CS 1678/2078 Homework 2

Alaa Alghwiri

Written Responses (Part 1)

Given that $f_*(x) = 6x + 4\cos(3x+2) - x^2 + 10\ln(\frac{|x|}{10} + 1) + 7$, find the following:

Problem 1

In this part, I will need to find $\phi(x) = [?]^T$: Based on the given function and in order,

$$\phi(x) = \begin{bmatrix} x & cos(3x+2) & x^2 & \ln(\frac{|x|}{10} + 1) & 1 \end{bmatrix}^T$$

Problem 2

In this part i will need to find the optimal weights that corresponds to the features in part 1:

$$w^* = \begin{bmatrix} 6 & 4 & -1 & 10 & 7 \end{bmatrix}^T$$

Problem 3

In this part, i will need to evaluate the same requirements for part 1 and 2 but for the following function:

$$f_*(x) = 6x \times 4\cos(3x+2) \times x^2 \times 10\ln(\frac{|x|}{10} + 1) \times 7$$

The relationship between features in this case is multiplicative and not additive. Mathematically, i can apply a trick by using the natural log on $f_*(x)$ and this will convert the relationship between features into additive relationship:

$$\ln f_*(x) = \ln(6x \times 4\cos(3x+2) \times x^2 \times 10\ln(\frac{|x|}{10}+1) \times 7)$$

$\ln f_*(x) = \ln(6x) + \ln(4cos(3x+2)) + \ln(x^2) + \ln(10\ln(\frac{|x|}{10}+1)) + ln(7)$

$\ln f_*(x) = \ln(6) + \ln(x) + \ln(4) + \ln(cos(3x+2)) + \ln(x^2) + \ln(10) + \ln\ln(\frac{|x|}{10}+1) + \ln(7)$

According to this and in order:

$\phi(x) = \begin{bmatrix} 1 & \ln(x) & 1 & \ln(cos(3x+2)) & \ln(x^2) & 1 & \ln\ln(\frac{|x|}{10}+1) & 1 \end{bmatrix}^T$

And

$w^* = \begin{bmatrix} \ln(6) & 1 & \ln(4) & 1 & 1 & \ln(10) & 1 & \ln(7) \end{bmatrix}^T$

In this case, I can simplify the multiplication terms as follow: $6 \times 4 \times 7 \times 10 = 1680$, then we get the following:

$$1680 \times x^3 \times \cos(3x+2) \times \ln(\frac{|x|}{10}+1)$$

Distribute $168 \times x^3 \times \cos(3x+2)$ inside the natural log, this will results to:

$$1680 \times x^3 \times \cos(3x+2) \times \ln(\frac{|x|}{10}+1) + 168 \times x^3 \times \cos(3x+2)$$

Problem 4

In this problem, we are looking for:

$\frac{\partial}{\partial \hat{y}} g(\hat{y}, y) = \frac{\partial}{\partial \hat{y}} \left(\frac{1}{2m} \sum_{i=1}^{m} (\hat{y}_i - y_i)^2 \right)$

Since differentiation is linear, we can move the derivative inside the summation:

$\frac{1}{2m} \sum_{i=1}^{m} \frac{\partial}{\partial \hat{y}_i} (\hat{y}_i - y_i)^2$

$\frac{\partial}{\partial \hat{y}_i} (\hat{y}_i - y_i)^2 = 2(\hat{y}_i - y_i)$

$\frac{1}{2m} \sum_{i=1}^{m} 2(\hat{y}_i - y_i)$

And finally, the derivative evaluates to:

$\frac{1}{m} \sum_{i=1}^{m} (\hat{y}_i - y_i)$

And this is obviously: $\mathbb{E}[\hat{y} - y]$

In a vector form:

$$\frac{1}{m} \begin{bmatrix} \hat{y_1} - y_1 & \hat{y_2} - y_2 & ... & \hat{y_m} - y_m \end{bmatrix}^T$$

Problem 5

$\hat{y} = f(X, w) = Xw$

Now, i need to find:

$\frac{\partial}{\partial w} f(X, w) = \frac{\partial}{\partial w} \sum_{i=1}^{m} X_i w_i$

Solution:

$= \frac{\partial}{\partial w} \sum_{i=1}^{m} X_i w_i$

$= \sum_{i=1}^{m} \frac{\partial}{\partial w} X_i w_i$

And finally:

$\frac{\partial \hat{y}_i}{\partial w_i} = \sum_{i=1}^{m} X_i$

In a matrix representation:

$\begin{bmatrix} X_1 & X_2 & X_3 & .... & X_m \end{bmatrix}$

Problem 6

In this question, i will evaluate the gradient for the loss function with respect to the weight w. And this should be expressed in matrices/vectors without summation.

$\nabla l(w) = \frac{\partial}{\partial w} g(f(X, w), y)$

$\frac{\partial}{\partial w} g(f(X, w), y) = \frac{\partial}{\partial w} \frac{1}{2m} (X_i w_i - y_i)^2$

$= \frac{1}{m} (X_i w_i - y_i) x_i$

$\nabla l(w) = x^T \nabla g(\hat{y}, y)$