第四章习题

4.1A-7

(a) $\begin{bmatrix} -2 & -2 \\ -2 & -3 \end{bmatrix} \prec 0$, 负定。 (b) $\begin{bmatrix} 2 & 1 & 0 \\ 1 & 2 & 1 \\ 0 & 1 & 3 \end{bmatrix} \succ 0$， 正定。

4.1B-5

(a) $\nabla f(x) = \begin{bmatrix} 2x_1 + x_2 - 6 \\ x_1 - 8 \end{bmatrix}$; $\nabla f(x_0) = 0$; 是平稳点;

$\nabla^2 f(x) = \begin{bmatrix} 2 & 1 \\ 1 & 0 \end{bmatrix}$ 不定；平稳点是鞍点。

(b) $\nabla f(x) = \begin{bmatrix} 20x_1 \\ \frac{12}{x_2} \end{bmatrix}$; $\nabla f(x_0) = \begin{bmatrix} 20 \\ 6 \end{bmatrix}$; 不是平稳点；$d = \begin{bmatrix} 20 \\ 6 \end{bmatrix}$.

4.1B-7（参考例4.1）

$\nabla f(x) = c + Dx$, 令 $\nabla f(x) = 0$, 得到平稳点 $x_0 = -D^{-1}c$. 由于 $\nabla^2 f(x) = D < 0$, 所以 x_0 为全局极小值点。最优目标函数值为

$$\begin{aligned} f(x_0) &= c^T x_0 + \frac{1}{2} x_0^T D x_0 \\ &= -c^T D^{-1} c + \frac{1}{2} c^T (D^{-1})^{\mathsf{T}} D D^{-1} c \\ &= -\frac{1}{2} c^{\mathsf{T}} D^{-1} c. \end{aligned}$$

4.2A-2（参考讲义90页）

$g(x)$ 在点 x 处的雅可比矩阵为 $J(x) = I_n$, I_n 表示 $n \times n$ 的单位矩阵。

4.2B-1

(a) （参考例4.6）
定义拉格朗日函数

$$L(x, \lambda) = x_1^2 + 2x_2 + 10x_3^2 - \lambda_1(x_1 - x_2^2 + x_3 - 5) - \lambda_2(x_1 + 5x_2 + x_3 - 11/4).$$

求梯度，并令梯度等于零，

$$\begin{aligned}
\frac{\partial L}{\partial x_1} &= 2x_1 - \lambda_1 - \lambda_2 = 0 \\
\frac{\partial L}{\partial x_2} &= 2 + 2\lambda_1 x_2 - 5\lambda_2 = 0 \\
\frac{\partial L}{\partial x_3} &= 20x_3 - \lambda_1 - \lambda_2 = 0 \\
\frac{\partial L}{\partial \lambda_1} &= -(x_1 - x_2^2 + x_3 - 5) = 0 \\
\frac{\partial L}{\partial \lambda_2} &= -(x_1 + 5x_2 + x_3 - 11/4) = 0.
\end{aligned}$$

有两组解 $(x_1, x_2, x_3, \lambda_1, \lambda_2)$ 分别取值:

- $(x^*, \lambda^*) = (105/22,\ -1/2,\ 21/44,\ 503/44,\ -83/44)$.
- $(\bar{x}^*, \bar{\lambda}^*) = (505/22,\ -9/2,\ 101/44,\ -2503/44,\ 4523/44)$.

验证海森矩阵（同时考虑定理4.5），对于第一组解，

$$\nabla^2 L(x, \lambda^*) = \begin{bmatrix} 2 & 0 & 0 \\ 0 & 2\lambda_1^* & 0 \\ 0 & 0 & 20 \end{bmatrix} > 0,$$

则定理 4.5 中的（2）成立，所以 $(x^*, \lambda^*) = (105/22,\ -1/2,\ 21/44,\ 503/44,\ -83/44)$ 为严格极小值点。

对于第二组解，

$$\nabla^2 L(x, \bar{\lambda}^*) = \begin{bmatrix} 2 & 0 & 0 \\ 0 & 2\bar{\lambda}_1^* & 0 \\ 0 & 0 & 20 \end{bmatrix} \text{（不定）}.$$

