最优化理论与算法 (数值优化) －－第 1 章作业参考解答

1. (该练习的目的是提高你的建模技巧，同时熟悉利用计算机求解线性优化问题) 一个原油精练场有 8 百万桶 (barrel) 原油 A 和 5 百万桶原油 B 进行下个月的生产. 这些资源可以用来生产汽油，其售价为 38 美元 / 桶；或者生产家用加热油，其售价为 33 美元 / 桶. 这里有三种生产过程，其特征如下：

	process 1	process 2	process 3
input crude A	3	1	5
input crude B	5	1	3
output gasoline	4	1	3
onput heating oil	3	1	4
cost	\$51	\$11	\$40

所有的量均以桶为单位. 例如, 第一个过程而言, 利用 3 桶原油 A 和 5 桶原油 B 可以生产 4 桶汽油和 3 桶民用燃料油. 表格中的成本指总的成本 (即原油成本和生产过程的成本). 将此问题表述成线性规划, 其可以帮助管理者来极大化下个月的净利润.

请利用 Excel 或 Matlab 在计算机上求解此问题.

解：设下个月利用第一个过程生产 x 次，第二个过程生产 y 次，第三个过程生产 z 次，均以百万桶为单位. 则利润为

$$\begin{aligned} f(x,y,z) &= (38\times 4+33\times 3-51)x+(38+33-11)y+(38\times 3+33\times 4-40)z \\ &= 200x+60y+206z \end{aligned}$$

其数学模型为

$$\begin{aligned} \text{minimize} \quad & 200x+60y+206z \\ \text{subject to} \quad & 3x+y+5z\le 8 \\ & 5x+y+3z\le 5 \\ & x,y,z\ge 0. \end{aligned}$$

最优解是 $x=0, y=0.5, z=1.5$ 百万.

2. 利用你所能想到的方法求解下列问题

$$\begin{aligned} \texttt{minimization} \quad & (x_1-2)^2+(x_2-1)^2 \\ \texttt{Subject to} \quad & x_1^2-x_2\le 0, \\ & x_1+x_2\le 2. \end{aligned}$$

提示：可用图解法，即在平面上画出约束集及目标函数的等值线，并观察求解.

3. 设 a 为给定的 $n-$ 维向量， A 为给定的 $n\times n$ 对称矩阵. 计算函数 $f_1(x)=a^Tx$ 和 $f_2(x)=x^TAx$ 的梯度和 Hessian 阵.

解答：设 $a=(a_1,\cdots,a_n)^T,\quad A=(a_{ij})_{n\times n},\quad i,j=1,2,\cdots,n$, 则 $f_1(x)=a^Tx=a_1x+a_2x+\cdots+a_nx$

所以 $\nabla f_1(x)=(a_1,\cdots,a_n)^T=a,\quad \nabla^2 f_1(x)=\theta_{n\times n}$($n$ 阶零方阵)

$f_x(x)=x^TAx=a_{11}x_1^2+a_{22}x_2^2+\cdots+a_{nn}x_n^2+2a_{12}x_1x_2+\cdots+2a_{1n}x_1x_n+2a_{23}x_2x_3+\cdots+2a_{2n}x_2x_n+\cdots+2a_{n-1n}x_{n-1}x_n$

所以 $\frac{\partial f_2}{\partial x_1}=2a_{11}x_1+2a_{12}x_2+\cdots+2a_{1n}x_n=2(a_{11},\cdots,a_{1n})x.$

$\frac{\partial f_2}{\partial x_2}=2a_{21}x_1+2a_{22}x_2+\cdots+2a_{2n}x_n=2(a_{21},\cdots,a_{2n})x.$

$\frac{\partial f_2}{\partial x_2}=2a_{21}x_1+2a_{22}x_2+\cdots+2a_{2n}x_n=2(a_{21},\cdots,a_{2n})x.$

$\cdots$

$\frac{\partial f_2}{\partial x_n}=2a_{n1}x_1+2a_{n2}x_2+\cdots+2a_{nn}x_n=2(a_{n1},\cdots,a_{nn})x.$

所以 $\nabla f_2(x)=2Ax,\quad$ 又 $\frac{\partial^2 f_2}{\partial x_i\partial x_j}=2a_{ij}$ 所以 $\nabla^2 f_2(x)=2A$

4. 写出函数 $\cos(1/x)$ 在非零点 x 附近的二阶 Taylor 展式；写出函数 $\cos x$ 在任一点 x 附近的三阶 Taylor 展式；并针对具体值 $x=1$ 计算二阶展开？

解答： $(cos\frac{1}{x})'=-sin\frac{1}{x}(-\frac{1}{x^2})=\frac{1}{x^2}sin\frac{1}{x}$

$(cos\frac{1}{x})''=-\frac{2}{x^3}sin\frac{1}{x}+\frac{1}{x^2}[cos\frac{1}{x}(-\frac{1}{x^2})]=-\frac{2}{x^3}sin\frac{1}{x}-\frac{1}{x^4}cos\frac{1}{x}$

所以 $cos\frac{1}{x}$ 在非零点 x_0 附近的二阶 Taylor 展开式为：

$(cos\frac{1}{x})=cos\frac{1}{x_0}+\frac{1}{x^2}sin\frac{1}{x_0}(x-x_0)-\frac{1}{2}(\frac{2}{x_0^3}sin\frac{1}{x_0}+\frac{1}{x_0^4}cos\frac{1}{x_0})(x-x_0)^2$

所以 $cosx$ 在任意一点 x_0 附近的三阶 Taylor 展开式为：

$cosx=cosx_0-sinx_0(x-x_0)-\frac{1}{2}cosx_0(x-x_0)^2+\frac{1}{6}sinx_0(x-x_0)^3.$

最优化理论与算法 (数值优化) – – 第 2 章作业参考解答

1. 第 2 章 2.2　将下面的问题转化成标准形

$$\begin{aligned} \text{minimize} \quad & |x|+|y|+|z| \\ \text{subject to} \quad & x+y\le 1, \\ & 2x+z=3. \end{aligned}$$

解：令 $x^+ = \max\{x,0\}$,　$x^- = \max\{-x,0\}$, 则 $x^+ \ge 0$,　$x^- \ge 0$, 且 $x^+\cdot x^- = 0$,　$x = x^+ - x^-$,　$|x| = x^+ + x^-$. 由此可将原问题化为

$$\begin{aligned} \text{minimize} \quad & x^+ + x^- + y^+ + y^- + z^+ + z^- \\ \text{subject to} \quad & x^+ - x^- + y^+ - y^- + s = 1, \\ & 2x^+ - 2x^- + z^+ - z^- = 3, \\ & x^+, x^-, y^+, y^-, z^+, z^-, s \ge 0, \\ & x^+x^- = 0, y^+y^- = 0, y^+y^- = 0. \end{aligned}$$

如用单纯形法解此问题，则互补条件自动满足，从而此时便可化为线性规划的标准形式

$$\begin{aligned} \text{minimize} \quad & x^+ + x^- + y^+ + y^- + z^+ + z^- \\ \text{subject to} \quad & x^+ - x^- + y^+ - y^- + s = 1, \\ & 2x^+ - 2x^- + z^+ - z^- = 3, \\ & x^+, x^-, y^+, y^-, z^+, z^-, s \ge 0. \end{aligned}$$

2. 第 2 章 2.3　一类逐段线性函数能够表示为 $f(x) = \texttt{Maximum}(c_1^Tx + d_1, c_2^Tx + d_2, \cdots, c_p^Tx + d_p)$. 针对这样的函数，考虑问题

$$\begin{aligned} \text{minimize} \quad & f(x) \\ \text{subject to} \quad & Ax = b, \\ & x \ge 0. \end{aligned}$$

说明如何将此问题转化成线性规划问题.

3. 第 2 章 2.4　给出

$$\begin{aligned} & x_1 + \tfrac{8}{3}x_2 \le 4, \\ & x_1 + x_2 \le 2, \\ & 2x_1 \le 3, \\ & x_1 \ge 0, x_2 \ge 0. \end{aligned}$$

所对应标准形系统的所有基本解.

解：引入松弛变量 x_3, x_4, x_5, 将原问题化为标准形

$$\begin{cases} x_1 + \frac{8}{3}x_2 + x_3 = 4 \\ x_1 + x_2 + x_4 = 2 \\ 2x_1 + x_5 = 3 \\ x_i \geq 0, \quad i = 1,2,3,4,5 \end{cases}$$

新问题的系数矩阵为 $A = [\, a_1 \quad a_2 \quad a_3 \quad a_4 \quad a_5 \,] = \begin{bmatrix} 1 & \frac{8}{3} & 1 & 0 & 0 \\ 1 & 1 & 0 & 1 & 0 \\ 2 & 0 & 0 & 0 & 1 \end{bmatrix}$.

序号	基B	基本解	可行性
1	$[\, a_1 \quad a_2 \quad a_3 \,]$	$(\frac{3}{2}, \frac{1}{2}, \frac{7}{6}, 0, 0)$	可行
2	$[\, a_1 \quad a_2 \quad a_4 \,]$	$(\frac{3}{2}, \frac{15}{16}, 0, -\frac{7}{16}, 0)$	不可行
3	$[\, a_1 \quad a_2 \quad a_5 \,]$	$(\frac{3}{2}, \frac{6}{5}, 0, 0, \frac{7}{5})$	可行
4	$[\, a_1 \quad a_3 \quad a_4 \,]$	$(\frac{3}{2}, 0, \frac{5}{2}, \frac{1}{2}, 0)$	可行
5	$[\, a_1 \quad a_3 \quad a_5 \,]$	$(2, 0, 2, 0, -1)$	不可行
6	$[\, a_1 \quad a_4 \quad a_5 \,]$	$(4, 0, 0, -2, -5)$	不可行
7	$[\, a_2 \quad a_3 \quad a_5 \,]$	$(0, 2, -\frac{4}{3}, 0, 3)$	不可行
8	$[\, a_2 \quad a_4 \quad a_5 \,]$	$(0, \frac{3}{2}, 0, \frac{1}{2}, 3)$	可行
9	$[\, a_3 \quad a_4 \quad a_5 \,]$	$(0, 0, 4, 2, 3)$	可行

4. 第 2 章 2.6　考虑问题

$$\begin{aligned} \text{minimize} \quad & c_1x_1 + c_2x_2 + \cdots + c_nx_n \\ \text{subject to} \quad & a_{11}x_1 + a_{12}x_2 + \cdots + a_1nx_n \leq b_1 \\ & a_{21}x_1 + a_{22}x_2 + \cdots + a_2nx_n \leq b_2 \\ & \vdots \\ & a_{m1}x_1 + a_{m2}x_2 + \cdots + a_{mn}x_n \leq b_m \\ & x_1 \geq 0, x_2 \geq 0, \cdots, x_n \geq 0, \end{aligned}$$

这种情况下, 约束集完全由线性不等式确定. 为了将其化为标准形, 引入松弛变量 $y_1, y_2, \cdots, y_m$, 该问题可以另外表示为

$$\begin{aligned} \text{minimize} \quad & c_1x_1 + c_2x_2 + \cdots + c_nx_n \\ \text{subject to} \quad & a_{11}x_1 + a_{12}x_2 + \cdots + a_1nx_n + y_1 = b_1 \\ & a_{21}x_1 + a_{22}x_2 + \cdots + a_2nx_n + y_2 = b_2 \\ & \vdots \\ & a_{m1}x_1 + a_{m2}x_2 + \cdots + a_{mn}x_n + y_m = b_m \\ & x_1 \geq 0, x_2 \geq 0, \cdots, x_n \geq 0, \\ & y_1 \geq 0, y_2 \geq 0, \cdots, y_m \geq 0. \end{aligned}$$

通过考虑有 $n+m$ 个未知数 $x_1, x_2, \cdots, x_n, y_1, y_2, \cdots, y_m$ 的问题，原问题转化成了标准形．描述线性等式约束的 $m \times (n+m)$ 矩阵具有特殊形式 $[A, I]$(即可将列剖分成几个集合；前 n 列由原

来的矩阵 A 组成，后 m 列由 $m\times m$ 的单位矩阵组成).

上述的两个线性规划问题，一个在 E^n ，另一个在 E^{n+m}. 证明这两个问题的极点之间存在一一对应.

证明. 设原问题的可行集为 S, 化为标准形后其可行集为 $\bar{S}$, 即

$$S=\{x\in\mathbb{R}^n \mid Ax\le b, x\ge 0\},\quad \bar{S}=\{(x,y)\in\mathbb{R}^{n+m}\mid Ax+y=b, x\ge 0, y\ge 0\}.$$

首先设 x 为 S 的一个极点，令 $y=b-Ax$ ，下面证明 (x,y) 是 $\bar{S}$ 的极点. 如不然，则存在 $(x^1,y^1),\quad (x^2,y^2)\in\bar{S}$ 且 $(x^1,y^1)\ne(x^2,y^2)$ 及 $\alpha\in(0,1)$ 使得

$$(x,y)=\alpha(x^1,y^1)+(1-\alpha)(x^2,y^2). \tag{1}$$

一方面，因为 $(x^1,y^1),\quad (x^2,y^2)\in\bar{S}$ 且 $(x^1,y^1)\ne(x^2,y^2)$ ，所以 $y^1=b-Ax^1\ge 0,\quad y^2=b-Ax^2\ge 0$, 从而 $x^1,x^2\in S$ ，且 $x^1\ne x^2$(否则， $y^1=y^2$ ，这与 $(x^1,y^1)\ne(x^2,y^2)$ 矛盾); 另一方面，由 (1) 有 $x=\alpha x^1+(1-\alpha)x^2$ ，这与 x 为 S 的极点矛盾！

反之，设 (x,y) 是 S 的极点，则 x 也应该为 S 的极点. 否则，存在 $x^1,x^2\in S$, 且 $x^1\ne x^2$, 及 $\alpha\in(0,1)$, 使得

$$x=\alpha x^1+(1-\alpha)x^2. \tag{2}$$

令

$$y^1=b-Ax^1,\quad y^2=b-Ax^2, \tag{3}$$

则一方面易见 $(x^1,y^1),\quad (x^2,y^2)\in\bar{S}$ 且 $(x^1,y^1)\ne(x^2,y^2)$ ，另外，由 (2) 和 (3) 有 $(x,y)=\alpha(x^1,y^1)+(1-\alpha)(x^2,y^2)$ ，这与 (x,y) 是 $\bar{S}$ 的极点矛盾！

综上可知两个问题的极点之间存在一一对应关系.

5. 第 2 章 2.7　利用转轴求解联立方程组：

$$\begin{aligned}3x_1+2x_2&=5\\ 5x_1+x_2&=9.\end{aligned}$$

解：构造初始基得增广表格为

$$\begin{array}{cccc|c} e_1 & e_2 & a_1 & a_2 & b\\ \hline 1 & 0 & \boxed{3} & 2 & 5\\ 0 & 1 & 5 & 1 & 9\end{array}\quad\rightarrow\quad\begin{array}{cccc|c} e_1 & e_2 & a_1 & a_2 & b\\ \hline 1/3 & 0 & 1 & 2/3 & 5/3\\ -5/3 & 1 & 0 & \boxed{-7/3} & 2/3\end{array}\quad\rightarrow\quad\begin{array}{cccc|c} e_1 & e_2 & a_1 & a_2 & b\\ \hline -1/7 & 2/7 & 1 & 0 & 13/7\\ 5/7 & -3/7 & 0 & 1 & -2/7\end{array}$$

解为 $x_1=13/7,\quad x_2=-2/7$.

利用转轴求解联立方程组：

$$\begin{aligned}x_1+2x_2+x_3&=7\\ 2x_1-x_2+2x_3&=9\\ x_1+x_2+3x_3&=12.\end{aligned}$$

解：构造初始基得增广表格为

e_1	e_2	e_3	a_1	a_2	a_3	b
1	0	0	$\boxed{1}$	2	1	7
0	1	0	2	-1	2	9
0	0	1	1	1	3	12

$\rightarrow$

e_1	e_2	e_3	a_1	a_2	a_3	b
1	0	0	1	2	1	7
-2	1	0	0	$\boxed{-5}$	0	-5
-1	0	1	0	-1	2	5

$\rightarrow$

e_1	e_2	e_3	a_1	a_2	a_3	b
$1/5$	$2/5$	0	1	0	1	5
$2/5$	$-1/5$	0	0	1	0	1
$-3/5$	$-1/5$	1	0	0	$\boxed{2}$	6

$\rightarrow$

e_1	e_2	e_3	a_1	a_2	a_3	b
$1/2$	$1/2$	$-1/2$	1	0	0	2
$2/5$	$-1/5$	0	0	1	0	1
$-3/10$	$-1/10$	$1/2$	0	0	1	3

解为 $x_1 = 2, x_2 = 1, x_3 = 3$.

