

Home Work 1

Question (1) Math Review (25 points)

Question (1_1): $f(x, y, z) = 3x^2 + \sin(y)z$, I need to find partial derivatives for each x,y, and z:

Solution:

With respect to x: $\frac{\partial f(x,y,z)}{\partial x} = 6x$, everything else is constant ($\sin(y)z$) as they are not function of x.

With respect to y: $\frac{\partial f(x,y,z)}{\partial y} = z \cos(y)$, the first term ($3x^2$) derivative is zero in this case.

With respect to z: $\frac{\partial f(x,y,z)}{\partial z} = \sin(y)$, the first term ($3x^2$) derivative is also zero in this case and ($\sin(y)$) is treated as constant when deriving for z.

Question (1_2): In this question, I will need to find the $\nabla f(x, y, z)$:

Using the results from the previous question,

$$\nabla f(x, y, z) = \begin{bmatrix} 6x \\ z \cos(y) \\ \sin(y) \end{bmatrix}$$

Question (1_3): Here we are replicating 1 and 2 but $f(x) = 3x_1^2 + \sin(x_2)x_3$:

$$\nabla f(x_1, x_2, x_3) = \begin{bmatrix} 6x_1 \\ x_3 \cos(x_2) \\ \sin(x_2) \end{bmatrix}$$

Question (1_4_A):In this part, I will need to get the derivative for $\|x\|_2^2$:

$$\|x\|_2^2 = \sum_{i=1}^n x_i^2 = x_1^2 + x_2^2 + \dots + x_n^2$$

$$\text{In this case: } \frac{\partial \|x\|_2^2}{\partial x} = \frac{\partial(x_1^2+x_2^2+\dots+x_n^2)}{\partial x_1} + \frac{\partial(x_1^2+x_2^2+\dots+x_n^2)}{\partial x_2} + \dots + \frac{\partial(x_1^2+x_2^2+\dots+x_n^2)}{\partial x_n}$$

$$\frac{\partial \|x\|_2^2}{\partial x} = \begin{bmatrix} 2x_1 \\ 2x_2 \\ \dots \\ 2x_n \end{bmatrix}. \text{ This represents the partial derivative for each } x.$$

Question (1_4_B):In this part, I will need to get the derivative for $\|x\|_2$:

$$\|x\|_2 = \sqrt{\sum_{i=1}^n x_i^2} = \sqrt{(x_1^2)} + \sqrt{(x_2^2)} + \dots + \sqrt{(x_n^2)}$$

To get this derivative, $\frac{\partial \|x\|_2}{\partial x}$, I will be using the chain rule assuming that a new function $h(z)$ $= \sqrt{z}$, where $z(x) = \sum_{i=1}^n x_i^2$.

$$\frac{\partial h(z)}{\partial x} = \frac{\partial h(z)}{\partial z} * \frac{\partial z(x)}{\partial x}$$

$$\frac{\partial h(z)}{\partial x} = \frac{1}{2\sqrt{\sum_{i=1}^n x_i^2}} * \sum_{i=1}^n 2x_i$$

$$\frac{\partial h(z)}{\partial x} = \frac{x}{\|x\|_2}$$

Question (1_4_C):In this part, I will need to get the derivative for $\|x\|_1$:

$$\|x\|_1 = \sum_{i=1}^n |x_i| = |x_1| + |x_2| + \dots + |x_n|$$

$\frac{\partial \|x\|_1}{\partial x}$: depends on the value of x_i . If x_i positive, it would be 1, if negative, it would be -1, and undefined when x_i equals zero. To represent the derivative of this function, the sign function can be used in this case:

$\text{sgn}(x)$

$$\frac{\partial \|x\|_1}{\partial x} = \begin{bmatrix} \text{sgn}(x_1) \\ \text{sgn}(x_2) \\ \dots \\ \text{sgn}(x_n) \end{bmatrix}$$

Question (1_4_D): In this part, I will need to get the derivative for $\|x\|_\infty$:

$\|x\|_\infty = \max |X_i|$, This entails many cases under the hood and can be represented using the sign function:

$\frac{\partial \|x\|_\infty}{\partial x_i} = \text{sgn}(x_i)$ Accordingly, $\frac{\partial \|x\|_\infty}{\partial x_i} = [0, \dots, \text{sgn}(x_i), \dots, 0]^T$. Those components that are not achieving the maximum, the derivative is 0.

Question (1_5): In this part, I will need to get the derivative of $f(x) = e^{\frac{-1}{2}\|x\|_2^2}$.

For this function, the chain rule will be used assuming a new function $h(z) = e^z$, where $z = \frac{-1}{2}\|x\|_2^2$. In this case, we would be looking for:

$$\frac{\partial h(z)}{\partial x} = \frac{\partial h(z)}{\partial z} * \frac{\partial z(x)}{x}$$

$$\frac{\partial h(z)}{\partial x} = e^z * -x_i$$

$$\frac{\partial h(z)}{\partial x} = -x_i * e^{\frac{-1}{2}\|x\|_2^2} = [-x_1 e^{\frac{-1}{2}\|x_1\|_2^2}, -x_2 e^{\frac{-1}{2}\|x_2\|_2^2}, \dots, -x_n e^{\frac{-1}{2}\|x_n\|_2^2}]$$

$$\begin{bmatrix} -x_1 e^{\frac{-1}{2}\|x_1\|_2^2} \\ -x_2 e^{\frac{-1}{2}\|x_2\|_2^2} \\ \dots \\ -x_n e^{\frac{-1}{2}\|x_n\|_2^2} \end{bmatrix}$$

