Assignment 1

a) Derivation of gradient of objective function

objective function is given as.

$$J = \sum_{i=1}^{n} L(x_i, y_i | w, b) = \frac{-1}{n} \sum_{i=1}^{n} \left(y \log \hat{y} + (1-y) \log (1-\hat{y}) \right) + \alpha \sum_{j=1}^{d} w_j^2$$

where,

$$\hat{y} = \frac{1}{1+e^{-(xw+b)}}$$

$$\therefore J = \frac{-1}{n} \sum_{i=1}^{n} \left[y \log\left(\frac{1}{1+e^{-(xw+b)}}\right) + (1-y) \log\left(1 - \left(\frac{1}{1+e^{-(xw+b)}}\right)\right)\right] + \alpha \sum_{j=1}^{d} w_j^2$$

$\frac{\partial J}{\partial w}$ $\nabla J = \left[\frac{\partial J}{\partial w}, \frac{\partial J}{\partial b} \right]$

let $z = e^{-(xw+b)}$

$\therefore$ using chain rule $\frac{\partial J}{\partial z} \frac{\partial z}{\partial w}$

$$\therefore \frac{\partial}{\partial z} \left[\frac{-1}{n} \sum_{i=1}^{n} \left[y . \log\left(\frac{1}{1+e^{-z}}\right) + (1-y) \log\left(1 - \frac{1}{1+e^{-z}}\right)\right] \right] + \alpha \sum_{j=1}^{d} w_j^2$$

$$\therefore \frac{\partial J}{\partial z} = \left[\frac{-1}{n} \sum_{i=1}^{n} \left(y \left(\frac{1}{1+e^{-z}}\right) + (1-y) . \left(- \frac{1}{1+e^{-z}} \right) \right) \right] + 2\alpha \sum_{j=1}^{d} w_j$$

$$\frac{\partial z}{\partial w} = \left[\frac{\partial}{\partial w} . e^{-(xw+b)} \right] \Rightarrow e^{-(xw+b)} \times (-x)$$

$$\therefore \frac{\partial J}{\partial w} = \frac{1}{n} \sum_{i=1}^{n} (y - \hat{y}) x + 2\alpha \sum_{j=1}^{d} w_j$$

Similarly for $\frac{\partial J}{\partial b}$

$$\frac{\partial J}{\partial b} = \frac{1}{n} \sum_{i=1}^{n} (\hat{y} - y) + 0 .$$