6. 第 2 章 2.12 将下面的问题转化成标准形并且求解

   (a) 利用单纯形过程，求解

$$\begin{aligned} \text{maximize} \quad & -x_1 + x_2 \\ \text{subject to} \quad & x_1 - x_2 \leq 2, \\ & x_1 + x_2 \leq 6, \\ & x_1 \geq 0, x_2 \geq 0. \end{aligned}$$

   (b) 画出问题在 x_1, x_2 空间的图形解释，并标明单纯形步骤的迭代路径.

   (c) 针对下面问题重复

$$\begin{aligned} \text{maximize} \quad & x_1 + x_2 \\ \text{subject to} \quad & -2x_1 + x_2 \leq 1, \\ & x_1 - x_2 \leq 1, \\ & x_1 \geq 0, x_2 \geq 0. \end{aligned}$$

解： a) 引入松弛变量 x_3, x_4 ，化为标准形

$$\begin{aligned} \text{minimize} \quad & x_1 - x_2 \\ \text{subject to} \quad & x_1 - x_2 + x_3 = 2, \\ & x_1 + x_2 + x_4 = 6, \\ & x_i \geq 0, i = 1, 2, 3, 4. \end{aligned}$$

写出初始表，其已是第一张单纯形表

	x_1	x_2	x_3	x_4	b
x_3	1	-1	1	0	2
x_4	1	$\boxed{1}$	0	1	6
$c^T(r^T)$	1	-1	0	0	0

$\rightarrow$

	x_1	x_2	x_3	x_4	b
x_3	2	0	1	1	8
x_2	1	1	0	1	6
r^T	2	0	0	1	6

因为 $r_j \geq 0$, 所以得最优解 $x^* = (0, 6)$, 最优值为 $f^* = -6$. 从而原问题最优值为 6.

b)(此处为一曲线图) 单纯形步骤的迭代路径为由 0 到 A, 其中 A 即为最优解.

c) 引入松弛变量 x_3, x_4 ，化为标准形

$$\begin{aligned} \text{minimize} \quad & -x_1 - x_2 \\ \text{subject to} \quad & -2x_1 - x_2 + x_3 = 1, \\ & x_1 - x_2 + x_4 = 1, \\ & x_i \geq 0, i = 1,2,3,4. \end{aligned}$$

写出初始表，其已是第一张单纯形表

	x_1	x_2	x_3	x_4	b
x_3	-2	1	1	0	1
x_4	$\boxed{1}$	-1	0	1	1
$c^T(r^T)$	-1	-1	0	0	1

$\rightarrow$

	x_1	x_2	x_3	x_4	b
x_3	0	-1	1	2	3
x_1	1	-1	0	1	1
r^T	0	-2	0	1	2

因为 x_2 对应列无正元素，所以原问题是无界的.

7. 第 2 章 2.13 解：将原问题化为标准形得初始表为：

	a_1	a_2	a_3	a_4	a_5	a_6	a_7	b
x_5	1	3	0	1	1	0	0	4
x_6	$\boxed{2}$	1	0	0	0	1	0	3
x_7	0	1	4	1	0	0	1	3
$c^T(r^T)$	-2	-4	-1	-1	0	0	0	0

$\rightarrow$

	a_1	a_2	a_3	a_4	a_5	a_6	a_7	b
x_4	0	$\boxed{5/2}$	0	1	1	$-1/2$	0	5/2
x_1	1	1/2	0	0	0	1/2	0	3/2
x_6	0	1	4	1	0	0	1	3
r^T	0	-3	-1	-1	0	1	0	3

$\rightarrow$

	a_1	a_2	a_3	a_4	a_5	a_6	a_7	b
x_2	0	1	0	2/5	2/5	$-1/5$	0	1
x_1	1	0	0	$-1/5$	$-1/5$	3/5	0	1
x_6	0	0	$\boxed{4}$	3/5	$-2/5$	1/5	1	2
r^T	0	0	-1	1/5	6/5	2/5	0	6

$\rightarrow$

	a_1	a_2	a_3	a_4	a_5	a_6	a_7	b
x_2	0	1	0	2/5	2/5	$-1/5$	0	1
x_1	1	0	0	$-1/5$	$-1/5$	3/5	0	1
x_3	0	0	1	3/20	$-1/10$	1/20	1/4	1/2
r^T	0	0	0	7/20	11/10	9/20	1/4	13/2

至此所有 $r_j \geq 0$, 得到最优解 $x^* = (1, 1, 1/2, 0)$.

(a) 设 b 的改变量为 $\Delta b = (\Delta b_1, \Delta b_2, \Delta b_3)$, 要使最优基不变，则需要 $B^{-1}(b + \Delta b) \geq 0$, 即

$$\begin{bmatrix} 2/5 & -1/5 & 0 \\ -1/5 & 3/5 & 0 \\ -1/10 & 1/20 & 1/4 \end{bmatrix} \begin{bmatrix} 4 + \Delta b_1 \\ 3 + \Delta b_2 \\ 3 + \Delta b_3 \end{bmatrix} \geq 0 \tag{1}$$

当 b 只有一个分量改变时，由 (1) 可得其各分量的变化范围为

$$\Delta b_1 \in [-5/2, 5],\ \Delta b_2 \in [-5/3, 5],\ \Delta b_3 \in [-2/3, +\infty).$$

当 b 不只一个分量改变时， Δb 应满足 (1). 当 b 的三个分量变化相同时，其变化范围应为 $\epsilon_b = \Delta b_1 \bigcap \Delta b_2 \bigcap \Delta b_3 = [-2/3, 5]$.

(b) 设 c 的改变量为 $\Delta c = (\Delta c_1, \Delta c_2, \Delta c_3, \Delta c_4)$, 要使最优基不变，则需要 $c_D - c_B B^{-1} D \geq 0$, 即

$$(-1 + \Delta c_4, 0, 0, 0) - (-4 + \Delta c_2, -2 + \Delta c_1, -1 + \Delta c_3) \begin{bmatrix} 2/5 & 2/5 & -1/5 & 0 \\ -1/5 & -1/5 & 3/5 & 0 \\ 3/20 & -1/10 & 1/20 & 1/4 \end{bmatrix} \geq 0$$

当 c 只有一个分量改变时，由上式可得其各分量的变化范围为

$$\Delta c_1 \in [-7/4, 3/4],\ \Delta c_2 \in [-9/4, 7/8],\ \Delta c_3 \in [-11, 1],\ \Delta c_4 \in [-7/20, +\infty).$$

当 c 不只一个分量改变时， Δc 应满足上式. 当 c 的三个分量变化相同时，其变化范围应为 $\epsilon_c = \Delta c_1 \bigcap \Delta c_2 \bigcap \Delta c_3 \bigcap \Delta c_4 = [7/20, 3/4]$.

(c) 对于 b 小的改变，由 (a) 可知其最优解将不会改变.

(d) 对于 c 小的改变，由 (b) 可知其最优解将不会改变，但其最优费用会变，即当 c 增加时，最优费用增加； c 减少时，最优费用减少.

8. 第 2 章 2.16　利用两阶段单纯形过程求解

(a)

$$\begin{aligned} \text{minimize} \quad & -3x_1 + x_2 + 3x_3 - x_4 \\ \text{subject to} \quad & x_1 + 2x_2 - x_3 + x_4 = 0, \\ & 2x_1 - 2x_2 + 3x_3 + 3x_4 = 9, \\ & x_1 - x_2 + 2x_3 - x_4 = 6, \\ & x_i \geq 0, i = 1, 2, 3, 4. \end{aligned}$$

(b)

$$\begin{aligned} \text{minimize} \quad & x_1 + 6x_2 - 7x_3 + x_4 + 5x_5 \\ \text{subject to} \quad & 5x_1 - 4x_2 + 13x_3 - 2x_4 + x_5 = 20, \\ & x_1 - x_2 + 5x_3 - x_4 + x_5 = 8, \\ & x_i \geq 0, i = 1, 2, 3, 4, 5. \end{aligned}$$

解： a) 第 I 阶段：引入人工变量 x_5, x_6, x_7 ，构造辅助问题，得

$$\begin{aligned} \text{minimize} \quad & x_5 + x_6 + x_7 \\ \text{subject to} \quad & x_1 + 2x_2 - x_3 + x_4 + x_5 = 0, \\ & 2x_1 - 2x_2 + 3x_3 + 3x_4 + x_6 = 9, \\ & x_1 - x_2 + 2x_3 - x_4 + x_7 = 6, \\ & x_i \geq 0, i = 1, 2, 3, \cdots, 7. \end{aligned}$$

则辅助问题的初始单纯形表为

	x_1	x_2	x_3	x_4	x_5	x_6	x_7	b
x_5	1	2	−1	1	1	0	0	0
x_6	2	−2	3	3	0	1	0	9
x_7	1	−1	2	−1	0	0	1	6
c^T	0	0	0	0	1	1	1	0

将最后一行与基变量 x_5, x_6, x_7 对应的系数化为 0，得第一张单纯形表

	x_1	x_2	x_3	x_4	x_5	x_6	x_7	b
x_5	$\boxed{1}$	2	−1	1	1	0	0	0
x_6	2	−2	3	3	0	1	0	9
x_7	1	−1	2	−1	0	0	1	6
r^T	−4	1	−4	−3	0	0	0	−15

$\rightarrow$

	x_1	x_2	x_3	x_4	x_5	x_6	x_7	b
x_1	1	2	−1	1	1	0	0	0
x_6	0	−6	$\boxed{5}$	1	−2	1	0	9
x_7	0	−3	3	−2	−1	0	1	6
r^T	0	9	−8	1	4	0	0	−15

$\rightarrow$

	x_1	x_2	x_3	x_4	x_5	x_6	x_7	b
x_1	1	4/5	0	6/5	3/5	1/5	0	9/5
x_3	0	−6/5	1	1/5	−2/5	1/5	0	9/5
x_7	0	$\boxed{3/5}$	0	−13/5	1/5	−3/5	1	3/5
r^T	0	−3/5	0	13/5	4/5	8/5	0	−3/5

$\rightarrow$

	x_1	x_2	x_3	x_4	x_5	x_6	x_7	b
x_1	1	0	0	14/3	1/3	1	−4/3	1
x_3	0	0	1	−5	0	−1	2	3
x_2	0	1	0	−13/3	1/3	−1	5/3	1
r^T	0	0	0	0	1	1	1	0

由此得原问题的基本可行解 $(1, 1, 3, 0)$.

第 II 阶段：以上述基本可行解作为初始基本可行解，利用单纯形法求解原问题. 首先得到如下初始表

	x_1	x_2	x_3	x_4	b
x_1	1	0	0	14/3	1
x_3	0	0	1	−5	3
x_2	0	1	0	−13/3	1
c^T	−3	1	3	−1	0

.

将最后一行与基变量对应的系数化为零后，得第一张单纯形表

	x_1	x_2	x_3	x_4	b
x_1	1	0	0	$14/3$	1
x_3	0	0	1	-5	3
x_2	0	1	0	$-13/3$	1
r^T	0	0	0	$97/3$	-7

已达最优，所以 $x^* = (1, 1, 3, 0)$, 最优值 $f^* = 7$.

b) 第 I 阶段：引入人工变量 x_6, x_7 ，构造辅助问题

$$\begin{aligned} \text{minimize} \quad & x_6 + x_7 \\ \text{subject to} \quad & 5x_1 - 4x_2 + 13x_3 - 2x_4 + x_5 + x_6 = 20, \\ & x_1 - x_2 + 5x_3 - x_4 + x_5 + x_7 = 8, \\ & x_i \geq 0, i = 1, 2, \cdots, 7. \end{aligned}$$

辅助问题的初始单纯形表为

	x_1	x_2	x_3	x_4	x_5	x_6	x_7	b
x_6	5	-4	13	-2	1	1	0	20
x_7	1	-1	5	-1	1	0	1	8
c^T	0	0	0	0	0	1	1	0

.

将最后一行与基变量对应的系数化为零后，得第一张单纯形表

	x_1	x_2	x_3	x_4	x_5	x_6	x_7	b
x_6	5	-4	13	-2	1	1	0	20
x_7	1	-1	5	-1	$\boxed{1}$	0	1	8
r^T	-6	5	-18	3	-2	0	0	-28

(注：如选 13 为转轴元，则计算比较繁琐.)

$\rightarrow$

	x_1	x_2	x_3	x_4	x_5	x_6	x_7	b
x_6	4	-3	$\boxed{8}$	-1	0	1	-1	12
x_5	1	-1	5	-1	1	0	1	8
r^T	-4	3	-8	1	0	0	2	-12

$\rightarrow$

	x_1	x_2	x_3	x_4	x_5	x_6	x_7	b
x_3	$1/2$	$-3/8$	1	$-1/8$	0	$1/8$	$-1/8$	$3/2$
x_5	$-3/2$	$7/8$	0	$-3/8$	1	$-5/8$	$13/8$	$1/2$
r^T	0	0	0	0	0	1	1	0

至此得基本可行解 $(0, 0, 3/2, 0, 1/2)$.

第 II 阶段：以上述基本可行解作为初始基本可行解，利用单纯形法求解原问题．首先得到如下初始表

	x_1	x_2	x_3	x_4	x_5	b
x_3	$1/2$	$-3/8$	1	$-1/8$	0	$3/2$
x_5	$-3/2$	$7/8$	0	$-3/8$	1	$1/2$
c^T	1	6	-7	1	5	0

将最后一行与基变量对应的系数化为零后，得第一张单纯形表

	x_1	x_2	x_3	x_4	x_5	b
x_3	$1/2$	$-3/8$	1	$-1/8$	0	$3/2$
x_5	$-3/2$	$\boxed{7/8}$	0	$-3/8$	1	$1/2$
r^T	12	-1	0	2	0	8

$\rightarrow$

	x_1	x_2	x_3	x_4	x_5	b
x_3	$-1/7$	0	1	$-2/7$	$3/7$	$12/7$
x_2	$-12/7$	1	0	$-3/7$	$8/7$	$4/7$
r^T	$72/7$	0	0	$11/7$	$8/7$	$60/7$

因为 $r_j \geq 0$, 所以已达最优解 $x^* = (0, 4/7, 12/7, 0, 0)$, 且 $f^* = -60/7$.

9. 第 2 章 2.18 假定给系统 $Ax = b, x \geq 0$ 应用单纯形过程，且所得表格 (忽略费用行) 形如

$$
\begin{array}{|ccc|ccc|ccc|ccc|c|}
\hline
x_1 & & x_k & x_{k+1} & \cdots & x_n & y_1 & \cdots & y_k & y_{k+1} & \cdots & y_m & \\
\hline
1 & & & & & & & & & 0 & \cdots & 0 & \bar{b}_1 \\
& \ddots & & & R_1 & & & S_1 & & \vdots & & \vdots & \vdots \\
& & 1 & & & & & & & 0 & \cdots & 0 & \bar{b}_k \\
\hline
0 & \cdots & 0 & & & & & & & 1 & & & 0 \\
\vdots & & \vdots & & R_2 & & & S_2 & & & \ddots & & \vdots \\
0 & \cdots & 0 & & & & & & & & & 1 & 0 \\
\hline
\end{array}
$$

对应地有 $m - k$ 个基变量在零水平.

(a) 证明 R_2 中的任何非零元素都可作为转轴元以消去一个人工基变量，这样将产生一个类似的表格，但 k 会增加 1.

(b) 假设 (a) 中的过程被重复，直到 $R_2 = \mathbf{0}$. 证明原始系统是冗余的，并说明可以删除底端的这些行，然后继续第 II 阶段.

