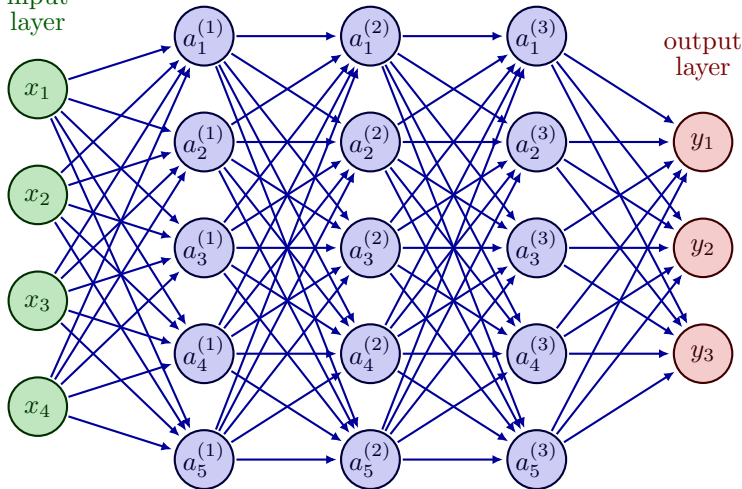



input  
layer

hidden layers

output  
layer



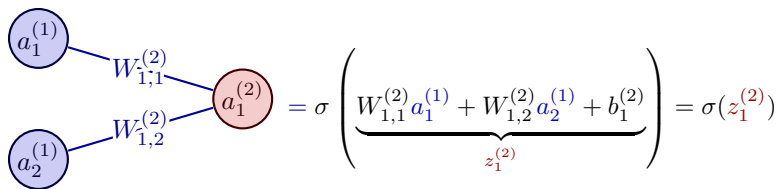


A diagram showing a green input node labeled  $a_1^{(0)}$  connected to two blue output nodes labeled  $a_1^{(1)}$  and  $a_2^{(1)}$ . The connection to  $a_1^{(1)}$  is labeled  $W_{1,1}$  and the connection to  $a_2^{(1)}$  is labeled  $W_{2,1}$ .

$$\begin{aligned}
 &= \sigma \left( \underbrace{W_{1,1}^{(1)} a_1^{(0)} + b_1^{(1)}}_{z_1^{(1)}} \right) = \sigma(z_1^{(1)}) \\
 &= \sigma \left( \underbrace{W_{2,1}^{(1)} a_1^{(0)} + b_2^{(1)}}_{z_2^{(1)}} \right) = \sigma(z_2^{(1)})
 \end{aligned}$$

$$\begin{bmatrix} a_1^{(1)} \\ a_2^{(1)} \end{bmatrix} = \sigma \left( \begin{bmatrix} W_{1,1}^{(1)} \\ W_{2,1}^{(1)} \end{bmatrix} \begin{bmatrix} a_1^{(0)} \end{bmatrix} + \begin{bmatrix} b_1^{(1)} \\ b_2^{(1)} \end{bmatrix} \right) = \sigma \left( \begin{bmatrix} z_1^{(1)} \\ z_2^{(1)} \end{bmatrix} \right)$$

$$\mathbf{a}^{(1)} = \sigma \left( \mathbf{W}^{(1)} \cdot \mathbf{a}^{(0)} + \mathbf{b}^{(1)} \right) = \sigma(\mathbf{z}^{(1)})$$



$$\begin{bmatrix} a_1^{(2)} \end{bmatrix} = \sigma \left( \begin{bmatrix} W_{1,1}^{(2)} & W_{1,2}^{(2)} \end{bmatrix} \begin{bmatrix} a_1^{(1)} \\ a_2^{(1)} \end{bmatrix} + \begin{bmatrix} b_1^{(2)} \end{bmatrix} \right) = \sigma \left( \begin{bmatrix} z_1^{(2)} \end{bmatrix} \right)$$

$$\mathbf{a}^{(2)} = \sigma \left( \mathbf{W}^{(2)} \cdot \mathbf{a}^{(1)} + \mathbf{b}^{(2)} \right) = \sigma(\mathbf{z}^{(2)})$$

$$\mathbf{a}^{(2)} = \sigma \left( W_{1,1}^{(2)} \sigma \left( W_{1,1}^{(1)} \mathbf{a}_1^{(0)} + b_1^{(1)} \right) + W_{1,2}^{(2)} \sigma \left( W_{1,1}^{(1)} \mathbf{a}_1^{(0)} + b_1^{(1)} \right) + b_1^{(2)} \right)$$