（1）对于优化问题 $\begin{array}{ll} \min & f(x_1,x_2) \\ s.t. & 4-(x_1^2+x_2^2)\geq 0 \\ & x_1\geq 0 \\ & 1-x_1\geq 0 \\ & x_2\geq 0 \end{array}$ ，$p=(-1,b)^T$ 是该问题在点 $(1,\sqrt{3})^T$ 处的可行方向，则常数 b 的取值范围是 $\underline{b<1/\sqrt{3}}$ 。

解：令 $x=(1,\sqrt{3})^T+\alpha(-1,b)^T=(1-\alpha,\sqrt{3}+\alpha b)^T$，在 $\alpha>0$ 充分小时，x 是可行点。

因此，$\begin{array}{l} 4-(1-\alpha)^2-(\sqrt{3}+\alpha b)^2\geq 0 \\ 1-\alpha\geq 0 \\ \alpha\geq 0 \\ \sqrt{3}+\alpha b\geq 0 \end{array}$

后面三个不等式在 $\alpha>0$ 充分小时显然成立。第一个不等式是：

$2(1-\sqrt{3}b)\alpha\geq\alpha^2(1+b^2)$，$2(1-\sqrt{3}b)\geq\alpha(1+b^2)$

所以有 $1-\sqrt{3}b>0$，$b<\dfrac{1}{\sqrt{3}}$。

(2) 向量 $(-1,r)^T$ 是问题 $\begin{array}{l} \max -2x_1+3x_2 \\ s.t.\ 4x_1+x_2\leq -4 \\ \quad -x_1+3x_2\geq 1 \\ \quad x_1\leq 0 \end{array}$ 在点 $(-1,0)^T$ 的下降方向，则 r 的范围是 _r<-2/3______。

解：目标函数是 $-2x_1+3x_2$, $g=(-2,3)^T$, $x=(-1,0)^T$, $p=(-1,r)^T$

$g^Tp=2+3r$.

当 $2+3r<0$ 时，是下降方向，当 $2+3r<0$ 时，是下降方向。

当 $2+3r=0$，$r=-\dfrac{2}{3}$，

$$x+\alpha p=(-1,0)^T+\alpha(-1,-\frac{2}{3})^T=(-1-\alpha,-\frac{2}{3}\alpha)^T$$

$$f(x+\alpha p)=-2(-1-\alpha)+3(-\frac{2}{3}\alpha)=2=f(x)$$

所以 $r=-\dfrac{2}{3}$ 时不是下降方向。

结论：$r<-\dfrac{2}{3}$ 时是下降方向。

证明题：

（1） $f(x)$ 为凸集 $D\subset R^n$ 上的函数，上图 $epi(f)=\{(x,y)\,|\,x\in D, y\in R, y\geq f(x)\}$，证明 $f(x)$ 为凸函数的充要条件是 $epi(f)$ 为凸集。

证明：充分性：对于任意的 $x_1,x_2\in D,\alpha\in[0,1]$，由 $epi(f)$ 的定义，$(x_1,f(x_1)),(x_2,f(x_2))\in epi(f)$，而 $epi(f)$ 为凸集，得

$\alpha(x_1,f(x_1))+(1-\alpha)(x_2,f(x_2))\in epi(f)$，即

$(\alpha x_1+(1-\alpha)x_2,\alpha f(x_1)+(1-\alpha)f(x_2))\in epi(f)$，

因此 $\alpha f(x_1)+(1-\alpha)f(x_2)\geq f(\alpha x_1+(1-\alpha)x_2)$，从而 $f(x)$ 为凸函数。

必要性：对于任意的 $(x_1,y_1),(x_2,y_2)\in epi(f),\alpha\in[0,1]$，有

$y_1\geq f(x_1), y_2\geq f(x_2)$

于 $f(x)$ 为凸函数，有 $f(\alpha x_1+(1-\alpha)x_2)\leq\alpha f(x_1)+(1-\alpha)f(x_2)\leq\alpha y_1+(1-\alpha)y_2$，

所以 $(\alpha x_1+(1-\alpha)x_2,\alpha y_1+(1-\alpha)y_2)\in epi(f)$，即

$\alpha(x_1,y_1)+(1-\alpha)(x_2,y_2)\in epi(f)$，

$epi(f)$ 为凸集。

(2)考虑规划问题：$\begin{array}{ll}\min & f(x)\\ s.t. & c_i(x)\leq 0\end{array}$，其中，$f(x),c_i(x)(i=1,2,\cdots,m):R^n\to R$ 是凸函数，证明：（1）该问题的可行域是凸集；（2）该问题的最优解的集合 A 是凸集。

