第四章作业

4.1B-6

（a）$\nabla \mathrm{f}(\mathrm{x}) = \binom{2x_1+x_2-6}{x_1-8}$,

$\nabla \mathrm{f}(x_0) = \binom{0}{0}$,

x_0是平稳点。

（b） $\nabla \mathrm{f}(\mathrm{x}) = \binom{20x_1}{12/x_2}$,

$\nabla \mathrm{f}(x_0) = \binom{20}{6}$,

x_0不是平稳点。

极大化问题能改善解的移动方向可取梯度方向：

$\mathrm{d} = (20,6)$。 注释：从目标函数在该点的泰勒展开可得到。

（c，d）类似。

4.2B-1

题目修改：去掉变量 x_3 和 x_4 ，改为不等式约束问题。

（a） 求 KKT 点。 （b）求最优解。

解： $\mathrm{f}(\mathrm{x}) = 5x_1 + 3x_2$

$g_1(\mathrm{x}) = x_1 + 2x_2 - 6$

$g_2(\mathrm{x}) = 3x_1 + x_2 - 9$

$g_3(\mathrm{x}) = -x_1$

$g_4(\mathrm{x}) = -x_2$

（a）（定理 4.6）得：

1） $g_i(x) \le 0,\ i = 1,2,3,4$

2） $\nabla \mathrm{f}(x_0) - \sum_{i=1}^{4} \lambda_i \nabla g_i(x_0) = 0 \ \Rightarrow$

$$5 - \lambda_1 - 3\lambda_2 + \lambda_3 = 0$$

$$3 - 2\lambda_1 - \lambda_2 + \lambda_4 = 0$$

3）$\lambda_i g_i = 0$，$i = 1\ ,2\ ,3, 4$

4）$\lambda_i \ge 0\ \ i = 1$，2，3，4

分类讨论：（λ_i要么等于 0，要么大于 0）。

（1） λ_1 λ_2 同时为零， 则 λ_3 λ_4 为负，矛盾。

（2） $\lambda_1 > 0$，$\lambda_2 > 0$， 则 $g_1(\mathrm{x}) = 0,\ g_2(\mathrm{x}) = 0,$

得$x_1 = 12/5,\ x_2 = 9/5$。

则 $\lambda_3 = 0$，$\lambda_4 = 0$， $\lambda_1 = 4/5$，$\lambda_2 = 7/5$，

（3） $\lambda_1 = 0$，$\lambda_2 > 0$，$\lambda_3 > 0$，则

$x_1 = 0,\ x_2 = 9,$ 不满足第一个约束。

（4） $\lambda_1 > 0$，$\lambda_2 = 0$，则 $\lambda_4 > 0$ ，

$x_2 = 0,\ x_1 = 6$ 不满足第二个约束。

所以 KKT 点为（12/5, 9/5）。

（b）起作用的约束为 $g_1(\mathrm{x}),\ g_2(\mathrm{x})$，又$\nabla g_1(\mathrm{x}) = \binom{1}{2}$，和 $\nabla g_2(\mathrm{x}) = \binom{3}{1}$线性无关，则对于定理 7.7（补充材料），不存在非零的 d，则定理 7.7 无条件成立，所以（12/5, 9/5）为局部极大值点。

4.2B-2

解：先求拉格朗日函数的平稳点

$$L(x,\lambda) = x_1^2 + 2x_2 + 10x_3^2 - \lambda_1(x_1 + x_2^2 + x_3 - 5) - \lambda_2(x_1 + 5x_2 + x_3 - 7)$$

$$\frac{\partial L}{\partial x} = 0, \quad \frac{\partial L}{\partial \lambda} = 0$$

有两组解：

$$x_1 = -5 - \frac{25\sqrt{17}}{11},\ x_2 = \frac{5+\sqrt{17}}{2},\ x_3 = -\frac{1}{2} - \frac{5\sqrt{17}}{22}$$
$$\lambda_1 = 35.34,\ \lambda_2 = -64.08.$$
$$x_1 = -5 + \frac{25\sqrt{17}}{11},\ x_2 = \frac{5-\sqrt{17}}{2},\ x_3 = -\frac{1}{2} + \frac{5\sqrt{17}}{22}$$
$$\lambda_1 = 10.12,\ \lambda_2 = -1.37.$$

约束的梯度：

$$\nabla g_1(x) = \begin{bmatrix} 1 \\ 2x_2 \\ 1 \end{bmatrix},\ \nabla g_2(x) = \begin{bmatrix} 1 \\ 5 \\ 1 \end{bmatrix}$$

考虑定理 7.6，对于两组解，满足条件的 d 都为

$$(w, 0, -w)^{\mathsf{T}},\ w \neq 0$$

由于

$$\nabla^2 L(x) = \begin{bmatrix} 2 & 0 & 0 \\ 0 & -2\lambda_1 & 0 \\ 0 & 0 & 20 \end{bmatrix}$$

则 $d^\mathsf{T}\nabla^2 L_1(x^*)d > 0$ 对于两组解都成立，所以两组解对应的点都是严格极小值点。通过比较目标函数值的大小，得到第二组解为问题的最优解。

