Problem 2 (cont'd)

$y = 0$

d.)

$$\max \left[\| \nabla_w Loss(x,y,w) \| \right]$$

occurs when $\nabla_w^2 Loss(x,y,w) = 0$

let's maximize w/r t W

~~which is equivalent to maximizing by w since $\phi(x) = const$~~

$$\frac{d}{dW}\left[\nabla_w Loss(x,y,w)\right]$$

$$= \frac{d}{dW}\left[2\phi(x)\left((\sigma(z))^2 - (\sigma(z))^3\right)\right]$$

$$= 2\phi(x)\left(2\sigma(z) - 3(\sigma(z))^2\right)\frac{d\sigma(z)}{dW}$$

$$\frac{d\sigma(z)}{dW} = \frac{d}{dW}\left[\frac{1}{1+e^{-z}}\right] = \frac{\phi(x)e^{-z}}{(1+e^{-z})^2} = \frac{\phi(x)(1-\sigma(z))}{\sigma(z)^3} = 0$$

$\sigma(z) \to \infty$ (known to be min)

or

$\phi(x) \to 0$ (null answer)

or

$\sigma(z) = 0 \Leftarrow 2\sigma(z) = 3\sigma(z)^2 \Rightarrow \sigma(z) = \frac{2}{3}$

$$\frac{1}{1+e^{-z}} = \frac{2}{3} \Rightarrow \frac{3}{2} - 1 = e^{-z} \Rightarrow e^{z} = 2$$

$w \cdot \phi(x) = 2$ or $\sigma(z) = \frac{2}{3}$