计算约束方程的梯度，

$$\nabla g_1(\bar{x}^*) = \begin{bmatrix} 1 \\ 2x_2 \\ 1 \end{bmatrix} = \begin{bmatrix} 1 \\ -9 \\ 1 \end{bmatrix},\ \nabla g_2(\bar{x}^*) = \begin{bmatrix} 1 \\ 5 \\ 1 \end{bmatrix}.$$

记定理4.5中的 d 为 $d = (d_1, d_2, d_3)^\mathsf{T}$. 则集合 D 中的 d 应该满足

$$\begin{bmatrix} 1 & -9 & 1 \end{bmatrix} \begin{bmatrix} d_1 \\ d_2 \\ d_3 \end{bmatrix} = 0, \quad \begin{bmatrix} 1 & 5 & 1 \end{bmatrix} \begin{bmatrix} d_1 \\ d_2 \\ d_3 \end{bmatrix} = 0,$$

即 $d \in D$, 当且仅当 d 具有形式： $d = (d_1, 0, -d_1)^\mathsf{T}$。验证

$$d^\mathsf{T} \nabla^2 L(x, \bar{\lambda}^*) d > 0,\ d_1 \neq 0.$$

所以 $(\bar{x}^*, \bar{\lambda}^*) = (505/22,\ -9/2,\ 101/44,\ -2503/44,\ 4523/44)$ 也是严格极小值点。

比较 $f(x^*) < f(\bar{x}^*)$, 最优解为 $(x^*) = (105/22,\ -1/2,\ 21/44)$.

(b) （参考讲义96页）

$$\Delta f = \lambda^* \Delta b = \begin{bmatrix} 503/44 & -83/44 \end{bmatrix} \begin{bmatrix} 0.01 \\ 0.02 \end{bmatrix} = 337/4400.$$

4.2B-3 （与4.2B-1方法一样）

(a) 平稳点 $(x^*, \lambda^*) = (3,\ 5/2,\ 1/2)$， 是极小值点。

(b) 平稳点 $(x^*, \lambda^*) = (12,\ -4,\ 80)$， 是极大值点。

4.2C-3 （参考讲义98页）

对于极大值问题，分别令 $L(x, s, \lambda)$ 关于 x, s, λ 的偏导数为零，并分析朗格朗日乘子的限制条件，得到 KKT条件为

1. $g_i(x) \leq 0$ $(i = 1, 2, \ldots, r)$; $g_i(x) \geq 0$ $(i = r+1, r+2, \ldots, p)$; $g_i(x) = 0$ $(i = p+1, p+2, \ldots, m)$.
2. $\nabla f(x) - \sum_{i=1}^{m} \nabla g_i(x) = 0.$
3. $\lambda_i g_i(x) = 0,\ i = 1, 2, \ldots, m.$
4. $\lambda_i \geq 0$ $(i = 1, 2, \ldots, r)$; $\lambda_i \leq 0$ $(i = r+1, r+2, \ldots, p)$.

对于极小值问题，

1. $g_i(x) \leq 0$ $(i = 1, 2, \ldots, r)$; $g_i(x) \geq 0$ $(i = r+1, r+2, \ldots, p)$; $g_i(x) = 0$ $(i = p+1, p+2, \ldots, m)$.
2. $\nabla f(x) - \sum_{i=1}^{m} \nabla g_i(x) = 0.$
3. $\lambda_i g_i(x) = 0,\ i = 1, 2, \ldots, m.$
4. $\lambda_i \leq 0$ $(i = 1, 2, \ldots, r)$; $\lambda_i \geq 0$ $(i = r+1, r+2, \ldots, p)$.

4.2C-5(a) （参考4.2C-3）

提示（不要漏了非负约束 x_1 ,x_2, $x_3 \geq 0$）

4.2C-9

(a) KKT条件

1. $g_1(x) = 3x_1 + 2x_2 - 8 = 0$, $g_2(x) = x_1 \geq 0$, $g_3(x) = x_2 \geq 0$.

2. $30x_1 - 3\lambda_1 - \lambda_2 = 0$, $8x_2 - 2\lambda_1 - \lambda_3 = 0$.

3. $\lambda_i g_i(x) = 0$, $i = 1, 2, 3$.

4. $\lambda_2 \geq 0$, $\lambda_3 \geq 0$.

