

# 深度学习汇报

汇报人：乔栋

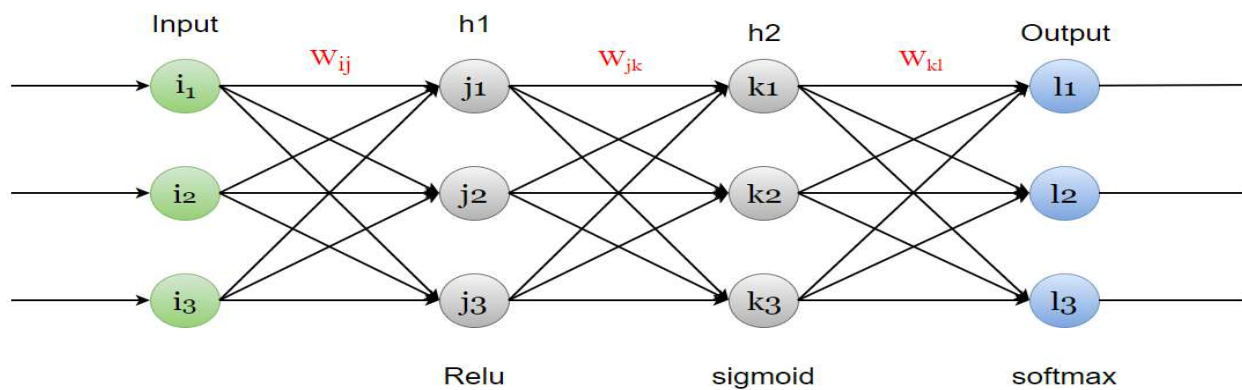
2018 / 04





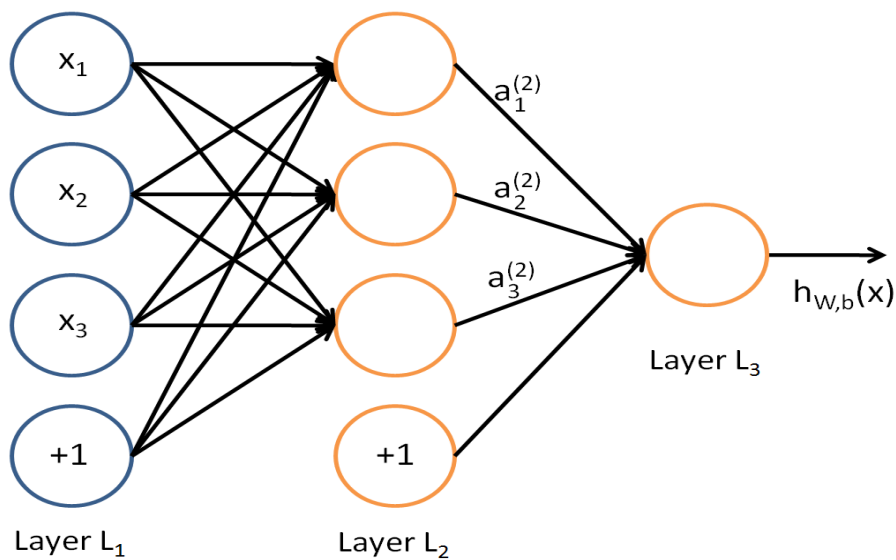
## 前向 / 反向传播算法

前向传播是为预测；  
反向传播是为更新权值。





### 前向传播算法



$$a_j^l = \sigma(z_j^l) = \sigma\left(\sum_{k=1}^m w_{jk}^l a_k^{l-1} + b_j^l\right)$$



中国地质大学  
China University of Geosciences

艰苦朴素 求真务实

## 反向传播算法

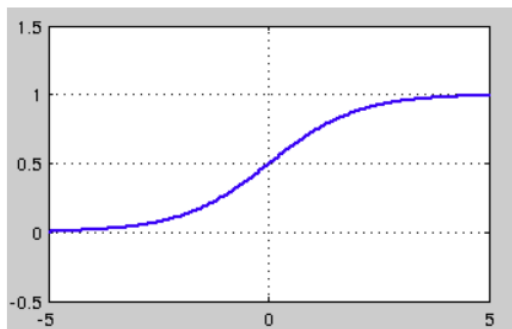
反向传播算法：梯度下降 + 偏导(链式法则)。反向传播算法的方向是从输出层到输入层，层层更新 $w$ 权值，更新的过程就是利用偏导求得每一个权值方向上的梯度，然后更新参数。