Question (1_6): In this part, I will need to get the two components of $f(A, x)$:

$$\text{In this case } A \text{ is a } 2 \times 3 \text{ matrix} = A_{2 \times 3} = \begin{bmatrix} A1,1 & A1,2 & A1,3 \\ A2,1 & A2,2 & A2,3 \end{bmatrix}$$

$$\text{And } x_{1 \times 3} = [x1 \quad x2 \quad x3]$$

$$f(A, x) = Ax^T = \begin{bmatrix} A1,1 & A1,2 & A1,3 \\ A2,1 & A2,2 & A2,3 \end{bmatrix} * \begin{bmatrix} x1 \\ x2 \\ x3 \end{bmatrix}$$

$$f(A, x) = Ax^T = \begin{bmatrix} A1,1 * x1 + A1,2 * x2 + A1,3 * x3 \\ A2,1 * x1 + A2,2 * x2 + A2,3 * x3 \end{bmatrix}$$

Question (1_7): In this part,I would need to get $\frac{\partial f(A,x)_1}{\partial x}$ and $\frac{\partial f(A,x)_2}{\partial x}$.

$$\frac{\partial f(A,x)_1}{\partial x} = \begin{bmatrix} A1, 1 \\ A1, 2 \\ A1, 3 \end{bmatrix}$$

$$\frac{\partial f(A,x)_2}{\partial x} = \begin{bmatrix} A2, 1 \\ A2, 2 \\ A2, 3 \end{bmatrix}$$

Question (1_8): In this part,I would need to get

$\frac{\partial f(A,x)}{\partial x}$: Using results from the previous question, The result will be:

$$\frac{\partial f(A,x)}{\partial x} = \begin{bmatrix} A1, 1 & A2, 1 \\ A1, 2 & A2, 2 \\ A1, 3 & A2, 3 \end{bmatrix}$$

Question (1_9): In this part,I would need to get the derivative of

$$\mathbb{E}[f(x)].$$

Given that: $\mathbb{E}[f(x)] = \sum \Pr(X = x)f(x)$, then:

$$\frac{\partial \mathbb{E}[f(x)]}{\partial x} = \frac{\partial}{\partial x} \sum \Pr(X = x)f(x), \text{ note that } \Pr(X = x) \text{ is constant with respect to } x.$$

$$\frac{\partial \mathbb{E}[f(x)]}{\partial x} = \sum \Pr(X = x) \frac{\partial}{\partial x} f(x).$$

Question (2): Linear Algebra Output

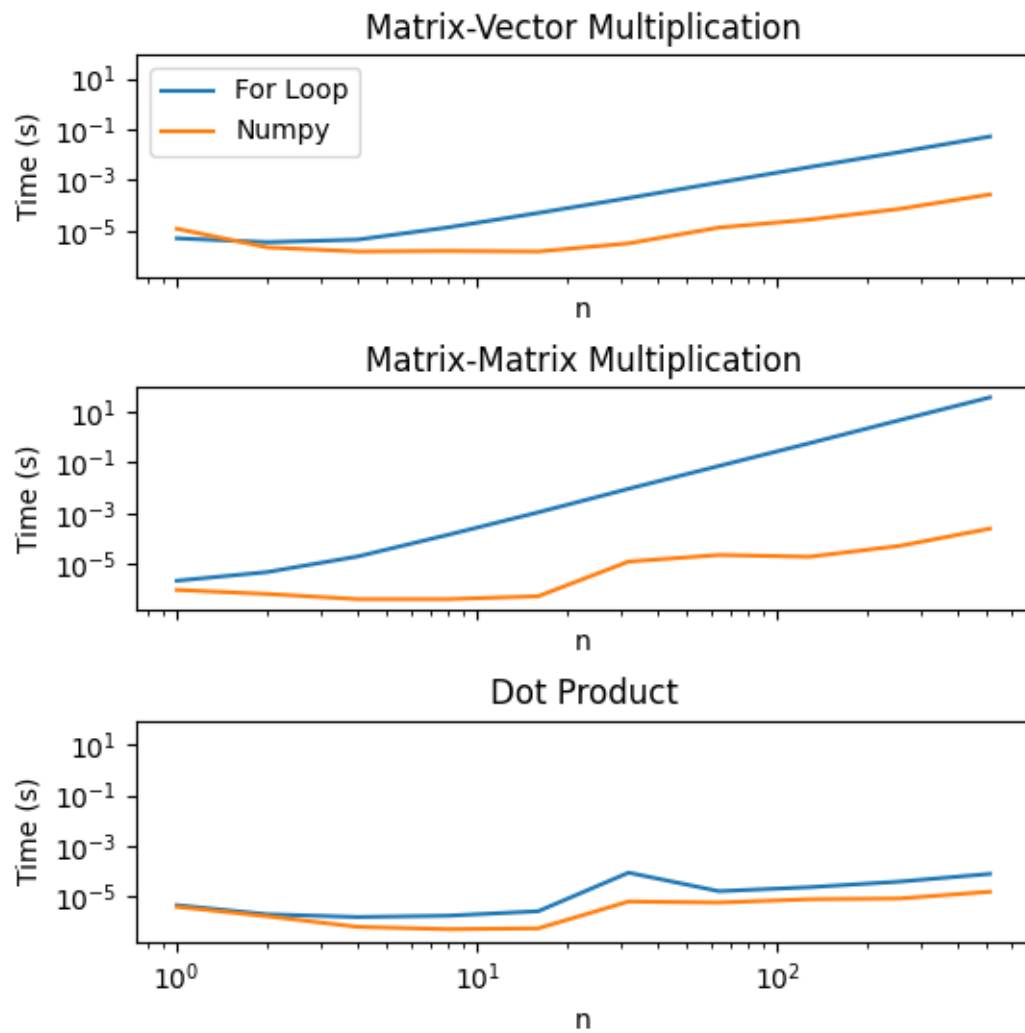


Figure 1: Performance comparison