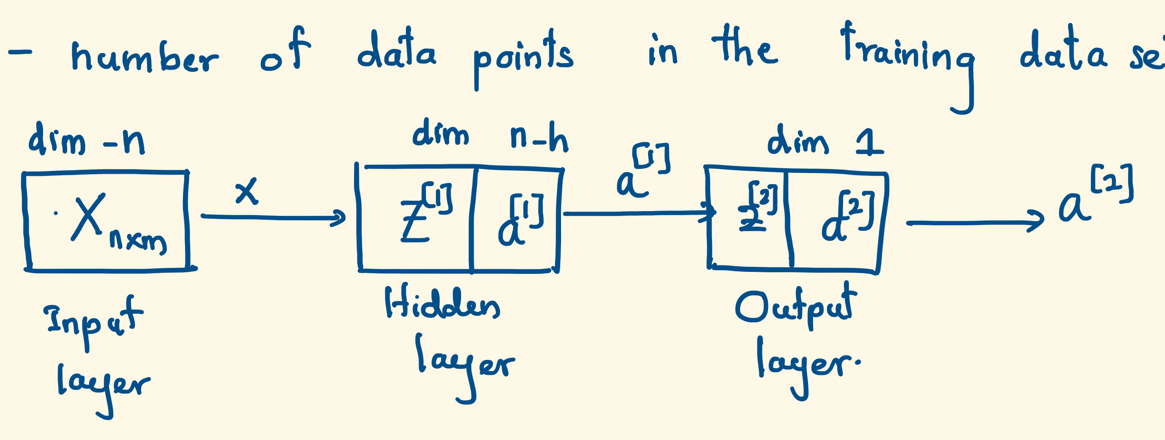
HW2

Monday, October 16, 2023 6:47 AM

One hidden Layer NN

Let n – dimension of the data set

m – number of data points in the training data set.



For 1 data point go through hidden layer

$$z^{[1]}_{(n\text{-}h,1)} = w^{[1]}_{(n\text{-}h,n)} \cdot x^{(1)}_{[n,1]} + b^{[1]}_{(n\text{-}h,1)}$$

$$a^{[1]}_{(n\text{-}h,1)} = \text{ReLU}\left(z^{[1]}_{(n\text{-}h,1)}\right)$$

Hence, For m data points go through hidden layer.

$$Z^{[1]}_{(n\text{-}h,m)} = W^{[1]}_{(n\text{-}h,m)} \, X_{[n,m]} + b^{[1]}_{(n\text{-}h,m)}$$