证明：（1）设 $D_i=\{x\,|\,c_i(x)\leq 0\}$，可行域 $D=\bigcap_{i=1}^{m}D_i$。

对于任意 $x_1,x_2\in D_i=\{x\,|\,c_i(x)\leq 0\}$，任意 $\alpha\in[0,1]$，有

$c_i(x_1)\leq 0, c_i(x_2)\leq 0$，根据 $c_i(x)$ 是凸函数

$c_i(\alpha x_1+(1-\alpha)x_2)\leq\alpha c_i(x_1)+(1-\alpha)c_i(x_2)\leq 0$，

因此 $\alpha x_1+(1-\alpha)x_2\in D_i$。

凸集的交集是凸集，因此 $D=\bigcap_{i=1}^{m}D_i$ 为凸集。

（2）设 A 为最优解的集合，若 A 不是空集，任取 $x_1,x_2\in A,\alpha\in[0,1]$，有 $x_1,x_2\in D$。由于 D 是凸集，$\alpha x_1+(1-\alpha)x_2\in D$。$f(x)$ 是凸函数，

$f(\alpha x_1+(1-\alpha)x_2)\leq\alpha f(x_1)+(1-\alpha)f(x_2)$。

x_1,x_2 都是最优解，因此 $f(x_1)=f(x_2)$ 为最优函数值。得

$f(\alpha x_1+(1-\alpha)x_2)\leq\alpha f(x_1)+(1-\alpha)f(x_2)=f(x_1)$

$f(\alpha x_1+(1-\alpha)x_2)=f(x_1)$，因此 $\alpha x_1+(1-\alpha)x_2$ 也是最优解，从而 A 为凸集

(3)设 z^*,s^* 分别为下列两个问题 $(I)\begin{array}{ll} & \min c^Tx\\ s.t. & Ax=b\\ & x\geq 0\end{array}$ $(II)\begin{array}{ll} & \min c^Tx\\ s.t. & Ax=b+d\\ & x\geq 0\end{array}$ 的最优值。y^* 是(I)的对偶问题的最优解，证明 $z^*+y^{*T}d\leq s^*$。

证明：（I）与（II）的对偶规划分别为 $(DI)\begin{array}{ll} & \max b^Ty\\ s.t. & A^Ty\leq c\end{array}$ $(DII)\begin{array}{ll} & \max (b+d)^Ty\\ s.t. & A^Ty\leq c\end{array}$

（I）的最优值与（DI）的最优值相同，得 $z^*=b^Ty^*$.

y^*是（DI）对偶规划的最优解，从而是（DII）的可行解。y^*在（DII）的目标函数值不大于最优值，$(b+d)^T y^* \le s^*$。因此$z^* + y^{*T} d \le s^*$。

（4）设$\bar{x}, \bar{y}$分别为下列两个问题

$$(I)\ \begin{array}{ll} & \min c^T x \\ s.t. & Ax \ge b \\ & x \ge 0 \end{array} \qquad (II)\ \begin{array}{ll} & \max y^T b \\ s.t. & y^T A \le c^T \\ & y \ge 0 \end{array}$$

的可行解。证明$c^T \bar{x} \ge \bar{y}^T b$。

证明：由题意，$A\bar{x} \ge b, \bar{x} \ge 0, \bar{y}^T A \le c^T, \bar{y} \ge 0,$

由$A\bar{x} - b \ge 0, \bar{y} \ge 0$得$\bar{y}^T (A\bar{x} - b) \ge 0$，$\bar{y}^T A\bar{x} \ge \bar{y}^T b$，

由$\bar{x} \ge 0, \bar{y}^T A - c^T \le 0$得$(\bar{y}^T A - c^T)\bar{x} \le 0$，$\bar{y}^T A\bar{x} \le c^T \bar{x}$，所以$c^T \bar{x} \ge \bar{y}^T b$。

计算题

一、（18%）（1）用单纯形方法求解下面的线性规划
$$\begin{array}{l} \min -3x_1 - 4x_2 \\ s.t.\ 2x_1 + x_2 \le 4 \\ \quad x_1 + 3x_2 \le 3 \\ \quad x_1, x_2 \ge 0 \end{array}$$

（2）写出该线性规划的影子价格向量；

（3）若在上面的线性规划中要求变量为整数，在相应的整数规划中，请对变量x_1写出对应的割平面方程。

解：（1）该线性规划的标准型为：
$$\begin{array}{l} \min -3x_1 - 4x_2 \\ s.t.\ 2x_1 + x_2 + x_3 = 4 \\ \quad x_1 + 3x_2 \quad + x_4 = 3 \\ \quad x_1, x_2, x_3, x_4 \ge 0 \end{array}$$

以p_3, p_4为基，$(0,0,4,3)^T$为初始基可行解，单纯形表为

c_j			-3	-4	0	0	θ_j
cB	B	b	p1	p2	p3	p4	
0	p3	4	2	1	1	0	4
0	p4	3	1	(3)	0	1	1
σ j			-3	-4	0	0	
0	p3	3	(5/3)	0	1	-1/3	9/5
-4	p2	1	1/3	1	0	1/3	3
σ j			-5/3	0	0	4/3	
-3	p1	9/5	1	0	3/5	-1/5	
-4	p2	2/5	0	1	-1/5	2/5	
σ j			0	0	1	1	

