CS 6375.502

ML Assignment - 2

1.1) Gradient Descent

The output of the perceptron is

$$o = \omega_0 + \omega_1(x_1 + x_1^2) + \dots + \omega_n(x_n + x_n^2)$$

Perceptron Inputs : $x_1, x_2 \dots\dots x_n$
Corresponding Weights : $\omega_0, \omega_1 \dots \omega_n$

First, we need to define the error function which is

$$E(\omega) = \frac{1}{2} \sum_{d \in D} (t_d - o_d)^2 \quad -①$$

where o_d = o [output function]

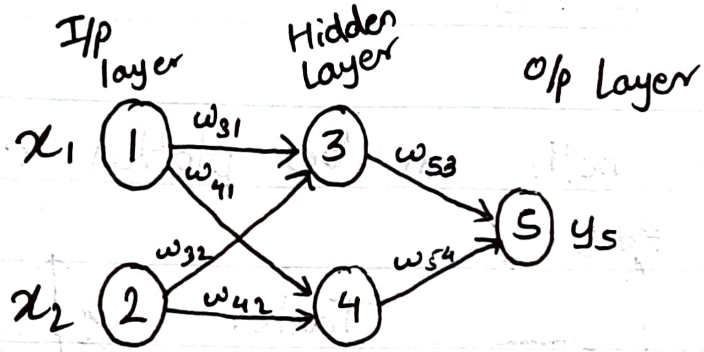
Considering $(t-o) = z$,
we need to find the derivative of E w.r.t (ω)
Then, the partial derivative is $\frac{\partial E}{\partial z} \cdot \frac{\partial z}{\partial \omega_i}$ -②

From ① and ②

$$= \frac{1}{2} \cdot (2)(t_d - o_d)(0 - x_i - x_i^2)$$

$$\therefore \frac{\partial E}{\partial \omega_i} = \frac{\cancel{2}}{\cancel{2}} \boxed{(t_d - o_d)(-x_i - x_i^2)}$$

1.2) Comparing Activation Function



a) Output of Neural Net

i) Output of Hidden Layer :

$$y_3 \rightarrow h(x_1 w_{31} + x_2 w_{32})$$
$$y_4 \rightarrow h(x_1 w_{41} + x_2 w_{42})$$

ii) Output of y_5

$$y_5 \rightarrow h(y_3 w_{53} + y_4 w_{54})$$

Final Output of Neural Net :-

$$y_5 = h\Big(w_{53} \cdot \big(h(x_1 w_{31} + x_2 w_{32})\big) + w_{54} \cdot \big(h(x_1 w_{41} + x_2 w_{42})\big)\Big)$$

b) With vector notation, output of the neural net is :-

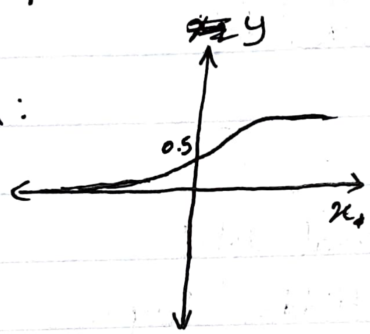
$$y_5 = \begin{pmatrix} w_{53} & w_{54} \end{pmatrix} \left[\begin{pmatrix} w_{31} & w_{32} \\ w_{41} & w_{42} \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} \right]$$

1.2) c) $h_s(x) = \dfrac{1}{1+e^{-x}}$ .... (Sigmoid)

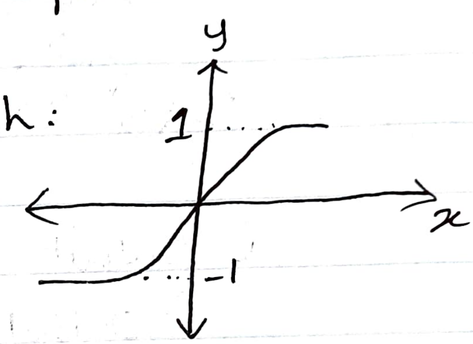
$h_t(x) = \dfrac{e^{x}-e^{-x}}{e^{x}+e^{-x}}$ ...... (Tanh)

The graph of both can be plotted as:

Sigmoid:



Tanh:



In order to show that the output of both functions is the same, we will rearrange the tanh function similar to the sigmoid function.

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

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

$$= \frac{e^{x}+e^{-x}}{e^{x}+e^{-x}} + \frac{(-2)\,e^{-x}}{e^{x}+e^{-x}}$$

$$= 1 - 2\left(\frac{1}{e^{+x}(e^{x}+e^{-x})}\right)$$

1.2) c) (continued)

$$\therefore \tanh(x) = 1 - 2\left(\frac{1}{1+e^{2x}}\right)$$

From the sigmoid function, we can infer that $\left(\frac{1}{1+e^{2x}}\right) = \sigma(-2x)$

Since $\sigma(-2x) = 1 - \sigma(2x)$,

$$\tanh(x) = 1 - 2(1 - \sigma(2x))$$
$$= 2\sigma(2x) - 1$$

Thus, we can see that the output of sigmoid & tanh functions are the same with parameters differing only by linear transformation & constants.