2.2

Propose that each layer has n input nodes.

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

$$\vec{a}^{(i)} = \begin{pmatrix} a_0^{(i)} & a_1^{(i)} & \dots & a_{n-1}^{(i)} \end{pmatrix}^T$$

$$b^{(i)} = \begin{pmatrix} b_0^{(i)} & b_1^{(i)} & \dots & b_{n-1}^{(i)} \end{pmatrix}^T$$

$$W^{(i)} = \begin{pmatrix} w_{0,0}^{(i)} & w_{0,1}^{(i)} & \cdots & w_{0,n-1}^{(i)} \\ w_{1,0}^{(i)} & w_{1,1}^{(i)} & \dots & w_{1,n-1}^{(i)} \\ \vdots & \vdots & \ddots & \vdots \\ w_{n-1,0}^{(i)} & w_{n-1,1}^{(i)} & \cdots & w_{n-1,n-1}^{(i)} \end{pmatrix}$$

$$i = 1,2,\dots,\mathrm{n}-1$$

In network 1:

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

In network 2:

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

$$\because \vec{a}^{(3)} = \vec{a}$$

$$\therefore W^{(3)}W^{(2)}W^{(1)}\vec{a}^{(0)} + W^{(3)}W^{(2)}\vec{b}^{(1)} + W^{(3)}\vec{b}^{(2)} + \vec{b}^{(3)} = \widetilde{W}\vec{a}^{(0)} + \tilde{b}$$

$$\therefore \begin{cases} \widetilde{W} = W^{(3)}W^{(2)}W^{(1)} \\ \tilde{b} = W^{(3)}W^{(2)}\vec{b}^{(1)} + W^{(3)}\vec{b}^{(2)} + \vec{b}^{(3)} \end{cases}$$

$$\therefore \begin{cases} \widetilde{W} = \begin{pmatrix} w_{0,0}^{(3)} & w_{0,1}^{(3)} & \cdots & w_{0,n-1}^{(3)} \\ w_{1,0}^{(3)} & w_{1,1}^{(3)} & \dots & w_{1,n-1}^{(3)} \\ \vdots & \vdots & \ddots & \vdots \\ w_{n-1,0}^{(3)} & w_{n-1,1}^{(3)} & \cdots & w_{n-1,n-1}^{(3)} \end{pmatrix} \begin{pmatrix} w_{0,0}^{(2)} & w_{0,1}^{(2)} & \cdots & w_{0,n-1}^{(2)} \\ w_{1,0}^{(2)} & w_{1,1}^{(2)} & \dots & w_{1,n-1}^{(2)} \\ \vdots & \vdots & \ddots & \vdots \\ w_{n-1,0}^{(2)} & w_{n-1,1}^{(2)} & \cdots & w_{n-1,n-1}^{(2)} \end{pmatrix} \begin{pmatrix} w_{0,0}^{(1)} & w_{0,1}^{(1)} & \cdots & w_{0,n-1}^{(1)} \\ w_{1,0}^{(1)} & w_{1,1}^{(1)} & \dots & w_{1,n-1}^{(1)} \\ \vdots & \vdots & \ddots & \vdots \\ w_{n-1,0}^{(1)} & w_{n-1,1}^{(1)} & \cdots & w_{n-1,n-1}^{(1)} \end{pmatrix} \\ \tilde{b} = \begin{pmatrix} w_{0,0}^{(3)} & w_{0,1}^{(3)} & \cdots & w_{0,n-1}^{(3)} \\ w_{1,0}^{(3)} & w_{1,1}^{(3)} & \dots & w_{1,n-1}^{(3)} \\ \vdots & \vdots & \ddots & \vdots \\ w_{n-1,0}^{(3)} & w_{n-1,1}^{(3)} & \cdots & w_{n-1,n-1}^{(3)} \end{pmatrix} \begin{pmatrix} w_{0,0}^{(2)} & w_{0,1}^{(2)} & \cdots & w_{0,n-1}^{(2)} \\ w_{1,0}^{(2)} & w_{1,1}^{(2)} & \dots & w_{1,n-1}^{(2)} \\ \vdots & \vdots & \ddots & \vdots \\ w_{n-1,0}^{(2)} & w_{n-1,1}^{(2)} & \cdots & w_{n-1,n-1}^{(2)} \end{pmatrix} \begin{pmatrix} b_0^{(1)} \\ b_0^{(1)} \\ \vdots \\ b_0^{(1)} \end{pmatrix} \\ \qquad + \begin{pmatrix} w_{0,0}^{(3)} & w_{0,1}^{(3)} & \cdots & w_{0,n-1}^{(3)} \\ w_{1,0}^{(3)} & w_{1,1}^{(3)} & \dots & w_{1,n-1}^{(3)} \\ \vdots & \vdots & \ddots & \vdots \\ w_{n-1,0}^{(3)} & w_{n-1,1}^{(3)} & \cdots & w_{n-1,n-1}^{(3)} \end{pmatrix} \begin{pmatrix} b_0^{(2)} \\ b_0^{(2)} \\ \vdots \\ b_0^{(2)} \end{pmatrix} + \begin{pmatrix} b_0^{(3)} \\ b_0^{(3)} \\ \vdots \\ b_0^{(3)} \end{pmatrix} \end{cases}$$