2.2 Formulating equivalent networks

In network 1:

$\vec{a}^{(1)} = W^{(1)}\vec{a}^{(0)} + \vec{b}^{(1)}$ ①

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

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

In network 2,

if network1 and network2 ~~are~~ equivalent:

$\vec{a}^{(3)}$ is output and $\vec{a}^{(0)}$ is input for network2.

so, $\vec{a}^{(3)} = \tilde{W}\vec{a}^{(0)} + \tilde{b}$

also
$$
\begin{aligned}
\vec{a}^{(3)} &= W^{(3)}\vec{a}^{(2)} + \vec{b}^{(3)} \\
&= W^{(3)}\left(W^{(2)}\vec{a}^{(1)} + \vec{b}^{(2)}\right) + \vec{b}^{(3)} \\
&= W^{(3)}\left[W^{(2)}\left(W^{(1)}\vec{a}^{(0)} + \vec{b}^{(1)}\right) + \vec{b}^{(2)}\right] + \vec{b}^{(3)} \\
&= W^{(3)}W^{(2)}W^{(1)}\vec{a}^{(0)} + W^{(3)}W^{(2)}\vec{b}^{(1)} + W^{(3)}\vec{b}^{(2)} + \vec{b}^{(3)}
\end{aligned}
$$

$\therefore$ $\tilde{W} = W^{(3)} \cdot W^{(2)} \cdot W^{(1)}$

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