(b) 验证 $x_0 = (0, 4)$ 不满足 KKT 条件， 所以不是最优解。

(c) 在点 $x_0 = (0, 4)$ 处起作用的约束为 $g_1(x)$ 和 $g_2(x)$. 则需要验证在点 x_0 处， 方向为 $d = (2, -3)$ 上的小邻域内， 约束 $g_1(x)$ 和 $g_2(x)$ 仍满足。 则

$$\nabla g_1(x)7 = \begin{bmatrix} 3 \\ 2 \end{bmatrix}, \quad \nabla g_2(x) = \begin{bmatrix} 1 \\ 0 \end{bmatrix},$$
$$g_1(x_1 + d) = g_1(x_0) + \nabla g_1(x)^{\mathsf{T}} d = 0,$$
$$g_2(x_1 + d) = g_2(x_0) + \nabla g_2(x)^{\mathsf{T}} d > 0,$$

起作用的约束仍满足，所以是可行方向。 又 $\nabla f(x)\Big|_{x_0} \times d = -96 < 0$, 则是下降方向。

(d) 用试探法，求解 KKT 条件得 $x^* = (1, 5/2)$, $\lambda^* = (10, 0, 0)$.

定义 $L_1(x) = L(x, \lambda^*) = 15x_1^2 + 4x_2^2 - 10(3x_1 + 2x_2 - 8)$, 则

$$\nabla^2 L_1(x) = \begin{bmatrix} 30 & 0 \\ 0 & 8 \end{bmatrix} > 0$$

所以 $x^* = (1, 5/2)$ 是唯一的严格极小值点，则是最优解。

4.3A-2

对于任意不同的两个列向量 $x_1, x_2 \in S$, 则有， $x_1 = Ay_1$, $x_2 = Ay_2$, $y_1, y_2 \in \mathbb{C}$. 所以，

$$\begin{aligned} x &= \alpha x_1 + (1-\alpha)x_2 \\ &= \alpha A y_1 + (1-\alpha) A y_2 \\ &= A(\alpha y_1 + (1-\alpha) y_2) \end{aligned}$$

由于 $y_1, y_2 \in \mathbb{C}$, $\mathbb{C}$ 是凸集， 所以 $(\alpha y_1 + (1-\alpha)y_2) \in \mathbb{C}$. 所以 $x = \alpha x_1 + (1-\alpha)x_2 \in S$. 所以 S 是凸集。

4.3A-5

假设： $x_0 = \begin{bmatrix} \bar{x}_0 \\ 0_{n-m} \end{bmatrix}$ 为基解 $\Longrightarrow$ x_0 不是 S 的极点。

记 $A = [A_1 \ A_2]$, $A_1 \in \mathbb{R}^{m \times m}$, $A_2 \in \mathbb{R}^{m \times (n-m)}$. 由 x_0 为基解，则可得

$$A x_0 = [A_1 \ A_2] \begin{bmatrix} \bar{x}_0 \\ 0_{n-m} \end{bmatrix} = A_1 \bar{x}_0 = b, \ \ \bar{x}_0 \text{ 存在且唯一.}$$

x_0 不是 S 的极点，则存在 $x_1, x_2 \in S$, 使得

$$\begin{aligned}
&x_0 = \alpha x_1 + (1-\alpha)x_2 \\
\Rightarrow &\begin{bmatrix} \bar{x}_0 \\ 0_{n-m} \end{bmatrix} = \alpha \begin{bmatrix} \bar{x}_1 \\ \hat{x}_1 \end{bmatrix} + (1-\alpha) \begin{bmatrix} \bar{x}_2 \\ \hat{x}_2 \end{bmatrix} \\
\Rightarrow &x_0 = \alpha \bar{x}_1 + (1-\alpha)\bar{x}_2, \quad 0 = \alpha \hat{x}_1 + (1-\alpha)\hat{x}_2.
\end{aligned}$$

记 $x_1 = \begin{bmatrix} \bar{x}_1 \\ \hat{x}_1 \end{bmatrix}$, $x_2 = \begin{bmatrix} \bar{x}_2 \\ \hat{x}_2 \end{bmatrix}$, $\bar{x}_1,\ \bar{x}_2 \in \mathbb{R}^m$, $\hat{x}_1,\ \hat{x}_2 \in \mathbb{R}^{n-m}$. 则有

$$\begin{aligned}
&A(\alpha x_1 + (1-\alpha)x_2) = b \\
\Rightarrow &\alpha [A_1\ A_2] \begin{bmatrix} \bar{x}_1 \\ \hat{x}_1 \end{bmatrix} + (1-\alpha)[A_1\ A_2] \begin{bmatrix} \bar{x}_2 \\ \hat{x}_2 \end{bmatrix} = b \\
\Rightarrow &A_1(\alpha \bar{x}_1 + (1-\alpha)\bar{x}_2) + A_2(\alpha \hat{x}_1 + (1-\alpha)\hat{x}_2) = b \\
\Rightarrow &A_1(\alpha \bar{x}_1 + (1-\alpha)\bar{x}_2) = b,
\end{aligned}$$

