Homework 7 by 20307100005.

Problem 1

$S_k = x_{k+1} - x_k \quad y_k = \nabla f(x_{k+1}) - \nabla f(x_k)$

$\because f$ is strong convex $\therefore$

$\therefore (\nabla f(x_{k+1}) - \nabla f(x_k))^T (x_{k+1} - x_k) \geq m \| x_{k+1} - x_k \|_2^2 \geq 0$

That is $S_k^T y_k \geq 0$

This is because of the fact that in Hw5, we showed the equivalence of these two conditions:

① $g(x) = f(x) - \frac{\mu}{2}\|x\|^2$ is convex for all x

② $(\nabla f(x) - \nabla f(y))^T (x-y) \geq \mu \|x-y\|^2, \ \forall x, y$

$\because f$ is strongly convex $\therefore \nabla^2 f(x) \succeq mI, \ m > 0.$

$\therefore$ for $g(x) = f(x) - \frac{m}{2}\|x\|^2$, $\nabla^2 g(x) = \nabla^2 f(x) - mI \succeq 0$

$\Rightarrow g(x)$ is convex

∎

Problem 2

Update equation H_k:

$$H_{k+1} = (I - \rho_k s_k y_k^T) H_k \underbrace{(I - \rho_k y_k s_k^T)}^{V_k} + \rho_k s_k s_k^T$$

$$B_{k+1} = H_{k+1}^{-1} = \left(V_k^T H_k V_k + \rho_k s_k s_k^T\right)^{-1}$$

According to the identity:

$$(A+UCV)^{-1} = A^{-1} - A^{-1}U(C^{-1} + VA^{-1}U)^{-1} VA^{-1}$$

$$B_{k+1} = (V_k^T H_k V_k)^{-1} - (V_k^T H_k V_k)^{-1} s_k \left[\rho_k^{-1} + s_k^T (V_k^T H_k V_k)^{-1} s_k\right]^{-1} s_k^T (V_k^T H_k V_k)^{-1}$$

- Now, want to derive $(V_k^T H_k V_k)^{-1}$

$V_k^{-1} = (I - \rho_k y_k s_k^T)^{-1}$ According to Sherman-Morrison Formula:

$$(A + uv^T)^{-1} = A^{-1} - \frac{A^{-1}uv^TA^{-1}}{1 + v^TA^{-1}u} \Rightarrow V_k^{-1} = I + \frac{\rho_k y_k s_k^T}{1 - \rho_k s_k^T y_k}$$

And $\rho_k = \frac{1}{y_k^T s_k}$ $\therefore$ $V_k^{-1} = I + \frac{y_k s_k^T}{y_k^T s_k}$

$$\therefore (V_k^T H_k V_k)^{-1} = V_k^{-1} H_k^{-1} (V_k^{-1})^T = \left(I + \frac{y_k s_k^T}{y_k^T s_k}\right) H_k^{-1} \left(I + \frac{s_k y_k^T}{y_k^T s_k}\right) = H_k^{-1} + \rho_k H_k^{-1} s_k y_k^T H_k^{-1}$$

$$\Rightarrow B_{k+1} = H_k^{-1} + \frac{y_k y_k^T}{y_k^T s_k} - \frac{H_k^{-1} s_k s_k^T H_k^{-1}}{s_k^T H_k^{-1} s_k} = B_k + \frac{y_k y_k^T}{y_k^T s_k} - \frac{B_k s_k s_k^T B_k}{s_k^T B_k s_k}$$

Problem 3

- For rank-1 update, we have:

$$\det(A+uv^T) = \det(A)(1+v^TA^{-1}u)$$

Now, $B_{k+1} = B_k + \dfrac{y_k y_k^T}{y_k^T s_k} - \dfrac{B_k s_k s_k^T B_k}{s_k^T B_k s_k}$

first: Let $u = y_k / \sqrt{y_k^T s_k}$ $\quad v^T = y_k^T / \sqrt{y_k^T s_k}$.

$$\det\left(B_k + \frac{y_k y_k^T}{y_k^T s_k}\right) = \det(B_k)\left(1 + \frac{y_k^T B_k^{-1} y_k}{y_k^T s_k}\right) = \det(B_k)\left(1 + \frac{y_k^T s_k}{y_k^T s_k}\right) = 2\det(B_k)$$

Second: Let $u = B_k s_k / \sqrt{s_k^T B_k s_k}$ $\quad v^T = s_k^T B_k / \sqrt{s_k^T B_k s_k}$

Let $\left(B_k + \dfrac{y_k y_k^T}{y_k^T s_k}\right) = A$ $\qquad A^{-1} = B_k^{-1} - \dfrac{B_k^{-1} y_k y_k^T B_k^{-1}}{1 + y_k^T B_k^{-1} y_k} \cdot \dfrac{1}{y_k^T s_k}$

$\because B_k s_k = y_k$ $\qquad \therefore A^{-1} = B_k^{-1} - \dfrac{B_k^{-1}}{2} = \dfrac{1}{2} B_k^{-1}$

$$\therefore \det(B_{k+1}) = 2\det(B_k)\left(1 - \frac{s_k^T B_k \frac{1}{2} B_k^{-1} B_k s_k}{s_k^T B_k s_k}\right) = 2\det(B_k)\left(1 - \frac{s_k^T B_k s_k}{2 s_k^T B_k s_k}\right)$$

$$= \det(B_k)$$

And $\dfrac{y_k^T s_k}{s_k^T B_k s_k} = \dfrac{s_k^T B_k^T s_k}{s_k^T B_k s_k} = 1$ $\qquad \therefore \det(B_{k+1}) = \det(B_k) \dfrac{y_k^T s_k}{s_k^T B_k s_k}$