在实际计算时，我们利用计算图来进行。



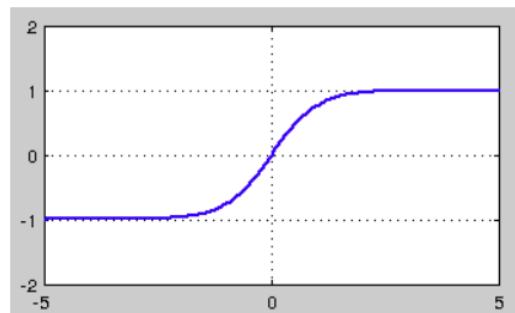
## 激活函数

$$f(x) = \frac{1}{1 + e^{-x}}$$



sigmoid

$$f(x) = \frac{1 - e^{-2x}}{1 + e^{-2x}}$$

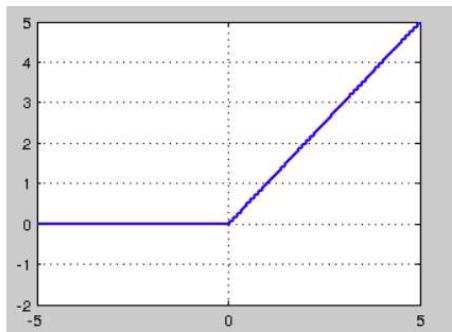


tanh



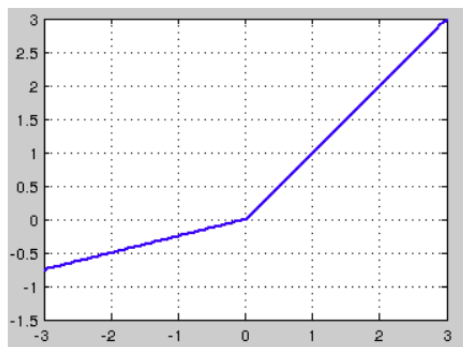
## 激活函数

$$f(x) = \begin{cases} x, & \text{if } x \geq 0 \\ 0, & \text{if } x < 0 \end{cases}$$
$$f(x) = \max(0, x)$$



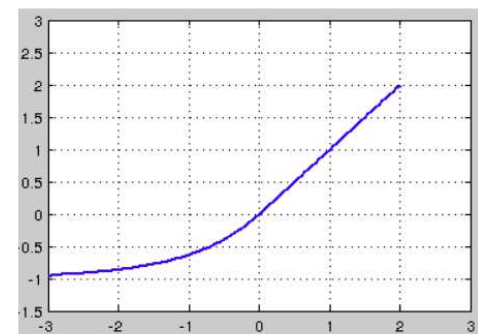
ReLU

$$f(x) = \begin{cases} x, & \text{if } x \geq 0 \\ \alpha x, & \text{if } x < 0 \end{cases}$$



Leaky-ReLU

$$f(x) = \begin{cases} x, & \text{if } x \geq 0 \\ \alpha(e^x - 1), & \text{if } x < 0 \end{cases}$$



ELU



中国地质大学  
China University of Geosciences

艰苦朴素 求真务实

1952

## 激活函数

- 尽量使用ReLU，的同时，注意学习率
- 尝试使用Leaky-ReLU / MaxOut / ELU
- 尝试使用tanh
- 不要使用sigmoid



## 损失函数

平方损失函数:

$$J(w, b) = \frac{1}{2m} \sum_{i=1}^m (Y - f(X))^2$$

交叉熵损失函数:

$$J(w, b) = \frac{1}{m} \sum_{i=1}^m L(\hat{y}^{(i)}, y^{(i)}) = -\frac{1}{m} \sum_{i=1}^m [y^{(i)} \log \hat{y}^{(i)} + (1 - y^{(i)}) \log (1 - \hat{y}^{(i)})]$$





## 参数调整

数据预处理相关参数:

- 数据集增强
- 数据标准化（归一化，Normalization）
- BN处理

训练相关参数:

- BGD, SGD
- 学习的次数
- 学习率（learning rate）
- 参数初初始化
- 正则化

网络相关参数:

- 网络层数
- 每层结点数



## 参数调整措施

欠拟合:

- 增加网络的复杂度（深度）
- 降低learning rate
- 优化数据集
- 增加网络的非线性度（ReLU）
- 采用batch normalization

过拟合:

- 丰富数据
- 增加网络的稀疏度
- 降低网络的复杂度（深度）
- L1 regularization
- L2 regularization
- 添加Dropout
- Early stopping
- 适当降低Learning rate
- 适当减少epoch的次数



中國地質大學  
China University of Geosciences

艱苦樸素 求真務實

地質大學

## 数据预处理

- 減均值
- 数据标准化
- PCA和白化



中國地質大學  
China University of Geosciences

艱苦樸素 求真務實

地質大學

## 样本标注工具

DataTurks / Pixorize / **Labelbox** / AnnoStation /  
Images Annotation Programme