2-5 Neural Network with Multiple Hidden Layers

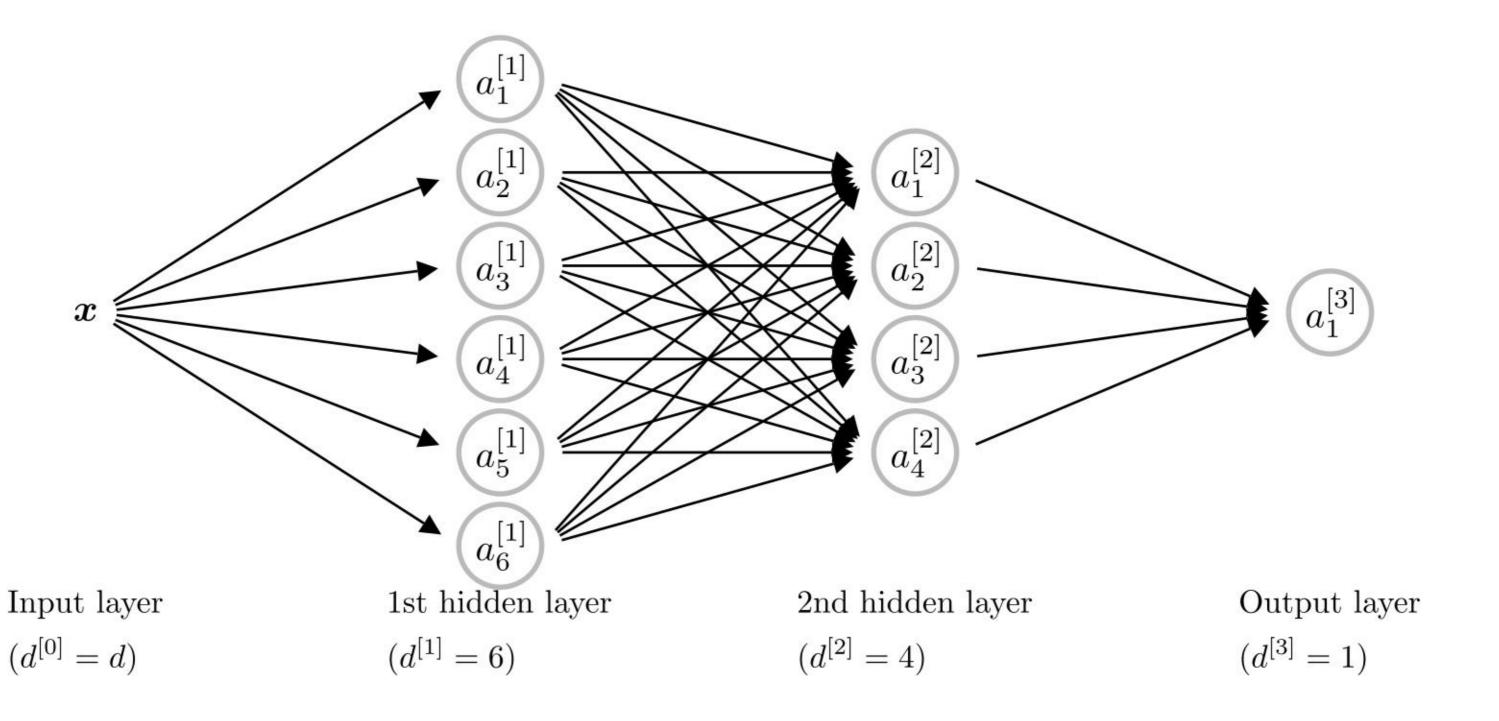
Zhonglei Wang
WISE and SOE, XMU, 2025

Contents

1. Forward propagation

2. Backpropagation

Example



Forward propagation

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

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

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

Backpropagation

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

$$\mathrm{d}\boldsymbol{Z}^{[l]} = \mathrm{d}\boldsymbol{A}^{[l]} \circ \sigma^{[l]'} \left(\boldsymbol{Z}^{[l]}\right)$$

$$\mathrm{d} oldsymbol{W}^{[l]} = \left(\mathrm{d} oldsymbol{Z}^{[l]}
ight)^{\mathrm{T}} \mathrm{d} oldsymbol{A}^{[l-1]}$$

$$\mathrm{d}m{b}^{[l]} = \left(\mathrm{d}m{Z}^{[l]}
ight)^{\mathrm{T}} \mathbf{1}$$

$$\mathbf{d}\boldsymbol{A}^{[l-1]} = \mathbf{d}\boldsymbol{Z}^{[l]}\boldsymbol{W}^{[l]}$$