原问题最优解为$(\frac{9}{5}, \frac{2}{5})^T$，最优值为-7。

（2）影子价格向量为$(c_B^T B^{-1})^T = \left((-3,-4) \begin{pmatrix} \frac{3}{5} & -\frac{1}{5} \\ -\frac{1}{5} & \frac{2}{5} \end{pmatrix} \right)^T = (-1,-1)^T$

（3）由单纯形表可以得到

$x_1+\frac{3}{5}x_3-\frac{1}{5}x_4=\frac{9}{5}$,

即 $x_1-x_4-1=\frac{4}{5}-\frac{3}{5}x_3-\frac{4}{5}x_4$

割平面方程为 $\frac{4}{5}-\frac{3}{5}x_3-\frac{4}{5}x_4\le 0$。

二、解：（1）利用消元法，得到以 p_1,p_2 为基矩阵的规范式为：

$$\min\ 2x_3-x_4+\frac{7}{2}$$

$$s.t.\ x_1+x_3-3x_4=\frac{5}{2}$$

$$x_2+4x_4=\frac{1}{2}$$

$$x_1,x_2,x_3,x_4\ge 0$$

（或者：以 $B=\begin{bmatrix}1&1\\2&0\end{bmatrix}$ 为基矩阵，在约束方程组的两端同时左乘 B^{-1}，得到约束方程组为

$$\begin{bmatrix}1&1\\2&0\end{bmatrix}^{-1}\begin{bmatrix}1&1&1&1\\2&0&2&-6\end{bmatrix}\begin{bmatrix}x_1\\x_2\\x_3\\x_4\end{bmatrix}=\begin{bmatrix}1&1\\2&0\end{bmatrix}^{-1}\begin{bmatrix}3\\5\end{bmatrix}$$，得到 $\begin{matrix}x_2+4x_4=\frac{1}{2}\\x_1+x_3-3x_4=\frac{5}{2}\end{matrix}$，

然后再将目标函数中基变量系数消为 0。
约束等式交换两行，不影响后面的计算和结果）

以 p_1,p_2 为基，$(\frac{5}{2},\frac{1}{2},0,0)^T$ 为初始基可行解，单纯形表为

c_j			0	0	2	-1	θ_j
cB	B	b	p1	p2	p3	p4	
0	P1	5/2	1	0	1	-3	
0	P2	1/2	0	1	0	(4)	1/8
σ j			0	0	2	-1	
0	P1	23/8	1	3/4	1	0	
-1	P4	1/8	0	1/4	0	1	
σ j			0	1/4	2	0	

原问题最优解为 $(\frac{23}{8},0,0,\frac{1}{8})^T$，最优函数值为 27/8。

（2）影子价格向量为 $(c_B^TB^{-1})^T=\left((1,4)\begin{pmatrix}1&1\\2&-6\end{pmatrix}^{-1}\right)^T$

$$=\left((1,4)\begin{pmatrix}\frac{3}{4}&\frac{1}{8}\\\frac{1}{4}&-\frac{1}{8}\end{pmatrix}^{-1}\right)^T=(\frac{7}{4},-\frac{3}{8})^T$$

（3）由单纯形表可以得到

因为 $\frac{23}{8}=2+\frac{7}{8}$，所以分枝后的两个线性规划为

$$\begin{array}{ll}
\min 2x_3-x_4+\frac{7}{2} & \min 2x_3-x_4+\frac{7}{2} \\
s.t.\ x_1+x_3-3x_4=\frac{5}{2} & s.t.\ x_1+x_3-3x_4=\frac{5}{2} \\
\quad x_2+4x_4=\frac{1}{2} \quad , & \quad x_2+4x_4=\frac{1}{2} \\
x_1\le 2 & x_1\ge 3 \\
\quad x_1,x_2,x_3,x_4\ge 0 & \quad x_1,x_2,x_3,x_4\ge 0
\end{array}$$

三、提示：先处理 $|x_4|\le 2$ 这个约束，最基本的处理方法：考虑两个约束：$x_4\le 2, x_4\ge -2$
也可以令 $2+x_4$（或 $2-x_4$）为一个新变量再处理。
然后必须用两阶段或者大 M 法来进行求解。
最优解是 $(3,0,0,2)^T$.

