

2-5 具有多个隐藏层的神经网络

王中雷

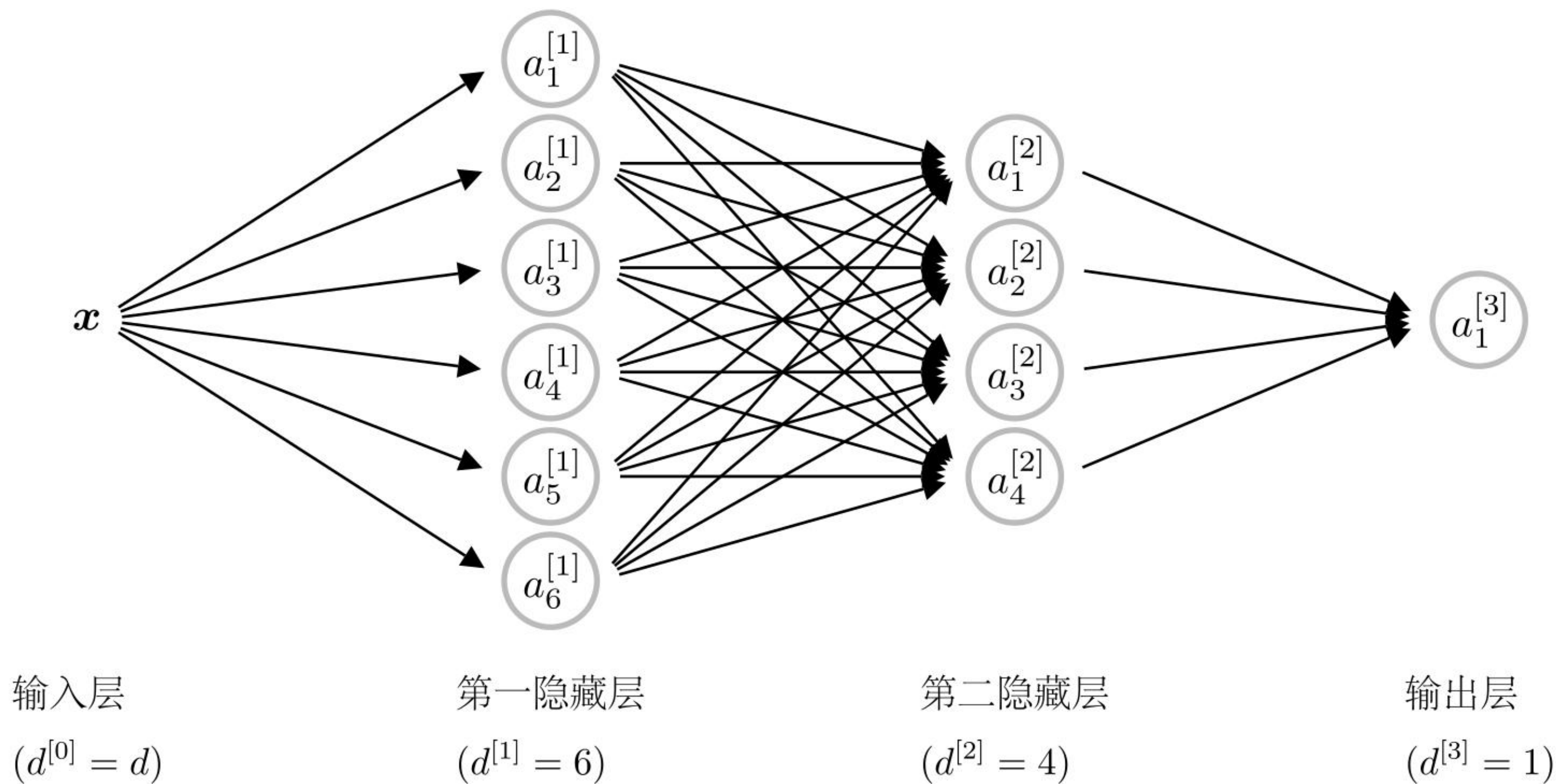
厦门大学王亚南经济研究院和经济学院, 2025

内容摘要

1. 前向传播

2. 后向传播

例子



前向传播

1. 记 $\mathbf{A}^{[0]} = \mathbf{X}$

2. 对于 $l = 1, \dots, L$,

$$\begin{aligned}\mathbf{Z}^{[l]} &= (\mathbf{b}^{[l]})^T + \mathbf{A}^{[l-1]} (\mathbf{W}^{[l]})^T \\ \mathbf{A}^{[l]} &= \sigma^{[l]}(\mathbf{Z}^{[l]})\end{aligned}$$

- $\sigma^{[l]}(z)$: 第 l 层的激活函数
- Python 中的广播机制用于计算每个维度的激活值

后向传播

1. $\mathbf{dA}^{[L]}$ 可以通过代价函数计算得到（默认给定）

2. 假设 $\mathbf{dA}^{[l]}$ 已知 $(l = L, \dots, 2)$

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

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

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

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