(c) 利用上面的方法求解线性规划

$$
\begin{aligned}
\text{minimize} \quad & 2x_1 + 6x_2 + x_3 + x_4 \\
\text{subject to} \quad & x_1 + 2x_2 + x_4 = 6, \\
& x_1 + 2x_2 + x_3 + x_4 = 7, \\
& x_1 + 3x_2 - x_3 + 2x_4 = 7, \\
& x_1 + x_2 + x_3 = 5, \\
& x_i \geq 0, i = 1, 2, 3, 4.
\end{aligned}
$$

解： c) 第 I 阶段，引入人工变量 y_1, y_2, y_3, y_4 ，构造辅助问题

$$
\begin{aligned}
\text{minimize} \quad & y_1 + y_2 + y_3 + y_4 \\
\text{subject to} \quad & x_1 + 2x_2 + x_4 + y_1 = 6, \\
& x_1 + 2x_2 + x_3 + x_4 + y_2 = 7, \\
& x_1 + 3x_2 - x_3 + 2x_4 + y_3 = 7, \\
& x_1 + x_2 + x_3 + y_4 = 5, \\
& x_i \geq 0, y_i \geq 0, i = 1, 2, 3, 4.
\end{aligned}
$$

辅助问题的初始表为

	x_1	x_2	x_3	x_4	y_1	y_2	y_3	y_4	b
y_1	1	2	0	1	1	0	0	0	6
y_2	1	2	1	1	0	1	0	0	7
y_3	1	3	-1	2	0	0	1	0	7
y_4	1	1	1	0	0	0	0	1	5
c^T	0	0	0	0	1	1	1	1	0

将最后一行与基变量对应的系数化为零，得第一张单纯形表

	x_1	x_2	x_3	x_4	y_1	y_2	y_3	y_4	b
y_1	1	2	0	1	1	0	0	0	6
y_2	1	2	1	1	0	1	0	0	7
y_3	1	$\boxed{3}$	-1	2	0	0	1	0	7
y_4	1	1	1	0	0	0	0	1	5
r^T	-4	-8	-1	-4	0	0	0	0	-25

$\rightarrow$

	x_1	x_2	x_3	x_4	y_1	y_2	y_3	y_4	b
y_1	$1/3$	0	$2/3$	$-1/3$	1	0	$-2/3$	0	$4/3$
y_2	$1/3$	0	$\boxed{5/3}$	$-1/3$	0	1	$-2/3$	0	$7/3$
x_2	$1/3$	1	$-1/3$	$2/3$	0	0	$1/3$	0	$7/3$
y_4	$2/3$	0	$4/3$	$-2/3$	0	0	$-1/3$	1	$8/3$
r^T	$-4/3$	0	$-11/3$	$4/3$	0	0	$8/3$	0	$-19/3$

$\rightarrow$

	x_1	x_2	x_3	x_4	y_1	y_2	y_3	y_4	b
y_1	$\boxed{1/5}$	0	0	$-1/5$	1	$-2/5$	$-2/5$	0	$2/5$
x_3	$1/5$	0	1	$-1/5$	0	$3/5$	$-2/5$	0	$7/5$
x_2	$2/5$	1	0	$3/5$	0	$1/5$	$1/5$	0	$14/5$
y_4	$2/5$	0	0	$-2/5$	0	$-4/5$	$1/5$	1	$4/5$
r^T	$-3/5$	0	0	$3/5$	0	$11/5$	$6/5$	0	$-6/5$

	x_1	x_2	x_3	x_4	y_1	y_2	y_3	y_4	b
x_1	1	0	0	−1	5	−2	−2	0	2
x_3	0	0	1	0	−1	1	0	0	1
$\to$ x_2	0	1	0	1	−2	1	1	0	2
y_4	0	0	0	0	−2	0	1	1	0
r^T	0	0	0	0	3	1	0	0	0

由 b) 可知第 4 个约束为冗余的，可删除．至此得基本可行解 $(2,2,1,0)$.

第 II 阶段：以上述基本可行解作为初始基本可行解，利用单纯形法求解原问题．首先得到如下初始表

	x_1	x_2	x_3	x_4	b
x_1	1	0	0	−1	2
x_3	0	0	1	0	1
x_2	0	1	0	1	2
c^T	2	6	1	1	0

将最后一行与基变量对应的系数化为零后，得第一张单纯形表

	x_1	x_2	x_3	x_4	b
x_1	1	0	0	−1	2
x_3	0	0	1	0	1
x_2	0	1	0	$\boxed{1}$	2
r^T	0	0	0	−3	−17

	x_1	x_2	x_3	x_4	b
x_1	1	1	0	0	4
$\to$ x_3	0	0	1	0	1
x_4	0	1	0	1	2
r^T	0	3	0	0	−11

因为 $r_j \geq 0$, 所以已达最优解 $x^* = (4,0,1,2)$, $f^* = 11$.

10. 第 2 章 2.19 利用修正单纯形法找出下列系统的一个基本可行解

$$
\begin{aligned}
&x_1 + 2x_2 - x_3 + x_4 = 3, \\
&2x_1 + 4x_2 + x_3 + 2x_4 = 12, \\
&x_1 + 4x_2 + 2x_3 + x_4 = 9, \\
&x_i \geq 0, i = 1,2,3,4.
\end{aligned}
$$

解：引入人工变量 x_5, x_6, x_7 ，构造辅助问题．得辅助问题的如下表格

	B^{-1}			x_5
x_5	1	0	0	3
x_6	0	1	0	12
x_7	0	0	1	9

$\lambda^T = (1,1,1), r_D^T = c_D - \lambda^T D = (-4,-10,-2.4)$.
让 x_2 进基，计算 $y_2 = B^{-1}a_2 = (2,4,4)$ ，得

	B^{-1}			x_B	y_2
x_5	1	0	0	3	$\boxed{2}$
x_6	0	1	0	12	4
x_7	0	0	1	9	4
λ^T	1	1	1	24	10

$\rightarrow$

	B^{-1}			x_B
x_2	1/2	0	0	3/2
x_6	-2	1	0	6
x_7	-2	0	1	3
λ^T	-4	1	1	9

计算相对费用系数，得 $r_1 = 1, r_3 = -7, r_4 = 1, r_5 = 5$. 选取 x_3 进基，计算 $y_3 = B^{-1}a_3 = (-1/2, 3, 4)$. 故有

	B^{-1}			x_B	y_3
x_2	1/2	0	0	3/2	$-1/2$
x_6	-2	1	0	6	3
x_7	-2	0	1	3	$\boxed{4}$
λ^T	-4	1	1	9	7

$\rightarrow$

	B^{-1}			x_B
x_2	1/4	0	1/8	15/8
x_6	$-1/2$	1	$-3/4$	15/4
x_3	$-1/2$	0	1/4	3/4
λ^T	$-1/2$	1	$-3/4$	15/4

计算相对费用系数，得 $r_1 = -3/4, r_4 = -3/4, r_5 = 3/2, r_7 = 7/4$. 选取 x_1 进基，计算 $y_1 = B^{-1}a_1 = (3/8, 3/4, -1/4)$. 故有

	B^{-1}			x_B	y_1
x_2	1/4	0	1/8	15/8	3/8
x_6	$-1/2$	1	$-3/4$	15/4	$\boxed{3/4}$
x_3	$-1/2$	0	1/4	3/4	$-1/4$
λ^T	$-1/2$	1	$-3/4$	15/4	3/4

$\rightarrow$

	B^{-1}			x_B
x_2	1/2	$-1/2$	1/2	0
x_1	$-2/3$	4/3	-1	5
x_3	$-2/3$	1/3	0	2
λ^T	0	0	0	0

计算相对费用系数，得 $r_4 = 0, r_5 = r_6 = r_7 = 1$. 至此得到辅助问题的最优基本可行解. 因为基变量不含人工变量，所以得原问题的基本可行解 $x = (5,0,2,0)$.

11. 第 2 章 2.21 下面的表格是求解极小化问题的一个中间阶段：

	y_1	y_2	y_3	y_4	y_5	y_6	y_0
	1	2/3	0	0	4/3	0	4
	0	$-7/3$	3	1	$-2/3$	0	2
	0	$-2/3$	-2	0	2/3	1	2
r^T	0	8/3	-11	0	4/3	0	-8

(a) 确定下一个转轴元.

(b) 给定当前基的逆

$$B^{-1} = [a_1, a_4, a_6]^{-1} = \frac{1}{3}\begin{bmatrix} 1 & 1 & -1 \\ 1 & -2 & 2 \\ -1 & 2 & 1 \end{bmatrix}$$

和对应的费用系数

$$c_B^T = (c_1, c_4, c_6) = (-1, -3, 1),$$

求原始问题.

解：

(a) 下一个转轴元为 3(第 2 行第 3 列)

(b) $B=(a_1,a_4,a_6)=(B^{-1})^{-1}=\begin{bmatrix}2&1&0\\1&0&1\\0&1&1\end{bmatrix}$. 由 $y_0=B^{-1}b$ 可得 $b=By_0=\begin{bmatrix}10\\6\\4\end{bmatrix}$.

由 $y_q=B^{-1}a_q$ 可得 $(a_2,a_3,a_5)=B(y_2,y_3,y_5)=\begin{bmatrix}-1&3&2\\0&-2&2\\-3&1&0\end{bmatrix}$ 所以

$$A=(a_1\quad a_2\quad a_3\quad a_4\quad a_5\quad a_6)=\begin{bmatrix}2&-1&3&1&2&0\\1&0&-2&0&2&1\\0&-3&1&1&0&1\end{bmatrix}.$$

由 $r_q=c_q-c_B^Ty_q$ 可得 $(c_2,c_3,c_5)=(25/3,-22,8/3)$. 所以原问题为

$$\begin{aligned}
\text{minimize}\quad & -x_1+(25/3)x_2-22x_3-3x_4+(8/3)x_5+x_6\\
\text{subject to}\quad & 2x_1-1x_2+3x_3+x_4+2x_5=10,\\
& x_1-2x_3+2x_5+x_6=6,\\
& -3x_2+x_3+x_4+x_6=4,\\
& x_i\geq 0, i=1,2,3,4,5,6.
\end{aligned}$$

最优化理论与算法 (数值优化) – – 第 3 章作业参考解答

第 4 章作业： 4.1 详细验证线性规划问题对偶的对偶是原始问题.

解：原问题和对偶问题分别如下：

$$
\begin{array}{llll}
\text{(P)} & \text{minimize} & c^T x & \\
& \text{subject to} & Ax = b, & \\
& & x \geq 0. &
\end{array}
\qquad
\begin{array}{lll}
\text{(D)} & \text{maximize} & \lambda^T b \\
& \text{subject to} & A^T \lambda \leq c.
\end{array}
$$

方法 I：将对偶问题 (D) 化成标准形，写出对偶，再转化成与原问题等价的线性规划．具体操作见讲义.

第 4 章作业： 4.5 构造一个原问题和对偶问题都没有可行解的例子.

解：构造原问题如下

$$
\begin{array}{ll}
\text{minimize} & -x_1 - x_2 \\
\text{subject to} & x_1 \geq 1, \\
& -x_2 \geq 1, \\
& x_1 \geq 0, \quad x_2 \geq 0.
\end{array}
$$

其对偶问题为

$$
\begin{array}{ll}
\text{maximize} & \lambda_1 + \lambda_2 \\
\text{subject to} & \lambda_1 \leq -1, \\
& -\lambda_2 \leq -1, \\
& \lambda_1 \geq 0, \quad \lambda_2 \geq 0.
\end{array}
$$

易见二者均无可行解.

第 4 章作业： 4.9 考虑原线性规划问题

$$
\begin{array}{ll}
\text{minimize} & c^T x \\
\text{subject to} & Ax = b, \\
& x \geq 0.
\end{array}
$$

假设该问题和它的对偶是可行的．设 $\bar{\lambda}$ 是对偶的一个最优解.

1. 如果给原问题的第 k 个方程乘以 $\mu \neq 0$ ，确定这个新问题对偶的一个最优解.
2. 假设，在原问题中，我们给第 r 个方程加上第 k 个方程的 μ 倍．对应对偶问题的最优解 λ 怎么样？
3. 假设，在原问题中，我们给 c 加上 A 的第 k 行的 μ 倍．对应对偶问题的最优解 λ 怎么样？

解：该问题的对偶问题为

$$\begin{aligned} \text{maximize} \quad & b^T\lambda \\ \text{subject to} \quad & A^T\lambda \le c. \end{aligned}$$

设原问题的最优解为 $\bar{x}$ ，对偶问题的最优解为 $\bar{\lambda}$ ，则其满足如下的互补松弛条件

$$\begin{aligned} A\bar{x} - b = 0 \\ \bar{x} \ge 0 \\ c - A^T\bar{\lambda} \ge 0 \\ \bar{x}^T(c - A^T\bar{\lambda}) = 0 \end{aligned} \tag{1}$$

以下设 $A = \begin{bmatrix} a^1 \\ \vdots \\ a^m \end{bmatrix}$. 则原问题的等式约束 $Ax = b$ 即为 $a^1x = b_1, \cdots, a^mx = b_m$. 对偶问题的约束 $c - A^T\lambda \ge 0$ 即 $c^T - \sum_{i=1}^m \lambda_i a^i \ge 0$.

1. 给原问题的第 k 个方程乘以 $\mu \neq 0$ 后，第 k 个约束变为 $\mu a^k x = \mu b_k$ ，易见原问题的最优解是新问题的最优解，但是对偶最优解变了．所得新问题的互补松弛条件为

$$\begin{aligned} c^T - \mu\lambda_k a^k - \textstyle\sum_{i=1, i\neq k}^m \lambda_i a^i \ge 0 \\ a^i x = b_i, i \neq k, \quad \mu a^k x = \mu b_k \\ x \ge 0 \\ (c^T - \mu\lambda_k a^k - \textstyle\sum_{i=1, i\neq k}^m \lambda_i a^i)x = 0. \end{aligned} \tag{2}$$

   因为 (1) ，故可验证原问题最优解 $\bar{x}$ 和 $\lambda = (\bar{\lambda}_1, \cdots, \frac{1}{\mu}\bar{\lambda}_k, \cdots, \bar{\lambda_m})^T$ 满足 (2). 由互补松弛定理， λ 即为新问题的对偶问题的最优解.

2. 给第 r 个方程加上第 k 个方程的 μ 倍后，第 r 个约束变为 $(a^r + \mu a^k)x = b_r + \mu b_k$ ，易见原问题的最优解是新问题的最优解，但是对偶最优解变了．所得新问题的 KKT 条件为

$$\begin{aligned} c^T - (a^r + \mu a^k)\lambda_r - \textstyle\sum_{i=1, i\neq r}^m \lambda_i a^i \ge 0 \\ a^i x = b_i, i \neq r, \quad (a^r + \mu a^k)x = b_r + \mu b_k \\ x \ge 0 \\ (c^T - (a^r + \mu a^k)\lambda_r - \textstyle\sum_{i=1, i\neq r}^m \lambda_i a^i)x = 0 \end{aligned} \tag{3}$$

   令 $\lambda_i = \bar{\lambda}_i, i \neq k, \lambda_k = \bar{\lambda}_k - \mu\bar{\lambda}_r$ ，同上可以验证 $\bar{x}$ 和 λ 满足 (3). 所以 λ 即为新问题对偶问题的最优解.

3. 给 c 加上 A 的第 k 行的 μ 倍后，目标函数变为 $c^Tx + \mu a^k x$. 所得新问题的 KKT 条件为

$$\begin{aligned} c^T + \mu a^k - \textstyle\sum_{i=1}^m \lambda_i a^i \ge 0 \\ a^i x = b_i, i = 1, \cdots, m \\ x \ge 0 \\ (c^T + \mu a^k - \textstyle\sum_{i=1}^m \lambda_i a^i)x = 0 \end{aligned} \tag{4}$$

   令 $\lambda_i = \bar{\lambda}_i, i \neq k, \lambda_k = \bar{\lambda}_k + \mu$ ，同上可以验证 $\bar{x}$ 和 λ 满足 (4). 所以 λ 即为新问题对偶问题的最优解.

第 4 章作业： 4.12 解： (a) 先将原问题化为标准形．由提示 $x=(2,0,0,1,0)$ 为基本可行解．可得与其对应的规范形数据如下：

	a_1	a_2	a_3	a_4	a_5	b
x_4	0	-1	3	1	-2	1
x_1	1	-1	1	0	-1	2

进而得到与此基本可行解对应的初始表格

	a_1	a_2	a_3	a_4	a_5	b
x_4	0	-1	3	1	-2	1
x_1	1	-1	1	0	-1	2
c^T	2	-1	0	0	0	0

通过初等行变换将基变量对应的相对费用系数变为零，得第一张单纯形表，并依次进行单纯形转轴

$\rightarrow$

	a_1	a_2	a_3	a_4	a_5	b
x_4	0	-1	$\boxed{3}$	1	-2	1
x_1	1	-1	1	0	-1	2
r^T	0	1	-2	0	2	-4

$\rightarrow$

	a_1	a_2	a_3	a_4	a_5	b
x_3	0	$-1/3$	1	$1/3$	$-2/3$	$1/3$
x_1	1	$-2/3$	0	$-1/3$	$-1/3$	$5/3$
r^T	0	$1/3$	0	$2/3$	$2/3$	$-10/3$

至此所有 $r_j \geq 0$, 得到最优解 $x^*=(1,1,1/2,0)$.
(b) 原问题的对偶问题为

$$\begin{aligned} \text{maximize} \quad & 3\lambda_1+2\lambda_2 \\ \text{subject to} \quad & 2\lambda_1+\lambda_2 \leq 2, \\ & -\lambda_1-\lambda_2 \leq -1, \\ & -\lambda_1+\lambda_2 \leq 0, \\ & \lambda_1 \geq 0,\ \lambda_2 \geq 0. \end{aligned}$$

由 (a) 的最后一张表可知对偶问题的最优解为 $\lambda^*=(2/3,2/3)$.

