Question 2.

Given that: 
$$\begin{cases} \vec{a}^{(1)} = W^{(1)} \vec{a}^{(0)} + \vec{b}^{(1)} \\ \vec{a}^{(2)} = W^{(2)} \vec{a}^{(1)} + \vec{b}^{(2)} \\ \vec{a}^{(3)} = W^{(3)} \vec{a}^{(2)} + \vec{b}^{(3)} \end{cases}$$

$$\vec{a}^{(3)} = W^{(3)} \vec{a}^{(2)} + \vec{b}^{(3)}$$

$$= W^{(3)} ( W^{(2)} \vec{a}^{(1)} + \vec{b}^{(2)} ) + \vec{b}^{(3)}$$

$$= W^{(3)} [ W^{(2)} ( W^{(1)} \vec{a}^{(0)} + \vec{b}^{(1)} ) + \vec{b}^{(2)} ] + \vec{b}^{(3)}$$

thus, $\vec{a}^{(3)} = W^{(3)} W^{(2)} W^{(1)} \vec{a}^{(0)}$

$$+ W^{(3)} W^{(2)} \vec{b}^{(1)} + W^{(3)} \vec{b}^{(2)} + \vec{b}^{(3)}$$

for $\vec{a}^{(3)} \approx \hat{W} \vec{a}^{(0)} + \hat{b}$

$$\begin{cases} \hat{W} \approx W^{(3)} W^{(2)} W^{(1)} \\ \hat{b} \approx W^{(3)} W^{(2)} \vec{b}^{(1)} + W^{(3)} \vec{b}^{(2)} + \vec{b}^{(3)} \end{cases}$$