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 $p_1=-g_1+\beta_0 p_0$，$\beta_0=\frac{g_1^T g_1}{g_0^T g_0}=9$，$p_1=(-12,-6)^T$,

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 $p_0=-g_0=(2,-2)^T$，

(1) 从 x_0 出发，沿 p_0 进行一维搜索，即求

$\min \quad \varphi_0(\alpha)=f(x_0+\alpha p_0)=20\alpha^2-8\alpha$

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 $\beta_0=g_1^T(g_1-g_0)/g_0^T g_0=\frac{1}{25}$， 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

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

的极小点，得 $\alpha_1 = \frac{5}{4}$. 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 $H_{k+1} = H_k - \frac{H_k y_k y_k^T H_k}{y_k^T H_k y_k} + \frac{s_k s_k^T}{y_k^T s_k}$.

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

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

$g_1 = (\frac{1}{3}, -\frac{1}{3})^T$，$s_0 = x_1 - x_0 = (-\frac{1}{3}, -\frac{1}{3})^T$, $y_0 = g_1 - g_0 = (-\frac{2}{3}, -\frac{4}{3})^T$，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 $\min f(x_1 + \alpha p_1) = \frac{1}{5}\alpha^2 - \frac{1}{6}\alpha + \frac{5}{144}$ 的极小点，得 $\alpha_1 = \frac{5}{12}$。$x_2 = x_1 + \alpha_1 p_1 = (0,0)^T$，$g_2 = (0,0)^T$

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：

$p_1 = u_1$，$\boldsymbol{p}_{k+1} = \boldsymbol{u}_{k+1} - \sum_{i=1}^{k} \frac{\boldsymbol{u}_{k+1}^T G \boldsymbol{p}_i}{\boldsymbol{p}_i^T G \boldsymbol{p}_i} \boldsymbol{p}_i (k = 1, 2, \cdots, n-1)$，证明 $\boldsymbol{p}_1, \boldsymbol{p}_2, \cdots, \boldsymbol{p}_n$ 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

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

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

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

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

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 $\boldsymbol{i} \in \{\boldsymbol{1, 2, \cdots, n}\}$，有 $p_i^T A x = \sum_{k=1}^{n} \alpha_k p_i^T A p_k$。

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

$p_i^T A x = \alpha_i p_i^T A p_i$， $\alpha_i = \dfrac{p_i^T A x}{p_i^T A p_i}$.

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

第四章计算题

一、用外罚函数法求解 $\begin{array}{l}\min\ x_1^2+3x_2^2\\ s.t.\ x_1+2x_2-1=0\end{array}$。

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 $P(x,\sigma)=x_1^2+3x_2^2+\frac{\sigma}{2}(x_1+2x_2-1)^2$)

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

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

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

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

二、用内罚函数法（对数罚函数）求解 $\begin{array}{l}\min\ x_1^2+5x_2^2\\ s.t.\ x_1+x_2-3\ge0\end{array}$。

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

令 $r\to0$， $x_1=\frac{5}{2},x_2=\frac{1}{2}.\ x\to x^*=(\frac{5}{2},\frac{1}{2})^T,f^*=\frac{15}{2}$。

三、用乘子法求解问题 $\begin{array}{l}\min\ x_1^2+2x_2^2\\ s.t.\ 2x_1+x_2-4=0\end{array}$。

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

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

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

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。

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

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。

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

$\lambda_2\ge 0$

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 $\lambda_2=\dfrac{3-p}{2}\geq 0$，解得 $p\leq 3$。

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

解：由 KT 条件，存在常数 λ_1,λ_2 使得

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

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

$\nabla f(x)=(2(x_1-1),2(x_2-2))^T$， $\nabla c_1(x)=(1,0)^T,\nabla c_2(x)=(1,1)^T$，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 $x_1=-\dfrac{1}{2},x_2=\dfrac{1}{2},\lambda_2=-3$，但因为 $(-\dfrac{1}{2},\dfrac{1}{2})$ 不满足可行条件 $x_1-2\geq 0$，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 $\begin{aligned}&\min\quad f(x)\\ &s.t.\quad x\geq 0\end{aligned}$ 的最优解的充要条件是 $x^*\geq 0,\nabla f(x^*)\geq 0,(\nabla f(x^*))^T x^*=0$。

证明：$\begin{aligned}&\min\quad f(x)\\ &s.t.\quad x\geq 0\end{aligned}$ 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

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

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 $x^* \geq 0, \nabla f(x^*) \geq 0$，

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