（b）

$$\Delta f = \lambda^* \Delta b = (10.12 \ \ -1.37)\begin{bmatrix} 0.01 \\ 0.02 \end{bmatrix} = 0.0738$$

4.2C-3

对于极大化问题，KKT 条件为：

$$\begin{aligned} g_i(x) \le 0,\ i = 1, 2, \cdots, r \\ g_i(x) \ge 0,\ i = r+1, \cdots, p \\ g_i(x) = 0,\ i = p+1, \cdots, m \\ \nabla f(x) - \sum_{i=1}^{m} \nabla g_i(x) = 0 \\ \lambda_i g_i(x) = 0,\ i = 1, \cdots, m \\ \lambda_i \ge 0,\ i = 1, 2, \cdots, r, \\ \lambda_i \le 0,\ i = r+1, \cdots, p \end{aligned}$$

对于极小化问题，KKT 条件和极大化问题前面 5 条一样，最后两条为

$$\begin{aligned} \lambda_i \le 0,\ i = 1, 2, \cdots, r, \\ \lambda_i \ge 0,\ i = r+1, \cdots, p \end{aligned}$$

4.2C-4 依据第三题的结果可以得到答案，注意不要漏掉非负性约束。

4.2C-7 （a）问题的 KKT 条件：

- $g_1(x) = 3x_1 + 2x_2 - 8 = 0$
  $g_2(x) = x_1 \ge 0$
  $g_3(x) = x_2 \ge 0$
- $30x_1 - 3\lambda_1 - \lambda_2 = 0$
  $8x_2 - 2\lambda_1 - \lambda_3 = 0$
- $\lambda_i g_i(x) = 0,\ i = 1, 2, 3$
- $\lambda_2 \ge 0,\ \lambda_3 \ge 0$

（b） 在该点(0,4)处起作用的约束为 g1(x), g2(x)。

$$\nabla g_1(x) = (3,\ 2)^{\mathsf{T}}, \quad \nabla g_2(x) = (1,\ 0)^{\mathsf{T}}$$

$$g_1(x+d) = g_1(x) + \nabla g_1(x)^{\mathsf{T}} d = 0$$

$$g_2(x+d) = g_2(x) + \nabla g_2(x)^{\mathsf{T}} d > 0$$

所以 d 为可行方向， 又

$$\nabla f(x_0)^{\mathsf{T}} d = (0\ 32)\begin{bmatrix}2\\3\end{bmatrix} < 0$$

所以 d 为下降方向。

（c） 由于 g1(x) , g2(x) 的梯度线性无关，由定理 4.6 知极值点必须满足 KKT 条件。 验证（0,4）不满足 KKT 条件，所以该点不是极值点，更不是最优解。

（d） 由定理 4.6 知， 如果起作用的约束线性无关则极值点必须是 KKT 点。
分析：第一个约束一定是起作用的约束，第二，三个约束至多有一个是起作用的约束（因为 x1，x2 不可能同时为零）。第一个约束的梯度和第二，三个约束的梯度都线性无关。
所以一定满足定理 4.6 的前提条件，所以极值点一定是 KKT 点。

（e） 略。

4.3B-4

只证明充分性，必要性参考讲义 111,112 页。$f(x)$ 为严格凸函数，则有

$$f(x+tu) > f(x) + t\nabla f^T(x)u$$

由泰勒展开，

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

, 由于 $\nabla^3 f(x) = 0$, 则 $o(||tu||^2) = 0$。由上述式对比有

$$\frac{t^2}{2}u^T\nabla^2 f(x)u > 0$$

即 $A > 0$。充分性得证。

4.3B-5

Let $\alpha_2 x_2 + \cdots + \alpha_k x_k = (1-\alpha_1)y$

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

Assume

$$
\begin{aligned}
f(\alpha_1 x_1 + \cdots + \alpha_k x_k) \le & \alpha_1 f(x_1) + \cdots + \alpha_m f(x_m) \\
& + (1-\alpha_1-\cdots-\alpha_m) f\left(\frac{\alpha_{m+1}x_{m+1} + \cdots \alpha_k x_k}{1-\alpha_1-\cdots-\alpha_m}\right)
\end{aligned}
$$

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

$$
\begin{aligned}
f(\alpha_1 x_1 + \cdots + \alpha_k x_k) \le & \alpha_1 f(x_1) + \cdots + \alpha_{m+1} f(x_{m+1}) \\
& + (1-\alpha_1-\cdots-\alpha_{m+1}) f\left(\frac{\alpha_{m+2}x_{m+2} + \cdots \alpha_k x_k}{1-\alpha_1-\cdots-\alpha_{m+1}}\right),
\end{aligned}
$$

Thus, the assumption holds. Let $m = k-1$, we complete the proof.

4.3C-1 （参考讲义 113,114 页）和例 4.14 类似。

4.3C-2 目标函数的海森为正定（目标函数求二阶偏导），所以目标函数为严格凸函数。 参考 113 页判断约束集为凸集。 所以该规划问题为凸规划，由定理 4.11 得 KKT 点即为全局极值点。（不要求求出该极值点）。