四、提示：把 mn 个变量对应的矩阵写出来。

对偶规划：
$$\begin{aligned}&\max \sum_{i=1}^{m} a_i y_i+\sum_{i=1}^{n} b_i z_i\\ &s.t.\ y_i+z_j\le c_{ij}(i=1,2,\cdots,m,\ j=1,2,\cdots,n)\end{aligned}$$

五、若线性规划
$$\begin{aligned}&\min x_1+2x_2+x_3\\ &s.t.\ 3x_1+4x_2 \qquad\ \ =5\\ &\quad\ x_1 \qquad\quad +2x_3=u\\ &\quad\ x_1,x_2,x_3\ge 0\end{aligned}$$
的最优解为 $(a,b,c)^T$，其对偶规划的最优解为 $(1/6,1/2)^T$。a,b,c,u 四个常数中，你可以确定哪些？如果有不能确定的常数，确定其范围。

解：$x^*=(a,b,c)^T, y^*=(\frac{1}{6},\frac{1}{2})^T$，

$$A^T y^*-\overline{c}=\begin{bmatrix}3 & 1\\ 4 & 0\\ 0 & 2\end{bmatrix}\begin{bmatrix}\frac{1}{6}\\ \frac{1}{2}\end{bmatrix}-\begin{bmatrix}1\\ 2\\ 1\end{bmatrix}=\begin{bmatrix}0\\ -\frac{4}{3}\\ 0\end{bmatrix}$$

（注：字母 c 在题目中已用，我们将价格向量用 $\overline{c}$ 表示）

由互补松弛性定理，$(A^T y^*-\overline{c})^T x^*=0$，得 $-\frac{4}{3}b=0$，$b=0$.

$x^*=(a,b,c)^T$ 可行，由第一个等式约束得 $3a+4b=5$，$a=\frac{5}{3}$.

c 的范围是 $c\ge 0$，由第二个约束等式，$\frac{5}{3}+4c=u$，$u\ge\frac{5}{3}$.

解答：

第三章计算题

一、用 Newton 方法求解问题 $\min\quad \boldsymbol{x}_1^4+\boldsymbol{x}_2^2-2\boldsymbol{x}_1-\boldsymbol{x}_2$，初始点 $\boldsymbol{x}^{(0)}=(1,1)^{\boldsymbol{T}}$. 求 $\boldsymbol{x}^{(1)}$.

解：$\nabla \boldsymbol{f}(\boldsymbol{x})=\begin{pmatrix}4\boldsymbol{x}_1^3-2\\2\boldsymbol{x}_2-1\end{pmatrix},\nabla^2\boldsymbol{f}(\boldsymbol{x})=\begin{pmatrix}12\boldsymbol{x}_1^2&0\\0&2\end{pmatrix}$,

$\boldsymbol{g}_0=\nabla\boldsymbol{f}(\boldsymbol{x}^{(0)})=\begin{pmatrix}2\\1\end{pmatrix},\boldsymbol{G}_0=\nabla^2\boldsymbol{f}(\boldsymbol{x}^{(0)})=\begin{pmatrix}12&0\\0&2\end{pmatrix}$,

由 Newton 法迭代公式,

$$\boldsymbol{x}^{(1)}=\boldsymbol{x}^{(0)}-(\boldsymbol{G}_0)^{-1}\boldsymbol{g}_0=\begin{pmatrix}1\\1\end{pmatrix}-\begin{pmatrix}12&0\\0&2\end{pmatrix}^{-1}\begin{pmatrix}2\\1\end{pmatrix}=\begin{pmatrix}1\\1\end{pmatrix}-\begin{pmatrix}1/6\\1/2\end{pmatrix}=\begin{pmatrix}5/6\\1/2\end{pmatrix}$$

二、用 FR 方法求解问题 $\min\quad \boldsymbol{x}_1^2+\frac{5}{2}\boldsymbol{x}_2^2-3\boldsymbol{x}_1\boldsymbol{x}_2+2\boldsymbol{x}_1-\boldsymbol{x}_2$，初始点 $\boldsymbol{x}^{(0)}=(1,1)^{\boldsymbol{T}}$.

解：$g(x)=(2x_1-3x_2+2,-3x_1+5x_2-1)^T$, $x^{(0)}=(1,1)^T$, $g_0=(1,1)^T$, $p_0=-g_0=(-1,-1)^T$,

一维搜索：$\min f(x^{(0)}+\alpha p_0)=\frac{1}{2}(\alpha^2-4\alpha+3)$ 的最优解为 $\alpha_0=2$,

$x^{(1)}=x^{(0)}+\alpha_0 p_0=(-1,-1)^T$,

$g^{(1)}=(3,-3)^T$,

根据 FR 方法，有 $p_1=-g_1+\beta_0 p_0$，$\beta_0=\frac{g_1^Tg_1}{g_0^Tg_0}=9$，$p_1=(-12,-6)^T$,

一维搜索：$\min f(x^{(1)}+\alpha p_1)=18\alpha^2-18\alpha-\frac{1}{2}$ 的最优解为 $\alpha_1=\frac{1}{2}$,

$x^{(2)}=x^{(1)}+\alpha p_1=(-7,-4)^T$，$g_2=(0,0)^T$。

因此最优解为 $x^*=x^{(2)}=(-3,-3)^T$ 最优值为 $f^*=-5$。

三、用 PRP 方法求解问题 $\min\quad \boldsymbol{x}_1^2+2\boldsymbol{x}_2^2-2\boldsymbol{x}_1\boldsymbol{x}_2-2\boldsymbol{x}_1$，初始点 $\boldsymbol{x}^{(0)}=(1,1)^{\boldsymbol{T}}$.

