ML Homework 3

Shreyas Shukla

Assignment 3

Question 3 (10 points):

Given the data:

$$\begin{bmatrix} x & y \\ 0 & 1 \\ 1 & 0 \\ 2 & 2 \\ 3 & -2 \end{bmatrix}$$

We assume the function to predict y from x is linear:

$$f(x) = ax + b$$

The goal is to determine the best values of a and b using the closed-form solution. The closed-form solution for linear regression is:

$$\hat{\beta} = (X^\top X)^{-1} X^\top y$$

where X is the matrix of input data, and y is the output vector.

For the given dataset, X and y are:

$$X = \begin{bmatrix} 1 & 0 \\ 1 & 1 \\ 1 & 2 \\ 1 & 3 \end{bmatrix}, \quad y = \begin{bmatrix} 1 \\ 0 \\ 2 \\ -2 \end{bmatrix}$$

First, calculate $X^\top X$:

$$X^\top X = \begin{bmatrix} 4 & 6 \\ 6 & 14 \end{bmatrix}$$

Then, calculate $X^\top y$:

$$X^\top y = \begin{bmatrix} 1 \\ -2 \end{bmatrix}$$

Now, compute the inverse of $X^\top X$:

$$(X^\top X)^{-1} = \frac{1}{(4)(14) - (6)(6)} \begin{bmatrix} 14 & -6 \\ -6 & 4 \end{bmatrix} = \frac{1}{8} \begin{bmatrix} 14 & -6 \\ -6 & 4 \end{bmatrix}$$

Multiplying this by $X^\top y$:

$$\hat{\beta} = \begin{bmatrix} 1.75 & -0.75 \\ -0.75 & 0.5 \end{bmatrix} \begin{bmatrix} 1 \\ -2 \end{bmatrix} = \begin{bmatrix} 3.25 \\ -1.75 \end{bmatrix}$$

Thus, the best values of a and b are:

$$\boxed{a = -1.75, \quad b = 3.25}$$

Question 7 (20 points):
We are given the probability density function:

$$p(x) = \begin{cases} x^2 - x + 1 & \text{for } 0 \le x \le b \\ 0 & \text{everywhere else} \end{cases}$$

The total probability must equal 1:

$$\int_0^b (x^2 - x + 1)\, dx = 1$$

First, compute the integral:

$$\int (x^2 - x + 1)\, dx = \frac{x^3}{3} - \frac{x^2}{2} + x$$

Now, evaluate the integral from 0 to b:

$$\left[\frac{b^3}{3} - \frac{b^2}{2} + b\right] = 1$$

This equation simplifies to:

$$\frac{b^3}{3} - \frac{b^2}{2} + b = 1$$

This cubic equation can be solved numerically (e.g., using Python or a calculator) to find the value of b.

Next, calculate the probability $P(0 \le x \le 0.3)$:

$$P(0 \le x \le 0.3) = \int_0^{0.3} (x^2 - x + 1)\, dx$$

Perform the integral:

$$\left[\frac{x^3}{3} - \frac{x^2}{2} + x\right]_0^{0.3} = \frac{(0.3)^3}{3} - \frac{(0.3)^2}{2} + 0.3$$

Substitute $x = 0.3$:

$$P(0 \le x \le 0.3) = 0.009 - 0.045 + 0.3 = 0.264$$

Thus, the probability is approximately:

$$\boxed{0.264}$$

Question 8 (10 points):
We want to show that:

$$\frac{1}{n}\sum_{i=1}^{n}\left(\phi(x_i)^{\top}w - y_i\right)^2 = \frac{1}{n}\left(\Phi w - y\right)^2$$

Start by expressing the sum in matrix form:

$$\sum_{i=1}^{n}\left(\phi(x_i)^{\top}w - y_i\right)^2 = \left(\Phi w - y\right)^{\top}\left(\Phi w - y\right)$$

This can be written as:

$$\frac{1}{n}\sum_{i=1}^{n}\left(\phi(x_i)^{\top}w - y_i\right)^2 = \frac{1}{n}\left(\Phi w - y\right)^{\top}\left(\Phi w - y\right)$$

For the derivative, we calculate:

$$\frac{d}{dw}\left(\frac{1}{n}\sum_{i}\left(\phi(x_i)^{\top}w - y_i\right)^2\right) = \frac{d}{dw}\left(\frac{1}{n}\left(\Phi w - y\right)^{\top}\left(\Phi w - y\right)\right)$$

The result simplifies to:

$$\frac{2}{n}\Phi^{\top}(\Phi w - y)$$