Exercise 1.写出基于Wolfe-Powell准则的非精确一维搜索算法中插值多项式$p^{(1)}(t), p^{(2)}(t)$的具体表达式。

设$p^{(1)}(t) = At^2 + Bt + C$，则有

$$\begin{cases} A\alpha_1^2 + B\alpha_1 + C & = \varphi_1 \\ 2A\alpha_1 + B & = \varphi_1' \\ A\alpha^2 + B\alpha + C & = \varphi \end{cases}$$

解得

$$\begin{cases} A = \frac{\varphi - \varphi_1 + \varphi_1'(\alpha_1 - \alpha)}{(\alpha - \alpha_1)^2} \\ B = \frac{\varphi_1'(\alpha^2 - \alpha_1)^2 + 2\alpha_1(\varphi_1 - \varphi)}{(\alpha - \alpha_1)^2} \\ C = \frac{\varphi_1\alpha(\alpha - 2\alpha_1) + \varphi_1'\alpha\alpha_1(\alpha_1 - \alpha) + \varphi\alpha_1^2}{(\alpha_1 - \alpha)^2} \end{cases}$$

设$p^{(2)}(t) = \tilde{A}t^2 + \tilde{B}t + \tilde{C}$，则有

$$\begin{cases} \tilde{A}\alpha^2 + \tilde{B}\alpha + C & = \varphi \\ 2\tilde{A}\alpha_1 + \tilde{B} & = \varphi_1' \\ 2\tilde{A}\alpha + \tilde{B} & = \varphi' \end{cases}$$

解得

$$\begin{cases} \tilde{A} = \frac{\varphi' - \varphi_1'}{2\alpha - 2\alpha_1} \\ \tilde{B} = \frac{\varphi_1'\alpha - \varphi'\alpha_1}{\alpha - \alpha_1} \\ \tilde{C} = \frac{\varphi(2\alpha - 2\alpha_1) + \varphi'(2\alpha\alpha_1 - \alpha^2) - \varphi_1'\alpha^2}{2\alpha - 2\alpha_1} \end{cases}$$

Exercise 2.证明基于Goldstein准则的非精确一维搜索算法的全局收敛性。

Goldstein准则：

$$\begin{cases} \varphi(\alpha) & \leq \varphi(0) + \rho\alpha\varphi'(0) & (92) \\ \varphi(\alpha) & \geq \varphi(0) + (1-\rho)\alpha\varphi'(0) & (93) \end{cases}$$

设$\forall k, g^{(k)} = \nabla f(x^k) \neq 0$和$f(x^{(k)})$有下界，则$f(x^{(k)}) - f(x^{(k+1)}) \to 0$，由(92)式，$-g^{(k)}s^{(k)} \to 0$。

（反证法）若$g^{(k)} \to 0$不成立，则$\exists \epsilon > 0$和子列$\{x^{(k)}\}_{k\in K} s.t. \|g^{(k)}\| \geq \epsilon$从而由

$$-{g^{(k)}}^\top s^{(k)} = \|g^{(k)}\|\|s^{(k)}\|\cos\theta_k \geq \epsilon\|s^{(k)}\|\sin\mu$$

以及$\theta_k \leq \frac{\pi}{2} - \mu, \forall k$得$\|s^{(k)}\| \to 0.$

又有

$$f(x^{(k)} + s^{(k)}) = f(x^{(k)}) + \nabla f(x^{(k)})^\top s^{(k)} + o(\|s^{(k)}\|)(\text{泰勒展开})$$

得

$$\lim_{k\to\infty}\frac{f(x^{(k)})-f(x^{(k+1)})}{-\nabla {f^{(k)}}^\top s^{(k)}}=1$$

由(93)式得

$$\frac{f(x^{(k)})-f(x^{(k+1)})}{-\nabla {f^{(k)}}^\top s^{(k)}}\le 1-\rho<1$$

矛盾!

