatrix} \right\},
$$

把约束放进 X 中

s.t.

则共轭函数

$$
\begin{aligned}
f^*(y_1,y_2,T) &= \max_{(x_1,x_2,S)\in\mathcal{X}}\{x_1y_1+x_2y_2+S\bullet T+x_2\}\\
&= \max_{x_1,x_2\in\mathbb{R}}\{x_1y_1+x_2y_2-x_2t_{11}-x_1t_{22}+(1+x_2)t_{23}+(1+x_2)t_{32}+x_2\}\\
&= \max_{x_1,x_2\in\mathbb{R}}\{(y_1-t_{22})x_1+(y_2-t_{11}+t_{23}+t_{32}+1)x_2+t_{23}+t_{32}\}\\
&= \begin{cases} t_{23}+t_{32} & y_1-t_{22}=0,\ y_2-t_{11}+t_{23}+t_{32}+1=0\\ +\infty & \text{其它}\end{cases}
\end{aligned}
$$

于是共轭对偶问题是

$$
\begin{aligned}
\max\quad & -t_{23}-t_{32}\\
\text{s.t.}\quad & y_1-t_{22}=0\\
& y_2-t_{11}+t_{23}+t_{32}=-1\\
& (y_1,y_2,T)\in\{0\}^2\times\mathcal{S}_+^2
\end{aligned}
$$

等价于

$$
\begin{aligned}
\max\quad & -t_{23}-t_{32}\\
\text{s.t.}\quad & t_{22}=0\\
& -t_{11}+t_{23}+t_{32}=-1\\
& T\in\mathcal{S}_+^2
\end{aligned}
$$

由 $t_{22}=0$ 可知必有 $t_{12}=t_{21}$ $t_{23}=t_{32}=0$，此时的目标函数值为 0。对偶问题显然可行，例如

$$
\begin{pmatrix}1&0&0\\0&0&0\\0&0&0\end{pmatrix},
$$

故对偶问题的最优值为 0。由 $1>0$，原始对偶问题不具有强对偶性。

4.11

原始问题 (4.29)、对偶问题 (4.30) 分别为：

$$
\begin{aligned}
\min\quad & c\bullet x\\
s.t.\quad & a^i\bullet x\geq b_i, i=1,2,\cdots,m.\\
& x\in\mathcal{K}.
\end{aligned}
\qquad
\begin{aligned}
\max\quad & b^Ty\\
s.t.\quad & \sum_{i=1}^m y_ia^i+s=c.\\
& s\in\mathcal{K}^*, y\in\mathbb{R}_+^m.
\end{aligned}
$$

弱对偶定理：若问题 (4.29) 和 (4.30) 都是可行的，则对问题 (4.29) 的任何可行解 x 与问题 (4.30) 的任何可行解 (y,s)，都有 $c\bullet x\geq b^Ty$ 成立。

证明： 由于

$$
c\bullet x-b^Ty=c\bullet x-\sum_{i=1}^m b_iy_i\geq c\bullet x-\sum_{i=1}^m y_ia^i\bullet x=s\bullet x\geq 0
$$

得证。 □

最优性定理：(1) 若问题 (4.29) 存在一个可行解 x^*，问题 (4.30) 存在一个可行解 (y^*, s^*) 且使得 $c \bullet x^* = b^T y^*$，则 x^* 为 (4.29) 的最优解，(y^*, s^*) 为 (4.29) 的最优解。

(2) 若问题 (4.29) 中存在 $x_0 \in \mathbb{E}$ 满足: $a^i \bullet x_0 > b_i, i = 1, 2, \cdots, m$, $x_0 \in ri(\mathcal{K})$ 且该问题下有界, 则该问题的一个可行解 x^* 为最优解的必要条件为问题 (4.30) 存在一个可行解 (y^*, s^*), 使得 $c \bullet x^* = b^T y^*$。

(3) 若问题 (4.30) 的目标值上有界, 存在 $s_0 \in ri(\mathcal{K}^*)$ 和 $y_0 \in \mathbb{R}^m_{++}$ 满足:, 则该问题的一个可行解 (y^*, s^*) 为最优解的必要条件为问题 (4.29) 存在一个可行解 x^*, 使得 $c \bullet x^* = b^T y^*$。