第 4 章作业： 4.13

1. 利用单纯形法求解

$$\begin{aligned} \text{minimize} \quad & 2x_1+3x_2+2x_3+2x_4 \\ \text{subject to} \quad & x_1+2x_2+x_3+2x_4=3, \\ & x_1+1x_2+2x_3+4x_4=5, \\ & x_i \geq 0, i=1,2,3,4. \end{aligned}$$

2. 利用 (a) 中所作工作和对偶单纯形法求解相同的问题，除过方程组的右端项分别变成 8 和 7.

解. a) 第一阶段：首先引入人工变量 x_5, x_6 ，构造辅助问题

$$
\begin{aligned}
\text{minimize} \quad & x_5 + x_6 \\
\text{subject to} \quad & x_1 + 2x_2 + x_3 + 2x_4 + x_5 = 3, \\
& x_1 + 1x_2 + 2x_3 + 4x_4 + x_6 = 5, \\
& x_i \geq 0, i = 1,2,3,4,5,6.
\end{aligned}
$$

利用单纯形法求解辅助问题. 首先得初始表

	x_1	x_2	x_3	x_4	x_5	x_6	x_B
x_5	1	2	1	2	1	0	3
x_6	1	1	2	4	0	1	5
c^T	0	0	0	0	1	1	0

将基变量 x_5, x_6 对应的最后一行的系数化为零，得第一张单纯形表

	x_1	x_2	x_3	x_4	x_5	x_6	x_B
x_5	1	2	1	2	1	0	3
x_6	1	1	2	$\boxed{4}$	0	1	5
r^T	-2	-3	-3	-6	0	0	-8

以 4 为转轴元，即 x_4 进基， x_6 出基，得

	x_1	x_2	x_3	x_4	x_5	x_6	x_B
x_5	$1/2$	$\boxed{3/2}$	0	0	1	$-1/2$	$1/2$
x_4	$1/4$	$1/4$	$1/2$	1	0	$1/4$	$5/4$
r^T	$-1/2$	$-3/2$	0	0	0	$3/2$	$-1/2$

再次转轴之后，得

	x_1	x_2	x_3	x_4	x_5	x_6	x_B
x_2	$1/3$	1	0	0	$2/3$	$-1/3$	$1/3$
x_4	$1/6$	0	$1/2$	1	$-1/6$	$1/3$	$7/6$
r^T	0	0	0	0	1	1	0

因为辅助问题的最优值为 0 ，所以原问题有可行解. 因为基变量不含人工变量，故 $x = (0, 1/3, 0, 7/6)$ 是一个基本可行解.

第 II 阶段：在第 I 阶段的最后一个单纯形表中删去人工变量所对应列和最后一行，可得原问题的一张初始表

	x_1	x_2	x_3	x_4	x_B
x_2	$1/3$	1	0	0	$1/3$
x_4	$1/6$	0	$1/2$	1	$7/6$
c^T	2	3	2	2	0

将基变量 x_2 和 x_4 下面的系数化为零，得第一张单纯形表

	x_1	x_2	x_3	x_4	x_B
x_2	$1/3$	1	0	0	$1/3$
x_4	$1/6$	0	$1/2$	1	$7/6$
r^T	$2/3$	0	1	0	$-10/3$

因为相对费用系数向量非负，故得最优解 $x^* = (0, 1/3, 0, 7/6)$，最优值 $f^* = 10/3$.

b) 由 a) 可知 $B = (a_2, a_4)$，所以 $B^{-1} = \begin{bmatrix} 2/3 & -1/3 \\ -1/6 & 1/3 \end{bmatrix}$. 右端项改为 8 和 7 后，只需将上表的最后一列改为 $B^{-1}\begin{bmatrix} 8 \\ 7 \end{bmatrix} = \begin{bmatrix} 3 \\ 1 \end{bmatrix}$, 即为

	x_1	x_2	x_3	x_4	x_B
x_2	$1/3$	1	0	0	3
x_4	$1/6$	0	$1/2$	1	1
r^T	$2/3$	0	1	0	-11

因为相对费用系数向量非负，故得最优解 $x^* = (0, 3, 0, 1)$，最优值 $f^* = 11$.

第 4 章作业： 4.14.

$$\begin{aligned} \text{minimize} \quad & 5x_1 - 3x_2 \\ \text{subject to} \quad & 2x_1 - x_2 + 4x_3 \leq 4, \\ & x_1 + x_2 + 2x_3 \leq 5, \\ & 2x_1 - x_2 + x_3 \geq 1, \\ & x_1 \geq 0, x_2 \geq 0, x_3 \geq 0. \end{aligned}$$

1. 以 1 为转轴元，利用单个转轴运算找到一个可行解.
2. 利用单纯形法求解问题.
3. 对偶问题是什么？
4. 对偶的解怎么样？

解.

1. 引入松驰变量 x_4, x_5 和盈余变量 x_6 后，得标准形问题

$$\begin{aligned} \text{minimize} \quad & 5x_1 - 3x_2 \\ \text{subject to} \quad & 2x_1 - x_2 + 4x_3 + x_4 = 4, \\ & x_1 + x_2 + 2x_3 + x_5 = 5, \\ & 2x_1 - x_2 + x_3 - x_6 = 1, \\ & x_1 \geq 0, x_2 \geq 0, x_3 \geq 0. \end{aligned}$$

得与等式约束对应的表格如下

x_1	x_2	x_3	x_4	x_5	x_6	b
2	-1	4	1	0	0	4
1	1	2	0	1	0	5
2	-1	$\boxed{1}$	0	0	-1	1

以 1 为转轴元转轴 (即从第三个方程解出 x_3，代入第一个和第二个方程消去 x_3) 后，得

x_1	x_2	x_3	x_4	x_5	x_6	b
-6	3	0	1	0	4	0
-3	3	0	0	1	2	3
2	-1	1	0	0	-1	1

该表对应着原问题的一个基本可行解.

2. 下面利用单纯形法求解原问题. 首先得初始表为

	x_1	x_2	x_3	x_4	x_5	x_6	x_B
x_4	-6	$\boxed{3}$	0	1	0	4	0
x_5	-3	3	0	0	1	2	3
x_3	2	-1	1	0	0	-1	1
c^T	5	-3	0	0	0	0	0

因为最后一行与基变量对应的系数已经为零，故上面的初始表即为第一张单纯形表. 以 3 为转轴元进行转轴后，得

	x_1	x_2	x_3	x_4	x_5	x_6	x_B
x_2	-2	1	0	1/3	0	4/3	0
x_5	$\boxed{3}$	0	0	-1	1	-2	3
x_3	0	0	1	1/3	0	1/3	1
r^T	-1	0	0	1	0	4	0

再以 3 为转轴元转轴后，得

	x_1	x_2	x_3	x_4	x_5	x_6	x_B
x_2	0	1	0	$-1/3$	2/3	0	2
x_1	1	0	0	$-1/3$	1/3	$-2/3$	1
x_3	0	0	1	1/3	0	1/3	1
r^T	0	0	0	2/3	1/3	10/3	1

因为相对费用系数向量非负，故得原问题最优解 $x^* = (1, 2, 1)$, 最优值 $f^* = -1$.

3. 原问题的对偶问题为

$$
\begin{array}{lll}
 & \text{maximize} & 4w_1 + 5w_2 + w_3 \\
\text{(D)} & \text{subject to} & 2w_1 + w_2 + 2w_3 \le 5, \\
 & & -w_1 + w_2 - w_3 \le -3, \\
 & & 4w_1 + 2w_2 + w_3 \le 0, \\
 & & w_1 \le 0, w_2 \le 0, w_3 \ge 0.
\end{array}
$$

4. 易于看到原问题引入松弛变量 x_4, x_5 和盈余变量 x_6 后，所得标准形问题的对偶问题也是该问题．从而标准形问题的最优乘子 $(\lambda^*)^T = c_B^T B^{-1}$ 即为对偶问题的最优解，其中 B 为最优解对应的基．而

$$\begin{aligned}
[\,r_4 \quad r_5 \quad r_6\,] &= [\,c_4 \quad c_5 \quad c_6\,] - c_B^T B^{-1} [\,a_4 \quad a_5 \quad a_6\,] \\
&= [\,0 \quad 0 \quad 0\,] - [\,\lambda_1^* \quad \lambda_2^* \quad \lambda_3^*\,] \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -1 \end{bmatrix} \\
&= [\,-\lambda_1^* \quad -\lambda_2^* \quad \lambda_3^*\,]
\end{aligned}$$

所以由 b) 中的最优单纯形表，得 $\lambda^* = (-2/3, -1/3, 10/3)$. 计算易得最优值为 -1.

最优化理论与算法 B(数值优化) ﹣﹣第 4 章及第 5 章部分作业参考解答

1. 说明函数 $f(x)=8x_1+12x_2+x_1^2-2x_2^2$ 仅有一个驻点，且其既不是最小点也不是最大点，而是一个鞍点．画出 f 的等值线.

   解答： $\nabla f = \begin{bmatrix} 8+2x_1 \\ 12-4x_2 \end{bmatrix}$, 令 $\triangle f = 0$ 则 $x^* = \begin{bmatrix} -4 \\ 3 \end{bmatrix}$ 为唯一解，从而仅有一个驻点。又 $\nabla^2 f = \begin{bmatrix} 2 & 0 \\ 0 & -4 \end{bmatrix}$ 是不定的，所以 x^* 为鞍点，它既不是最大值点，也不是最小值点。

   图形：双曲线的中心在 $(-4,3)$，渐近线不平行于坐标轴。

2. 考虑问题 $\min\|Ax-b\|^2$，其中 A 是 $m\times n$ 矩阵， b 是 m 维向量.

   (a) 给出问题的几何解释.

   (b) 写出最优性的必要条件．这也是一个充分条件吗？

   (c) 最优解唯一吗？理由是什么？

   (d) 你能给出最优解的一种闭合形式吗？可用规定任何你所需的假设.

   (e) 针对如下给出的 A 和 b:

   $$A=\begin{bmatrix} 1 & -1 & 0 \\ 0 & 2 & 1 \\ 0 & 1 & 0 \\ 1 & 0 & 1 \end{bmatrix}, \quad b=\begin{bmatrix} 2 \\ 1 \\ 1 \\ 0 \end{bmatrix}$$

   求解问题.

   解答： (a) 几何解释： b 到 A 的值空间 $R(A)$(即由 A 的列生成的子空间) 的最小距离。设 b 在 $R(A)$ 上的投影为 c，则满足 $Ax=c$ 的解 x^* 即为最优解。请同学们自己举一个例子，如三个方程，两个未知数的矛盾方程。

   (b) 令 $f(x)=\| Ax-b \|^2 = x^TA^TAx-2b^TAx+b^Tb, \quad \nabla f(x)=2A^TAx-2A^Tb$

   一阶必要条件为： $\nabla f=0, \quad$ 即 $A^TAx=A^Tb$

   又 $\nabla^2 f(x)=2A^TA$ 为半正定矩阵，所以 $f(x)$ 为凸函数，从而也为充分条件。

   (c) 最优解不一定唯一。当 $b\in R(A)$ 且 $\operatorname{rank}(A)<\min(m,n)$ 时， $Ax=b$ 有无穷多解，而此时满足 $Ax=b$ 的解均为最优解。

   (d) 如果设 A^TA 是正定的 (或者 A 是列满秩的)，则可给出最优解的一种闭合形式 $x^*=(A^TA)^{-1}A^Tb$.

   (e) 解 $A^TAx=A^Tb$ 即可。 $x^*=(3,\frac{3}{2},\frac{-5}{2})^T$.

3. 对标量 β 的每一个值，找到下列函数的所有驻点

   $$f(x,y)=x^2+y^2+\beta xy+x+2y.$$

   这些驻点中的哪些是全局极小点.

解答： $\nabla f(x,y)=\begin{bmatrix}2x+\beta y+1\\2y+\beta x+2\end{bmatrix}$， $\nabla^2 f(x,y)=\begin{bmatrix}2&\beta\\\beta&2\end{bmatrix}$.

令 $\nabla f(x,y)=0$, 即 $\begin{cases}2x+\beta y+1=0\cdots\cdots(1)\\2y+\beta x+2=0\cdots\cdots(2)\end{cases}$

由 $(1), x=-\frac{1}{2}-\frac{\beta}{2}y$ 代入 (2) 得 $(2-\frac{\beta^2}{2})y=\frac{\beta}{2}-2\cdots\cdots(3)$

当 $\beta=\pm 2$ 时， (3) 无解，从而无驻点；

当 $\beta\neq\pm 2$ 时， $y=\frac{\beta-4}{4-\beta^2}$， $x=-\frac{1}{2}-\frac{\beta}{2}y=\frac{2\beta-2}{4-\beta^2}$， $|\nabla^2 f(x,y)|=4-\beta^2$;

当 $\beta\in(-2,2)$ 时， (因为 $2>0$, 且 $|\nabla^2 f(x,y)|>0$), 此时 f 为凸函数，所有此时 $(\frac{2\beta-2}{4-\beta^2},\frac{\beta-4}{4-\beta^2})^T$ 为全局极小点；

当 $\beta<-2$ 或者 $\beta>2$ 时， $\nabla^2 f(x,y)$ 不定且有负特征值，所有 f 可以任意小，即没有极小点。

综上， $f(x,y)$ 的驻点中只有当 $\beta\in(-2,2)$ 时， $(\frac{2\beta-2}{4-\beta^2},\frac{\beta-4}{4-\beta^2})^T$ 为全局极小点。

4. 第 4 章作业： 4.8

说明如果非零向量 $p_0,p_1,\cdots,p_l$ 关于对称且正定的矩阵 A 是共轭的, 则这些向量是线性无关的. (该结论蕴含着 A 最多有 n 个共轭方向.)

解答：设存在数 $k_0,k_1,\cdots,k_l$, 使得 $k_0p_0+k_1p_1+\cdots+k_lp_l=0$, 则

$$p_i^TA(k_0p_0+k_1p_1+\cdots+k_lp_l)=0,\quad i=0,1,\cdots,l\quad\cdots\cdots(1)$$

又 $p_0,p_1,\cdots,p_l$ 关于 A 共轭，所以 $p_i^TAp_j=0,\quad(i\neq j)$. 由 (1) 式可以得到 $k_ip_i^TAp_i=0$, 又 A 正定，所以 $p_i^TAp_i>0$, 由此得到 $k_i=0,\quad i=0,1,\cdots,l$, 从而 $p_0,p_1,\cdots,p_l$ 线性无关。

5. 第 4 章作业： 4.9

考查函数 $f(x_1,x_2)=(x_1+x_2^2)^2$. 在点 $x_k=(1,0)$ 处考虑搜索方向 $p_k=(-1,1)$ ，证明 p_k 是一个下降方向，并且求出一维极小化问题

$$\min_{\alpha>0} f(x_k+\alpha p_k)$$

的所有极小点.

解答： $\nabla f=\begin{bmatrix}2(x_1+x_2^2)\\4x_2(x_1+x_2^2)\end{bmatrix}$， $\nabla f(x_k)^Tp_k=\begin{bmatrix}2\\0\end{bmatrix}^T\begin{bmatrix}-1\\1\end{bmatrix}=-2<0$, 故 p_k 是一个下降方向。

$x_k+\alpha p_k=\begin{bmatrix}1\\0\end{bmatrix}+\alpha\begin{bmatrix}-1\\1\end{bmatrix}=\begin{bmatrix}1-\alpha\\\alpha\end{bmatrix}$, 所以 $f(x_k+\alpha p_k)=(1-\alpha+\alpha^2)^2$, 又 $1-\alpha+\alpha^2=(\alpha-\frac{1}{2})^2+\frac{3}{4}\geq\frac{3}{4}$, 故当且仅当 $\alpha=\frac{1}{2}$ 时取等号。所以当 $\alpha=\frac{1}{2}$ 时， $f(x_k+\alpha p_k)$ 的值最小。

6. 第 5 章作业： 5.1 考虑将带精确线搜索的最速下降法应用到凸二次函数，即

$$f(x)=\frac{1}{2}x^TQx+b^Tx,$$

其中 Q 对称半正定，最优解为 x^* ，则利用本章给出的性质说明：如果初始点使得 x_0-x^* 与 Q 的一个特征向量平行，则最速下降法将在一步内找到解.