所以$g^{(k)}=\nabla f(x^{(k)})\to 0, k\to\infty.$

Exercise 3.将非线性方程组求根$F(x)=0$的牛顿迭代，用于求最优化问题$\min_{x\in\mathbb{R}^n} f(x)$，给出相应的迭代格式并说明理由。

牛顿迭代法的原理是取函数$F(x)$在$x^{(k)}$处泰勒展开的线性部分，作为函数的一阶近似，令其等于0并得到迭代格式，即:

$$F(x^{(k)})+DF(x^{(k)})(x-x^{(k)})=0$$

其中$DF(x^{(k)})$为F在$x^{(k)}$的Jacobian，其形式为

$$DF(x)=\begin{bmatrix}\frac{\partial F_1(x)}{\partial x_1} & \frac{\partial F_1(x)}{\partial x_2} & \cdots & \frac{\partial F_1(x)}{\partial x_n}\\ \frac{\partial F_2(x)}{\partial x_1} & \frac{\partial F_2(x)}{\partial x_2} & \cdots & \frac{\partial F_n(x)}{\partial x_n}\\ \vdots & \vdots & \ddots & \vdots\\ \frac{\partial F_n(x)}{\partial x_1} & \frac{\partial F_n(x)}{\partial x_2} & \cdots & \frac{\partial F_n(x)}{\partial x_n}\end{bmatrix}$$

解得迭代格式为

$$x^{(k+1)}=x^{(k)}-{DF(x^{(k)})}^{-1}F(x^{(k)})$$

同理，对于无约束最优化问题$min_{x\in\mathbb{R}^n f(x)}$, 类似地求解$\nabla f(x)=0$即可，得到迭代格式为

$$x^{(k+1)}=x^{(k)}-(D(\nabla f(x)))^{-1}\nabla f(x^{(k)})=x^{(k)}-(\nabla^2 f(x))^{-1}\nabla f(x^{(k)})$$

Exercise 4.证明对称秩一牛顿法具有遗传性和二次终止性

对于二次函数$f(x)=\frac{1}{2}x^\top Gx+c^\top x$,

$$\nabla f(x)=Gx+c,\nabla^2 f(x)=G'$$

由牛顿法$y^{(k)}=G_k s^{(k)}$, 正割条件为$H_{k+1}y^{(k)}=s^{(k)}$

遗传性：使用归纳法。

1. $k=1$时，由正割条件，$H_1 y^{(0)}=s^{(0)}$。

2. 假设遗传性对于k成立，即$H_k y^{(l)} = s^{(l)}, l = 0, 1, \ldots, k-1.$
3. 对于$k+1$, 只须证明$l < k$的情况（$l = k$时，由正割条件直接可得结论成立），即

$$H_{k+1} y^{(l)} = s^{(l)}, l = 0, 1, \ldots, k-1$$

   由对称秩一校正公式

$$H_{k+1} y^{(l)} = H_k y^{(l)} + \frac{(s^{(k)} - H_k y^{(k)})(s^{(k)} - H_k y^{(k)})^\top y^{(l)}}{(s^{(k)} - H_k y^{(k)})^\top y^{(k)}}$$

   其中

$$\begin{aligned}(s^{(k)} - H_k y^{(k)})^\top y^{(l)} &= {s^{(k)}}^\top y^{(l)} - {y^{(k)}}^\top H_k y^{(l)} \\ &= {s^{(k)}}^\top y^{(l)} - {y^{(k)}}^\top s^{(l)} \\ &= {s^{(k)}}^\top G s^{(l)} - {s^{(k)}}^\top G s^{(l)} \\ &= 0\end{aligned}$$

   故$H_{k+1} y^{(l)} = H_k y^{(l)} = s^{(l)}, l = 0, 1, \ldots, k-1.$

二次终止性：（这里需要假定$s_0, s_1, \ldots, s_{n-1}$线性无关）

由遗传性知

$$\begin{gathered} H_n y^{(l)} = s^{(l)}, l = 0, 1, \ldots, n-1 \\ \Downarrow \\ H_n G s^{(l)} = s^{(l)}, l = 0, 1, \ldots, n-1 \\ \Downarrow \\ (H_n G - I) s^{(l)} = 0, l = 0, 1, \ldots, n-1 \end{gathered}$$

