Theoretical Machine Learning
Assignment 2

Naman Mohan Singh

Theoretical

Problem 1

Prove the following, where X is a Random Variable, and Ω is the Sample Space:

(a) $X^{-1}(\Phi) = \Phi$

(b) $X^{-1}(\mathbb{R}) = \Omega$

Solution:

(a) $X^{-1}(\Phi) = \Phi$ By definition, the inverse image of the empty set under any function is the empty set. Formally, for any set A, $X^{-1}(A) = \{\omega \in \Omega \mid X(\omega) \in A\}$.

When $A = \Phi$:

$$X^{-1}(\Phi) = \{\omega \in \Omega \mid X(\omega) \in \Phi\}$$

Since no element can belong to the empty set(An Empty Event), $X^{-1}(\Phi)$ must be empty. Thus:

$$X^{-1}(\Phi) = \Phi$$

(b) $X^{-1}(\mathbb{R}) = \Omega$ By definition, $X : \Omega \to \mathbb{R}$. The inverse image of the entire real line $\mathbb{R}$ under X is the set of all elements in Ω that X maps to some real number. Since X is defined on the entire sample space Ω:

$$X^{-1}(\mathbb{R}) = \{\omega \in \Omega \mid X(\omega) \in \mathbb{R}\}$$

Since $X(\omega)$ is real for all $\omega \in \Omega$:

$$X^{-1}(\mathbb{R}) = \Omega$$

We can say the whole sample space is the output of the whole real line

Problem 2

We have a Random Variable X from the Exponential Distribution. Distribution Function of Exp(λ) is as -

$$F_X(x) = \begin{cases} 0 & x < 0 \\ 1 - e^{-\lambda x} & x \geq 0 \end{cases}$$

Find the support of X.

Solution:

The support of a random variable X is the set of all points x such that the probability density function $f_X(x)$ is positive. For the exponential distribution Exp(λ):

$$f_X(x) = \begin{cases} 0 & x < 0 \\ \lambda e^{-\lambda x} & x \geq 0 \end{cases}$$

The function $\lambda e^{-\lambda x}$ is positive for all $x \geq 0$. Therefore, the support of X is:

$$\text{Support of } X = [0, \infty)$$

Problem 3

You were given the expression of m and c is a linear regression in 2 variable case with $L2$-Norm as -

$$m = \frac{N(\sum xy) - (\sum x)(\sum y)}{N(\sum x^2) - (\sum x)^2}$$

$$c = \frac{(\sum y)(\sum x^2) - (\sum x)(\sum xy)}{N(\sum x^2) - (\sum x)^2}$$

Show how you get these values, and that it is indeed a minima for the loss function for $L2$ Norm.

Solution:

The objective of linear regression is to find the line $y = mx + c$ that minimizes the sum of squared errors (L2 norm) between the observed values y and the predicted values $\hat{y} = mx + c$.

The loss function L is given by:

$$L = \sum_{i=1}^{N} (y_i - (mx_i + c))^2$$

To minimize L, we take the partial derivatives with respect to m and c and set them to zero.

$$\frac{\partial L}{\partial m} = -2 \sum_{i=1}^{N} x_i (y_i - (mx_i + c)) = 0$$

$$\frac{\partial L}{\partial c} = -2 \sum_{i=1}^{N} (y_i - (mx_i + c)) = 0$$

Solving these equations simultaneously:

$$\sum y_i = m \sum x_i + Nc \tag{1}$$

$$\sum x_i y_i = m \sum x_i^2 + c \sum x_i \tag{2}$$

From Equation (1):

$$c = \frac{\sum y_i - m \sum x_i}{N}$$

Substituting c into Equation (2):

$$\sum x_i y_i = m \sum x_i^2 + \frac{\sum y_i - m \sum x_i}{N} \sum x_i$$

$$N \sum x_i y_i = mN \sum x_i^2 + (\sum y_i - m \sum x_i) \sum x_i$$

$$N \sum x_i y_i = mN \sum x_i^2 + \sum y_i \sum x_i - m(\sum x_i)^2$$

$$N \sum x_i y_i - \sum y_i \sum x_i = m(N \sum x_i^2 - (\sum x_i)^2)$$

$$m = \frac{N \sum x_i y_i - \sum x_i \sum y_i}{N \sum x_i^2 - (\sum x_i)^2}$$

For c:

$$c = \frac{\sum y_i - m \sum x_i}{N}$$

Substituting the value of m into the equation for c:

$$c = \frac{\sum y_i - \left(\frac{N \sum x_i y_i - \sum x_i \sum y_i}{N \sum x_i^2 - (\sum x_i)^2} \right) \sum x_i}{N}$$

$$c = \frac{\sum y_i (N \sum x_i^2 - (\sum x_i)^2) - \sum x_i (N \sum x_i y_i - \sum x_i \sum y_i)}{N(N \sum x_i^2 - (\sum x_i)^2)}$$

$$c = \frac{(\sum y_i)(N \sum x_i^2) - (\sum y_i)(\sum x_i)^2 - (\sum x_i)(N \sum x_i y_i) + (\sum x_i)^2 (\sum y_i)}{N(N \sum x_i^2 - (\sum x_i)^2)}$$

$$c = \frac{(\sum y_i)(\sum x_i^2) - (\sum x_i)(\sum x_i y_i)}{N(\sum x_i^2) - (\sum x_i)^2}$$

Thus, the expressions for m and c are indeed the solutions for minimizing the L2 loss function.

Took a whole lot of time to Write in Latex :P