Name: Hongda Li
Class: AMATH 515 Winter 2021
HW0: Theoretical Portion

Calculus Primer

For each of the given function find out the Gradient and the Hessian of the given matrix.

(a) $f(x) = \sin(x_1 + x_2 + x_3 + x_4)$ By definition of the Gradient we have:

$$\nabla f(x) = \begin{bmatrix} \partial_{x1} f(x) \\ \partial_{x2} f(x) \\ \partial_{x3} f(x) \\ \partial_{x4} f(x) \end{bmatrix} = \begin{bmatrix} \cos(x_1 + x_2 + x_3 + x_4) \\ \cos(x_1 + x_2 + x_3 + x_4) \\ \cos(x_1 + x_2 + x_3 + x_4) \\ \cos(x_1 + x_2 + x_3 + x_4) \end{bmatrix}$$

And the Hessian of the function is just taking derivative on the columns for each of the variables, giving us:

$$\nabla^2 f(x) = - \begin{bmatrix} \sin(x) & \sin(x) & \sin(x) & \sin(x) \\ \sin(x) & \sin(x) & \sin(x) & \sin(x) \\ \sin(x) & \sin(x) & \sin(x) & \sin(x) \\ \sin(x) & \sin(x) & \sin(x) & \sin(x) \end{bmatrix}$$

(b) $f(x) = \|x\|_2^2$ Then the gradient is:

$$f(x) = \begin{bmatrix} \partial_{x1} \|x\| \\ \partial_{x2} \|x\| \\ \partial_{x3} \|x\| \\ \partial_{x4} \|x\| \end{bmatrix} = 2 \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix}$$

And the Hessian is produced by taking the column wrt to each of the variables and then stack them up horizontally, giving us:

$$\nabla^2 f(x) = 2 \begin{bmatrix} \partial_{x1} x_1 & \partial_{x2} x_1 & \partial_{x3} x_1 & \partial_{x4} x_1 \\ \partial_{x1} x_2 & \partial_{x2} x_2 & \partial_{x3} x_2 & \partial_{x4} x_2 \\ \partial_{x1} x_3 & \partial_{x2} x_3 & \partial_{x3} x_3 & \partial_{x4} x_3 \\ \partial_{x1} x_4 & \partial_{x2} x_4 & \partial_{x3} x_4 & \partial_{x4} x_4 \end{bmatrix}$$

And interestingly all the non diagonal elements goes to zero and then we have:

$$\nabla^2 f(x) = 2I_4$$

The 4 by 4 identity matrix multiplied by 2 for the Hessian of the function.

(c) $f(x) = \ln(x_1 x_2 x_3 x_4) = \ln(x_1) + \ln(x_2) + \ln(x_3) + \ln(x_4)$ Just like that first example, there is no interaction for each of the term in the function, and hence the gradient is:

$$\nabla f(x) = \begin{bmatrix} \partial_{x_1} f(x) \\ \partial_{x_2} f(x) \\ \partial_{x_3} f(x) \\ \partial_{x_4} f(x) \end{bmatrix} = \begin{bmatrix} \frac{1}{x_1} \\ \frac{1}{x_2} \\ \frac{1}{x_3} \\ \frac{1}{x_4} \end{bmatrix}$$

Similar to the second example, because of lack of interactions between each of the terms, There will only be diagonal left after we take the gradient wrt to each of the variables, giving us:

$$\nabla^2 f(x) = - \begin{bmatrix} \frac{1}{x_1^2} & & & \\ & \frac{1}{x_2^2} & & \\ & & \frac{1}{x_3^2} & \\ & & & \frac{1}{x_4^2} \end{bmatrix}$$

And the empty entries are zeros.

Liner Algebra Primer

(a) For a Diagonal Matrix, the Eigenvalues are its diagonal, and hence, the matrix:

$$\begin{bmatrix} 1 & & & \\ \pi & 2 & & \\ 64 & -15 & 3 & \\ 321 & 0 & 0 & 5 \end{bmatrix}$$

Has Eigenvalues: $1, 2, 3, 5$.

(b) For any matrix that is in the form: vu^T, it will be rank 1 and all of it's columns are spanned by the vector v, therefore the matrix:

$$A = \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix} \begin{bmatrix} 1 & 1 & 1 \end{bmatrix}$$

And hence that nullspace of the matrix is:

$$\text{span}\left(\begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} -1 \\ 0 \\ 1 \end{bmatrix} \right)$$

(c) Let $A \in \mathbb{R}^{10\times 5}$ and let the vector $b \in \mathbb{R}^{10}$ Then:

- $A^T A \in \mathbb{R}^{5\times 5}$ and $A^T b \in \mathbb{R}^5$
- The system $Ax = b$ is over-determined, and the solution to the system resides in $\mathbb{R}^5$.
- There will always be solution to the system $A^T Ax = A^T b$ because the left and side and the right hand side of the equations resides in the span of the rows of A. If A is not full-rank, then $A^T A$ has null space, and therefore there will be infinitely many solutions for it, else it has one unique solution.
- If A has linear independent columns, then it's full-rank, which means that $Ax = b$ has an unique solution if $b \in \text{col}(A)$, else there is no solution to it assuming. And $A^T A = A^T b$ has a unique solution if A is full-rank.