解：$g(x)=(2x_1-2x_2-2,4x_2-2x_1)^T.g_0=(-2,2)^T\neq 0.$

$\therefore$ 取 $\mathrm{p}_0=-\mathrm{g}_0=(2,-2)^T$,

(1) 从 x_0 出发，沿 p_0 进行一维搜索，即求

$\min\quad \varphi_0(\alpha)=\mathrm{f}(\mathrm{x}_0+\alpha\mathrm{p}_0)=20\alpha^2-8\alpha$

的极小点，得步长 $\alpha_0=\frac{1}{5}$. 于是得到 $x_1=x_0+\alpha_0 p_0=(\frac{7}{5},\frac{3}{5})^T$, $g_1=(-\frac{2}{5},-\frac{2}{5})^T$.

由 PRP 公式得 $\beta_0=g_1^T(g_1-g_0)/g_0^Tg_0=\frac{1}{25}$，故

$p_1=-g_1+\beta_0 p_0=(\frac{12}{25},\frac{8}{25})^T$.

(2) 从 x_1 出发，沿 p_1 进行一维搜索，即求

$\min \quad \varphi_1(\alpha) = f(x_1 + \alpha p_1) = \dfrac{16}{125}\alpha^2 - \dfrac{40}{125}\alpha - \dfrac{44}{25}$

的极小点，得 $\alpha_1 = \dfrac{5}{4}$. 于是得到 $x_2 = x_1 + \alpha_1 p_1 = (2,1)^T$, 此时 $g_2 = (0,0)^T$.

故 $x^* = x_2 = (2,1)^T$, $f^* = 2$.

四、用 DFP 方法求解问题 $\min \quad x_1^2 + 2x_2^2$，初始点 $x^{(0)} = (1/2, 1/4)^T$. $H_0 = I$. DFP 矩阵修正公式为 $H_{k+1} = H_k - \dfrac{H_k y_k y_k^T H_k}{y_k^T H_k y_k} + \dfrac{s_k s_k^T}{y_k^T s_k}$.

解：$\nabla f(x) = (2x_1, 4x_2)^T$

$\nabla f(x_0) = (1,1)^T$，初始方向 $p_0 = -g_0 = (-1,-1)^T$，进行一维搜索，求解

$f(x_0 + \alpha p_0) = 3\alpha^2 - 2\alpha + \dfrac{3}{8}$ 的极小点，得步长 $\alpha_0 = \dfrac{1}{3}$，$x_1 = x_0 + \alpha_0 p_0 = (\dfrac{1}{6}, -\dfrac{1}{12})^T$，

$g_1 = (\dfrac{1}{3}, -\dfrac{1}{3})^T$，$s_0 = x_1 - x_0 = (-\dfrac{1}{3}, -\dfrac{1}{3})^T, y_0 = g_1 - g_0 = (-\dfrac{2}{3}, -\dfrac{4}{3})^T$，由 DFP 修正公式

$$H_1 = H_0 - \frac{H_0 y_0 y_0^T H_0}{y_0^T H_0 y_0} + \frac{s_0 s_0^T}{y_0^T s_0} = \begin{pmatrix} \frac{29}{30} & -\frac{7}{30} \\ -\frac{7}{30} & \frac{11}{30} \end{pmatrix},\quad p_1 = -H_1 g_1 = (-\frac{2}{5}, \frac{1}{5})^T。$$

从 x_1 出发，沿 p_1 进行一维搜索，求 $\min f(x_1 + \alpha p_1) = \dfrac{1}{5}\alpha^2 - \dfrac{1}{6}\alpha + \dfrac{5}{144}$ 的极小点，得 $\alpha_1 = \dfrac{5}{12}$。$x_2 = x_1 + \alpha_1 p_1 = (0,0)^T$，$g_2 = (0,0)^T$

第三章证明题：

一、设 G 为 n 阶正定对称矩阵，$u_1, u_2, \cdots, u_n \in R^n$ 线性无关。p_k 按如下方式生成：

$p_1 = u_1$，$\boldsymbol{p_{k+1} = u_{k+1} - \sum_{i=1}^{k} \frac{u_{k+1}^T G p_i}{p_i^T G p_i} p_i (k = 1, 2, \cdots, n-1)}$，证明 $\boldsymbol{p_1, p_2, \cdots, p_n}$ 关于 $\boldsymbol{G}$ 共轭。

证明思路：数学归纳法

二、设 A 为 n 阶对称正定矩阵，非零向量 $p_1, p_2, \cdots, p_n$ 关于矩阵 A 共轭，证明

$A^{-1} = \sum_{k=1}^{n} \dfrac{p_k p_k^T}{p_k^T A p_k}$。

证明：设 $B = \sum_{k=1}^{n} \dfrac{p_k p_k^T}{p_k^T A p_k}$，

由于非零向量 $p_1, p_2, \cdots, p_n$ 关于矩阵 A 共轭，所以 $p_k^T A p_i = 0 (k \neq i)$，

对于任意 $i (1 \le i \le n)$，

$$BAp_i = \sum_{k=1}^{n} \frac{p_k p_k^T A p_i}{p_k^T A p_k} = \frac{p_i p_i^T A p_i}{p_i^T A p_i} = p_i$$

因此$BA[p_1,p_2,\cdots,p_n]=[BAp_1,BAp_2,\cdots,BAp_n]=[p_1,p_2,\cdots,p_n]$。
非零向量$p_1,p_2,\cdots,p_n$关于矩阵A共轭，所以$p_1,p_2,\cdots,p_n$线性无关，$[p_1,p_2,\cdots,p_n]$可逆，所以有$BA=I,B=A^{-1}$.

