Machine learning Homework- Deep Learning

Abinav Ravi Venkatakrishnan - 03694216 and Abhijeet Parida - 03679676

December 15, 2018

1 Activation Function

Problem 1:

The matrix operation $w^T + b$ is essentially a linear operation. When we stack linear operations over other linear operations we essentially get a linear function. It is impossible to approximate complex functions with just linear operations, therefore non-linearity is introduced to overcome this problem.

Problem 2:

The sigmoid activation function is

$$\sigma(x) = \frac{1}{1 + e^{-x}}$$

The tanh activation is

$$\begin{aligned} tanh(x) &= \tfrac{e^{2x}-1}{e^{2x}+1} \\ tanh(\frac{x}{2}) &= \tfrac{e^{x}-1}{e^{x}+1} \\ tanh(\frac{x}{2}) &= \tfrac{1-e^{-x}}{1+e^{-x}} \\ tanh(\frac{x}{2}) &= (1 - e^{-x})\sigma(x) \end{aligned}$$

Problem 3:

2 Numerical Stability

Problem 4:

Problem 5:

Problem 6: