Assignment 1 Theoretical ML

Hardik Tiwari

May 16, 2024

Problem 1

(a) Proof

Given:

- A is an $m \times n$ matrix.
- $\mathbf{x}$ is an $n \times 1$ vector.

We need to prove:

$$\frac{d}{d\mathbf{x}}[A\mathbf{x}] = A$$

Solution:

The expression $A\mathbf{x}$ represents a linear transformation of the vector $\mathbf{x}$ by the matrix A. Each component of $A\mathbf{x}$ is a linear function of the components of $\mathbf{x}$.

To find the derivative, we consider the i-th component of $A\mathbf{x}$:

$$(A\mathbf{x})_i = \sum_{j=1}^{n} A_{ij} x_j$$

where A_{ij} is the element in the i-th row and j-th column of A, and x_j is the j-th component of $\mathbf{x}$.

Taking the partial derivative with respect to x_k, we get:

$$\frac{\partial (A\mathbf{x})_i}{\partial x_k} = A_{ik}$$

Thus, the derivative of $A\mathbf{x}$ with respect to $\mathbf{x}$ is:

$$\frac{d}{d\mathbf{x}}[A\mathbf{x}] = A$$

This completes the proof for part (a).

(b) Proof

Given:

- A is an $n \times n$ matrix.
- $\mathbf{x}$ is an $n \times 1$ vector.

We need to prove:

$$\frac{d}{d\mathbf{x}}[\mathbf{x}^T A\mathbf{x}] = \mathbf{x}^T(A + A^T)$$

Solution:

The expression $\mathbf{x}^T A\mathbf{x}$ is a scalar quadratic form in $\mathbf{x}$. To find the derivative, we use the properties of matrix calculus.

First, write the quadratic form explicitly:

$$\mathbf{x}^T A\mathbf{x} = \sum_{i=1}^{n}\sum_{j=1}^{n} x_i A_{ij} x_j$$

Taking the partial derivative with respect to x_k, we get:

$$\frac{\partial}{\partial x_k}(\mathbf{x}^T A\mathbf{x}) = \sum_{i=1}^{n} A_{ik} x_i + \sum_{j=1}^{n} A_{kj} x_j$$

Notice that:

$$\sum_{i=1}^{n} A_{ik} x_i = (A\mathbf{x})_k$$

and

$$\sum_{j=1}^{n} A_{kj} x_j = (A^T\mathbf{x})_k$$

Thus:

$$\frac{\partial}{\partial x_k}(\mathbf{x}^T A\mathbf{x}) = (A\mathbf{x})_k + (A^T\mathbf{x})_k$$

In vector notation, this can be written as:

$$\frac{d}{d\mathbf{x}}[\mathbf{x}^T A\mathbf{x}] = A\mathbf{x} + A^T\mathbf{x}$$

Transposing both sides, we get:

$$\frac{d}{d\mathbf{x}}[\mathbf{x}^T A\mathbf{x}] = \mathbf{x}^T(A + A^T)$$

This completes the proof for part (b).

Problem 2

Ans: The dimension of final result is m x n x k

Problem 3

(a)

We need to differentiate the following vector function with respect to the vector $\begin{pmatrix} x \\ y \end{pmatrix}$:

$$\mathbf{f}(x, y) = \begin{pmatrix} 2\sin^2(x)\cos(y) \\ x^2 + 3e^y \end{pmatrix}$$

The Jacobian matrix $\mathbf{J}$ of $\mathbf{f}$ with respect to $\begin{pmatrix} x \\ y \end{pmatrix}$ is given by:

$$\mathbf{J} = \begin{pmatrix} \cdot\frac{\partial f_1}{\partial x} & \frac{\partial f_1}{\partial y} \\ \frac{\partial f_2}{\partial x} & \frac{\partial f_2}{\partial y} \end{pmatrix}$$

Calculate the partial derivatives:

$$\frac{\partial f_1}{\partial x} = \frac{\partial}{\partial x}[2\sin^2(x)\cos(y)] = 4\sin(x)\cos(x)\cos(y)$$

$$\frac{\partial f_1}{\partial y} = \frac{\partial}{\partial y}[2\sin^2(x)\cos(y)] = -2\sin^2(x)\sin(y)$$

$$\frac{\partial f_2}{\partial x} = \frac{\partial}{\partial x}[x^2 + 3e^y] = 2x$$

$$\frac{\partial f_2}{\partial y} = \frac{\partial}{\partial y}[x^2 + 3e^y] = 3e^y$$

Therefore, the Jacobian matrix $\mathbf{J}$ is:

$$\mathbf{J} = \begin{pmatrix} 4\sin(x)\cos(x)\cos(y) & -2\sin^2(x)\sin(y) \\ 2x & 3e^y \end{pmatrix}$$

(b)

$$\textbf{Ans.J} = \begin{pmatrix} 6xy + yzw & 3x^2 + xzw & xyw & xyz \\ 2x\cos(x^2 + yw - z) & w\cos(x^2 + yw - z) & -\cos(x^2 + yw - z) & y\cos(x^2 + yw - z) \end{pmatrix}$$