三、用 PRP 共轭梯度法求解无约束优化问题$\min\ f(x)$，若一维搜索是精确的，且在求解过程中，每一步的梯度都是非零向量，证明 PRP 方法产生的搜索方向都是下降方向。

证明：

（1） $p_0=-g_0$，$p_0^T g_0=-g_0^T g_0<0$，所以p_0是下降方向。

（2）$k\ge 1$时， $p_k=-g_k+\beta_{k-1}p_{k-1}$，

其中$\beta_{k-1}=\dfrac{g_k^T(g_k-g_{k-1})}{g_{k-1}^T g_{k-1}}$，

由于一维搜索是精确的，所以$g_k^T p_{k-1}=0$

$$g_k^T p_k=-g_k^T g_k+\beta_{k-1}g_k^T p_{k-1}=-g_k^T g_k<0$$

所以p_k是下降方向。

四、设A为n阶对称正定矩阵，非零向量$p_1,p_2,\cdots,p_n$关于矩阵A共轭， $x\in R^n$. 证明

$x=\sum_{k=1}^n \dfrac{p_k^T Ax}{p_k^T Ap_k}p_k$。

证明：非零向量$p_1,p_2,\cdots,p_n$关于矩阵A共轭，因此$p_1,p_2,\cdots,p_n$线性无关，它们构成R^n的一组基底。

设x在这组基底下的线性表示为$x=\sum_{k=1}^n \alpha_k p_k$。

对任意$\boldsymbol{i}\in\{1,2,\cdots,\boldsymbol{n}\}$，有 $p_i^T Ax=\sum_{k=1}^n \alpha_k p_i^T Ap_k$。

$p_1,p_2,\cdots,p_n$关于矩阵A共轭，当$k\ne i$时 $p_j^T Ap_k=0$，因此

$p_i^T Ax=\alpha_i p_i^T Ap_i$， $\alpha_i=\dfrac{p_i^T Ax}{p_i^T Ap_i}$.

因此$x=\sum_{k=1}^n \dfrac{p_k^T Ax}{p_k^T Ap_k}p_k$

第四章计算题

一、用外罚函数法求解 $\begin{array}{l}\min\ \boldsymbol{x}_1^2+3\boldsymbol{x}_2^2\\ \boldsymbol{s.t.}\ \boldsymbol{x}_1+2\boldsymbol{x}_2-1=0\end{array}$。

解：增广目标函数是 $P(x,\sigma)=x_1^2+3x_2^2+\sigma(x_1+2x_2-1)^2$，

(注：可以是 $P(x,\sigma)=x_1^2+3x_2^2+\dfrac{\sigma}{2}(x_1+2x_2-1)^2$)

$$\frac{\partial P(x,\sigma)}{\partial x_1}=2x_1+2\sigma(x_1+2x_2-1)$$

$$\frac{\partial P(x,r)}{\partial x_2}=6x_2+4\sigma(x_1+2x_2-1)$$

求驻点：$\dfrac{\partial P(x,\sigma)}{\partial x_1}=\dfrac{\partial P(x,\sigma)}{\partial x_2}=0$，得 $x_1=\dfrac{3\sigma}{3+7\sigma},x_2=\dfrac{2\sigma}{3+7\sigma}$，

令 $\sigma\to+\infty$，$x_1\to\dfrac{3}{7},x_2\to\dfrac{2}{7}.\ x\to x^*=(\dfrac{3}{7},\dfrac{2}{7})^T,\ f^*=\dfrac{3}{7}$。

二、用内罚函数法（对数罚函数）求解 $\begin{array}{l}\min\ \boldsymbol{x}_1^2+5\boldsymbol{x}_2^2\\ \boldsymbol{s.t.}\ \boldsymbol{x}_1+\boldsymbol{x}_2-3\ge 0\end{array}$。

解：构造增广目标函数为 由 $B(x,r)=x_1^2+5x_2^2-r\ln(x_1+x_2-3)$，

$$\frac{dB(x,r)}{dx_1}=2x_1-\frac{r}{x_1+x_2-3},$$

$$\frac{dB(x,r)}{dx_2}=10x_2-\frac{r}{x_1+x_2-3},$$

令 $\dfrac{dB(x,r)}{dx_1}=\dfrac{dB(x,r)}{dx_2}=0$，得 $x_1=\dfrac{30+\sqrt{900+240r}}{24},x_2=\dfrac{30+\sqrt{900+240r}}{120}$（负号舍去)

