Machine Learning Homework day 6

Trinh Thi Huong Lien

October 2022

1 Proof that with logistic regression, loss binary crossentropy is the convex function with W

We have

$$-f(x) = ylog(\hat{y}) + (1-y)log(1-\hat{y}) \tag{1}$$

$$-f(x) = ylog(\frac{1}{1+e^{-\theta x}}) + (1-y)log(1-\frac{1}{1+e^{-\theta x}}) \tag{2}$$

$$-f(x) = ylog(\frac{e^{\theta x}}{1+e^{\theta x}}) + (1-y)log(\frac{1}{1+e^{\theta x}}) \tag{3}$$

$$-f(x) = ylog(e^{\theta x}) - ylog(1+e^{\theta x}) + (1-y)log(1) - (1-y)log(1+e^{\theta x}) \tag{4}$$

$$-f(x) = y(\theta x) - ylog(1+e^{\theta x}) - log(1+e^{\theta x}) + ylog(1+e^{\theta x}) \tag{5}$$

$$-f(x) = xy\theta - log(1+e^{\theta x}) \tag{6}$$

$$f(x) = log(1+e^{\theta x}) - xy\theta \tag{7}$$

$$\frac{\partial f}{\partial \theta} = \frac{1}{1+e^{\theta x}} xe^{\theta x} - xy \tag{8}$$

$$\frac{\partial f}{\partial \theta} = \frac{xe^{\theta x}}{1+e^{\theta x}} - xy \tag{9}$$

$$=> \frac{\partial^2 f}{\partial \theta^2} = x\frac{-1}{(1+e^{-\theta x})^2}(e^{-\theta x})(-x) \tag{10}$$

$$=> \frac{\partial^2 f}{\partial \theta^2} = \frac{x^2 e^{-\theta x}}{(1+e^{-\theta x})^2} \geq 0 \forall x \tag{11}$$

=> Loss binary cross entropy is the convex function with W