证明：(1) 显然成立。

(2) 根据定理 4.29 可知，存在问题 (4.30) 的可行解 (y^*, s^*) 使得 $b^T y^*$ 等于问题 (4.29) 的最优值，即 $c \bullet x^* = b^T y^*$ 成立。反之，显然成立。

(3) 根据定理 4.29 可知，存在问题 (4.29) 的可行解 x^* 使得 $c \bullet x^*$ 等于问题 (4.30) 的最优值，即 $c \bullet x^* = b^T y^*$ 成立。反之，显然成立。

4.12

(1) 证明:

$\forall \lambda, \eta \in \mathcal{Y}, t \in [0,1] \Rightarrow Q_0 + \sum_{i=1}^m \lambda_i Q_i \in \mathcal{S}^n_{++}, Q_0 + \sum_{i=1}^m \eta_i Q_i \in \mathcal{S}^n_{++}$，由 $\mathcal{S}^n_{++}$ 为凸集，于是有 $Q_0 + \sum_{i=1}^m (t\lambda_i + (1-t)\eta_i) Q_i = t(Q_0 + \sum_{i=1}^m \lambda_i Q_i) + (1-t)(Q_0 + \sum_{i=1}^m \eta_i Q_i) \in \mathcal{S}^n_{++}$，又由于 $t\lambda + (1-t)\eta_i \in \mathbb{R}^m_{++}$，可知 $t\lambda + (1-t)\eta_i \in \mathcal{Y}$，于是有 $\mathcal{Y}$ 为凸集。

另一方面，由 $\mathcal{S}^n_{++}$ 为开集，于是 $\forall \lambda \in \mathcal{Y} \Rightarrow Q_0 + \sum_{i=1}^m \lambda_i Q_i \in \mathcal{S}^n_{++} \Rightarrow \exists\, \varepsilon > 0,\ s.t.\ Q_0 + \sum_{i=1}^m \lambda_i Q_i + A \in \mathcal{S}^n_{++}, \forall \|A\| < \varepsilon$。取 $M = \max\{\|Q_i\|; 1 \leqslant i \leqslant m\}$,此时不妨假设 $M > 0$,否则的话,所有的 Q_i 均为 0 矩阵,则此时一定有 $Q_0 \in \mathcal{S}^n_{++}$,于是 $\mathcal{Y} = \mathbb{R}^m_{++}$,显然为开集。故 $M > 0$ 时由 $\forall \|\alpha\| < \varepsilon / mM$ 有 $\|\sum_{i=1}^m \alpha_i Q_i\| \leqslant \sum_{i=1}^m |\alpha_i| \|Q_i\| < \varepsilon$ 于是有 $Q_0 + \sum_{i=1}^m (\lambda_i + \alpha_i) Q_i = Q_0 + \sum_{i=1}^m \lambda_i Q_i + \sum_{i=1}^m \alpha_i Q_i \in \mathcal{S}^n_{++} \Rightarrow \lambda + \alpha \in \mathcal{Y} \Rightarrow \mathcal{Y}$ 为开集。

于是有 $\mathcal{Y}$ 为开凸集。

另证:

由于 $\lambda \mapsto Q_0 + \sum_{i=1}^m \lambda_i Q_i$ 为仿射函数，$\mathcal{S}^n_{++}$ 为凸开集，因此其原像 $\mathcal{Y}' = \{\lambda \in \mathbb{R}^m : Q_0 + \sum_{i=1}^m \lambda_i Q_i \in \mathcal{S}^n_{++}\}$ 为凸开集；又由于 $\mathbb{R}^m_{++}$ 为凸开集，因而 $\mathcal{Y} = \mathcal{Y}' \cap \mathbb{R}^m_{++}$ 为凸开集。

□

(2) 证明: 考虑原问题的拉格朗日函数

$$L(x, \lambda) = \frac{1}{2} x^T (Q_0 + \sum_{i=1}^m \lambda_i Q_i) x + (q^0 + \sum_{i=1}^m \lambda_i q^i)^T x + c_0 + \sum_{i=1}^m \lambda_i c_i.$$