解答：因为 $x_0 - x^*$ 与 Q 的一个特征向量平行，所以 $Q(x_0 - x^*) = \lambda(x_0 - x^*)$, λ 为与 $(x_0 - x^*)$ 平行的特征向量所属的特征值。

因为 x^* 为最优解，从而 $Qf(x^*) = Qx^* + b = 0$, 所以 $b = -Qx^*$, 从而 $\nabla f(x_0) = Qx_0 + b = Q(x_0 - x^*) = \lambda(x_0 - x^*)$.

case1： $\lambda = 0$, 则 $\nabla f(x_0) = 0$, 而 f 为凸函数，从而 x_0 即为最优解；

case2： $\lambda \neq 0$, 则 $x_0 = \frac{\nabla f_0^T \nabla f_0}{\nabla f_0^T Q \nabla f_0} = \frac{\lambda^2 (x-x^*)^T (x-x^*)}{\lambda^3 (x-x^*)^T (x-x^*)} = \frac{1}{\lambda}$;

所以 $x_1 = x_0 + x_0 p_0 = x_0 - \frac{1}{\lambda}\nabla f_0 = x_0 - \frac{1}{\lambda}\lambda(x_0 - x^*) = x^*$.

所以，最速下降法将在一步内找到解。

7. 第 5 章作业： 5.2

假设我们需求极小化

$$f(x_1, x_2) = 5x_1^2 + 5x_2^2 - x_1 x_2 - 11x_1 + 11x_2 + 11.$$

(a) 找到一个满足一阶必要条件的解.

(b) 说明该点是一个全局极小点.

(c) 针对该问题最速下降法的最坏速率将会是多少？

(d) 从 $(x_1, x_2) = (0, 0)$ 出发，最多需要多少步可用将目标函数值减少到 10^{-11}?

解答：

a) 令 $\nabla f = \begin{bmatrix} 10x_1 - x_2 - 11 \\ 10x_2 - x_1 + 11 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$, 可得满足一阶必要条件的解为 $x^* = \begin{bmatrix} 1 \\ -1 \end{bmatrix}$.

b) 因为 $\nabla^2 f = \begin{bmatrix} 10 & -1 \\ -1 & 10 \end{bmatrix}$ 正定，所以 f 为凸函数，从而 $x^* = (1, -1)^T$ 也是全局极小点。

c)$\nabla^2 f$ 的特征值为 $9, 11$ ，所以最坏速率 $v = (\frac{11-9}{11+9})^2 = 0.01$.

d)$f(x^*) = 0$, $f(0,0) = 11$, $f(x_k) - f(x^*) \leq r^k [f(x_0) - f(x^*)] \cdots\cdots (*)$, 即 $10^{-11} = 10^{-2k} \cdot 11$, $2k = lg11 + 11$, $k = \frac{lg11+11}{2}$, 令 $(*)$ 式中等号成立，则应取 $k = 7$ ，否则可能达不到要求。所以最多需要 7 步即可达到要求。

最优化理论与算法 B(数值优化) ––第 5 章部分作业参考解答

5.1 考虑将精确步长的最速下降法应用于二次函数，即

$$f(\mathbf{x}) = \frac{1}{2}\mathbf{x}^{\mathrm{T}}\mathbf{G}\mathbf{x} - \mathbf{b}^{\mathrm{T}}\mathbf{x},$$

其中 $\mathbf{G}$ 对称半正定，最优解 $\mathbf{x}^* = \mathbf{G}^{-1}\mathbf{b}$，利用最速下降法的迭代说明：如果初始点 $\mathbf{x}^{(0)}$ 满足 $\mathbf{x}^{(0)} - \mathbf{x}^*$ 平行于 $\mathbf{G}$ 的某个特征向量，则使用精确步长的最速下降法仅迭代一次即可找到解.

解答：因为 $\mathbf{x}^{(0)} - \mathbf{x}^*$ 与 $\mathbf{G}$ 的一个特征向量平行，所以 $\mathbf{G}(\mathbf{x}^{(0)} - \mathbf{x}^*) = \lambda(\mathbf{x}^{(0)} - \mathbf{x}^*)$， λ 为与 $\mathbf{x}^{(0)} - \mathbf{x}^*$ 平行的特征向量所属的特征值。

因为 $\mathbf{x}^*$ 为最优解，从而 $\mathbf{g}^* = \mathbf{G}\mathbf{x}^* - \mathbf{b} = \mathbf{0}$, 所以 $\mathbf{b} = \mathbf{G}\mathbf{x}^*$, 从而 $\mathbf{g}^{(0)} = \mathbf{G}\mathbf{x}^{(0)} - \mathbf{b} = \mathbf{G}(\mathbf{x}^{(0)} - \mathbf{x}^*) = \lambda(\mathbf{x}^{(0)} - \mathbf{x}^*)$.

case1： $\lambda = 0$, 则 $\mathbf{g}^{(0)} = \mathbf{0}$, 而 f 为凸函数，从而 $\mathbf{x}^{(0)}$ 即为最优解;

case2： $\lambda \neq 0$, 则 $\alpha_0 = \frac{\mathbf{g}^{(0)\mathrm{T}}\mathbf{g}^{(0)}}{\mathbf{g}^{(0)\mathrm{T}}\mathbf{G}\mathbf{g}^{(0)}} = \frac{\lambda^2(\mathbf{x}^{(0)}-\mathbf{x}^*)^{\mathrm{T}}(\mathbf{x}^{(0)}-\mathbf{x}^*)}{\lambda^3(\mathbf{x}^{(0)}-\mathbf{x}^*)^{\mathrm{T}}(\mathbf{x}^{(0)}-\mathbf{x}^*)} = \frac{1}{\lambda}$;

所以 $\mathbf{x}^{(1)} = \mathbf{x}^{(0)} + \alpha_0\mathbf{p}^{(0)} = \mathbf{x}^{(0)} - \frac{1}{\lambda}\mathbf{g}^{(0)} = \mathbf{x}^{(0)} - \frac{1}{\lambda}\lambda(\mathbf{x}^{(0)} - \mathbf{x}^*) = \mathbf{x}^*$.

所以，最速下降法将在一步内找到解。

5.2 假设我们需求极小化

$$f(x_1, x_2) = 5x_1^2 + 5x_2^2 - x_1x_2 - 11x_1 + 11x_2 + 11.$$

a) 找到一个满足一阶必要条件的解.

b) 说明该点是一个全局极小点.

c) 针对该问题最速下降法的最坏速率将会是多少？

d) 从 $(x_1, x_2) = (0, 0)$ 出发，最多需要多少步可用将目标函数值减少到 10^{-11}?

解： a) 令 $\nabla f = \begin{bmatrix} 10x_1 - x_2 - 11 \\ 10x_2 - x_1 + 11 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$, 可得满足一阶必要条件的解为 $x^* = (1, -1)$.

b) 因为 $\nabla^2 f = \begin{bmatrix} 10 & -1 \\ -1 & 10 \end{bmatrix}$ 正定，所以 f 为凸函数，从而 $x^* = (1, -1)$ 也是全局极小点。

c) $\nabla^2 f$ 的特征值为 $9, 11$ ，所以最坏速率 $v = (\frac{11-9}{11+9})^2 = 0.01$.

d) $f(x^*) = 0, \quad f(0,0) = 11,$

$$f(\mathbf{x}^{(k)}) - f(\mathbf{x}^*) \le r^k [f(x_0) - f(x^*)] \cdots\cdots (*)$$

即 $10^{-11} = 10^{-2k} \cdot 11, \quad 2k = \ln 11 + 11, \quad k = \frac{\ln 11 + 11}{2}$, 令 $(*)$ 式中等号成立, 则应取 $k = 7$, 否则可能达不到要求。所以最多需要 7 步即可达到要求。

5.7

(a) 如果 f 是连续可微的且严格凸，则

$$f(\mathbf{y}) > f(\mathbf{x}) + \nabla f(\mathbf{x})^{\mathrm{T}} (\mathbf{y} - \mathbf{x}), \quad \forall\ \mathbf{x} \ne \mathbf{y}.$$

并利用该事实证明：对于连续可微的严格凸函数， **曲率条件**(curvature condition)

$$\mathbf{s}^{(k)^{\mathrm{T}}} \mathbf{y}^{(k)} > 0$$

对任一向量 $\mathbf{x}^{(k)} \ne \mathbf{x}^{(k+1)}$ 成立.

(b) 给出一个单变量函数 f ，其满足 $f(0) = -1$ 和 $f(1) = -\frac{1}{4}$ ，并说明在这种情况下曲率条件不成立.

解. (a) 对于严格凸函数，由所给性质有

$$\begin{aligned} f^{(k)} &> f^{(k+1)} + \mathbf{g}^{(k+1)^{\mathrm{T}}} (\mathbf{x}^{(k)} - \mathbf{x}^{(k+1)}) \\ f^{(k+1)} &> f^{(k)} + \mathbf{g}^{(k)^{\mathrm{T}}} (\mathbf{x}^{(k+1)} - \mathbf{x}^{(k)}) \end{aligned}$$

将这两个不等式相加，得

$$(\mathbf{g}^{(k)} - \mathbf{g}^{(k+1)})^{\mathrm{T}} (\mathbf{x}^{(k+1)} - \mathbf{x}^{(k)}) < 0,$$

给该不等式两边同时乘以 -1 得 $\mathbf{s}^{(k)^{\mathrm{T}}} \mathbf{y}^{(k)} > 0$.

(b) 取 $f(x) = -x^2 + \frac{7}{4}x - 1$ ，则 $f'(x) = -2x + \frac{7}{4}$. 且 $f(0) = -1, f(1) = -\frac{1}{4}, f'(0) = \frac{7}{4}, f'(1) = -\frac{1}{4}$. 如果取 $x_k = 0, x_{k+1} = 1$ ，则有 $s_k^T y_k = (1-0)(-\frac{1}{4} - \frac{7}{4}) = -2 < 0$ ，所以曲率条件不成立.

5. 8 说明如果 $f(\mathbf{x})$ 是严格凸二次函数, 则函数 $h(\sigma) \stackrel{\text{def}}{=} f(\mathbf{x}^{(0)} + \sigma_1 \mathbf{p}^{(0)} + \cdots + \sigma_{k-1} \mathbf{p}^{(k-1)})$ 也是关于 $\sigma = (\sigma_0, \cdots, \sigma_{k-1})$ 的严格凸函数.

解答：因为 $f(\mathbf{x})$ 是严格凸二次函数，所以海森矩阵为常矩阵且正定，设为 $\mathbf{G}$. 令 $\mathbf{P}=\begin{bmatrix}\mathbf{p}^{(0)} & \mathbf{p}^{(1)} & \cdots & \mathbf{p}^{(k-1)}\end{bmatrix}$, 则 $h(\sigma)=f(\mathbf{P}\sigma+\mathbf{x}^{(0)})$ ，且

$$\nabla^2 h=\mathbf{P}^{\mathrm{T}}\mathbf{G}\mathbf{P}=\begin{bmatrix}{\mathbf{p}^{(0)}}^{\mathrm{T}}\mathbf{G}\mathbf{p}^{(0)} & \cdots & 0\\ 0 & {\mathbf{p}^{(1)}}^{\mathrm{T}}\mathbf{G}\mathbf{p}^{(1)} & 0\\ \cdots & \cdots & \cdots\\ 0 & \cdots & {\mathbf{p}^{(k-1)}}^{\mathrm{T}}\mathbf{G}\mathbf{p}^{(k-1)}\end{bmatrix},$$

又因为 $\mathbf{G}$ 正定，所以 ${\mathbf{p}^{(i)}}^{\mathrm{T}}\mathbf{G}\mathbf{p}^{(i)}>0,\quad i=0,1,\cdots,k-1$. 从而 $\nabla^2 h$ 也正定，故 $h(\sigma)$ 也是关于 σ 的严格凸函数。

5.9 利用 CG 的迭代形式直接证明共轭梯度法的性质对 $k=1$ 成立.

证明：由线性共轭梯度法，有 $\mathbf{g}^{(0)}=\mathbf{G}\mathbf{x}^{(0)}-\mathbf{b},\quad \mathbf{p}^{(0)}=-\mathbf{g}^{(0)}$,

$$\alpha_0=\frac{{\mathbf{g}^{(0)}}^{\mathrm{T}}\mathbf{g}^{(0)}}{{\mathbf{p}^{(0)}}^{\mathrm{T}}\mathbf{G}\mathbf{p}^{(0)}},\ \mathbf{x}^{(1)}=\mathbf{x}^{(0)}+\alpha_0\mathbf{p}^{(0)},\ \mathbf{g}^{(1)}=\mathbf{g}^{(0)}+\alpha_0\mathbf{G}\mathbf{p}^{(0)},$$

$$\beta_0=\frac{{\mathbf{g}^{(1)}}^{\mathrm{T}}\mathbf{g}^{(1)}}{{\mathbf{g}^{(0)}}^{\mathrm{T}}\mathbf{g}^{(0)}},\quad \mathbf{p}^{(1)}=-\mathbf{g}^{(1)}+\beta_0\mathbf{p}^{(0)}.$$

1) 将 β_0 代入，得

的选取即可知 $p_1^T A p_0=0$.

2) ${\mathbf{g}^{(1)}}^{\mathrm{T}}\mathbf{g}^{(0)}=(\mathbf{g}^{(0)}+\alpha_0\mathbf{G}\mathbf{p}^{(0)})^{\mathrm{T}}\mathbf{g}^{(0)}$. 将 $\mathbf{p}^{(0)}=-\mathbf{g}^{(0)}$ 和 α_0 代入，得

${\mathbf{g}^{(1)}}^{\mathrm{T}}\mathbf{g}^{(0)}=(\mathbf{g}^{(0)}-\alpha_0\mathbf{G}\mathbf{g}^{(0)})^{\mathrm{T}}\mathbf{g}^{(0)}={\mathbf{g}^{(0)}}^{\mathrm{T}}\mathbf{g}^{(0)}-\frac{{\mathbf{g}^{(0)}}^{\mathrm{T}}\mathbf{g}^{(0)}}{{\mathbf{g}^{(0)}}^{\mathrm{T}}\mathbf{G}\mathbf{g}^{(0)}}{\mathbf{g}^{(0)}}^{\mathrm{T}}\mathbf{G}\mathbf{g}^{(0)}=0$

3) ${\mathbf{p}^{(1)}}^{\mathrm{T}}\mathbf{g}^{(1)}=(-\mathbf{g}^{(1)}+\beta_0\mathbf{p}^{(0)})^{\mathrm{T}}\mathbf{g}^{(1)}=-{\mathbf{g}^{(1)}}^{\mathrm{T}}\mathbf{g}^{(1)}-\beta_0{\mathbf{g}^{(0)}}^{\mathrm{T}}\mathbf{g}^{(1)}$. 由 2) 得 ${\mathbf{p}^{(1)}}^{\mathrm{T}}\mathbf{g}^{(1)}=-{\mathbf{g}^{(1)}}^{\mathrm{T}}\mathbf{g}^{(1)}$.

4) 因为 $\mathbf{g}^{(1)}=\mathbf{g}^{(0)}+\alpha_0\mathbf{G}\mathbf{p}^{(0)}=\mathbf{g}^{(0)}-\alpha_0\mathbf{G}\mathbf{g}^{(0)}$, 所以 $\mathbf{g}^{(1)}\in[\mathbf{g}^{(0)},\mathbf{G}\mathbf{g}^{(0)}]$, 从而 $[\mathbf{g}^{(0)},\mathbf{g}^{(1)}]\subset[\mathbf{g}^{(0)},\mathbf{G}\mathbf{g}^{(0)}]$.

又因为 $\mathbf{G}\mathbf{g}^{(0)}=\frac{\mathbf{g}^{(1)}-\mathbf{g}^{(0)}}{-\alpha_0}$, 所以 $\mathbf{G}\mathbf{g}^{(0)}\in[\mathbf{g}^{(0)},\mathbf{g}^{(1)}]$. 从而 $[\mathbf{g}^{(0)},\mathbf{g}^{(1)}]\supset[\mathbf{g}^{(0)},\mathbf{G}\mathbf{g}^{(0)}]$.

综上有 $[\mathbf{g}^{(0)},\mathbf{g}^{(1)}]=[\mathbf{g}^{(0)},\mathbf{G}\mathbf{g}^{(0)}]$.

5) 因为 $\mathbf{p}^{(0)}=-\mathbf{g}^{(0)}$, $\mathbf{p}^{(1)}=-\mathbf{g}^{(1)}+\beta_0\mathbf{p}^{(0)}$ ，所以 $\mathbf{p}^{(1)}=-\mathbf{g}^{(1)}-\beta_0\mathbf{g}^{(0)}$. 从而 $[\mathbf{p}^{(0)},\mathbf{p}^{(1)}]\supset[\mathbf{g}^{(0)},\mathbf{g}^{(1)}]$.

此外， $\mathbf{g}^{(0)}=-\mathbf{p}^{(0)},\mathbf{g}^{(1)}=\beta_0\mathbf{p}^{(0)}-\mathbf{p}^{(1)}$ ，所以 $[\mathbf{g}^{(1)},\mathbf{g}^{(0)}]\supset[\mathbf{p}^{(0)},\mathbf{p}^{(1)}]$.