令 $r\to 0$，$x_1=\dfrac{5}{2},x_2=\dfrac{1}{2}.\ x\to x^*=(\dfrac{5}{2},\dfrac{1}{2})^T,\ f^*=\dfrac{15}{2}$。

三、用乘子法求解问题 $\begin{array}{l}\min\ \boldsymbol{x}_1^2+2\boldsymbol{x}_2^2\\ \boldsymbol{s.t.}\ 2\boldsymbol{x}_1+\boldsymbol{x}_2-4=0\end{array}$。

解：增广 Lagrange 函数 $M(x_1,x_2,\lambda,\sigma)=x_1^2+2x_2^2-\lambda(2x_1+x_2-4)+\dfrac{\sigma}{2}(2x_1+x_2-4)^2$，

令 $$\begin{aligned}\frac{\partial M}{\partial x_1}&=2x_1-2\lambda+2\sigma(2x_1+x_2-4)=0\\ \frac{\partial M}{\partial x_2}&=4x_2-\lambda+\sigma(2x_1+x_2-4)=0\end{aligned}$$

得 $x_1=\dfrac{4\lambda+16\sigma}{4+9\sigma},x_2=\dfrac{\lambda+4\sigma}{4+9\sigma}$

将 λ 换为 λ_k 再由乘子迭代公式 $\lambda_{k+1}=\lambda_k-\sigma(2x_1+x_2-4)$ 得到

$\lambda_{k+1} = \dfrac{4}{4+9\sigma}\lambda_k + \dfrac{16\sigma}{4+9\sigma}$。

在 $\sigma > 0$ 时，$\{\lambda_k\}$ 收敛，设 $\lambda_k \to \lambda^*$，得

$\lambda^* = \dfrac{4}{4+9\sigma}\lambda^* + \dfrac{16\sigma}{4+9\sigma}$，$\lambda^* = \dfrac{16}{9}$。

$x_1 = \dfrac{16}{9}, x_2 = \dfrac{4}{9}$，所以 $x^* = (\dfrac{16}{9}, \dfrac{4}{9})^T, f^* = \dfrac{32}{9}.$

四、求解二次规划（不使用代入法）

$$\begin{aligned}&\min\ 3x_1^2 + 2x_1x_2 + x_1x_3 + 2.5x_2^2 + 2x_2x_3 + 2x_3^2 - 8x_1 - 3x_2 - 3x_3\\ &s.t.\ x_1 + x_3 = 3\\ &\quad\ \ x_2 + x_3 = 0\end{aligned}$$ 。

解：$G = \begin{bmatrix} 6 & 2 & 1 \\ 2 & 5 & 2 \\ 1 & 2 & 4 \end{bmatrix}$，$c = \begin{bmatrix} -8 \\ -3 \\ -3 \end{bmatrix}$，$A = \begin{bmatrix} 1 & 0 \\ 0 & 1 \\ 1 & 1 \end{bmatrix}$，$b = \begin{bmatrix} 3 \\ 0 \end{bmatrix}$，

根据等式约束二次规划最优解的充要条件有

$$\begin{bmatrix} G & -A \\ A^T & 0 \end{bmatrix}\begin{bmatrix} x \\ \lambda \end{bmatrix} = \begin{bmatrix} -c \\ b \end{bmatrix},$$

即 $\begin{bmatrix} 6 & 2 & 1 & -1 & 0 \\ 2 & 5 & 2 & 0 & -1 \\ 1 & 2 & 4 & -1 & -1 \\ 1 & 0 & 1 & 0 & 0 \\ 0 & 1 & 1 & 0 & 0 \end{bmatrix}\begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ \lambda_1 \\ \lambda_2 \end{bmatrix} = \begin{bmatrix} 8 \\ 3 \\ 3 \\ 3 \\ 0 \end{bmatrix}$，解得 $\begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ \lambda_1 \\ \lambda_2 \end{bmatrix} = \begin{bmatrix} 2 \\ -1 \\ 1 \\ 3 \\ -2 \end{bmatrix}$。

最优解为 $x^* = (2,-1,1)^T$。

五、已知 $x^* = (1,3)^T$ 是求下面问题的 KT 点，确定常数 p 的取值范围。

$$\begin{aligned}&\min\ px_1^2 + x_2^2\\ &s.t.\ c_1(x) = x_1 + x_2 - 4 = 0\\ &\quad\ \ c_2(x) = -3x_1 + x_2 \geq 0\end{aligned}$$

解：点 $x^* = (1,3)^T$ 处的有效集为 $I^* = \{1,2\}$，

$\nabla f(x) = (2px_1, 2x_2)^T$，$\nabla f(x^*) = (2p, 6)^T$，$\nabla c_1(x) = (1,1)^T$，$\nabla c_2(x) = (-3,1)^T$，

