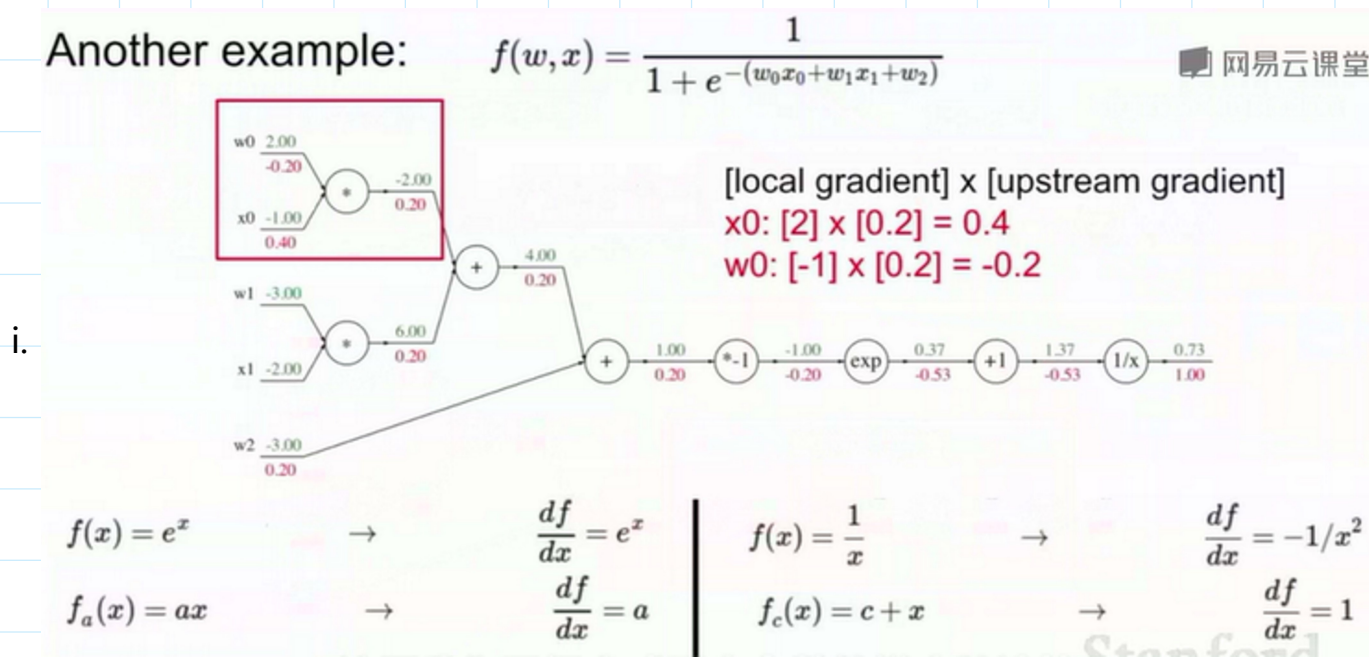


L4 反向传播和神经网络

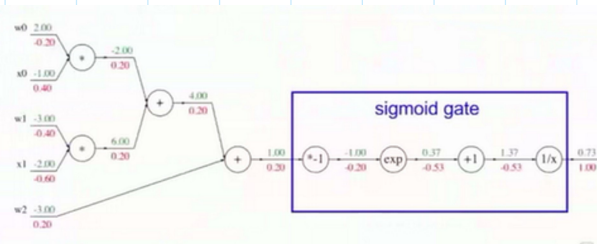
2020年3月29日 15:50

1. 反向传播

a. 计算图

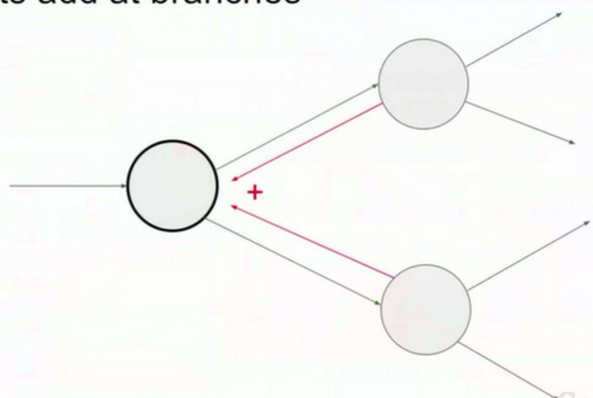


ii. 可以将多个计算节点整合为一个e.g.:



