Math 164 - Assignment #5, Fall 2024

- **Due Date:** 11:59 pm on November 8st, 2024

---

1. Consider the following algorithm for minimizing a function f :

$$\boldsymbol{x}^{(k+1)} = \boldsymbol{x}^{(k)} + \alpha_k \boldsymbol{d}^{(k)},$$

where $\alpha_k = \arg\min_\alpha f\left(\boldsymbol{x}^{(k)} + \alpha \boldsymbol{d}^{(k)}\right)$. Let $\boldsymbol{g}^{(k)} = \nabla f\left(\boldsymbol{x}^{(k)}\right)$ (as usual).

Suppose that f is quadratic with Hessian $\mathbf{Q}$. We choose $\mathbf{d}^{(k+1)} = \gamma_k \boldsymbol{g}^{(k+1)} + \boldsymbol{d}^{(k)}$, and we wish the directions $\boldsymbol{d}^{(k)}$ and $\boldsymbol{d}^{(k+1)}$ to be $\boldsymbol{Q}$-conjugate. Find a formula for γ_k in terms of $\boldsymbol{d}^{(k)}$, $\boldsymbol{g}^{(k+1)}$, and $\boldsymbol{Q}$.

2. Represent the function

$$f(\boldsymbol{x}) = \frac{5}{2} x_1^2 + x_2^2 - 3x_1x_2 - x_2 - 7$$

In the form $f(\boldsymbol{x}) = \frac{1}{2}\boldsymbol{x}^\top Q \boldsymbol{x} - \boldsymbol{x}^\top \boldsymbol{b} + c$. Then, use the conjugate gradient algorithm to find a minimizer with $\boldsymbol{d}^{(0)} = \nabla f(\boldsymbol{x}^{(0)})$, where $\boldsymbol{x}^{(0)} = \mathbf{0}$.

3. Consider the DFP algorithm applied to the quadratic function

$$f(\boldsymbol{x}) = \frac{1}{2}\boldsymbol{x}^\top \boldsymbol{Q}\boldsymbol{x} - \boldsymbol{x}^\top \boldsymbol{b}$$

where $\boldsymbol{Q} = \boldsymbol{Q}^\top > 0$.

   (a) Write down a formula for α_k in terms of $\boldsymbol{Q}, \boldsymbol{g}^{(k)}$, and $\boldsymbol{d}^{(k)}$.

   (b) Show that if $\boldsymbol{g}^{(k)} \neq 0$, then $\alpha_k > 0$.

4. Let $f : \mathbb{R}^n \to \mathbb{R}$ be such that $f \in \mathcal{C}^1$. Consider an optimization algorithm applied to this f, of the usual form $\boldsymbol{x}^{(k+1)} = \boldsymbol{x}^{(0)} + a_k \boldsymbol{d}^{(k)}$, where $\alpha_k \geq 0$ is chosen according to line search. Suppose that $\boldsymbol{d}^{(k)} = -\boldsymbol{H}_k \boldsymbol{g}^{(k)}$, where $\boldsymbol{g}^{(k)} = \nabla f(\boldsymbol{x}^{(k)})$ and $\boldsymbol{H}_k$ is symmetric.

   (a) Show that if $\boldsymbol{H}_k$ satisfies the following conditions whenever the algorithm is applied to a quadratic, then the algorithm is quasi-Newton:

      i. $\boldsymbol{H}_{k+1} = \boldsymbol{H}_k + \boldsymbol{U}_k$.

      ii. $\boldsymbol{U}_k \Delta \boldsymbol{g}^{(k)} = \Delta \boldsymbol{x}^{(k)} - \boldsymbol{H}_k \Delta \boldsymbol{g}^{(k)}$.

      iii. $\boldsymbol{U}_k = \boldsymbol{a}^{(k)} (\Delta \boldsymbol{x}^{(k)})^\top + \boldsymbol{b}^{(k)} (\Delta \boldsymbol{g}^{(k)})^\top \boldsymbol{H}_k$, where $\boldsymbol{a}^{(k)}$ and $\boldsymbol{b}^{(k)}$ are in $\mathbb{R}^n$.

   (b) Which (if any) among the rank-one, DFP, and BFGS algorithms satisfy the three conditions in part a (whenever the algorithm is applied to a quadratic)? For those that do, specify the vectors $\boldsymbol{a}^{(k)}$ and $\boldsymbol{b}^{(k)}$.

5. Given a function $f : \mathbb{R}^n \to \mathbb{R}$, consider an algorithm $\boldsymbol{x}^{(k+1)} = \boldsymbol{x}^{(k)} - \alpha_k \boldsymbol{H}_k \boldsymbol{g}^{(k)}$ for finding the minimizer of f, where $\boldsymbol{g}^{(k)} = \nabla f\left(\boldsymbol{x}^{(k)}\right)$ and $\boldsymbol{H}_k \in \mathbb{R}^{n \times n}$ is symmetric. Suppose that $\boldsymbol{H}_k = \phi \boldsymbol{H}_k^{\mathrm{DFP}} + (1 - \phi)\boldsymbol{H}_k^{\mathrm{BFGS}}$, where $\phi \in \mathbb{R}$, and $\boldsymbol{H}_k^{\mathrm{DFP}}$ and $\boldsymbol{H}_k^{\mathrm{BFGS}}$ are matrices generated by the DFP and BFGS algorithms, respectively.

   (a) Show that the algorithm above is a quasi-Newton algorithm. Is the above algorithm a conjugate direction algorithm?

   (b) Suppose that $0 \leq \phi \leq 1$. Show that if $\boldsymbol{H}_0^{\mathrm{DFP}} > 0$ and $\boldsymbol{H}_0^{\mathrm{BFGS}} > 0$, then $\boldsymbol{H}_k > 0$ for all k. What can you conclude from this about whether or not the algorithm has the descent property?