# Continuous Optimization: Assignment 5

Due on May 28, 2024

**Honglu Ma** **Hiroyasu Akada** **Mathivathana Ayyappan**

# Exercise 1

Consider quadratic function $f(x) = \frac{1}{2}x^\top Qx + b^\top x$ where $Q$ is a symmetric positive definite matrix. Using Newton's method we obtain

$$\begin{aligned} x^{(1)} &= x^{(0)} - (\nabla^2 f(x^{(0)}))^{-1} \nabla f(x^{(0)}) \\ &= x^{(0)} - Q^{-1}(Qx^{(0)} + b) \\ &= x^{(0)} - Q^{-1}Qx^{(0)} + Q^{-1}b \\ &= x^{(0)} - x^{(0)} + Q^{-1}b \\ &= Q^{-1}b \end{aligned}$$

$\nabla f(x^{(1)}) = -QQ^{-1}b + b = 0 \Rightarrow x^{(1)}$ is the global minimizer

# Exercise 2

## (a)

$Q \in \mathbb{S}_{++}(n)$ with eigenvectors $(v_1, \dots, v_n)$. We want to show $v_i^\top Q v_j = 0$ for any $i \neq j$. We know $Q = U\Lambda U^\top$ where

$$U = \begin{pmatrix} | & & | \\ v_1 & \cdots & v_n \\ | & & | \end{pmatrix}$$

$U$ is an orthogonal matrix i.e. $UU^\top = I$.
For any $i \neq j$, we have

$$\begin{aligned} v_i^\top Q v_j &= v_i^\top U \Lambda U^\top v_j \\ &= v_i^\top \begin{pmatrix} | & & | \\ v_1 & \cdots & v_n \\ | & & | \end{pmatrix} \begin{pmatrix} \lambda_1 & & 0 \\ & \ddots & \\ 0 & & \lambda_n \end{pmatrix} \begin{pmatrix} | & & | \\ v_1 & \cdots & v_n \\ | & & | \end{pmatrix}^\top v_j \\ &= \begin{pmatrix} 0 & \cdots & v_i^\top v_i & \cdots & 0 \end{pmatrix} \begin{pmatrix} \lambda_1 & & 0 \\ & \ddots & \\ 0 & & \lambda_n \end{pmatrix} \begin{pmatrix} 0 \\ \vdots \\ v_j^\top v_j \\ \vdots \\ 0 \end{pmatrix} \\ &= \begin{pmatrix} 0 & \cdots & \lambda_i v_i^\top v_i & \cdots & 0 \end{pmatrix} \begin{pmatrix} 0 \\ \vdots \\ v_j^\top v_j \\ \vdots \\ 0 \end{pmatrix} \\ &= 0 \end{aligned}$$

## (b)

$\{d^{(0)}, \cdots, d^{(k)}\}$ are conjugate directions w.r.t. $Q$. We have $d_i^\top Q d_j = 0$ for any $i \neq j$.

Exercise 2 continued on next page... 

**Proof by Contradiction:**
Assume that $\{d^{(0)}, \cdots, d^{(k)}\}$ is not a linearly independent set. Then there exists a non-zero vector $d_j$ such that $d_j$ can be expressed as a linear combination of some other vectors in the set i.e. $d_j = \sum_{i \neq j} \delta_i d_i$. Take $d_m$ that is part of the linear combination. We have

$$\begin{aligned}
d_j^\top Q d_m &= \left(\sum_{i \neq j} \delta_i d_i\right)^\top Q d_m \\
&= \delta_m d_m^\top Q d_m + \left(\sum_{i \neq j, i \neq m} \delta_i d_i\right)^\top Q d_m \\
&= \delta_m d_m^\top Q d_m && \text{Property of Conjugate Directions} \\
&\neq 0 && Q \in \mathbb{S}_{++}(n)
\end{aligned}$$

which is a contradiction.

### (c)

For all $i = 0, \cdots, k-1$, we have

$$\begin{aligned}
\langle d^{(i)}, \nabla f(x^{(k)}) \rangle &= \langle d^{(i)}, Qx^{(k)} - b \rangle \\
&= \langle Qx^{(k)}, d^{(i)} \rangle - \langle b, d^{(i)} \rangle \\
&= \langle Q \left( x^{(i+1)} + \sum_{j=i+1}^{k-1} \tau_j d^{(j)} \right), d^{(i)} \rangle - \langle b, d^{(i)} \rangle \\
&= \langle Qx^{(i+1)}, d^{(i)} \rangle + \sum_{j=i+1}^{k-1} \langle \tau_j d^{(j)}, Qd^{(i)} \rangle - \langle b, d^{(i)} \rangle \\
&= \langle Qx^{(i+1)}, d^{(i)} \rangle - \langle b, d^{(i)} \rangle && \text{Property of Conjugate Directions} \\
&= \langle Qx^{(i+1)} - b, d^{(i)} \rangle \\
&= \langle \nabla f(x^{(i+1)}), d^{(i)} \rangle \\
&= 0 && \text{Exact line search optimal condition}
\end{aligned}$$

# Exercise 3

We expand the BFGS update formula:

$$\begin{aligned}
H_{k+1} &= (I - \rho_k s^{(k)} (y^{(k)})^\top) H_k (I - \rho_k y^{(k)} (s^{(k)})^\top) + \rho_k s^{(k)} (s^{(k)})^\top \\
&= (H_k - \rho_k s^{(k)} (y^{(k)})^\top H_k)(I - \rho_k y^{(k)} (s^{(k)})^\top) + \rho_k s^{(k)} (s^{(k)})^\top \\
&= H_k - \rho_k \left( H_k y^{(k)} \right) (s^{(k)})^\top - \rho_k s^{(k)} \left( (y^{(k)})^\top H_k \right) + \rho_k^2 y^{(k)} (s^{(k)})^\top \left( \left( H_k y^{(k)} \right) (s^{(k)})^\top \right) + \rho_k s^{(k)} (s^{(k)})^\top
\end{aligned}$$

by setting parentheses as above, only matrix vector multiplications are evaluated.