因为$s_0, s_1, \ldots, s_{n-1}$线性无关，所以$(s_0, s_1, \ldots, s_{n-1})$可逆，所以

$$H_n G - I = 0 \Rightarrow H_n = G^{-1} = (\nabla^2 f(x))^{-1}$$

由迭代格式

$$x^{(n+1)} = x^{(n)} - H_n \nabla f(x^{(n)}) = x^{(n)} - G^{-1} \nabla f(x^{(n)})$$

又因为

$$\begin{gathered} \nabla f(x^{(n+1)}) - \nabla f(x^{(n)}) = G(x^{(n+1)} - x^{(n)}) \\ \Downarrow \\ x^{(n+1)} = x^{(n)} - G^{-1} \nabla f(x^{(n)}) + G^{-1} \nabla f(x^{(n+1)}) \end{gathered}$$

所以

$$G^{-1} \nabla f(x^{(n+1)}) = 0 \Rightarrow \nabla f(x^{(n+1)}) = 0$$

即有限步终止且$H_n = [\nabla^2 f(x^*)]^{-1}.$

Exercise 5.利用秩一校正的求逆公式（sherman-Morrison定理），由$H_{k+1}^{(DFP)}$推导$B_{k+1}^{(DFP)}$.

$$
\begin{aligned}
(A+uv^\top)^{-1} &= A^{-1} - \frac{A^\top uv^\top A^{-1}}{1+v^\top A^{-1}u} \\
H_{k+1}^{(DFP)} &= H_k + \frac{s^{(k)}s^{(k)\top}}{s^{(k)\top}y^{(k)}} - \frac{H_k y^{(k)} y^{(k)\top} H_k}{y^{(k)\top} H_k y^{(k)}} \\
B_{k+1}^{(DFP)} &= B_k + (1+\frac{s^{(k)\top}B_k s^{(k)}}{y^{(k)\top}s^{(k)}})\frac{y^{(k)}y^{(k)\top}}{y^{(k)\top}s^{(k)}} - \frac{B_k s^{(k)} y^{(k)\top} + y^{(k)} s^{(k)\top} B_k}{y^{(k)\top}s^{(k)}}
\end{aligned}
$$

为方便书写，忽视所有的角标k, 并令$M = H + \frac{ss^\top}{s^\top y}$，则有

$$
M^{-1} = H^{-1} - \frac{H^{-1}ss^\top H^{-1}}{s^\top y + s^\top Bs} = B - \frac{Bss^\top B}{s^\top y + s^\top Bs} \quad (5.1)
$$

$$
B_{k+1}^{(DFP)} = (H_{k+1}^{(DFP)})^{-1} = M^{-1} + \frac{M^{-1}Hyy^\top HM^{-1}}{y^\top Hy - y^\top HM^{-1}Hy} \quad (5.2)
$$

将(5.1)代入(5.2)第二个等号右边第二项并展开，消去分子分母常数和$BH = I$得

$$
\begin{aligned}
\frac{M^{-1}Hyy^\top HM^{-1}}{y^\top Hy - y^\top HM^{-1}Hy} &= \frac{yy^\top(s^\top y + s^\top Bs)}{y^\top ss^\top y} - \frac{ys^\top B}{s^\top y} - \frac{Bsy^\top}{y^\top s} + \frac{Bss^\top B}{s^\top y + s^\top Bs} \\
&= (1+\frac{s^\top Bs}{y^\top s})\frac{yy^\top}{y^\top s} - \frac{ys^\top B + Bsy^\top}{y^\top s} + \frac{Bss^\top B}{s^\top y + s^\top Bs} \quad (5.3)
\end{aligned}
$$

