DSL251
Data Analytics and Visualization
Homework 2

Amay Dixit
12340220

`Google Colab Notebook Link`
https://colab.research.google.com/drive/1SiAznpjyErzbzoyA8r2L4zOqpUfBzYaz?usp=sharing

Solution to Question 5.5

We are given the following functions:

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

(a) Dimensions of $\frac{\partial f_i}{\partial x}$:

For $f_1(x)$:
The function $f_1(x) = \sin(x_1)\cos(x_2)$ is a scalar function, where $x_1, x_2 \in \mathbb{R}$. We compute the partial derivatives:

$$\frac{\partial f_1}{\partial x_1} = \cos(x_1)\cos(x_2), \quad \frac{\partial f_1}{\partial x_2} = -\sin(x_1)\sin(x_2)$$

Thus, the Jacobian J_1 is:

$$J_1 = \begin{bmatrix} \frac{\partial f_1}{\partial x_1} & \frac{\partial f_1}{\partial x_2} \end{bmatrix} = \begin{bmatrix} \cos(x_1)\cos(x_2) & -\sin(x_1)\sin(x_2) \end{bmatrix} \in \mathbb{R}^{1\times 2}.$$

For $f_2(x, y)$:
The function $f_2(x, y) = x^\top y$ is a scalar function, where $x, y \in \mathbb{R}^n$. We compute the partial derivatives:

$$\frac{\partial f_2}{\partial x} = y^\top \quad \text{and} \quad \frac{\partial f_2}{\partial y} = x^\top$$

Thus, the Jacobian J_2 is:

$$J_2 = \begin{bmatrix} \frac{\partial f_2}{\partial x} & \frac{\partial f_2}{\partial y} \end{bmatrix} = \begin{bmatrix} y^\top & x^\top \end{bmatrix} \in \mathbb{R}^{1\times 2n}.$$

For $f_3(x)$:

The function $f_3(x) = xx^\top$ is a matrix, where $x \in \mathbb{R}^n$. The dimensions of the Jacobian depend on the components of $f_3(x)$. We compute the partial derivatives with respect to each x_i:

$$\frac{\partial f_3}{\partial x_1} = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} \begin{bmatrix} x_1 & 0 & \cdots & 0 \end{bmatrix}, \quad \frac{\partial f_3}{\partial x_2} = \begin{bmatrix} 0 \\ x_1 \\ \vdots \\ x_n \end{bmatrix} \begin{bmatrix} 0 & x_2 & \cdots & 0 \end{bmatrix}, \quad \text{and so on.}$$

Thus, the Jacobian matrix J_3 is:

$$J_3 = \begin{bmatrix} \frac{\partial f_3}{\partial x_1} & \frac{\partial f_3}{\partial x_2} & \cdots & \frac{\partial f_3}{\partial x_n} \end{bmatrix} \in \mathbb{R}^{n^2 \times n}.$$

(b) Compute the Jacobians:

For $f_1(x)$:
We already computed the Jacobian matrix J_1 in part (a) as:

$$J_1 = \begin{bmatrix} \cos(x_1)\cos(x_2) & -\sin(x_1)\sin(x_2) \end{bmatrix} \in \mathbb{R}^{1 \times 2}.$$

For $f_2(x, y)$:
We already computed the Jacobian J_2 in part (a) as:

$$J_2 = \begin{bmatrix} y^\top & x^\top \end{bmatrix} \in \mathbb{R}^{1 \times 2n}.$$

For $f_3(x)$:
The Jacobian matrix J_3 is the matrix of partial derivatives of each component $f_3(x)$ with respect to each element of x. Each component $f_3(x)$ is the outer product $xx^\top$, and the derivative with respect to each x_i is a matrix with all zeros except for a column where x_i is repeated. Therefore, the Jacobian is:

$$J_3 = \begin{bmatrix} \frac{\partial f_3}{\partial x_1} & \frac{\partial f_3}{\partial x_2} & \cdots & \frac{\partial f_3}{\partial x_n} \end{bmatrix} \in \mathbb{R}^{n^2 \times n}.$$

Solution to Question 5.8

Part (a):
We are given the function:

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

where $x, \mu \in \mathbb{R}^D$, $S \in \mathbb{R}^{D \times D}$.
Using the chain rule, the derivative of f with respect to x is:

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

First, we compute each partial derivative:
1. $\frac{df}{dz} = \frac{d}{dz} \exp\left(-\frac{1}{2}z\right) = -\frac{1}{2} \exp\left(-\frac{1}{2}z\right)$, with dimensions 1×1.

2. $z = g(y) = y^\top S^{-1} y$, so:

$$\frac{dz}{dy} = \frac{d}{dy}\left(y^\top S^{-1} y\right) = 2S^{-1}y, \quad \text{with dimensions} \quad D \times D.$$

3. $y = h(x) = x - \mu$, so:

$$\frac{dy}{dx} = \frac{d}{dx}(x - \mu) = I_D, \quad \text{with dimensions} \quad D \times D.$$

Now, combining these terms:

$$\frac{df}{dx} = -\frac{1}{2}\exp\left(-\frac{1}{2}z\right) \cdot 2S^{-1}y \cdot I_D = -\exp\left(-\frac{1}{2}z\right) S^{-1}y$$

Thus, the final derivative is:

$$\frac{df}{dx} = -\exp\left(-\frac{1}{2}z\right) S^{-1}y, \quad \text{with dimensions} \quad D \times 1.$$

Part (b):
We are given the function:

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

Here, $\operatorname{tr}(A)$ represents the trace of matrix A, which is the sum of the diagonal elements A_{ii}.

We first compute the derivative of the trace term:

$$\frac{d}{dx}\operatorname{tr}(xx^\top) = \frac{d}{dx}\left(\sum_{i=1}^{D}\sum_{j=1}^{D} x_i x_j\right) = 2x$$

Thus:

$$\frac{df}{dx} = 2x.$$

Part (c):
We are given the function:

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

The derivative of f with respect to x is:

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

First, we compute each partial derivative:
1. $\frac{df}{dz} = \frac{d}{dz}\tanh(z) = 1 - \tanh^2(z)$, with dimensions $M \times M$.
2. $\frac{dz}{dx} = \frac{d}{dx}(Ax + b) = A$, with dimensions $M \times N$.

Now, combining these terms:

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

Thus, the final derivative is:

$$\frac{df}{dx} = (1 - \tanh^2(z)) \cdot A, \quad \text{with dimensions} \quad M \times N.$$