Please identify anyone, whether or not they are in the class, with whom you discussed your homework. This problem is worth 1 point, but on a multiplicative scale (multiplied by the rest). Collaboration is encouraged.

1. Consider the following family of Quasi-Newton methods in Algorithm 1

Algorithm 1 The family of Quasi-Newton methods.

Input: Choose $\boldsymbol{x}^{(0)} \in \mathbb{R}^n$
Input: A symmetric positive definite matrix $\boldsymbol{H}_0 \in \mathbb{R}^{n \times n}$.
Output: An approximation to $\boldsymbol{x}^* := \arg\min_{\boldsymbol{x} \in \mathbb{R}^n} f(\boldsymbol{x})$.

for $k = 0, 1, 2, \ldots$ until converging **do**
  Compute $\boldsymbol{g}^{(k)} = \nabla f(\boldsymbol{x}^{(k)})$.
  Compute $\boldsymbol{d}^{(k)} = -\boldsymbol{H}_k \boldsymbol{g}^{(k)}$.
  Find $\alpha_k = \arg\min_{\alpha \geq 0} f\left(\boldsymbol{x}^{(k)} + \alpha \boldsymbol{d}^{(k)}\right)$.
  Update $\boldsymbol{x}^{(k+1)} = \boldsymbol{x}^{(k)} + \alpha_k \boldsymbol{d}^{(k)}$.
  Update $\boldsymbol{H}_{k+1}$.
end for

Suppose that the function we wish to minimize is a standard quadratic function:

$$f(\boldsymbol{x}) := \frac{1}{2}\boldsymbol{x}^\top \boldsymbol{Q} \boldsymbol{x} - \boldsymbol{b}^\top \boldsymbol{x} + c, \quad \text{where } \boldsymbol{Q} \in \mathbb{R}^{n \times n}, \boldsymbol{Q} = \boldsymbol{Q}^\top, \boldsymbol{Q} > 0.$$

The essense of quasi-Newton methods is that following equations are satisfied

$$\boldsymbol{H}_{k+1} \Delta \boldsymbol{g}^{(i)} = \Delta \boldsymbol{x}^{(i)}, \quad \text{for } i = 0, 1, \ldots, k, \tag{$*$}$$

where $\Delta \boldsymbol{x}^{(i)} := \boldsymbol{x}^{(i+1)} - \boldsymbol{x}^{(i)}$, and $\Delta \boldsymbol{g}^{(i)} := \boldsymbol{g}^{(i+1)} - \boldsymbol{g}^{(i)}$.

(a) (10 points) Show that, if the update formula of $\{\boldsymbol{H}_k\}$ satisfies the following three conditions while maintaining the symmetry of $\boldsymbol{H}_k$, then $\{\boldsymbol{H}_k\}$ satisfies ($*$). (Hint: prove for k first then use induction for the rest)

  (i) $\boldsymbol{H}_{k+1} = \boldsymbol{H}_k + \boldsymbol{U}_k$ for some $\boldsymbol{U}_k \in \mathbb{R}^{n \times n}$.

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

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

(b) (5 points) Verify that Rank-one (SR1), Davidon-Fletcher-Powell (DFP), and Broyden-Fletcher-Goldfarb-Shanno (BFGS) update rules all satisfy (i) and (ii).

(c) (5 points) Find $\boldsymbol{a}^{(k)}$ and $\boldsymbol{b}^{(k)}$ for SR1, DFP, and BFGS, respectively to satisfy (iii).