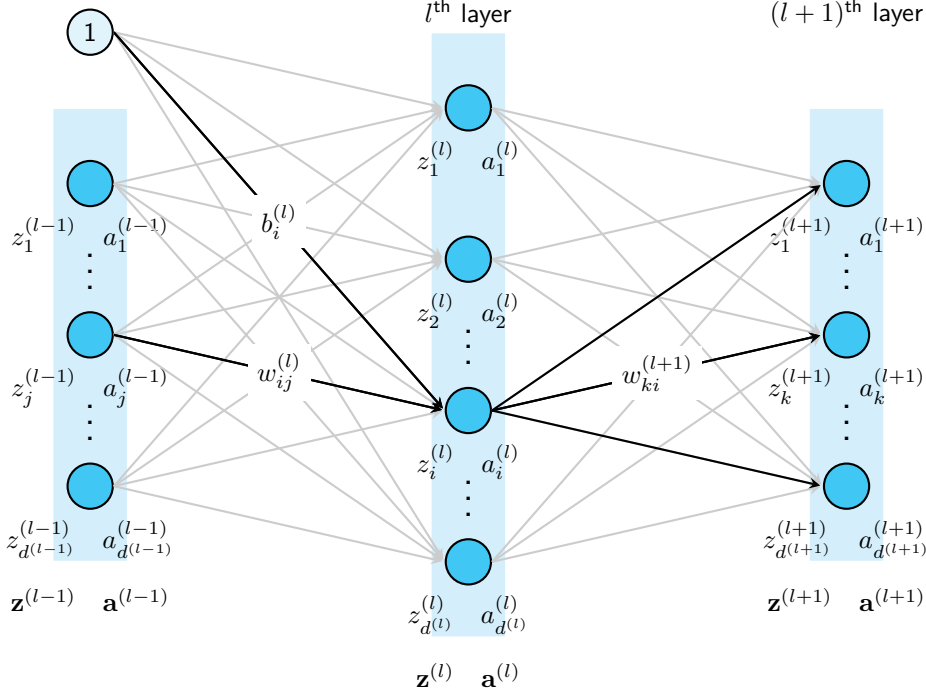


$(l-1)^{\text{th}}$ layer

l^{th} layer

$(l+1)^{\text{th}}$ layer



$$\mathbf{W}^{(l)} \in \mathbb{R}^{d^{(l)} \times d^{(l-1)}}$$

$$\mathbf{b}^{(l)} \in \mathbb{R}^{d^{(l)} \times 1}$$

$$z_i^{(l)} = \mathbf{w}_i^{(l)T} \mathbf{a}^{(l-1)} + b_i^{(l)}$$

$$\mathbf{z}^{(l)} = \mathbf{W}^{(l)T} \mathbf{a}^{(l-1)} + \mathbf{b}^{(l)}$$

$$\mathbf{a}^{(l)} = f(\mathbf{z}^{(l)})$$