$x^* = (1,3)^T$ 是问题的 KT 点，所以存在乘子 λ_1, λ_2，使得下面条件成立。

$\nabla f(x^*) - \lambda_1 \nabla c_1(x^*) - \lambda_2 \nabla c_2(x^*) = 0,$

$\lambda_2 \geq 0$

根据由 $\begin{pmatrix} 2p \\ 6 \end{pmatrix} - \lambda_1 \begin{pmatrix} 1 \\ 1 \end{pmatrix} - \lambda_2 \begin{pmatrix} -3 \\ 1 \end{pmatrix} = 0$，可以得

$\lambda_1 = \dfrac{9+p}{2}, \lambda_2 = \dfrac{3-p}{2}$。

由 $\lambda_2=\dfrac{3-p}{2}\ge 0$，解得 $p\le 3$。

六、求下面问题的可行域内的 KT 点。$\begin{aligned}\min\ & (x_1-1)^2+(x_2-2)^2\\ s.t.\ & c_1(x)=x_1-2\ge 0\\ & c_2(x)=x_1+x_2=0\end{aligned}$ 。

解：由 KT 条件，存在常数 λ_1,λ_2 使得

$\nabla f(x^*)-\lambda_1 c_1(x^*)-\lambda_2 c_2(x^*)=0$，

$\lambda_1 c_1(x^*)=0,\quad \lambda_1\ge 0$。

$\nabla f(x)=(2(x_1-1),2(x_2-2))^T$， $\nabla c_1(x)=(1,0)^T,\nabla c_2(x)=(1,1)^T$, 所以

$\begin{pmatrix}2(x_1-1)\\2(x_2-2)\end{pmatrix}-\lambda_1\begin{pmatrix}1\\0\end{pmatrix}-\lambda_2\begin{pmatrix}1\\1\end{pmatrix}=0$。

根据 $\lambda_1(x_1^*-2)=0$, 可以得到 λ_1=0 或 $x_1^*=2$

情形 I

λ_1=0, 此时由上式可得 x_1=-$\dfrac{1}{2}$, x_2=$\dfrac{1}{2}$, λ_2=-3, 但因为 (-$\dfrac{1}{2}$,$\dfrac{1}{2}$) 不满足可行条件 $x_1-2\ge 0$，所以舍去

情形 II

x_1^*=2,

由 $x_1^*+x_2^*=0$ 解出 x_2^*=-2, 再由

$\begin{pmatrix}2x_1^*-2\\2x_2^*-4\end{pmatrix}-\lambda_1^*\begin{pmatrix}1\\0\end{pmatrix}-\lambda_2^*\begin{pmatrix}1\\1\end{pmatrix}=\begin{pmatrix}2-\lambda_1^*-\lambda_2^*\\-8-\lambda_2^*\end{pmatrix}=\begin{pmatrix}0\\0\end{pmatrix}$

得到 λ_1=10>0, λ_2=-8.

所以求得 KT 点为 $(2,-2)$。

第四章证明题

一、 $f:R^n\to R$ 为可微凸函数，证明 x^* 为优化问题 $\begin{matrix}\min & f(x)\\ s.t. & x\ge 0\end{matrix}$ 的最优解的充要条件是 $x^*\ge 0,\nabla f(x^*)\ge 0,(\nabla f(x^*))^T x^*=0$。

证明：$\begin{matrix}\min & f(x)\\ s.t. & x\ge 0\end{matrix}$ 是凸规划问题，KT 条件是最优解的充要条件。

问题有 n 个约束，约束函数的梯度为 $e_i(i=1,2,\cdots,n)$，显然线性无关，该问题的 KT 条件是存在乘子向量 $\lambda=(\lambda_1,\lambda_2,\cdots,\lambda_n)$，使得

$\nabla f(x^*)-\sum_{i=1}^{n}\lambda_i e_i=0$， $\lambda_i x_i=0$， $\lambda_i\ge 0$。

结合解得可行性条件得最优解的充要条件是

$x^* \geq 0$，$\lambda = \nabla f(x^*)$，$\lambda_i x_i = 0$，$\lambda_i \geq 0$。

充分性：若 $x^* \geq 0, \nabla f(x^*) \geq 0, (\nabla f(x^*))^T x^* = 0$，

令 $\lambda = \nabla f(x^*)$，则有 $x^* \geq 0, \lambda = \nabla f(x^*) \geq 0, \lambda^T x^* = 0$。

x^*, λ 两个非负向量的内积为零，显然有 $\lambda_i x_i = 0 (i = 1, 2, \cdots, n)$。

必要性：若 $x^* \geq 0$，$\lambda = \nabla f(x^*)$，$\lambda_i x_i = 0$，$\lambda_i \geq 0$。

则 $x^* \geq 0, \nabla f(x^*) \geq 0$，

$(\nabla f(x^*))^T x^* = \sum_{i=1}^{n} \lambda_i x_i = 0$。