

2-5 Neural Network with Multiple Hidden Layers

Zhonglei Wang

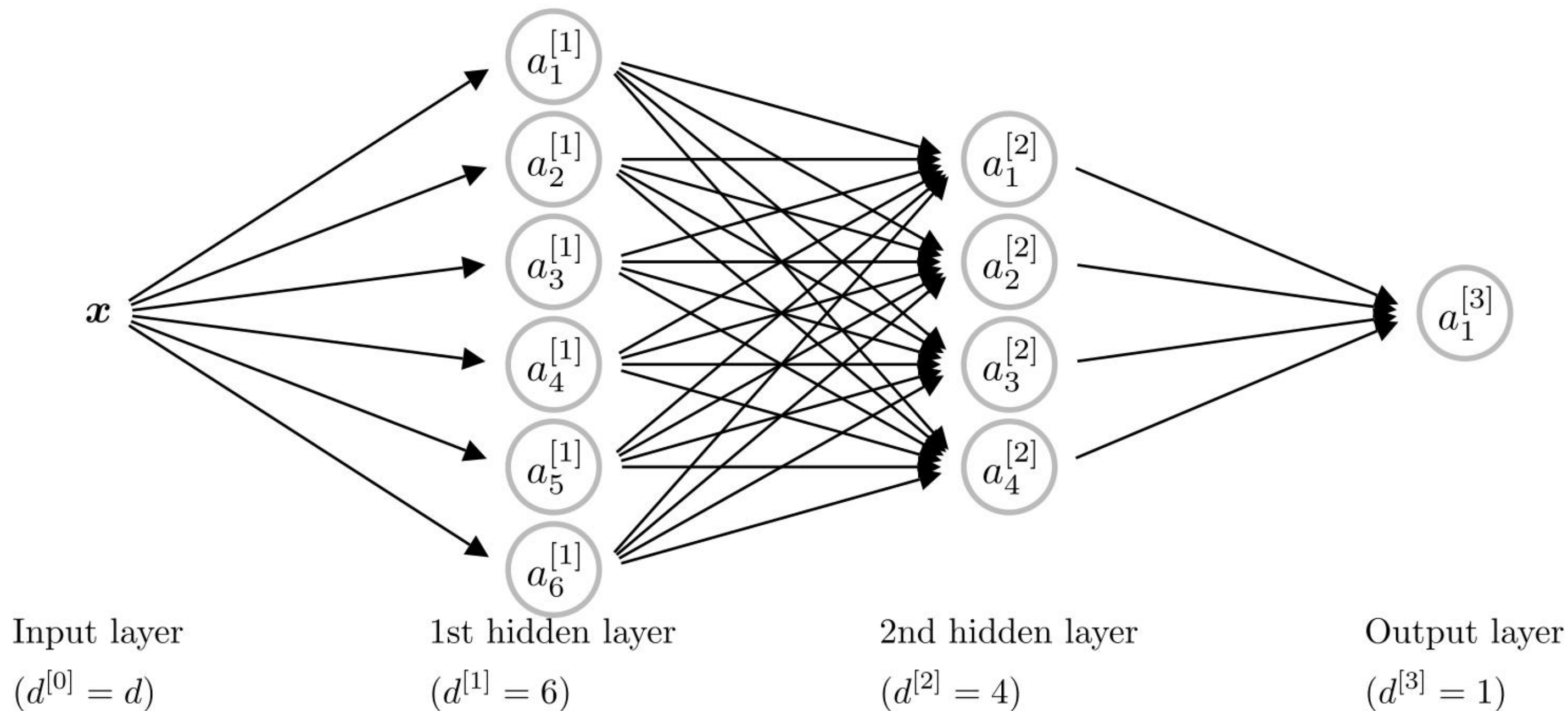
WISE and SOE, XMU, 2025

Contents

1. Forward propagation

2. Backpropagation

Example



Forward propagation

1. Let $\mathbf{A}^{[0]} = \mathbf{X}$
2. For $l = 1, \dots, L$,

$$\mathbf{Z}^{[l]} = \left(\mathbf{b}^{[l]}\right)^{\mathrm{T}} + \mathbf{A}^{[l-1]} \left(\mathbf{W}^{[l]}\right)^{\mathrm{T}}$$
$$\mathbf{A}^{[l]} = \sigma^{[l]} \left(\mathbf{Z}^{[l]}\right)$$

- $\sigma^{[l]}(z)$: activation function for the l th layer

Backpropagation

1. $\mathbf{dA}^{[L]}$ can be obtained from the cost function
2. Assume $\mathbf{dA}^{[l]}$ is available ($l = L, \dots, 2$)

$$\mathbf{dZ}^{[l]} = \mathbf{dA}^{[l]} \circ \sigma^{[l]'}(\mathbf{Z}^{[l]})$$

$$\mathbf{dW}^{[l]} = \left(\mathbf{dZ}^{[l]}\right)^T \mathbf{dA}^{[l-1]}$$

$$\mathbf{db}^{[l]} = \left(\mathbf{dZ}^{[l]}\right)^T \mathbf{1}$$

$$\mathbf{dA}^{[l-1]} = \mathbf{dZ}^{[l]} \mathbf{W}^{[l]}$$