综上，有 $[\mathbf{p}^{(0)},\mathbf{p}^{(1)}]=[\mathbf{g}^{(0)},\mathbf{g}^{(1)}]$. 再由 4) ，有 $[\mathbf{p}^{(0)},\mathbf{p}^{(1)}]=[\mathbf{g}^{(0)},\mathbf{G}\mathbf{g}^{(0)}]$.

最优化理论与算法 B(数值优化) ––第 6 章部分作业参考解答

6.1 设 $f(x) = 10(x_2 - x_1^2)^2 + (1 - x_1)^2$. 考虑信赖域法在点 $\mathbf{x}^{(k)} = (0, -1)$ 的信赖域子问题，并假定信赖二次模型中 $\mathbf{B}^{(k)} = \nabla^2 f(0, -1)$.

(a) 完整写出信赖域法在点 $\mathbf{x}^{(k)} = (0, -1)$ 的二次子问题，并画出子问题目标函数的等值线.

(b) 针对该子问题，画出信赖域半径从 $\Delta = 0$ 变到 $\Delta = 2$ 时，信赖域子问题解族的示意图.

(c) 对 $\Delta = 1$ ，求出该信赖域子问题的 Cauchy 点.

在点 $\mathbf{x}^{(k)} = (0, 0.5)$ 重复上面的工作.

解：$\mathbf{g}(\mathbf{x}) = \begin{bmatrix} 40x_1^3 - 40x_1x_2 + 2x_1 - 2 \\ -20x_1^2 + 20x_2 \end{bmatrix}$, $\mathbf{G}(\mathbf{x}) = \begin{bmatrix} 120x_1^2 - 40x_2 + 2 & -40x_1 \\ -40x_1 & 20 \end{bmatrix}$,
考虑信赖域法在点 $\mathbf{x}^{(k)} = (0, -1)$ 的信赖域子问题：

(a) 首先

$$f(\mathbf{x}^{(k)}) = 11, \quad \mathbf{g}^{(k)} = \begin{bmatrix} -2 \\ -20 \end{bmatrix}, \quad \mathbf{G}^{(k)} = \begin{bmatrix} 42 & 0 \\ 0 & 20 \end{bmatrix}.$$

所以原问题在 $\mathbf{x}^{(k)}$ 的局部模型，也即信赖域子问题为

$$\begin{aligned} &\text{minimize} \quad m_k(\mathbf{s}) = \tfrac{1}{2}(42s_1^2 + 20s_2^2) - 2s_1 - 20s_2 + 11 \\ &\text{subject to} \quad s_1^2 + s_2^2 \le \Delta^2. \end{aligned}$$

(b) 因为点 $\mathbf{x}^{(k)} = (0, -1)$ 处的海森矩阵正定，所以信赖域子问题的目标函数的等值线是一族同心椭圆，先求 $m_k(\mathbf{s})$ 的驻点，即得中心 $\mathbf{s}^* = (\frac{1}{21}, 1)$(以 s_1 为横坐标，以 s_2 为纵坐标). 当信赖域半径从 $\Delta = 0$ 变到 $\Delta = 2$ 时，可用图解法求解信赖域子问题，有三种情况：

I) 当 $\Delta = 0$ 时，解为 $\mathbf{s}^{(k)} = (0, 0)$;

II) 当 $0 < \Delta < \|\mathbf{s}^*\|_2 = \sqrt{1 + \frac{1}{21^2}}$ 时，解 $\mathbf{s}^{(k)}$ 是以 $(0, 0)$ 为中心，以 Δ 为半径的圆与 $m_k(\mathbf{s})$ 等值线的切点;

III) 当 $\Delta \geq \|\mathbf{s}^*\|_2$ 时，信赖域子问题的解即为 $\mathbf{s}^*$.

从而，当信赖域半径从 $\Delta = 0$ 变到 $\Delta = 2$ 时，信赖域子问题解族的示意图是一条连接 $(0,0)$ 与 $\mathbf{s}^*$ 的曲线，为了示意图更准确，取 $\Delta = 1$，用图解法得到子问题的解，三点就可以较精确地确定出该曲线.

(c) 该点处的最速下降方向 $-\mathbf{g}^{(k)} = (2,20)$，信赖域子问题的 Cauchy 点即 $m_k(\mathbf{s})$ 沿最速下降方向，且限制在信赖域内的极小点，可用图解法求，也可用解析法求。解析法即求解问题 $\text{minimize}_{\|-\alpha\mathbf{g}^{(k)}\|\leq 1} m_k(-\alpha\mathbf{g}^{(k)})$. 将 $\mathbf{g}^{(k)}$ 代入，即

$$\text{minimize}_{0\leq\alpha\leq 1/\sqrt{2+20^2}} 4084\alpha^2 - 404\alpha + 11$$

求得目标函数的无约束极小点是 $\alpha^* = \frac{404}{2\times 4084} = 0.0495$，其显然在约束区间 $[0, 1/\sqrt{2+20^2}] = [0, 0.0499]$. 故该信赖域子问题的 Cauchy 点 $\mathbf{s}_{\mathrm{C}}^{(k)} = -\alpha^*\mathbf{g}^{(k)} = \frac{404}{2\times 4084}(2,20) = \frac{101}{1021}(1,10)$.

最优化理论与算法 B： 第 8 章部分作业参考解答

8.3 考虑下面的问题

$$\begin{array}{ll} \text{minimum}_{\mathbf{x}\in\mathbb{R}^2} & (x_1-\frac{9}{4})^2+(x_2-2)^2 \\ \text{subject to} & -x_1^2+x_2\geq 0 \\ & x_1+x_2\leq 6 \\ & x_1\geq 0, x_2\geq 0. \end{array}$$

(a) 写出 KKT 最优性条件，并验证点 $\mathbf{x}^*=(\frac{3}{2},\frac{9}{4})$ 满足这些条件.

(b) 给出 $\mathbf{x}^*$ 处 KKT 条件的一个几何解释.

(c) 证明 $\mathbf{x}^*$ 是该问题的最优解.

解:

(a) $\mathcal{L}(\mathbf{x},\lambda)=(x_1-\frac{9}{4})^2+(x_2-2)^2-\lambda_1(-x_1^2+x_2)-\lambda_2(6-x_1-x_2)-\lambda_3x_1-\lambda_4x_2.$

$$\nabla_x\mathcal{L}(\mathbf{x},\lambda)=\begin{bmatrix}2(x_1-\frac{9}{4})+2\lambda_1x_1+\lambda_2-\lambda_3\\ 2(x_2-2)-\lambda_1+\lambda_2-\lambda_4\end{bmatrix}.$$

KKT 条件:

$$\begin{array}{rl} 2(1+\lambda_1)x_1+\lambda_2-\lambda_3-\frac{9}{2} & =0 \\ 2x_2-\lambda_1+\lambda_2-4 & =0 \\ -x_1^2+x_2 & \geq 0 \\ 6-x_1-x_2 & \geq 0 \\ x_1\geq 0, x_2 & \geq 0 \\ \lambda_1(-x_1^2+x_2) & =0 \\ \lambda_2(6-x_1-x_2) & =0 \\ \lambda_3x_1=0, \lambda_4x_2 & =0 \\ \lambda_i\geq 0, i=1,2, & 3,4. \end{array}$$

对 $\mathbf{x}^*=(\frac{3}{2},\frac{9}{4})$，有 $\mathcal{I}(\mathbf{x}^*)=\{1\}$. 代入上边的条件，得 $\lambda_1=\frac{1}{2},\lambda_2=\lambda_3=\lambda_4=0$ 符合条件. 从而 $\mathbf{x}^*$ 满足 KKT 条件.

(b) 在 $\mathbf{x}^*$ 处，积极约束为 $c_1(\mathbf{x})=-x_1^2+x_2$，在 $\mathbf{x}^*$ 处的梯度为 $\nabla c_1(\mathbf{x}^*)=(-3,1)$，而 $\nabla f(\mathbf{x}^*)=(-3/2,1/2)$，易见 $\nabla f(\mathbf{x}^*)=\frac{1}{2}\nabla c_1(\mathbf{x}^*)$. 所以 $\mathbf{x}^*$ 处目标函数与积极约束的梯度共线，且同方向.

(c) 目标函数显然为凸函数. $c_1(\mathbf{x})=-x_1^2+x_2$ 为凹函数，易见其余为线性约束，从而也是凸函数. 所以此问题为凸规划. 从而 KKT 点即是全局最优解. 所以 $\mathbf{x}^*$ 是该问题的最优解.

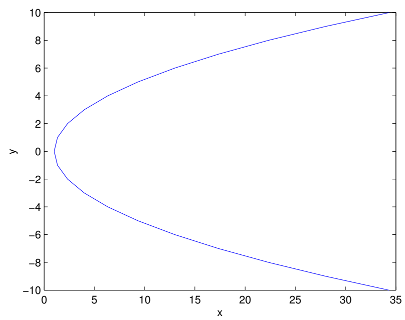
8.6 考虑找抛物线上哪个点离原点最近 (在 Euclidean 范数意义下) 的问题. 我们可以将该问题表述为

$$\text{minimize } f(x,y) = x^2 + y^2 \quad \text{subject to } \ (x-1)^3 = y^2.$$

(a) 图解法求解该问题. 消去 y 求解问题，所得到的函数有极小点吗？给出结果不一致的可能原因。消去 x 求解问题会得到怎样的结果？

(b) 对于该问题，找到所有的 KKT 点. LICQ 成立吗？这些点中的那些是解？

解： (a) 图解法：该问题的可行集如下图 所示， 目标函数的等值线是以圆



点为中心的同心圆，故问题的最优解 $x^* = 1, y^* = 0$.

用 $y^2 = (x-1)^3$ 代入，得到 $\min x^2 + (x-1)^3$，此时问题无界；这种消元法在将 $y^2 = (x-1)^3$ 代入时忽略了 $x \geq 1$ 这个暗含的条件.

用 $x = y^{2/3} + 1$ 代入，得到 $\min(y^{2/3}+1)^2 + y^2$, 得到最优解 $y^* = 0$，于是原问题的最优解 $x^* = 1, y^* = 0$.

(b) 该问题的 Lagrange 函数为 $\mathcal{L}(x,y;\lambda) = x^2 + y^2 - \lambda((x-1)^3 - y^2)$.

KKT 条件为：

$$\begin{aligned} 2x - 3\lambda(x-1)^2 &= 0 \\ 2y + 2\lambda y &= 0 \\ (x-1)^3 - y^2 &= 0. \end{aligned}$$

由第二个方程入手，分情况讨论该方程组的解。

$y = 0$, 则 $x = 1$, 此时第一个方程 $2x - 3\lambda(x-1)^2 = 0$ 不满足；

$\lambda = -1$, 此时 $2x + 3(x-1)^2 = 0$ 无解。

所以该问题没有 KKT 点！

在该问题的最优解 $x^* = 1, y^* = 0$ 处, 约束梯度 $\mathbf{a} = (0,0)$, 不满足 LICQ，故不能保证 $x^* = 1, y^* = 0$ 一定满足 KKT 条件。

注记：该问题的 (a) 说明消元时需要谨慎，否则可能会忽略掉一些隐含条件；从而得出错误的结论； (b) 说明约束规范的重要性，即对于一个优化问

题，当最优解处约束规范不成立时，即使是最优解，也可能不满足 KKT 条件。

8.7 考虑问题

$$\min_{\mathbf{x}\in\mathbb{R}^2} f(\mathbf{x}) = -2x_1 + x_2 \quad \text{subject to} \quad \begin{cases} (1-x_1)^3 - x_2 & \geq & 0. \\ x_2 + 0.25x_1^2 - 1 & \geq & 0. \end{cases}$$

最优解是 $\mathbf{x}^* = (0,1)$, 其中两个约束都是积极的. LICQ 在该点成立吗？KKT 条件满足吗？

解. $\nabla f(\mathbf{x}) = \begin{bmatrix} -2 \\ 1 \end{bmatrix}$, $\nabla c_1(\mathbf{x}) = \begin{bmatrix} -3(1-x_1)^2 \\ -1 \end{bmatrix}$, $\nabla c_2(\mathbf{x}) = \begin{bmatrix} 0.5x_1 \\ 1 \end{bmatrix}$.

在 $x^* = (0,1)$ 处，有

$$\nabla c_1(\mathbf{x}^*) = \begin{bmatrix} -3 \\ -1 \end{bmatrix}, \quad \nabla c_2(\mathbf{x}^*) = \begin{bmatrix} 0 \\ 1 \end{bmatrix},$$

易见 $\nabla c_1(\mathbf{x}^*)$ 和 $\nabla c_2(\mathbf{x}^*)$ 线性无关，所以 LICQ 成立.

此外，对于 $\mathbf{x}^* = (0,1)$ ，有 $\mathcal{I}(\mathbf{x}^*) = \{1,2\}$ ，易见 $\nabla f(\mathbf{x}^*) = \frac{2}{3}\nabla c_1(\mathbf{x}^*) + \frac{5}{3}\nabla c_2(\mathbf{x}^*)$. 所以 $\mathbf{x}^*$ 满足 KKT 条件.

8.9 利用一阶和二阶最优性条件找到函数 $f(\mathbf{x}) = x_1x_2$ 在单位圆 $x_1^2 + x_2^2 = 1$ 上的极小值点. 从几何上说明该问题.

解. 该问题的 Lagrange 函数为

$$\mathcal{L}(\mathbf{x},\lambda) = x_1x_2 - \lambda(x_1^2 + x_2^2 - 1),$$

KKT 条件为：

$$\begin{aligned} x_2 - 2\lambda x_1 &= 0 \\ x_1 - 2\lambda x_2 &= 0 \\ x_1^2 + x_2^2 &= 1. \end{aligned}$$

解上述系统，得 KKT 点：

$$\begin{array}{ll} \mathbf{x}^{(1)} = (\frac{\sqrt{2}}{2}, \frac{\sqrt{2}}{2}), & \lambda^{(1)} = \frac{1}{2}, \\ \mathbf{x}^{(2)} = (-\frac{\sqrt{2}}{2}, -\frac{\sqrt{2}}{2}), & \lambda^{(2)} = \frac{1}{2}, \\ \mathbf{x}^{(3)} = (\frac{\sqrt{2}}{2}, -\frac{\sqrt{2}}{2}), & \lambda^{(3)} = -\frac{1}{2}, \\ \mathbf{x}^{(4)} = (-\frac{\sqrt{2}}{2}, \frac{\sqrt{2}}{2}), & \lambda^{(4)} = -\frac{1}{2}, \end{array}$$

对 $i = 1$ ，计算易得：

$$\nabla^2_{xx}\mathcal{L}(\mathbf{x}^{(1)}, \lambda^{(1)}) = \begin{bmatrix} -1 & 1 \\ 1 & -1 \end{bmatrix},$$

$$G(\mathbf{x}^{(1)}, \lambda^{(1)}) = \{\mathbf{p} \in \mathbb{R}^2 \mid \mathbf{p}^{\mathrm{T}}\nabla c(\mathbf{x}^{(1)}) = 0\} = \{k(1,-1) \mid k \in \mathbb{R}\}.$$

因为对 $k \neq 0$ ，有 $k^2(1,-1)^{\mathrm{T}}\nabla^2_{xx}\mathcal{L}(\mathbf{x}^{(1)},\lambda^{(1)})(1,-1) = -4k^2 < 0$ ，所以，在 $\mathbf{x}^{(1)}$ 处不满足二阶必要条件．所以 $\mathbf{x}^{(1)}$ 不是局部解．同理可验证 $\mathbf{x}^{(2)}$ 不是局部解．

对 $i=3$ ，计算易得：

$$\nabla^2_{xx}\mathcal{L}(\mathbf{x}^{(3)},\lambda^{(3)}) = \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix},$$

$$G(\mathbf{x}^{(3)},\lambda^{(3)}) = \{\mathbf{p}\in\mathbb{R}^2 \mid \mathbf{p}^{\mathrm{T}}\nabla c(\mathbf{x}^{(3)}) = 0\} = \{k(1,1) \mid k\in\mathbb{R}\}.$$

因为对 $k \neq 0$ ，有 $k^2(1,1)^T\nabla^2_{xx}\mathcal{L}(\mathbf{x}^{(3)},\lambda^{(3)})(1,1) = 4k^2 > 0$ ，所以，在 $\mathbf{x}^{(3)}$ 处满足二阶充分条件．所以 $\mathbf{x}^{(3)}$ 是局部解．同理可验证 $\mathbf{x}^{(4)}$ 是局部解．

