1. Consider the following general iterative algorithm ($\mathcal{A}$) for finding the minimum of $f : \mathbb{R}^n \to \mathbb{R}$.

   i. Pick a starting point $x^{(0)} \in \mathbb{R}^n$.

   ii. For $k \geq 0$, let
   $$x^{(k+1)} = x^{(k)} + \alpha_k d^{(k)}$$
   where $d^{(0)}, d^{(1)}, \ldots$ are given vectors in $\mathbb{R}^n$ and α_k is chosen to minimize $f(x^{(k)} + \alpha d^{(k)})$; that is,
   $$\alpha_k = \arg\min_{\alpha \geq 0} f(x^{(k)} + \alpha d^{(k)}).$$

   Show that for each $k \geq 0$, the vector $x^{(k+1)} - x^{(k)}$ is orthogonal to $\nabla f(x^{(k+1)})$ (assuming that the gradient exists).
   Hint: Consider the FONC for $\phi_k(\alpha) = f(x^{(k)} + \alpha d^{(k)})$.

2. *Conjugate direction algorithm* –

   (a) Let $Q \in \mathbb{R}^{n \times n}$ be symmetric and positive definite and let $d^{(0)}, d^{(1)}, \ldots, d^{(n-1)} \in \mathbb{R}^n$ be non-zero and Q-conjugate. Show that $d^{(0)}, d^{(1)}, \ldots, d^{(n-1)}$ are linearly independent.

   (b) Show that in the basic conjugate direction algorithm,
   $$g^{(k+1)\top} d^{(i)} = 0$$
   for all $0 \leq k \leq n-1$ and $0 \leq i \leq k$.

3. *Conjugate gradient algorithm* ("Pick the directions as you go") – Use 2.(b) to show that in the conjugate gradient algorithm where
   $$d^{(k+1} = d^{(k)} + \beta_k d^{(k)}$$
   with $\beta_k = \frac{g^{(k+1)^\top} Q d^{(k)}}{d^{(k)^\top} Q d^{(k)}}$ and $g^{(0)} = \nabla f(x^{(0)})$ for some fixed starting point $x^{(0)}$, the directions $d^{(0)}, d^{(1)}, \ldots, d^{(n-1)}$ are Q-conjugate.

4. *Non-linear least squares* – Consider the following problem:
   $$\text{minimize} \sum_{i=1}^{m} r_i(x)^2$$
   where $r_i : \mathbb{R}^n \to \mathbb{R}$, $i = 1, \ldots, m$ are given functions. Solve using Newton's method.

   Hint: Define $r = [r_1 \ \cdots \ r_m]$; write the objective function as $f(x) = r(x)^\top r(x)$ and compute the gradient and Hessian of f.

   Note: You may refer to the textbook (Chapter 9) to answer this question. But make sure you understand the idea and do not just copy the answer. Points will be awarded only for answers with justification for each step.

5. *Newton's descent* – Prove that Newton's descent algorithm defined by

$$x^{(k+1)} = x^{(k)} - \alpha_k F(x^{(k)})^{-1} \nabla f(x^{(k)})$$

where

$$\alpha_k = \arg\min_{\alpha \geq 0} f(x^{(k)} - \alpha F(x^{(k)})^{-1} \nabla f(x^{(k)}))$$

has the descent property, namely, for all $k \geq 0$,

$$f(x^{(k+1)}) < f(x^{(k)}).$$

6. *Quasi-Newton* – Use the following "rank one algorithm" to minimize $f(x_1, x_2) = \frac{1}{2}x_1^2 + \frac{1}{3}x_2^2 + 1$.

   i. Set $k = 0$; select $x^{(0)}$ and a real symmetric positive definite H_0.

   ii. If $\nabla f(x^{(k)}) = 0$, stop; else let $d^{(k)} = -H_k \nabla f(x_k)$.

   iii. Compute

$$\begin{aligned} \alpha_k &= \arg\min_{\alpha \geq 0} f(x^{(k)} + \alpha d^{(k)}) \\ x^{(k+1)} &= x^{(k)} + \alpha_k d^{(k)} \end{aligned}$$

   iv. Compute

$$\begin{aligned} a_k &= \nabla f(x^{(k+1)}) - \nabla f(x^{(k)}) \\ b_k &= \alpha_k d^{(k)} - H_k a_k \\ H_{k+1} &= H_k + \frac{b_k b_k^\top}{a_k^\top b_k}. \end{aligned}$$

   v. Set $k = k + 1$; go to step (ii).

   Start at $x^{(0)} = [2 \;\; 3]^\top$ and let H_0 be the 2×2 identity matrix.

   Hint: Since the objective function is quadratic, recall that

$$\alpha_k = \arg\min_{\alpha \geq 0} f(x^{(k)} + \alpha d^{(k)}) = -\frac{\nabla f(x^{(k)})^\top d^{(k)}}{d^{(k)\top} Q d^{(k)}}.$$

7. *Line search* – Consider a unimodal function f of one variable on the interval $[a_0, b_0]$. Fix $\rho < 1/2$ and $N > 0$. Our goal is to progressively narrow the range until the minimizer of f is "boxed in" with sufficient accuracy after N steps. Consider the following algorithm based on binary search.

   - For $i = 0, 1, 2, \ldots, N$, do:

     1) Set

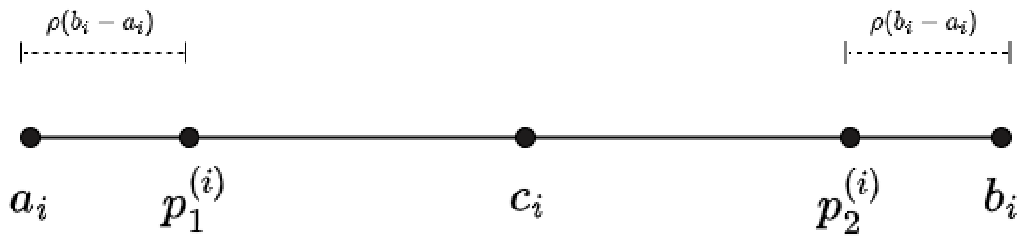
$$\begin{aligned} c_i &= \frac{1}{2}a_i + \frac{1}{2}b_i \\ p_1^{(i)} &= (1 - \rho)a_i + \rho b_i \\ p_2^{(i)} &= \rho a_i + (1 - \rho)b_i. \end{aligned}$$

2) If $f(p_1^{(i)}) \leq f(p_2^{(i)})$
   - set $a_{i+1} = a_i$ and $b_{i+1} = c_i$.

   Else
   - set $a_{i+1} = c_i$ and $b_{i+1} = b_i$.

- Output $[a_{N+1}, b_{N+1}]$ as the interval containing the minimizer of f.

Note that in the i-th step, c_i is the midpoint of the line joining a_i and b_i, $p_1^{(i)}$ is to the left of c_i, and $p_2^{(i)}$ is to the right of c_i (see figure).



i. Show that this algorithm may not always output the interval containing the minimizer (by constructing an appropriate unimodal function f)?

ii. Exhibit an example function f for which the algorithm converges to the minimizer? You do not need to provide proof of convergence.

8. Show that for $x \geq 0$, we have
$$\left(\frac{x}{6}\right)^{3/2} \geq \frac{x-2}{4}.$$

Hint: Consider $f(x) = \left(\frac{x}{6}\right)^{3/2} - \frac{x-2}{4}$. Verify that $f'(x)$ is an increasing function of x, that f attains the minimum at a certain x^*, and that $f(x^*) = 0$. Why would verifying these conditions imply the given inequality?