ROHIT RAGHUWANSHI(12341820)

QUESTION NO 5.5

Functions

1. $f_1(x) = \sin(x_1)\cos(x_2), \quad x \in \mathbb{R}^2$
2. $f_2(x, y) = x^\top y, \quad x, y \in \mathbb{R}^n$
3. $f_3(x) = xx^\top, \quad x \in \mathbb{R}^n$

Part a: Dimensions of

$\frac{\partial f_i}{\partial x}$ 1. For $f_1(x)$: - $x \in \mathbb{R}^2$ implies x_1 and x_2 are scalars. - Dimension of $\frac{\partial f_1}{\partial x}$: $\mathbb{R}^2$ (1 output, 2 inputs).

2. For $f_2(x, y)$: - $x, y \in \mathbb{R}^n$ implies $x^\top y$ is a scalar. - Dimension of $\frac{\partial f_2}{\partial x}$: $\mathbb{R}^n$ (1 output, n inputs).

3. For $f_3(x)$: - $xx^\top$ results in an $n \times n$ matrix. - Dimension of $\frac{\partial f_3}{\partial x}$: $\mathbb{R}^{n\times n}$ (n outputs, n inputs).

Part b: Compute the Jacobians

1. For $f_1(x)$:

$$J_{f_1} = \begin{bmatrix} \frac{\partial f_1}{\partial x_1} & \frac{\partial f_1}{\partial x_2} \end{bmatrix} = \begin{bmatrix} \cos(x_1)\cos(x_2) & -\sin(x_1)\sin(x_2) \end{bmatrix}$$

Dimension: 1×2.

2. For $f_2(x, y)$:

$$J_{f_2} = \begin{bmatrix} \frac{\partial f_2}{\partial x_1} & \dots & \frac{\partial f_2}{\partial x_n} \\ \frac{\partial f_2}{\partial y_1} & \dots & \frac{\partial f_2}{\partial y_n} \end{bmatrix} = \begin{bmatrix} y^\top & x^\top \\ x^\top & 0 \end{bmatrix}$$

Dimension: $1 \times n$.

3. For $f_3(x)$:

$$J_{f_3} = \frac{\partial (xx^\top)}{\partial x} = 2x\mathrm{d}x^\top \quad \text{(using the product rule)}$$

Dimension: $n \times n$.

QUESTION NO 5.8

Part a

Compute the derivative $\frac{df}{dx}$ using the chain rule.

The function is given as:

$$f(z) = \exp\left(-\frac{1}{2}z\right), \quad z = g(y) = S^{-1}y, \quad y = h(x) = x - \mu$$

Step 1: Understanding the Components

- $x, \mu \in \mathbb{R}^D$
- $S \in \mathbb{R}^{D \times D}$
- $y = h(x)$: dimension D
- $z = g(y)$: dimension D
- $f(z)$: dimension $\mathbb{R}$

Step 2: Applying the Chain Rule

$$\frac{df}{dx} = \frac{df}{dz} \cdot \frac{dz}{dy} \cdot \frac{dy}{dx}$$

Step 3: Compute Each Partial Derivative

1. $\frac{df}{dz} = -\frac{1}{2}\exp\left(-\frac{1}{2}z\right)$ Dimension: $\mathbb{R}$
2. $\frac{dz}{dy} = S^{-1}$ Dimension: $\mathbb{R}^{D \times D}$
3. $\frac{dy}{dx} = I$ Dimension: $\mathbb{R}^{D \times D}$

Step 4: Final Derivative

$$\frac{df}{dx} = \frac{1}{2}f(z)S^{-1}$$

Part b

Compute the derivative $\frac{df}{dx}$ for

$$f(x) = \mathrm{tr}(xx^\top + \sigma^2 I), \quad x \in \mathbb{R}^D$$

Step 1: Understanding the Components

- $xx^\top$: matrix dimension $D \times D$
- I: identity matrix, dimension $D \times D$
- $f(x)$ has dimension $\mathbb{R}$.

Step 2: Applying the Derivative

$$\frac{df}{dx} = \frac{d}{dx}\mathrm{tr}(xx^\top) = 2x^\top$$

Dimension: $\mathbb{R}^D$

Part c

Compute the derivative $\frac{df}{dx}$ using the chain rule for

$$f = \tanh(z), \quad z = Ax + b, \quad x \in \mathbb{R}^N, \quad A \in \mathbb{R}^{M\times N}, \quad b \in \mathbb{R}^M$$

Step 1: Understanding the Components

- x: dimension N
- z: dimension M
- f: dimension M

Step 2: Applying the Chain Rule

$$\frac{df}{dx} = \frac{df}{dz} \cdot \frac{dz}{dx}$$

Step 3: Compute Each Partial Derivative

1. $\frac{df}{dz} = \mathrm{diag}(1 - \tanh^2(z))$ Dimension: $M \times M$
2. $\frac{dz}{dx} = A$ Dimension: $M \times N$

Step 4: Final Derivative

$$\frac{df}{dx} = \mathrm{diag}(1 - \tanh^2(z))A$$

Dimension: $M \times N$