Homework 4: CS5984 Assignment 4

Richa Sinha

December 7, 2017

1 Loss function derivation

Given an autoencoders with binary input which uses cross entropy error and uses sigmoid activation layer for the output.

Now, we know that the Error function E is given by:

$$E = -\sum_{k}(x_k log(\hat{x_k}) + (1 - x_k) log(1 - \hat{x_k}))$$

where x_k and $\hat{x_k}$ are input and output respectively.

Now the output $\hat{x_k}$ is given by:

$$\hat{x_k} = \frac{1}{1 + e^{-s_k}}$$

where s_k is given as :

$$s_k = \sum_{j} h_j * W_{jk}$$

Now,

$$\frac{\partial E}{\partial W_{jk}} = \frac{\partial E}{\partial \hat{x_k}} * \frac{\partial \hat{x_k}}{\partial s_k} * \frac{\partial s_k}{\partial W_{jk}}$$

Calculating the above terms we get:

$$\frac{\partial E}{\partial \hat{x_k}} = \frac{\hat{x_k} - x_k}{\hat{x_k}(1 - \hat{x_k})}$$

$$\frac{\partial \hat{x_k}}{\partial s_k} = \hat{x_k}(1 - \hat{x_k})$$

$$\frac{\partial s_k}{\partial W_{jk}} = h_j$$

Combining them we get:

$$\frac{\partial E}{\partial W_{jk}} = h_j(\hat{x_k} - x_k)$$

In the given case, there is no hidden layer, the output becomes

$$\frac{\partial E}{\partial W_k} = (\hat{x_k} - x_k)$$