当 $\lambda \in \mathcal{Y}$ 时，$L(x, \lambda)$ 关于 x 为凸函数，$\min_x L(x, \lambda)$ 在 $\dfrac{\partial}{\partial x} L(x, \lambda) = 0$ 处取得，容易解得此时 $x^*(\lambda) = -(Q_0 + \sum_{i=1}^m \lambda_i Q_0)^{-1} (q^0 + \sum_{i=1}^m \lambda_i q^i)$，于是

$$\min_x L(x, \lambda) = L(x^*, \lambda) = -\frac{1}{2} (q^0 + \sum_{i=1}^m \lambda_i q^i)^T (Q_0 + \sum_{i=1}^m \lambda_i Q_i)^{-1} (q^0 + \sum_{i=1}^m \lambda_i q^i) + c_0 + \sum_{i=1}^m \lambda_i c_i.$$

这恰好是 $P(\lambda)$ 的表达式。另一方面，注意到

$$\begin{aligned} \frac{\partial}{\partial \lambda_j} P(\lambda) = &\frac{1}{2} (q^0 + \sum_{i=1}^m \lambda_i q^i)^T (Q_0 + \sum_{i=1}^m \lambda_i Q_i)^{-1} Q_j (Q_0 + \sum_{i=1}^m \lambda_i Q_i)^{-1} (q^0 + \sum_{i=1}^m \lambda_i q^i) \\ &- q_j^T (Q_0 + \sum_{i=1}^m \lambda_i Q_i)^{-1} (q^0 + \sum_{i=1}^m \lambda_i q^i) + c_j \end{aligned}$$

由于已知对偶问题的最优解 $\bar\lambda$ 在 $\mathcal{Y}$ 上可达，可知 $\bar\lambda$ 是 $\mathcal{Y}$ 的内点，所以必有 $\nabla P(\lambda)\big|_{\lambda=\bar\lambda}=0$。于是 $\frac{1}{2}x^*(\bar\lambda)^TQ_jx^*(\bar\lambda)+q_j^Tx^*(\bar\lambda)+c_j=0, j=1,\ldots,m$，于是在 $(x^*(\bar\lambda),\bar\lambda)$ 处互补松弛条件成立，由拉格朗日对偶原理，可知原问题的全局最优值在 $x^*(\bar\lambda)$ 处取到，$x^*(\bar\lambda)$ 即为题目中的 $\bar x$。 □

(3)

充分条件为 $Q_0+\sum_{i=1}^m\overline{\lambda_i}Q_i$ 可逆。则对于 $\overline{\lambda_i}\neq 0$，有 $0=\frac{\partial}{\partial\lambda_i}P(\lambda)=g_i(\overline{x})$；对于 $\overline{\lambda_i}=0$，有 $0\geq\frac{\partial}{\partial\lambda_i}P(\lambda)=g_i(\overline{x})$。于是，对任何的可行解 x，有 $f(\overline{x})=f(\overline{x})+\sum_{i=1}^m\overline{\lambda_i}g_i(\overline{x})=L(\overline{x},\overline{\lambda})\leq L(x,\overline{\lambda})=f(x)+\sum_{i=1}^m\overline{\lambda_i}g_i(x)\leq f(x)$。

定理 4.21 设原问题 (4.16) 和共轭对偶问题 (4.17) 都是可行的. 当 $x\in\mathcal{X}\cap\mathcal{K}$ 和 $y\in\mathcal{Y}\cap\mathcal{K}^*$, 则有

$$0\leq x\bullet y\leq f(x)+f^*(y)$$

且 $f(\bar x)+f^*(\bar y)=0$ 的充分必要条件是

强对偶 充要条件

$$\bar x\bullet\bar y=0\ \text{且}\ \bar y\in\partial f(\bar x).$$

当上等式成立时, $\bar x$ 和 $\bar y$ 分别为原问题和共轭对偶问题的最优解.