将(5.1), (5.3)代入(5.2)即可得

$$
B_{k+1}^{(DFP)} = B_k + (1+\frac{s^{(k)\top}B_k s^{(k)}}{y^{(k)\top}s^{(k)}})\frac{y^{(k)}y^{(k)\top}}{y^{(k)\top}s^{(k)}} - \frac{B_k s^{(k)} y^{(k)\top} + y^{(k)} s^{(k)\top} B_k}{y^{(k)\top}s^{(k)}}
$$

Exercise 6.共轭梯度法性质定理：设目标函数$f(x) = \frac{1}{2}x^\top Gx + c^\top x$,则采用精确一维搜索的共轭梯度法经$m \le n$步迭代后终止，且对所有的$1 \le k \le m$成立下列关系：

$$
\begin{aligned}
&d^{(k)\top}Gd^{(j)} = 0, j = 0,1,\dots,k-1 && (6.1)\\
&g^{(k)\top}g^{(j)} = 0, j = 0,1,\dots,k-1 && (6.2)\\
&d^{(k)\top}g^{(k)} = -g^{(k)\top}g^{(k)} && (6.3)\\
&span\{g^{(0)},g^{(1)},\dots,g^{(k)}\} = span\{g^{(0)},Gg^{(0)},\dots,G^k g^{(0)}\} && (6.4)\\
&span\{d^{(0)},d^{(1)},\dots,d^{(k)}\} = span\{d^{(0)},Gg^{(0)},\dots,G^k g^{(0)}\} && (6.5)
\end{aligned}
$$

共轭梯度法步骤中得到的的条件：

$$\begin{aligned}
&g^{(k+1)^\top} d^{(k)} = 0 && (\text{由精确一维搜索})(6.6)\\
&\alpha_k = \frac{g^{(k)^\top} g^{(k)}}{d^{(k)^\top} G d^{(k)}} && (\text{由精确一维搜索})(6.7)\\
&g^{(k+1)} = g^{(k)} + \alpha_k G d^{(k)} && (\nabla f(x^{(k+1)})\text{直接展开})(6.8)\\
&\beta_k = \frac{g^{(k+1)^\top} g^{(k+1)}}{g^{(k)^\top} g^{(k)}} && (6.9)\\
&d^{(k+1)} = -g^{(k+1)} + \beta_k d^{(k)} && (6.10)
\end{aligned}$$

证明(6.3)**:**

$$\begin{aligned}
d^{(k)^\top} g^{(k)} &= -g^{(k)^\top} g^{(k)} + \beta_k d^{(k-1)^\top} g^{(k)} && (by(6.10))\\
&= -g^{(k)^\top} g^{(k)} + 0 && (by(6.6))\\
&= -g^{(k)^\top} g^{(k)}
\end{aligned}$$

证明(6.1)**与**(6.2)**:**

1. k = 1时直接验证可得结论成立
2. 假设(6.1)与(6.2)对k成立
3. (6.8)式左右两边转置后同右乘$g^{(j)}$得

$$\begin{aligned}
g^{(k+1)} g^{(j)} &= g^{(k)^\top} g^{(j)} - \alpha_k d^{(k)^\top} G g^{(j)}\\
&= g^{(k)^\top} g^{(j)} - \alpha_k d^{(k)^\top} G(d^{(j)} - \beta_{j-1} d^{(j-1)})\\
&= g^{(k)^\top} g^{(j)} - \alpha_k d^{(k)^\top} G d^{(j)}
\end{aligned}$$

$j = k$时，将(6.7)式代入即可得上式为0；$j < k$时，由归纳假设得上式为0。

综上，(6.2)成立。

4. 由(6.10)式，

$$\begin{aligned}
d^{(k+1)^\top} G d^{(j)} &= (-g^{(k+1)} + \beta_k d^{(k)}) G d^{(j)}\\
&= -g^{(k+1)} G d^{(j)} + \beta_k d^{(k)} G d^{(j)}\\
&= g^{(k+1)^\top} (g^{(j)} - g^{(j+1)})/\alpha_k + \beta_k d^{(k)} G d^{(j)}
\end{aligned}$$

$j = k$时，由$(6.2), (6.7), (6.9)$得上式为0；$j < k$时，由归纳假设得上式为0.