$$a^{[1]}_{(n\text{-}h,m)} = \text{ReLU}\left(Z^{[1]}_{n\text{-}h,m}\right)$$

Loss: $L = \frac{1}{2}\left(y - a^{[2]}\right]^2$

Cost: $\text{cost} = \frac{1}{m}\left[\sum L\right]$

Backpropagation

loss $L = \frac{1}{2}\left(a^{[2]} - y\right]^2$

output layer activation f^n : Bipolar Sigmoid $(x) = (1-e^{-x})/(1+e^{-x})$

hidden layer activation f^n : ReLU $(x) = \max(0, x)$

Let $a = \text{Bi-Sig}(z)$

$$\frac{da}{dz} = \frac{d}{dz}\left[\frac{1 - e^{-z}}{1 + e^{-z}}\right)$$

$$= \frac{(1+e^{-z})\,e^{-z} - (1-e^{-z})(-e^{-z})}{(1+e^{-z})^2}$$

$$= \frac{e^{-z} \cancel{+1} + e^{-z} \cancel{-1}}{(1+e^{-z})^2}$$

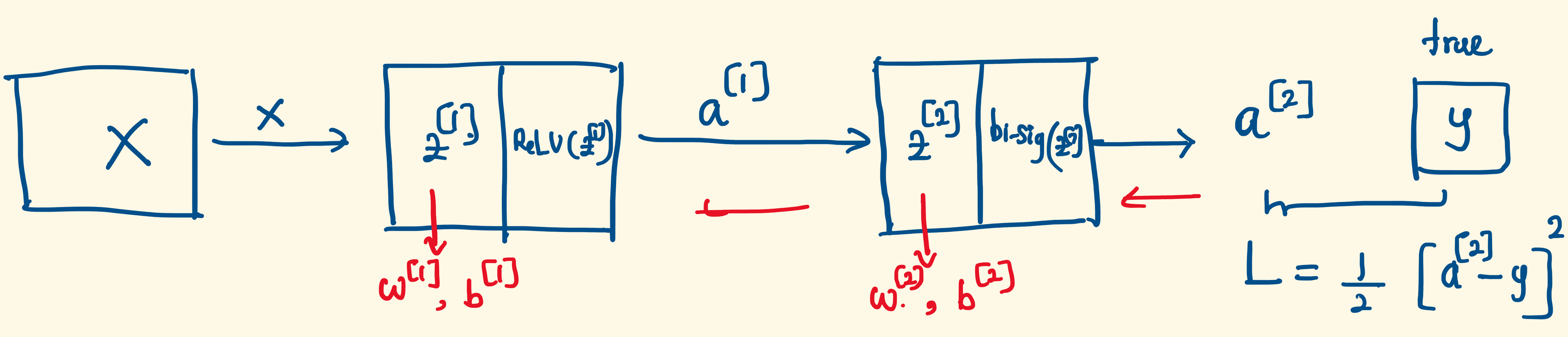
$$= \frac{2e^{-z}}{(1+e^{-z})^2}$$

Let $a = \text{ReLU}(z)$

$$\frac{da}{dz} = \begin{cases} 1 & \text{if } z \geq 0 \\ 0 & \text{if } z < 0 \end{cases}$$

Let $L = \frac{1}{2}\left[a - y\right]^2$

$$\frac{dL}{da} = [a - y]$$



$$da^{[2]} := \frac{dL}{da^{[2]}} = \left(a^{[2]} - y\right)$$

$L = \frac{1}{2}\left[a^{[2]} - y\right]^2$

$$dz^{[2]} := \frac{dL}{dz^{[2]}} = \frac{dL}{da^{[2]}} \cdot \frac{da^{[2]}}{dz^{[2]}}$$

$a^{[2]} = \text{bi-sig}\left[z^{[2]}\right)$

$$= \underbrace{\left(a^{[2]} - y\right)}_{da^{[2]}} \frac{d}{dz^{[2]}}\left[\text{bipolar-sigmoid}\left(z^{[2]}\right)\right)$$

$$\boxed{dw^{[2]}} := \frac{dL}{dw^{[2]}} = \frac{\partial L}{\partial z^{[2]}} \cdot \frac{\partial z^{[2]}}{\partial w^{[2]}}$$

$z^{[2]} = w^{[2]} a^{[1]} + b^{[2]}$

$$= \frac{1}{m} dz^{[2]} \cdot a^{[1]}$$

np dot

$$\boxed{db^{[2]}} := \frac{dL}{db^{[2]}} = \frac{\partial L}{\partial z^{[2]}} \cdot \frac{\partial z^{[2]}}{\partial b^{[2]}}$$

$$= dz^{[2]}$$

$$da^{[1]} := \frac{dL}{da^{[1]}} = \underbrace{\frac{dL}{da^{[2]}} \cdot \frac{da^{[2]}}{dz^{[2]}}} \cdot \frac{dz^{[2]}}{da^{[1]}}$$

$a^{[1]} = \text{ReLU}\left[z^{[1]}\right)$

$z^{[1]} = w^{[1]}x + b^{[1]}$

$$= dz^{[2]} \cdot w^{[2]}$$

$$dz^{[1]} := \frac{dL}{dz^{[1]}} = \frac{dL}{da^{[2]}} \cdot \frac{da^{[2]}}{dz^{[2]}} \cdot \frac{dz^{[2]}}{da^{[1]}} \cdot \frac{da^{[1]}}{dz^{[1]}}$$

$$= da^{[1]} \cdot \frac{d}{dz^{[1]}}\left[\text{ReLU}\left(z^{[1]}\right)\right]$$

$$dw^{[1]} := \frac{\partial L}{\partial w^{[1]}} = \underbrace{\frac{dL}{da^{[2]}} \cdot \frac{da^{[2]}}{dz^{[2]}} \cdot \frac{dz^{[2]}}{da^{[1]}} \cdot \frac{da^{[1]}}{dz^{[1]}}} \cdot \frac{\partial z^{[1]}}{\partial w^{[1]}}$$

$$= dz^{[1]} \cdot X$$

$$db^{[1]} := \frac{\partial L}{\partial b^{[1]}} = \frac{dL}{da^{[2]}} \cdot \frac{da^{[2]}}{dz^{[2]}} \cdot \frac{dz^{[2]}}{da^{[1]}} \cdot \frac{da^{[1]}}{dz^{[1]}} \cdot \frac{\partial z^{[1]}}{\partial b^{[1]}}$$

$$= dz^{[1]}$$

Parameter update.

$w_2 = w_2 + dw_2 * \gamma$

$w_1 = w_1 + dw_1 * \gamma$

$b_2 = b_2 + db_2 * \gamma$

$b_1 = b_1 + db_1 * \gamma$