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

$= (4,5,6) + 2\alpha\left(3 - h(1,2)\right)\nabla_\theta h$

$= (4,5,6) + 2\alpha\left(3 - \sigma(21)\right)\left(\frac{\partial h}{\partial b}, \frac{\partial h}{\partial w_1}, \frac{\partial h}{\partial w_2}\right)$

$= (4,5,6) + \alpha\left(6 - \frac{2}{1+e^{-21}}\right)\cdot\frac{e^{-21}}{(1+e^{-21})^2}(1,1,2)$

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

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

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

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

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

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