所以 $\bar{x}_0$ 不唯一，矛盾。 所以假设的命题不成立。 所以，基解 $\Longrightarrow$ 极点。

4.3B-4（参考讲义中定理4.11的证明，这里只证明充分性）

由 $f(x)$ 为严格凸函数，有

$$f(x+tu) > f(x) + t\nabla f(x)^\mathsf{T} u.$$

由泰勒公式，

$$f(x+tu) = f(x) + t\nabla f(x)^\mathsf{T} u + \frac{t^2}{2} u^\mathsf{T} \nabla^2 f(x) u + o(||tu||^2),$$

由于 $\nabla^3 f(x) = 0$, 则 $o(||tu||^2) = 0$. 对比以上式子得

$$\frac{t^2}{2} u^\mathsf{T} \nabla^2 f(x) u > 0, \text{ 对于任意的非零 } u \text{ 成立.}$$

则 $\nabla^2 f(x) = D > 0$.

4.3B-5

令 $\alpha_2 x_2 + \cdots + \alpha_k x_k = (1-\alpha_1)y$. 由 $f(x)$ 是凸函数，得

$$\begin{aligned}
f(\alpha_1 x_1 + \cdots + \alpha_k x_k) &\le \alpha_1 f(x_1) + (1-\alpha_1) f(y) \\
&= \alpha_1 f(x_1) + (1-\alpha_1) f(\frac{\alpha_2 x_2 + \cdots + \alpha_k x_k}{1-\alpha_1})
\end{aligned}$$

假设

$$f(\alpha_1 x_1 + \cdots + \alpha_k x_k) \le \alpha_1 f(x_1) + \cdots + \alpha_m f(x_m) + (1-\alpha_1-\cdots-\alpha_m) f(\frac{\alpha_{m+1} x_{m+1} + \cdots + \alpha_k x_k}{1-\alpha_1-\cdots-\alpha_m})$$

令

$$\frac{\alpha_{m+2}x_{m+2}+\cdots+\alpha_k x_k}{1-\alpha_1-\cdots-\alpha_m}=(1-\frac{\alpha_{m+1}}{1-\alpha_1-\cdots-\alpha_m})\bar{y}$$

则

$$f(\frac{\alpha_{m+1}x_{m+1}+\cdots+\alpha_k x_k}{1-\alpha_1-\cdots-\alpha_m})\leq\frac{\alpha_{m+1}}{1-\alpha_1-\cdots-\alpha_m}f(x_{m+1})+(1-\frac{\alpha_{m+1}}{1-\alpha_1-\cdots-\alpha_m})f(\bar{y}).$$

$$f(\bar{y})=f(\frac{\alpha_{m+1}x_{m+2}+\cdots+\alpha_k x_k}{1-\alpha_1-\cdots-\alpha_{m+1}}).$$

所以

$$f(\alpha_1x_1+\cdots+\alpha_kx_k)\leq\alpha_1f(x_1)+\cdots+\alpha_{m+1}f(x_{m+1})+(1-\alpha_1-\cdots-\alpha_{m+1})f(\frac{\alpha_{m+2}x_{m+2}+\cdots+\alpha_kx_k}{1-\alpha_1-\cdots-\alpha_{m+1}})$$

成立。

令 $m+1=k$, 并利用 $\alpha_1+\cdots+\alpha_k=1$. 则得证。

4.3C-1(a) （参考讲义112页以及例4.15）

4.3C-2(a)

令 $x_2^2=\bar{x}_2$. 则线性规划问题为

$$\begin{aligned}\max f(x)&=x_1^3-\bar{x}_2+x_1x_3^2\\ \text{s.t.}\quad x_1+\bar{x}_2+x_3&=5\\ 5x_1^2-\bar{x}_2-x_3&\geq 2\\ x_1,\bar{x}_2,x_3&\geq 0.\end{aligned}$$