1.

$$h(x_1,x_2)=\sigma(b+w_1x_1+w_2x_2)$$

$$\theta^0=(b,w_1,w_2)=(4,5,6)$$

$$(x_1,x_2,y)=(1,2,3)$$

$$\theta^1=\theta^0-\alpha\nabla_\theta Loss=\theta^0+2\alpha\left[\frac{1}{m}\sum_{i=1}^{m}(y^i-h(x_1^i,x_2^i))\nabla_\theta h\right]$$

$$\nabla_\theta h=\begin{bmatrix}\frac{\partial h}{\partial b}\\ \frac{\partial h}{\partial w_1}\\ \frac{\partial h}{\partial w_2}\end{bmatrix}=\begin{bmatrix}\sigma'(b+w_1x_1+w_2x_2)\\ x_1\sigma'(b+w_1x_1+w_2x_2)\\ x_2\sigma'(b+w_1x_1+w_2x_2)\end{bmatrix}$$

$$\theta^1=\begin{bmatrix}4\\5\\6\end{bmatrix}+2\alpha(3-\sigma(4+5\cdot1+6\cdot2))\begin{bmatrix}\sigma'(4+5\cdot1+6\cdot2)\\1\cdot\sigma'(4+5\cdot1+6\cdot2)\\2\sigma'(4+5\cdot1+6\cdot2)\end{bmatrix}$$

$$=\begin{bmatrix}4\\5\\6\end{bmatrix}+(6\alpha-2\alpha\sigma(21))\cdot\begin{bmatrix}\sigma'(21)\\\sigma'(21)\\2\sigma'(21)\end{bmatrix}$$

$$=\begin{bmatrix}4\\5\\6\end{bmatrix}+\sigma(21)(1-\sigma(21))\begin{bmatrix}1\\2\end{bmatrix}\cdot(6\alpha-2\alpha\sigma(21))$$

$$\frac{e^x}{(e^x+1)^2}$$

2.(a) $\sigma = \frac{1}{1+e^{-x}} \Rightarrow \frac{d\sigma}{dx} = \frac{-(-e^{-x})}{(1+e^{-x})^2} = \frac{e^{-x}}{(1+e^{-x})^2}$

$$= \sigma \cdot \frac{e^{-x}}{1+e^{-x}}$$

$$= \sigma(1-\sigma)$$

$$\frac{d^2\sigma}{dx^2} = \frac{d(\sigma(1-\sigma))}{dx} =$$

$$\frac{d^2\sigma}{dx^2} = \frac{d}{dx}\left(\sigma(1-\sigma)\right) = \frac{d\sigma}{dx}(1-\sigma) + \sigma\frac{d(1-\sigma)}{dx}$$

$$= \sigma(1-\sigma)^2 + \sigma(-\sigma(1-\sigma))$$

$$= \sigma(1-\sigma)^2 - \sigma^2(1-\sigma)$$

$$= \left[(1-\sigma) - \sigma\right]\sigma(1-\sigma)$$

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

$$\frac{d^3\sigma}{dx^3} = \frac{d}{dx}\left(\frac{d^2\sigma}{dx^2}\right) =$$

$$\frac{d^3\sigma}{dx^3} = \frac{d}{dx}\left(\frac{d^2\sigma}{dx^2}\right) = \frac{d(1-2\sigma)}{dx}\sigma(1-\sigma) + (1-2\sigma)\frac{d}{dx}\left(\frac{d\sigma}{dx}\right)$$

$$= -2\sigma(1-\sigma)\sigma(1-\sigma) + (1-2\sigma)(1-2\sigma)\sigma(1-\sigma)$$

$$= -2\sigma^2(1-\sigma)^2 + (1-2\sigma)^2\sigma(1-\sigma)$$

$$= \left[(1+2\sigma)^2 - 2\sigma(1-\sigma)\right]\sigma(1-\sigma)$$

#

(b)

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

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

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

~~$f(x) =$~~

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

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

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

$$\Rightarrow 2\sigma(x) = \tanh\left(\frac{x}{2}\right) + 1$$

$$\sigma(x) = \frac{1}{2}\left(\tanh\left(\frac{x}{2}\right)+1\right) \#$$

3.