2. From network 1 in Figure 2a, we can give:

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

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

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

For the network 2 in Figure 2b:

If network 1 and network 2 are said to be equivalent, $a^{(0)}$ is the input of network 2 and $a^{(3)}$ is the output of network 2

Then, we can give:

$$\vec{a}^{(3)} = \tilde{w}\, \vec{a}^{(0)} + \tilde{b} \quad ④$$

From Eq. ① ② ③, we can give:

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

contrast with Eq. ④, we can get:

$$\tilde{w} = w^{(3)} w^{(2)} w^{(1)}$$

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