Diego Escobedo	May 2, 2020
6.046 Problem Set 8	
Collaborators: *Mikael Nida*	

Problem 1

(A) We want to find:

$$\frac{\partial f_{A,b}}{\partial x_1} = \frac{\partial}{\partial x_1}\frac{1}{2}x^T(A^TA)x - \frac{\partial}{\partial x_1}b^TAx \quad \text{Apply product rule}$$
$$\frac{\partial f_{A,b}}{\partial x_1} = \frac{1}{2}(\frac{\partial}{\partial x_1}x^T)(A^TA)x + \frac{1}{2}x^T(A^TA)(\frac{\partial}{\partial x_1}x) - b^TA(\frac{\partial}{\partial x_1}x) \quad \text{Eliminate the empty terms}$$
$$\frac{\partial f_{A,b}}{\partial x_1} = \frac{1}{2}\begin{pmatrix}1 & 0\end{pmatrix}\begin{pmatrix}2 & 0\\ 0 & 8\end{pmatrix}\begin{pmatrix}x_1\\ x_2\end{pmatrix} + \frac{1}{2}\begin{pmatrix}x_1 & x_2\end{pmatrix}\begin{pmatrix}2 & 0\\ 0 & 8\end{pmatrix}\begin{pmatrix}1\\ 0\end{pmatrix} - \begin{pmatrix}1 & 3\end{pmatrix}\begin{pmatrix}1 & 2\\ -1 & 2\end{pmatrix}\begin{pmatrix}1\\ 0\end{pmatrix}$$
$$\frac{\partial f_{A,b}}{\partial x_1} = x_1 + x_1 + 2$$
$$\frac{\partial f_{A,b}}{\partial x_1} = 2x_1 + 2$$

And then, for the second one:

$$\frac{\partial f_{A,b}}{\partial x_2} = \frac{\partial}{\partial x_2}\frac{1}{2}x^T(A^TA)x - \frac{\partial}{\partial x_2}b^TAx \quad \text{Apply product rule}$$
$$\frac{\partial f_{A,b}}{\partial x_2} = \frac{1}{2}(\frac{\partial}{\partial x_2}x^T)(A^TA)x + \frac{1}{2}x^T(A^TA)(\frac{\partial}{\partial x_2}x) - b^TA(\frac{\partial}{\partial x_2}x) \quad \text{Eliminate the empty terms}$$
$$\frac{\partial f_{A,b}}{\partial x_2} = \frac{1}{2}\begin{pmatrix}0 & 1\end{pmatrix}\begin{pmatrix}2 & 0\\ 0 & 8\end{pmatrix}\begin{pmatrix}x_1\\ x_2\end{pmatrix} + \frac{1}{2}\begin{pmatrix}x_1 & x_2\end{pmatrix}\begin{pmatrix}2 & 0\\ 0 & 8\end{pmatrix}\begin{pmatrix}0\\ 1\end{pmatrix} - \begin{pmatrix}1 & 3\end{pmatrix}\begin{pmatrix}1 & 2\\ -1 & 2\end{pmatrix}\begin{pmatrix}0\\ 1\end{pmatrix}$$
$$\frac{\partial f_{A,b}}{\partial x_2} = 4x_2 + 4x_2 - 8$$
$$\frac{\partial f_{A,b}}{\partial x_2} = 8x_2 - 8$$

(B) To be a critical point, we have to show that both the derivatives are equal to 0. If we set $2x_1 + 2 = 0$, we see that $x_1 = -1$ is the only solution, and if we set $8x_2 - 8 = 0$, we see that $x_2 = 1$ is the only solution. So, we see that any critical point must have $x_2 = 1$ and $x_1 = -1$.

(C) The first step here is to find the Hessian. The Hessian here can be given by $\begin{bmatrix}\frac{\partial f_{A,b}}{\partial x_1x_1} & \frac{\partial f_{A,b}}{\partial x_1x_2}\\ \frac{\partial f_{A,b}}{\partial x_2x_1} & \frac{\partial f_{A,b}}{\partial x_2x_2}\end{bmatrix}$

which is numerically$\left(\begin{smallmatrix} 2 & 0 \\ 0 & 8 \end{smallmatrix}\right)$. Thus we need to find the largest α such that $x^T \left(\begin{smallmatrix} 2 & 0 \\ 0 & 8 \end{smallmatrix}\right) x \geq \alpha$ is true for all x. We are told that x is a unit norm matrix. It is fairly easy to see that since the hessian is diagonal with one value larger than the other, we can simply take the unit norm as a vector with just a single one to try to minimize this quantity. Using $x = \left(\begin{smallmatrix} 1 \\ 0 \end{smallmatrix}\right)$, we get that $f_{A,b}$ is α strongly convex where $\alpha = 2$. By a similar argument, we can see that it is β smooth where $\beta = 8$.

(D) We know teh formula or convergence given α strong convexity and β smoothness is $O(\frac{\beta}{\alpha} log \frac{f(x_0) - f(x^*)}{\epsilon})$. In this case, since α, β are constants, and we know the optimal solution from linear algebra $x^* = \left(\begin{smallmatrix} -1 \\ 1 \end{smallmatrix}\right)$, and thus $f_{A,b}(x^*) = -5$, we get that this gradient descent will converge in $O(log \frac{f(x_0)+5}{\epsilon})$.

Problem 2

Sorry I been swamped recently and struggling w some issues :(