综上，(6.1)成立。

证明(6.4)

由(6.10)式知，存在可逆方阵

$$Q = \begin{bmatrix}
-1 & B_0 & 0 & \dots & 0 & 0\\
0 & -1 & B_1 & \dots & 0 & 0\\
0 & 0 & -1 & \dots & 0 & 0\\
\vdots & \vdots & \vdots & \ddots & \vdots & \vdots\\
0 & 0 & 0 & \dots & -1 & B_{k-1}\\
0 & 0 & 0 & \dots & 0 & -1
\end{bmatrix}$$

使得$(d^{(0)}, d^{(1)}, \dots, d^{(k)})Q = (g^{(0)}, g^{(1)}, \dots, g^{(k)})$，所以

$$span\{g^{(0)}, g^{(1)}, \ldots, g^{(k)}\} = span\{d^{(0)}, d^{(1)}, \ldots, d^{(k)}\}(6.11)$$

1. $k=0$时，直接由定义得结论成立。
2. 假设结论对k成立，即$span\{g^{(0)}, g^{(1)}, \ldots, g^{(k)}\} = span\{g^{(0)}, Gg^{(0)}, \ldots, G^k g^{(0)}\}$.
3. 对于$k+1$，由(6.8)式和归纳假设，
$$g_{k+1} \in span\{g^{(0)}, Gg^{(0)}, \ldots, G^{k+1} g^{(0)}\}$$

   $d^{(0)}, d^{(1)}, \ldots, d^{(k)}$是一组共轭方向，由共轭方向法基本定理得
$$g^{(k+1)} \perp span\{d^{(0)}, d^{(1)}, \ldots, d^{(k)}\}(6.12)$$

   所以由$(6.11), (6.12)$和归纳假设
$$g^{(k+1)} \notin span\{d^{(0)}, d^{(1)}, \ldots, d^{(k)}\} = span\{g^{(0)}, Gg^{(0)}, \ldots, G^k g^{(0)}\}$$

   所以结论对$k+1$成立，即
$$span\{g^{(0)}, g^{(1)}, \ldots, g^{(k+1)}\} = span\{g^{(0)}, Gg^{(0)}, \ldots, G^{k+1} g^{(0)}\}$$

(6.5)式证明与(6.4)式同理。

Exercise 8.证明折线法（信赖域方法）子问题模型的函数单调性。

$$\begin{aligned}
s_C^{(k)} &= -\frac{\|g^{(k)}\|^2}{{g^{(k)}}^\top B_k g^{(k)}} g^{(k)} \\
s_N^{(k)} &= -B_k^{-1} g^{(k)} \\
x^{(k+1)} &= x^{(k)} + s_C^{(k)} + \lambda(s_N^{(k)} - s_C^{(k)})
\end{aligned}$$

(1)证明沿着Cauchy点$x_C^{(k+1)}$和牛顿点$x_N^{(k+1)}$的连线，到$x^{(k)}$的距离单调增加.

$$\begin{aligned}
L(\lambda) &= \|s_C^{(k)} + \lambda(s_N^{(k)} - s_C^{(k)})\|^2 \\
&= (s_C^{(k)} + \lambda(s_N^{(k)} - s_C^{(k)}))^\top (s_C^{(k)} + \lambda(s_N^{(k)} - s_C^{(k)})) \\
&= {s_C^{(k)}}^\top s_C^{(k)} + 2\lambda {s_C^{(k)}}^\top (s_N^{(k)} - s_C^{(k)})) + \lambda^2 (s_N^{(k)} - s_C^{(k)}))^\top (s_N^{(k)} - s_C^{(k)})) \\
L(\lambda)' &= 2{s_C^{(k)}}^\top (s_N^{(k)} - s_C^{(k)})) + 2\lambda(s_N^{(k)} - s_C^{(k)}))^\top (s_N^{(k)} - s_C^{(k)})) \\
&\geq 2{s_C^{(k)}}^\top (s_N^{(k)} - s_C^{(k)})) \\
&= 2\frac{\|g^{(k)}\|^2}{{g^{(k)}}^\top B_k g^{(k)}} g^{(k)} B_k^{-1} g^{(k)} (1 - \frac{\|g^{(k)}\|^2 g^{(k)}}{{g^{(k)}}^\top B_k g^{(k)} B_k^{-1} g^{(k)}})
\end{aligned}$$

