# CS5370: Assignment 3

Nakka Chakradhar - EE16BTECH1102

October 17, 2019

1. Consider a deep network $F(W, b, x)$ with $tanh$ activation. The loss function for the network be MSE (Mean Squared error)

   We know that $tanh(0) = 0$ since $tanh(x) = (e^x - e^{-x})/(e^x + e^{-x})$

   Now if we initialise all weights and biases to be 0, the resulting deep network would boil down to $F(0, 0, x)$ which is always 0.

   Since all weights are zero, the following gradient calculations for back-propagation would also be 0 causing no change in the parameters over the epochs.

   Hence 0 is the stationary point for the error function.

2. Answer to question 2

   $E = \frac{\lambda_1 w_1}{2} + \frac{\lambda_2 w_2}{2}$

   From the above expression we can say the following :

   $\frac{\partial E}{\partial w_1} = \lambda_1 w_1$

   $\frac{\partial E}{\partial w_2} = \lambda_2 w_2$

   $\frac{\partial^2 E}{\partial w_1^2} = \lambda_1$

   $\frac{\partial^2 E}{\partial w_2^2} = \lambda_2$

   $\frac{\partial^2 E}{\partial w_1 \partial w_2} = 0$

   From these, the Hessian of E turns to be as follows

   $Hess(E) = \begin{pmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{pmatrix}$

   From the above matrix, since its a diagonal matrix, we can conclude that the eigenvalues of Hessian of the given error function are indeed $\lambda_1$ and $\lambda_2$