目标函数的等值线是双曲线，以横轴和纵轴为渐近线．且在第 I 和第 III 卦限目标函数的值趋于正无穷；在第 II 和第 IV 卦限目标函数的值趋于负无穷．

8.10 设 K 是 $\mathbb{R}^n$ 中的凸集，设 $f(\cdot): K \to \mathbb{R}$. 证明 f 是 K 上的凸函数当且仅当对任一整数 $k \geq 2$ ，下面结论成立：

$$\lambda_1,\cdots,\lambda_k \geq 0, \sum_{j=1}^{k}\lambda_j = 1, \mathbf{x}_1,\cdots,\mathbf{x}_k \in K \Rightarrow f\left(\sum_{j=1}^{k}\lambda_j\mathbf{x}_j\right) \leq \sum_{j=1}^{k}\lambda_j f(\mathbf{x}_j)$$

证明．“必要性”：用数学归纳法证．若 f 是 S 上的凸函数，结论对 $k=2$ 显然成立．假设 $k=n$ 时结论成立．以下我们证明对于 $k=n+1$ 时结论仍然成立．

对任意的 $\mathbf{x}_1,\ldots,\mathbf{x}_{n+1}\in S$, $\lambda_1,\ldots,\lambda_{n+1}\geq 0$ 且 $\sum_{j=1}^{n+1}\lambda_j = 1$. 若 $\lambda_{n+1} = 1$, 此时结论显然成立．若 $0 \leq \lambda_{n+1} < 1$,　则

$$\begin{aligned}
f(\sum_{j=1}^{n+1}\lambda_j\mathbf{x}_j) &= f\left((1-\lambda_{n+1})\sum_{j=1}^{n}\frac{\lambda_j}{1-\lambda_{n+1}}\mathbf{x}_j + \lambda_{n+1}\mathbf{x}_{n+1}\right) \\
&\leq (1-\lambda_{n+1})f\left(\sum_{j=1}^{n}\frac{\lambda_j}{1-\lambda_{n+1}}\mathbf{x}_j\right) + \lambda_{n+1}f(\mathbf{x}_{n+1})(\text{ 凸函数定义}) \\
&\leq (1-\lambda_{n+1})\sum_{j=1}^{n}\frac{\lambda_j}{1-\lambda_{n+1}}f(\mathbf{x}_j) + +\lambda_{n+1}f(\mathbf{x}_{n+1}) \\
&= \sum_{j=1}^{n+1}\lambda_j f(\mathbf{x}_j)
\end{aligned}$$

其中第二个不等式是因为 $\frac{\lambda_j}{1-\lambda_{n+1}} \geq 0$ ，且 $\sum_{j=1}^{n}\frac{\lambda_j}{1-\lambda_{n+1}} = 1$ ，从而由归纳假设 $k=n$ 时结论成立得到的．由归纳法，必要性得证．

“充分性”：取 $k=2$, 即为凸函数的定义.

8.14 写出问题

$$\begin{array}{ll} \text{minimize} & \frac{1}{2}\sigma x_1^2+\frac{1}{2}x_2^2+x_1 \\ \text{subject to} & x_1\geq 0 \end{array}$$

的 Lagrange 对偶问题。分 $\sigma=1$ 和 $\sigma=-1$ 两种情况讨论对偶问题的解是否是原问题最优解处的 Lagrange 乘子，并解释两种情况下的结果。

解：问题的 Lagrange 对偶为

$$\max_{\lambda\geq 0}\min_{\mathbf{x}\in\mathbb{R}^2}[\frac{1}{2}\sigma x_1^2+\frac{1}{2}x_2^2+x_1-\lambda x_1].$$

若 $\sigma=1$, 则 $\min\limits_{\mathbf{x}\in\mathbb{R}^2}[\frac{1}{2}\sigma x_1^2+\frac{1}{2}x_2^2+x_1-\lambda x_1]$ 的最优解为 $\mathbf{x}^*(\lambda)=(\lambda-1,0)$ 。所以对偶问题即 $\max\limits_{\lambda\geq 0}-\frac{1}{2}(\lambda-1)^2$ ，其解为 $\lambda^*=1$.

该问题的 KKT 条件即

$$\begin{bmatrix} x_1+1 \\ x_2 \end{bmatrix}-\begin{bmatrix} \lambda \\ 0 \end{bmatrix}=0,$$

$$x_1\geq 0,$$

$$x_1\lambda=0.$$

解之，得到得到问题的最优解 $\mathbf{x}^*=(0,0)$ ， Lagrange 乘子为 1. 可见对偶问题的解是原问题最优解处得 Lagrange 乘子。

同样的步骤可知 $\sigma=-1$ 时对于任何 λ 值，对偶问题目标值都是 $-\infty$. 没有最优解。原问题无界。

两种情况下结果的解释：当 $\sigma=1$ 时，该问题是凸规划，且满足 Slater 约束规范，故强对偶定理成立。这时，对偶问题的最优解即原问题的 Lagrange 乘子，且在 Lagrange 函数中固定对偶最优解 λ^* ，所得函数的无约束极小点是原问题的最优解，即 $\mathbf{x}^*=\mathbf{x}^*(\lambda^*)=(0,0)$.

8.15 针对 (方) 阵 $\mathbf{M}\in\mathbb{R}^{n\times n}$ ，定义迹 $\text{trace}(\mathbf{M})=\sum_{j=1}^n m_{jj}$ ，并对两个矩阵 $\mathbf{A},\mathbf{B}\in\mathbb{R}^{k\times l}$ 定义

$$\mathbf{A}\bullet\mathbf{B}:=\sum_{i=1}^k\sum_{j=1}^l a_{ij}b_{ij},$$

其中 a_{ij},b_{ij} 分别为 $\mathbf{A}$ 和 $\mathbf{B}$ 的第 (i,j) 个元素．证明

(a) $\mathbf{A}\bullet\mathbf{B}=\text{trace}(\mathbf{A}^{\text{T}}\mathbf{B})$.

(b) $\text{trace}(\mathbf{MN})=\text{trace}(\mathbf{NM})$.

证明： (a)

$$\mathbf{A}^{\mathrm{T}}\mathbf{B} = \begin{bmatrix} a_{11} & a_{21} & \cdots & a_{k1} \\ a_{12} & a_{22} & \cdots & a_{k2} \\ \cdots & \cdots & \cdots & \\ a_{1l} & a_{2l} & \cdots & a_{kl} \end{bmatrix} \begin{bmatrix} b_{11} & b_{12} & \cdots & b_{1l} \\ b_{21} & b_{22} & \cdots & a_{2l} \\ \cdots & \cdots & \cdots & \\ a_{k1} & a_{k2} & \cdots & a_{kl} \end{bmatrix}$$

$$= \begin{bmatrix} a_{11}b_{11}+a_{21}b_{21}+\cdots+a_{k1}b_{k1} & \cdots & \cdots & \cdots \\ \cdots & a_{12}b_{12}+\cdots+a_{k2}b_{k2} & \cdots & \cdots \\ \cdots & \cdots & \cdots & \cdots \\ \cdots & \cdots & \cdots & a_{1l}b_{1l}+\cdots+a_{kl}b_{kl} \end{bmatrix}.$$

所以

$$\begin{aligned} \mathrm{trace}(\mathbf{A}^{\mathrm{T}}\mathbf{B}) &= a_{11}b_{11}+\cdots+a_{k1}b_{k1}+a_{12}b_{12}+\cdots+a_{k2}b_{k2}+\cdots+a_{1l}b_{1l}+\cdots+a_{kl}b_{kl} \\ &= \Sigma_{i=1}^{k}\Sigma_{j=1}^{l}a_{ij}b_{ij} \\ &= \mathbf{A}\bullet\mathbf{B}. \end{aligned}$$

(b) $\mathrm{trace}(\mathbf{MN}) = \Sigma_{i=1}^{n}\Sigma_{j=1}^{n}m_{ij}n_{ji} = \Sigma_{i=1}^{n}\Sigma_{j=1}^{n}n_{ij}m_{ji} = \mathrm{trace}(\mathbf{NM})$.

8.17 考虑矩阵

$$\mathbf{M} = \begin{bmatrix} \mathbf{A} & \mathbf{B} \\ \mathbf{B}^{\mathrm{T}} & \mathbf{C} \end{bmatrix},$$

其中 $\mathbf{A},\mathbf{C}$ 是对称矩阵, 且 $\mathbf{A}$ 是正定的. 证明 $\mathbf{M}\succeq\mathbf{0}$ 当且仅当 $\mathbf{C}-\mathbf{B}^{\mathrm{T}}\mathbf{A}^{-1}\mathbf{B}\succeq\mathbf{0}$.

解：设 $\mathbf{M}\in\mathbb{R}^{n\times n},\mathbf{C}\in\mathbb{R}^{m\times m}$ ，考虑非奇异矩阵 $\mathbf{T} = \begin{bmatrix} \mathbf{I} & -\mathbf{A}^{-1}\mathbf{B} \\ \mathbf{0} & \mathbf{I} \end{bmatrix}$, 则 $\mathbf{T}^{\mathrm{T}}\mathbf{M}\mathbf{T} = \begin{bmatrix} \mathbf{A} & 0 \\ 0 & \mathbf{C}-\mathbf{B}^{\mathrm{T}}\mathbf{A}^{-1}\mathbf{B} \end{bmatrix}$. 即对 $\mathbf{M}$ 作初等行变换和相应的初等列变换，即其化成对角块矩阵.

“ $\Rightarrow$ ” 设 $\mathbf{M}\succeq\mathbf{0}$. 则 $\forall\mathbf{x}\in\mathbb{R}^{m}$, 令 $\mathbf{y} = \begin{bmatrix} \mathbf{0} \\ \mathbf{x} \end{bmatrix}$ ，则

$$\begin{aligned} & \mathbf{x}^{\mathrm{T}}(\mathbf{C}-\mathbf{B}^{\mathrm{T}}\mathbf{A}^{-1}\mathbf{B})\mathbf{x} \\ = \ & [\mathbf{0}\ \mathbf{x}^{\mathrm{T}}] \begin{bmatrix} \mathbf{A} & 0 \\ 0 & \mathbf{C}-\mathbf{B}^{\mathrm{T}}\mathbf{A}^{-1}\mathbf{B} \end{bmatrix} \begin{bmatrix} \mathbf{0} \\ \mathbf{x} \end{bmatrix} \\ = \ & \mathbf{y}^{\mathrm{T}}\mathbf{T}^{\mathrm{T}}\mathbf{M}\mathbf{T}\mathbf{y} \\ = \ & (\mathbf{T}\mathbf{y})^{\mathrm{T}}\mathbf{M}(\mathbf{T}\mathbf{y}) \geq 0, \end{aligned}$$

最后一个不等式是因为 $\mathbf{M}\succeq\mathbf{0}$ 。故 $\mathbf{C}-\mathbf{B}^{\mathrm{T}}\mathbf{A}^{-1}\mathbf{B}\succeq\mathbf{0}$.

“ $\Leftarrow$ ” 设 $\mathbf{C}-\mathbf{B}^{T}\mathbf{A}^{-1}\mathbf{B}\succeq\mathbf{0}$, $\mathbf{A}\succeq\mathbf{0}$. 则 $\forall\mathbf{y}\in\mathbb{R}^{n\times n}$, 记 $\mathbf{T}^{-1}\mathbf{y} = \begin{bmatrix} \mathbf{x}_1 \\ \mathbf{x}_2 \end{bmatrix}$, 其中 $\mathbf{x}_1\in\mathbb{R}^{n-m},\mathbf{x}_2\in\mathbb{R}^{m}$. 于是

$$\begin{aligned} \mathbf{y}^{\mathrm{T}}\mathbf{M}\mathbf{y} = \ & \mathbf{y}^{\mathrm{T}}\mathbf{T}^{-\mathrm{T}} \begin{bmatrix} \mathbf{A} & 0 \\ 0 & \mathbf{C}-\mathbf{B}^{\mathrm{T}}\mathbf{A}^{-1}\mathbf{B} \end{bmatrix} \mathbf{T}^{-1}\mathbf{y} \\ & [\mathbf{x}_1^{\mathrm{T}}\quad\mathbf{x}_2] \begin{bmatrix} \mathbf{A} & 0 \\ 0 & \mathbf{C}-\mathbf{B}^{\mathrm{T}}\mathbf{A}^{-1}\mathbf{B} \end{bmatrix} \begin{bmatrix} \mathbf{x}_1 \\ \mathbf{x}_2 \end{bmatrix} \\ & = \mathbf{x}_1^{\mathrm{T}}\mathbf{A}\mathbf{x}_1 + \mathbf{x}_2^{\mathrm{T}}\mathbf{C}-\mathbf{B}^{\mathrm{T}}\mathbf{A}^{-1}\mathbf{B}\mathbf{x}_2 \end{aligned}$$

最后一个不等式是因为 $\mathbf{C}-\mathbf{B}^{T}\mathbf{A}^{-1}\mathbf{B}\succeq\mathbf{0}$, $\mathbf{A}\succeq\mathbf{0}$ 正定。故 $\mathbf{M}\succeq\mathbf{0}$.

最优化理论与算法 B： 第 9 章部分作业参考解答

9.1 下面的问题有解吗？给出解释.

$$\begin{array}{llll}(a) & \text{minimize } x_1+x_2 & \text{subject to} & x_1^2+x_2^2=2, 0\le x_1\le 1, 0\le x_2\le 1;\\ (b) & \text{minimize } x_1+x_2 & \text{subject to} & x_1^2+x_2^2\le 1, x_1+x_2=3;\\ (c) & \text{minimize } x_1x_2 & \text{subject to} & x_1^2+x_2^2=2.\end{array}$$

解： (a) 只有一个可行点 $(1,1)$, 从而 (a) 有解，且为 $(1,1)$.

(b) 的可行集为 ϕ, 从而 (b) 无解.

在 (c) 中，令 $x_1=\sqrt{2}\cos\theta,\quad x_2=\sqrt{2}\sin\theta,\quad \theta\in[0,2\pi]$, 则

$$x_1x_2 = 2\cos\theta\sin\theta=\sin(2\theta)$$

显然当 $\theta=\frac{1}{4}\pi$ 或 $\frac{7}{4}\pi$ 时， x_1x_2 最小，且为 -1. 所以 (c) 有解 $x^*=(1,-1)$ 或 $(-1,1)$.

9.2 考虑点 $\mathbf{x}^{(0)}$ 到多面集 $\{\mathbf{x}\mid \mathbf{A}^{\mathrm{T}}\mathbf{x}=\mathbf{b}\}$(其中 $\mathbf{A}$ 列满秩) 的最短距离问题，该问题可以表述为二次规划

$$\begin{array}{ll}\text{minimize} & \frac{1}{2}(\mathbf{x}-\mathbf{x}^{(0)})^{\mathrm{T}}(\mathbf{x}-\mathbf{x}^{(0)}),\\ \text{subject to} & \mathbf{A}^{\mathrm{T}}\mathbf{x}=\mathbf{b}.\end{array}$$

证明最优乘子

$$\boldsymbol{\lambda}^*=(\mathbf{A}^{\mathrm{T}}\mathbf{A})^{-1}(\mathbf{b}-\mathbf{A}^{\mathrm{T}}\mathbf{x}^{(0)}),$$

最优解

$$\mathbf{x}^*=\mathbf{x}^{(0)}+\mathbf{A}(\mathbf{A}^{\mathrm{T}}\mathbf{A})^{-1}(\mathbf{b}-\mathbf{A}^{\mathrm{T}}\mathbf{x}^{(0)}).$$

说明当 $\mathbf{A}=\mathbf{a}$ 是一个列向量的特殊情况下，从 $\mathbf{x}^{(0)}$ 到超平面 $\mathbf{a}^{\mathrm{T}}\mathbf{x}=b$ 的最短距离是

$$\frac{|b-\mathbf{a}^{\mathrm{T}}\mathbf{x}^{(0)}|}{\|\mathbf{a}\|}.$$

证明：

$$\mathcal{L}(x,\lambda)=\frac{1}{2}x^Tx-x_0^Tx+\frac{1}{2}x_0^Tx_0-\lambda^T(Ax-b)$$

$$\nabla_x\mathcal{L}(x,\lambda)=x-x_0-A^T\lambda$$

KKT 条件是：

$$\begin{aligned} x - x_0 - A^T\lambda &= 0, \\ Ax - b &= 0. \end{aligned}$$

因为 A 行满秩，故该条件的第一个方程等价于 $Ax - Ax_0 - AA^T\lambda = 0$. 将 $Ax = b$ 代入, 得 $AA^T\lambda = b - Ax_0$. 再由 A 行满秩, 有 AA^T 可逆. 从而最优乘子 $\lambda^* = (AA^T)^{-1}(b - Ax_0)$. 最优解 $x^* = x_0 + A^T\lambda^* = x_0 + A^T(AA^T)^{-1}(b - Ax_0)$.