其中

$$
\begin{aligned}
1-\frac{\|g^{(k)}\|^2 g^{(k)}}{g^{(k)\top}B_k g^{(k)} B_k^{-1} g^{(k)}} &= 1-\frac{\|g^{(k)}\|^4}{g^{(k)\top}B_k g^{(k)} g^{(k)\top} B_k^{-1} g^{(k)}} \\
&= 1-\frac{((\sqrt{B_k}g^{(k)})^\top((\sqrt{B_k^{-1}}g^{(k)}))^2}{g^{(k)\top}B_k g^{(k)} g^{(k)\top} B_k^{-1} g^{(k)}} \\
&\geq 1-\frac{\|\sqrt{B_k}g^{(k)}\|^2\|\sqrt{B_k^{-1}}g^{(k)}\|^2}{g^{(k)\top}B_k g^{(k)} g^{(k)\top} B_k^{-1} g^{(k)}} \\
&= 1-\frac{g^{(k)\top}B_k g^{(k)} g^{(k)\top} B_k^{-1} g^{(k)}}{g^{(k)\top}B_k g^{(k)} g^{(k)\top} B_k^{-1} g^{(k)}} \\
&= 0
\end{aligned}
$$

所以$L(\lambda)' \geq 0, \lambda \in [0,1]$，具有单调性。

(2)证明沿着Cauchy点$x_C^{(k+1)}$和牛顿点$x_N^{(k+1)}$的连线，子问题模型函数值单调减少。

$$
\begin{aligned}
h(\lambda) &= q^{(k)}(s_C^{(k)}+\lambda(s_N^{(k)}-s_C^{(k)})) \\
&= f(x^{(k)}) + g^{(k)\top}(s_C^{(k)}+\lambda(s_N^{(k)}-s_C^{(k)})) \\
&\quad + \frac{1}{2}(s_C^{(k)}+\lambda(s_N^{(k)}-s_C^{(k)}))^\top B_k(s_C^{(k)}+\lambda(s_N^{(k)}-s_C^{(k)})) \\
h(\lambda)' &= g^{(k)\top}(s_N^{(k)}-s_C^{(k)}) + s_C^{(k)\top}B_k(s_N^{(k)}-s_C^{(k)}) + \lambda(s_N^{(k)}-s_C^{(k)})^\top B_k(s_N^{(k)}-s_C^{(k)}) \\
&\leq g^{(k)\top}(s_N^{(k)}-s_C^{(k)}) + s_C^{(k)\top}B_k(s_N^{(k)}-s_C^{(k)}) + (s_N^{(k)}-s_C^{(k)})^\top B_k(s_N^{(k)}-s_C^{(k)}) \\
&= (g^{(k)\top} + s_C^{(k)\top}B_k + (s_N^{(k)}-s_C^{(k)})^\top B_k)(s_N^{(k)}-s_C^{(k)}) \\
&= (g^{(k)\top} + s_N^{(k)\top}B_k)(s_N^{(k)}-s_C^{(k)}) \\
&= 0\cdot(s_N^{(k)}-s_C^{(k)}) \\
&= 0
\end{aligned}
$$

所以沿着Cauchy点$x_C^{(k+1)}$和牛顿点$x_N^{(k+1)}$的连线，子问题模型函数值单调减少。