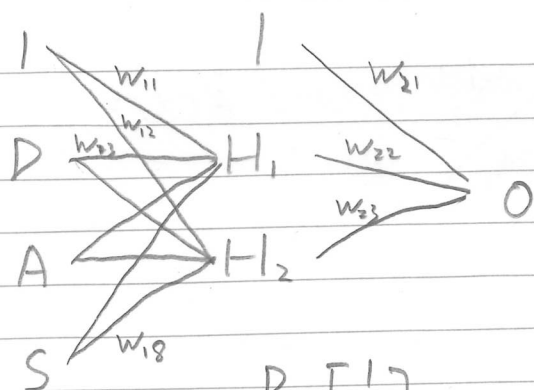


# Week 9 Q2.



$$W_{1i} = \begin{bmatrix} 1 \\ -1 \\ 1 \\ -1 \\ -2 \\ 2 \\ 3 \\ -2 \end{bmatrix}$$

$$W_{2j} = \begin{bmatrix} -1 \\ 1 \\ 1 \\ 1 \end{bmatrix}$$

Input:  $D = \begin{bmatrix} 1 \\ 2 \\ 2 \end{bmatrix}$   $O = [1]$

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

$$H_1 = 1 + 1 \cdot 1 + (-2) \cdot 2 + 3 \cdot 2 = 4$$

$$\sigma(H_1) = 0.98$$

$$H_2 = -1 + (-1) \cdot 1 + (2) \cdot 2 + (-2) \cdot 2 = -2$$

$$\sigma(H_2) = 0.12$$

forward  $O = -1 + 0.98 + 0.12 = 0.1$

$$\sigma(O) = 0.52 = z$$

loss  $\rightarrow L(w) = -[y \log(z) + (1-y) \log(1-z)]$  Binary Cross Entropy Loss

$$\frac{dL}{dz} = -\left[\frac{y}{z} - \frac{1-y}{1-z}\right] = \frac{z-y}{z(1-z)}$$

$$z = \sigma(h) = \frac{1}{1+e^{-h}} \leftarrow \text{Sigmoid function.}$$

$$\frac{dz}{dh} = z(1-z)$$

$$\frac{\partial O}{\partial W_{21}} = 1 \text{ for bias.}$$

$$\frac{\partial L}{\partial W_{21}} = \frac{\partial L}{\partial z} \cdot \frac{\partial z}{\partial O} \cdot \frac{\partial O}{\partial W_{21}} = \frac{z-y}{z(1-z)} \cdot z(1-z) = z-y = 0.52-1 = -0.48 \text{ bias}$$

$$\frac{\partial L}{\partial W_{22}} = 0.98 \cdot (-0.48) = -0.47 \leftarrow (\text{Current value}) \times (\text{backward gradient})$$

$$\frac{\partial L}{\partial W_{23}} = 0.12 \cdot (-0.48) = -0.06$$

Eg:  $\frac{\partial L}{\partial W_{13}} = \frac{\partial H_1}{\partial W_{13}} \cdot \frac{\partial \sigma(H_1)}{\partial H_1} \cdot \frac{\partial L}{\partial \sigma(H_1)}$

$$\frac{d\sigma(H_1)}{dH_1} = 0.98(1-0.98) = 0.02$$

derivative of Sigmoid

$$\frac{d\sigma(H_2)}{dH_2} = 0.12(1-0.12) = 0.11$$

$$\frac{\partial L}{\partial W_{1i}} = \begin{bmatrix} (0.02)(-0.47) \\ (0.11)(-0.06) \\ (1)(-0.0094) \\ (1)(-0.0066) \\ (2)(-0.0094) \\ (2)(-0.0066) \\ (2)(-0.0094) \\ (2)(-0.0066) \end{bmatrix} = \begin{bmatrix} -0.0094 \\ -0.0066 \\ -0.0094 \\ -0.0066 \\ -0.0188 \\ -0.0132 \\ -0.0188 \\ -0.0132 \end{bmatrix}$$

bias

$$w' = w - \eta \frac{\partial L}{\partial w}$$

Learning rate

choose a LR, then update