特别地，当 A 为行向量时，$\{x \mid Ax = b\}$ 表示超平面. 此时 $AA^T = \|A\|_2^2$. 所以 $x^* = x_0 + \frac{b - Ax_0}{\|A\|^2}A^T$. x_0 到超平面 $\{x \mid Ax = b\}$ 的最短距离为：

$$\|x^* - x_0\|_2 = \sqrt{(x^* - x_0)^T(x^* - x_0)} = \sqrt{\frac{(b - Ax_0)^2}{\|A\|_2^4}AA^T} = \frac{|b - Ax_0|}{\|A\|}.$$

9.3 考虑例 9.2

$$\begin{array}{ll} \text{minimize}_{\mathbf{x}\in\mathbb{R}^2} & q(\mathbf{x}) = (x_1 - 1)^2 + (x_2 - 2.5)^2, \\ \text{subject to} & x_1 - 2x_2 + 2 \geq 0, \\ & -x_1 - 2x_2 + 6 \geq 0, \\ & -x_1 + 2x_2 + 2 \geq 0, \\ & x_1 \geq 0, \\ & x_2 \geq 0. \end{array}$$

以 $\mathbf{x}^{(0)} = (2, 0), \mathcal{W}_0\{3, 5\}$ 的迭代结果如下：

k	$\mathbf{x}^{(k)}$	$\mathcal{W}_k$	$\mathbf{s}^{(k)}$	α_k	p	$\boldsymbol{\lambda}^{(k)}$	q
0	$(2, 0)$	$\{3, 5\}$	$(0, 0)$			$(-2, -1)$	3
1	$(2, 0)$	$\{5\}$	$(-1, 0)$	1			
2	$(1, 0)$	$\{5\}$	$(0, 0)$			-5	5
3	$(1, 0)$	$\emptyset$	$(0, 2.5)$	0.6	1		
4	$(1, 1.5)$	$\{1\}$	$(0.4, 0.2)$	1			
5	$(1.4, 1.7)$	$\{1\}$	$(0, 0)$			0.8	

保持初始点不变，工作集依次为 $\mathcal{W}_0 = \{3\}$， $\mathcal{W}_0 = \{5\}$ 和 $\mathcal{W}_0 = \emptyset$，完成积极集法的前两次迭代.

$\mathcal{W}_0 = \{5\}$

k	$\mathbf{x}^{(k)}$	$\mathcal{W}_k$	$\mathbf{s}^{(k)}$	α_k	p	$\boldsymbol{\lambda}^{(k)}$	q
0	$(2, 0)$	$\{5\}$	$(-1, 0)$	1			
1	$(1, 0)$	$\{5\}$	$(0, 0)$			-5	5
2	$(1, 0)$	$\emptyset$	$(0, 2.5)$	0.6	1		
3	$(1, 1.5)$	$\{1\}$	$(0.4, 0.2)$	1			
4	$(1.4, 1.7)$	$\{1\}$	$(0, 0)$			0.8	

$\mathcal{W}_0 = \{3\}$

k	$\mathbf{x}^{(k)}$	$\mathcal{W}_k$	$\mathbf{s}^{(k)}$	α_k	p	$\boldsymbol{\lambda}^{(k)}$	q
0	$(2, 0)$	$\{3\}$	$(0.2, 0.1)$	1			
1	$(2.2, 0.1)$	$\{3\}$	$(0, 0)$			-2.4	3
2	$(2.2, 0.1)$	$\emptyset$	$(-1.2, 2.4)$	2/3	1		
3	$(1.4, 1.7)$	$\{1\}$	$(0, 0)$			0.8	

$\mathcal{W}_0 = \emptyset$

k	$\mathbf{x}^{(k)}$	$\mathcal{W}_k$	$\mathbf{s}^{(k)}$	α_k	p	$\boldsymbol{\lambda}^{(k)}$	q
0	$(2,0)$	$\emptyset$	$(-1,2.5)$	$2/3$	1		
1	$(4/3,5/3)$	$\{1\}$	$(0.0667,0.333)$	1			
3	$(1.4,1.7)$	$\{1\}$	$(0,0)$			0.8	

9.14 用基本逐步二次规划法求解问题

$$\text{minimize } x_1 + x_2 \text{ subject to } x_2 \geq x_1^2,$$

(a) 取 $\mathbf{x}^{(0)} = \mathbf{0}, \lambda^{(0)} = 0$; 得到怎样的结果？解释方法失败的原因；

(b) 取 $\mathbf{x}^{(0)} = \mathbf{0}, \lambda^{(0)} = 1$; 得到怎样的结果？收敛快吗？

解：该问题的 Lagrange 函数

$$\mathcal{L}(\mathbf{x}, \lambda) = x_1 + x_2 - \lambda(x_2 - x_1^2).$$

对于 $\mathbf{x}^{(k)}$ ，有 $c(\mathbf{x}^{(k)}) = x_2^{(k)} - (x_1^{(k)})^2$,

$$\mathbf{a}^{(k)} = \begin{bmatrix} -2x_1^{(k)} \\ 1 \end{bmatrix}, \quad \mathbf{g}^{(k)} = \begin{bmatrix} 1 \\ 1 \end{bmatrix}, \quad \mathbf{W}^{(k)} = \nabla^2 \mathcal{L}(\mathbf{x}^{(k)}, \lambda^{(k)}) = \begin{bmatrix} 2\lambda^{(k)} & 0 \\ 0 & 0 \end{bmatrix}.$$

在 SQP 中，二次规划子问题为

$$\begin{array}{ll} \text{minimize} & \frac{1}{2}\mathbf{s}^{\mathrm{T}}\mathbf{W}^{(k)}\mathbf{s} + {\mathbf{g}^{(k)}}^{\mathrm{T}}\mathbf{s} \\ \text{subject to} & c(\mathbf{x}^{(k)}) + {\mathbf{a}^{(k)}}^{\mathrm{T}}\mathbf{s} \geq 0, \end{array}$$

当 $k = 0, \mathbf{x}^{(k)} = (0,0)$ 时，

$$c(\mathbf{x}^{(0)}) = 0, \quad \mathbf{a}^{(0)} = (0,1), \quad \mathbf{g}^{(0)} = (1,1), \quad \mathbf{W}^{(0)} = \begin{bmatrix} 2\lambda^{(0)} & 0 \\ 0 & 0 \end{bmatrix}.$$

将上述向量代入后，得

$$\begin{array}{ll} \text{minimize} & \lambda_0 s_1^2 + s_1 + s_2 \\ \text{subject to} & s_2 \geq 0. \end{array}$$

(a) 若取 $\lambda^{(0)} = 0$, 则二次规划子问题显然无界；故初值选取不当。

(b) 若取 $\lambda^{(0)} = 1$, 解上述二次规划子问题得解 $\mathbf{s}^{(0)} = (-0.5, 0)$, 对应的 Lagrange 乘子 $\lambda^{(1)} = 1$. 于是 $\mathbf{x}^{(1)} = \mathbf{x}^{(0)} + \mathbf{s}^{(0)} = (-0.5, 0)$.

第二次迭代：

$$c(\mathbf{x}^{(1)}) = -0.25, \quad \mathbf{a}^{(0)} = (-1,1), \quad \mathbf{g}^{(0)} = (1,1), \quad \mathbf{W}^{(1)} = \begin{bmatrix} 2 & 0 \\ 0 & 0 \end{bmatrix}.$$

将上述向量代入二次规划子问题后，得

$$\begin{array}{ll} \text{minimize} & s_1^2+s_1+s_2 \\ \text{subject to} & -0.25-s_1+s_2\geq 0. \end{array}$$

解上述二次规划子问题得解 $\mathbf{s}^{(1)}=(0,0.25)$, 对应的 Lagrange 乘子 $\lambda^{(2)}=1$. 于是 $\mathbf{x}^{(2)}=\mathbf{x}^{(1)}+\mathbf{s}^{(1)}=(-0.5,0.25)$.

易验证这是原问题的最优解。故从该初始点出发，由基本 SQP 法迭代两次即可得到问题的解。

9.14 考虑约束 $x_1^2+x_2^2=1$. 写出在下列点 $(0,0),(0,1),(0.1,0.02),-(0.1,0.02)$ 处的线性化约束. 解. 令 $c(x)=x_1^2+x_2^2-1, x^1=(0,0), x^2=(0,1)$ ，则 $\nabla c(x)=\begin{bmatrix}2x_1\\2x_2\end{bmatrix}$.

$$c(x^1)=-1, c(x^2)=0, \nabla c(x^1)=\begin{bmatrix}0\\0\end{bmatrix}, \nabla c(x^2)=\begin{bmatrix}0\\2\end{bmatrix},$$

所以约束 $x_1^2+x_2^2=1$ 在 x^1 处的线性化约束为 $c(x^1)+\nabla c(x_1)^T(x-x^1)=0$ ，即 $-1=0$;
约束 $x_1^2+x_2^2=1$ 在 x^2 处的线性化约束为 $c(x^2)+\nabla c(x_2)^T(x-x^2)=0$ ，即 $x_2=1$; 类似可得约束 $x_1^2+x_2^2=1$ 在其它点的线性化约束.

最优化理论与算法 B：第 10 章部分作业参考解答

10.1 考虑问题

$$\begin{array}{ll} \text{minimize} & f(x)=\frac{1}{2}(x_1^2-x_2^2)-3x_2 \\ \text{subject to} & x_2=0. \end{array}$$

(a) 计算最优解和 Lagrange 乘子.

(b) 对 $k=0,1,2$ 和惩罚参数 $\frac{1}{\mu_k}=10^{k+1}$，计算二次惩罚法迭代的点和增广 Lagrange 函数法迭代的点 (其中取 $\lambda^0=0$).

(c) 针对该问题，画出图形从几何上解释方法，并在该图形上画出这两种方法对 $k=0,1,2$ 的迭代.

(d) 假设在乘子法中将惩罚参数取为常数. 对于 μ 的哪些值，增广 Lagrange 函数将有一个极小点？对于 μ 的哪些值，方法将是收敛的？

解.

(a) 显然原问题的最优解为 $x^*=(0,0)$. 而 $\nabla f(x^*)=(0,-3),\nabla c(x^*)=(0,1)$，所以最优解 x^* 对应的 Lagrange 乘子为 $\lambda^*=-3$.

(b) **二次惩罚法**：二次惩罚函数 $P(x,\mu)=\frac{1}{2}(x_1^2-x_2^2)-3x_2+\frac{1}{2\mu}x_2^2$.
令

$$\nabla P(x,\mu)=\begin{bmatrix} x_1 \\ -x_2-3+\frac{1}{\mu}x_2 \end{bmatrix}=\begin{bmatrix} 0 \\ 0 \end{bmatrix},$$

解得 $x_1(\mu)=0,x_2(\mu)=\frac{3\mu}{1-\mu}$. 又因为 $\mu_k=10^{-(k+1)}<1$，所以

$$\nabla^2 P(x,\mu_k)=\begin{bmatrix} 1 & 0 \\ 0 & \frac{1}{\mu_k}-1 \end{bmatrix}$$

正定. 所以 $(0,\frac{3\mu_k}{1-\mu_k})$ 是 $P(x,\mu_k)$ 的极小值点.

当 $k=0$ 时，　$\mu_0=0.1$，　$x^0=(0,1/3)$ ；
当 $k=1$ 时，　$\mu_1=0.01$，　$x^1=(0,1/33)$ ；

当 $k=2$ 时， $\mu_2=0.001$, $x^2=(0,1/333)$ ；

乘子法:
增广 Lagrange 函数 $\mathcal{L}_A(x,\lambda,\mu)=\frac{1}{2}(x_1^2-x_2^2)-3x_2+\frac{1}{2\mu}x_2^2-\lambda x_2$. 令

$$\nabla\mathcal{L}_A(x,\lambda,\mu)=\begin{bmatrix} x_1 \\ -x_2-3+\frac{1}{\mu}x_2-\lambda \end{bmatrix}=\begin{bmatrix} 0 \\ 0 \end{bmatrix},$$

解得 $x_1(\mu)=0, x_2(\mu)=\frac{\mu}{1-\mu}(3+\lambda)$. 又因为 $\mu_k=10^{-(k+1)}<1$ ，所以

$$\nabla^2\mathcal{L}_A(x,\lambda_k,\mu_k)=\begin{bmatrix} 1 & 0 \\ 0 & \frac{1}{\mu_k}-1 \end{bmatrix}$$

正定. 所以 $(0,\frac{\mu_k(3+\lambda_k)}{1-\mu_k})$ 是 $\mathcal{L}_A(x,\lambda_k,\mu_k)$ 的极小值点.

当 $k=0$ 时， $\mu_0=0.1$, $\lambda_0=0$, $x^0=(0,1/3)$ ；
当 $k=1$ 时， $\mu_1=0.01$, $\lambda_1=\lambda_0-\frac{c(x^0)}{\mu_0}=-\frac{10}{3}$, $x^1=(0,-1/297)$ ；
当 $k=2$ 时， $\mu_2=0.001$, $\lambda_2=\lambda_1-\frac{c(x^1)}{\mu_1}=-\frac{890}{297}$, $x^2=(0,\frac{1}{999\times 297})$ ；
由此可见，乘子法迭代收敛到最优解的速率要比二次惩罚法快.

(c) 图示略。

(d) 因为

$$\nabla^2\mathcal{L}_A(x,\lambda_k,\mu)=\begin{bmatrix} 1 & 0 \\ 0 & \frac{1}{\mu}-1 \end{bmatrix},$$

所以当 $0<\mu<1$ 时， $\nabla^2\mathcal{L}_A(x,\lambda_k,\mu)$ 正定. 从而 $(0,\frac{\mu(3+\lambda_k)}{1-\mu})$ 是 $\mathcal{L}_A(x,\lambda_k,\mu)$ 的极小值点.

此时，由乘子法的迭代

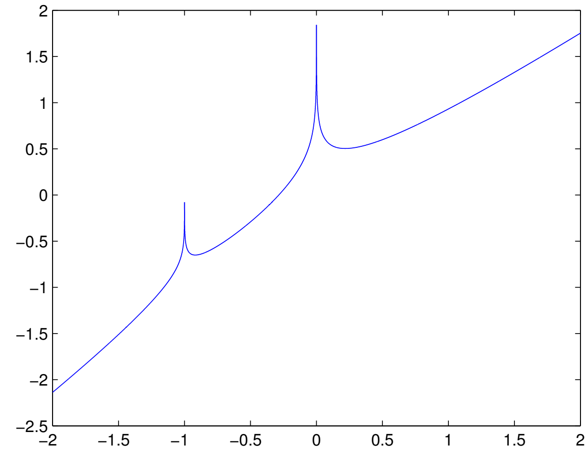
$$\lambda_{k+1}=\lambda_k-\frac{x_2^k}{\mu}=\lambda_k-\frac{1}{\mu}\frac{\mu}{1-\mu}(\lambda_k+3)=\frac{-\mu}{1-\mu}\lambda_k-\frac{3}{1-\mu}$$

易见 $|\frac{-\mu}{1-\mu}|<1$ ，即 $0<\mu<1/2$ 时， $\{\lambda_k\}$ 是收敛的，且 $\lambda_k\to -3(k\to\infty)$. 此时 $(0,\frac{\mu(3+\lambda_k)}{1-\mu})\to(0,0)(k\to\infty)$. 所以对于此问题而言，乘子法不用 $\mu_k\to 0$ ，只需 $0<\mu<1/2$ ，乘子法即是收敛的. 从而可以避免出现病态问题.

10.2 考虑标量极小化问题:

$$\min_x x,\quad \text{subject to } x^2\geq 0, x+1\geq 0,$$

该问题的解是 $x^*=-1$. 写出针对该问题的 $P(x;\mu)$ ，并找到它的局部极小点. 证明对数障碍函数的全局极小点收敛到约束全局极小点 $x^*=-1$ ，但是非全局极小点的局部极



小点收敛到零，其不是约束极小点.

解. 该问题的对数障碍函数为 $P(x,\mu)=x-\mu\ln x^2-\mu\ln(x+1)$. 令 $P'(x,\mu)=1-\frac{2\mu}{x}-\frac{\mu}{x+1}=0$，解得

$$\begin{aligned} x_1(\mu) &= \frac{3\mu-1+\sqrt{9\mu^2+2\mu+1}}{2}, \\ x_2(\mu) &= \frac{3\mu-1-\sqrt{9\mu^2+2\mu+1}}{2}. \end{aligned}$$

因为 $P''(x,\mu)=\frac{2\mu}{x^2}+\frac{\mu}{(x+1)^2}>0$，所以 $P(x,\mu)$ 的驻点 $x_1(\mu)$ 和 $x_2(\mu)$ 即为其局部极小点. 易见当 $\mu_k\downarrow 0$ 时， $x_1(\mu_k)\to 0, x_2(\mu_k)\to -1$. 上面是 $P(x,0.1)$ 的图形.