对偶的结论？

min $f(x)$
s.t. $x\in\mathcal{X}\cap\mathcal{K}$

定理 4.23 (*Fenchel* 定理/强对偶定理) 对于原优化问题 (4.16), 假设 $\mathcal{X}$ 为非空闭凸集和 $\mathcal{K}$ 为非空闭凸锥, $f:\mathcal{X}$ 为凸连续函数.

min $f^*(y)$
$y\in\mathcal{Y}\cap\mathcal{K}^*$

当 (4.16) 下有界且 $\mathrm{ri}(\mathcal{K})\cap\mathrm{ri}(\mathcal{X})\neq\emptyset$, 则共轭对偶问题 (4.17) 最优解可达且与原优化问题强对偶.

对称地当 $\mathcal{Y}$ 为非空闭凸集, $f^*:\mathcal{Y}$ 为凸连续函数, 共轭对偶问题下有界且 $\mathrm{ri}(\mathcal{K}^*)\cap\mathrm{ri}(\mathcal{Y})\neq\emptyset$ 时, 则原优化问题的最优解可达且与共轭对偶问题强对偶.

有界且内点集非空

以上两个定理表明, 优化问题 (4.17) 提供给问题 (4.16) 一个下界, 因此称它们为对偶是合理的. 同时定理4.21给出了判断优化问题 (4.16) 和 (4.17) 具有强对偶 $v_p + v_d = 0$ 的充分必要条件.

上面两个定理都在得到原问题和共轭对偶问题可行解的假设下, 判定是否具有强对偶条件. 我们同样关注, 在没有对优化问题 (4.16) 求解之前, 如何根据问题的定义域集合 $\mathcal{X}$ 和约束集合 $\mathcal{K}$ 所具有的特性, 得到具有强对偶的结论?

min $f(x)$
s.t. $x \in \mathcal{X} \cap K$.

定理 4.23 (*Fenchel 定理/强对偶定理*) *对于原优化问题 (4.16), 假设 $\mathcal{X}$ 为非空闭凸集和 $\mathcal{K}$ 为非空闭凸锥, $f : \mathcal{X}$ 为凸连续函数.*

min $f^*(y)$
$y \in \mathcal{Y} \cap K^*$

当 (4.16) 下有界且 $\mathrm{ri}(\mathcal{K}) \cap \mathrm{ri}(\mathcal{X}) \neq \emptyset$, 则共轭对偶问题 (4.17) 最优解可达且与原优化问题强对偶.

对称地当 $\mathcal{Y}$ 为非空闭凸集, $f^* : \mathcal{Y}$ 为凸连续函数, 共轭对偶问题下有界且 $\mathrm{ri}(\mathcal{K}^*) \cap \mathrm{ri}(\mathcal{Y}) \neq \emptyset$ 时, 则原优化问题的最优解可达且与共轭对偶问题强对偶.

证明: 由定理的条件, 再根据定理3.11的结论, 两组结论中只需证明一组即可, 另一组结论证明类似.

首先由引理3.8知 $f^* : \mathcal{Y}$ 存在. 由定理4.21, 有

$$f(x) + f^*(y) \geq x \bullet y \geq 0, \forall\, x \in \mathcal{X} \cap \mathcal{K}, \forall y \in \mathcal{Y} \cap \mathcal{K}^*,$$

$$f^*(y) \geq -f(x), \forall\, x \in \mathcal{X} \cap \mathcal{K}, \forall y \in \mathcal{Y} \cap \mathcal{K}^*.$$

因 (4.16) 下有界, 记其下确界为 v_p, 有

$$f^*(y) \geq -v_p, \forall y \in \mathcal{Y} \cap \mathcal{K}^*. \tag{4.18}$$

记

$$\mathcal{C} = \left\{ \begin{pmatrix} x \\ \mu \end{pmatrix} \mid x \in \mathcal{X}, \mu \in \mathbb{R}, f(x) \leq \mu \right\},$$

$$\mathcal{D} = \left\{ \begin{pmatrix} x \\ \mu \end{pmatrix} \mid x \in \mathcal{K}, \mu \in \mathbb{R}, v_p \geq \mu \right\}.$$

由假设条件得到 $\mathcal{C}$ 和 $\mathcal{D}$ 为非空闭凸集. 由定理3.4,

$$\mathrm{ri}(\mathcal{C}) = \left\{ \begin{pmatrix} x \\ \mu \end{pmatrix} \mid x \in \mathrm{ri}(\mathcal{X}), \mu \in \mathbb{R}, f(x) < \mu \right\},$$