Homework 5

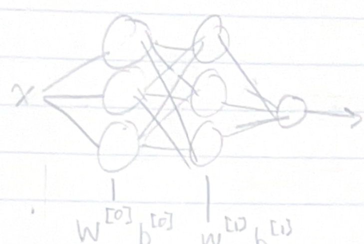
1) $a^{[0]} = x \qquad a^{[1]} = g(W^{[1]}x + b^{[1]}) \qquad g(W^{[1]}x + b^{[1]}) = W^{[1]}x + b^{[1]} = a^{[1]}$

$$a^{[2]} = g(\underbrace{W^{[2]}W^{[1]}}_{A}x + \underbrace{W^{[2]}b^{[1]} + b^{[2]}}_{C}) \Rightarrow g(Ax + C) = Ax + C$$

$$\boxed{a^{[r]} = Ax + C}$$

$r = 3$

2).



$J = \frac{1}{2}(a^{[r]} - y)^2$

$$J = \frac{1}{2}((Ax + C) - y)^2$$

$$\Rightarrow \frac{1}{2}\left[W^{[2]}x + b^{[2]} - y\right]^2$$

$$\frac{\partial}{\partial W} = (W^{[2]}x + b^{[2]} - y)(x)$$

$$\frac{\partial}{\partial b} = (W^{[2]}x + b^{[2]} - y)$$

3). $\sigma' = \frac{e^{-x}}{(1+e^{-x})^2} \Rightarrow w\left(\frac{e^{-(wa+b)}}{(1+e^{-(wa+b)})^2}\right)$

a).

* this function is maximized when $a, b = 0$

$$\Rightarrow w\left(\frac{1}{(1+1)^2}\right) = w\left(\frac{1}{4}\right)$$

$\boxed{\text{therefore } w \geq 4 \text{ yields a result } \geq 1}$

b) $\sigma' = \sigma(wa+b)(1-\sigma(wa+b))$ $\qquad \cancel{\sigma = \frac{1}{1+e^{-x}}}$

$$= \frac{1}{1+e^{-(wa+b)}}\left(1 - \frac{1}{1+e^{-(wa+b)}}\right)$$

$$h = \frac{1}{1+e^{-(wa+b)}} - \frac{1}{(1+e^{-(wa+b)})^2}$$

$\uparrow$

$$w\left(\frac{1}{1+e^{-(wa+b)}} - \frac{1}{(1+e^{-(wa+b)})^2}\right) \geq 1$$

$$h \geq \frac{1}{w} + \frac{1}{1+e^{-(wa+b)^2}}$$

$w(w-4)$

$wh(1-h) \geq 1$

$wh - wh^2 \geq 1$

$wh^2 - wh + 1 \leq 0 \rightarrow$ quadratic formula

$$h = \frac{w \pm \sqrt{w^2 - 4w}}{2w} = \frac{1}{2} \pm \frac{1}{2w}\sqrt{w^2 - 4w} = \frac{1}{2} \pm \frac{1}{2w}\sqrt{w^2\left(1 - \frac{4}{w}\right)}$$

$$\boxed{= \frac{1}{2} \pm \frac{1}{2}\sqrt{1 - \frac{4}{w}}}$$

4). $p(y|x,\theta) = \prod_{k=1}^{K} \frac{e^{z_k^{[r]}}}{\sum_{j=1}^{K} e^{z_j^{[r]}}}$

$$J = \log(p(y|x,\theta))$$

$$J = \sum_{k=1}^{K} \log\left(\frac{e^{z_k^{[r]}}}{\sum_{j=1}^{K} e^{z_j^{[r]}}}\right) = \sum_{k=1}^{K} e^{z_k^{[r]}} - \sum_{j=1}^{K} z_j^{[r]}$$