Logistic

Pham Van Hong – 11201638

Problem 1:

$$h\theta(x) = g(\theta^T x) = \frac{1}{1 + e^{-\theta^T x}}$$

$$g(z) = \frac{1}{(1 + e^{-z})}$$

We have:

$$L(\theta) = \log L(\theta) = \sum_{i=1}^{N} y^i . \log h\left(x^i\right) + \left(1 - y^i\right) . \log\left(1 - h\left(x^i\right)\right)$$

$$\frac{\partial L(\theta)}{\partial \theta_J} = \left(y . \frac{1}{g(\theta^T x)} - (1 - y) . \frac{1}{1 - g(\theta^T x)}\right) . \frac{\partial g(\theta^T x)}{\partial \theta_i}$$

$$= \left(y . \frac{1}{g(\theta^T x)} - (1 - y) . \frac{1}{1 - g(\theta^T x)}\right) . g(\theta^T x) . \left(1 - g(\theta^T x)\right) . \theta^T x$$

$$= \ (y . (1 - g(\theta^T x) - (1 - y) . g(\theta^T x)) . x_J$$

$$= (y - h\theta(x)) . x_J$$

Problem 5:

A, Let $t_i \ mean\ the\ corresponding\ target\ value\ where\ t = (t_1, t_2, t_3, \dots, t_n)^T$

We have:

$$\frac{\partial L}{\partial w} = x_i . (t_i - y_i)$$

$$\frac{\partial t_i}{\partial w} = x_i . t_i . (1 - t_i)$$

$$\frac{\partial^2 L}{\partial w^2} = x_i . \frac{\partial t_i}{\partial w}$$

$$= x_i^2 . t_i . (1 - t_i) \geq 0$$

So, the loss binary-crossentropy with logistic model is convex

B, MSE:

$$L = \frac{1}{N} . \sum_{n=1}^{N} (t_i - y_i)^2$$

We derivate:

$$\frac{\partial L}{\partial w} = \frac{\partial L}{\partial t} . \frac{\partial t}{\partial w}$$

$$= \ -2 . (y - t) . x . t . (1 - t)$$

$$= -2.x.(y.t - t^2).(1-t)$$
$$= -2.x.(y.t - y.t^2 - t^2 + t^3$$

The second derivate:

$$\frac{\partial^2 L}{\partial w^2} = -2.x.(y.\frac{\partial t^2}{\partial w} - y.\frac{\partial t^2}{\partial t}.\frac{\partial t}{\partial w} - \frac{\partial t^2}{\partial t}.\frac{\partial t}{\partial w} + \frac{\partial t^3}{\partial t}.\frac{\partial t}{\partial w}$$
$$= -2.x.\left(y.x.t.(1-t) - y.2.t.x.t.(1-t) - 2t.x.t.(1-t) + 3.t^2.x.t.(1-y)\right)$$
$$= -2.x^2.t.(1-t).(y - 2yt - 2t + 3t^2)$$

Because $x^2.t.(1-t) \geq 0, consider\ only: f(t) = -2.(y - 2yt - 2t + 3t^2)$

$$f(t) = \begin{cases} 4t - 6t^2 = 2t.(2t - 3t) \quad when\ y = 0\ (1) \\ -2 + 8t - 6t^2 = -2(3t-1).(t-1)\ when\ y = 1\ (2) \end{cases}$$

In the case of (1), $f(t) \leq 0\ when \frac{2}{3} \leq t \leq 1$

In the case of (2), $f(t) \leq 0\ when\ 0 \leq t \leq \frac{1}{3}$

So, the loss mean square error with logistic model is NOT convex.