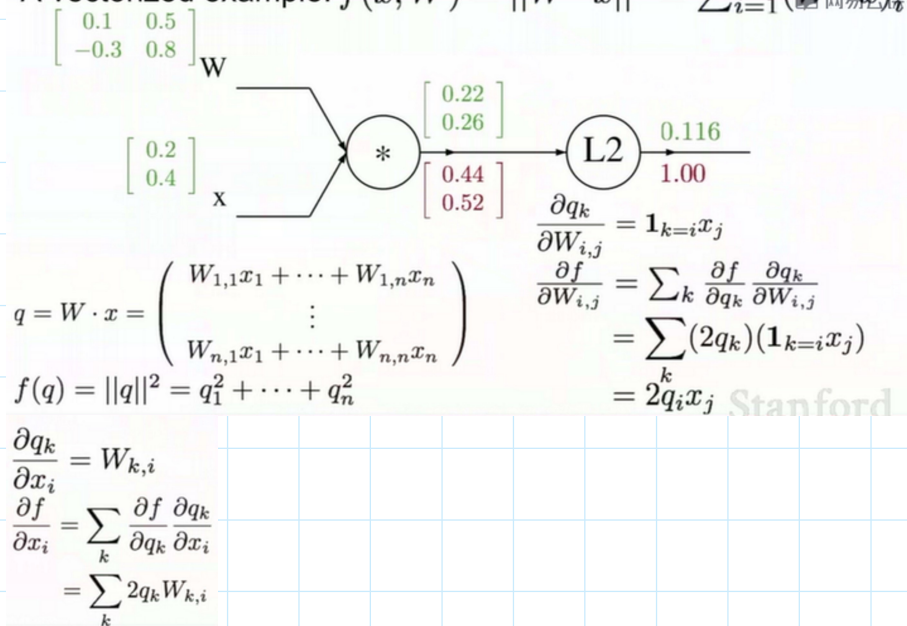
iii. 梯度回流

Gradients add at branches



iv. 向量梯度

A vectorized example: $f(x, W) = \|W \cdot x\|^2 = \sum_{i=1}^n (W_{i,1}x_1 + \dots + W_{i,n}x_n)^2$



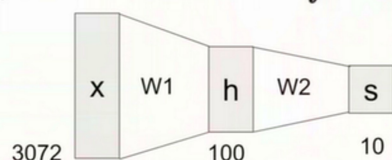
其中 $1_{k=i}$ 为指示函数

- b. 完整性检查：梯度向量应和变量大小相同
(同型矩阵)

2. 神经网络

- a. 定义：有简单函数构成的在顶层堆叠在一起的一组函数（多阶段分层计算）

(Now) 2-layer Neural Network $f = W_2 \max(0, W_1 x)$



or 3-layer Neural Network

$$f = W_3 \max(0, W_2 \max(0, W_1 x))$$

可以堆很多层（深度神经网络）

h即

$$\max(0, W_1 x)$$

- b. 常见激活函数：

Activation functions

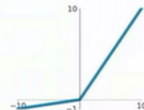
Sigmoid

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



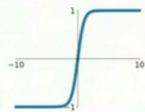
Leaky ReLU

$$\max(0.1x, x)$$



tanh

$$\tanh(x)$$



Maxout

$$\max(w_1^T x + b_1, w_2^T x + b_2)$$

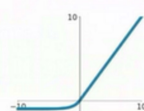
ReLU

$$\max(0, x)$$



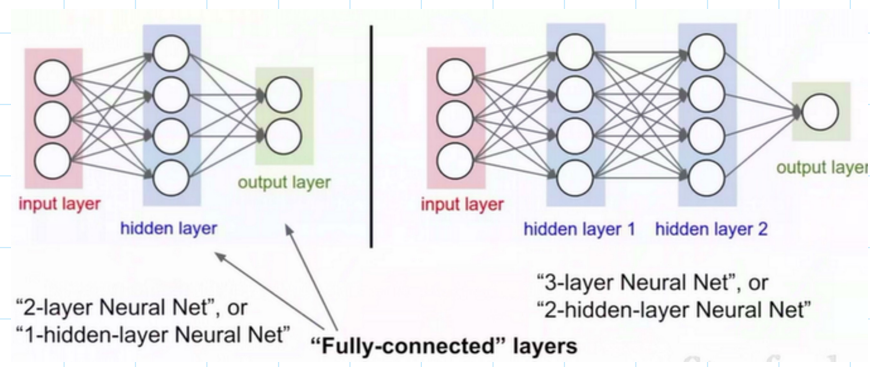
ELU

$$\begin{cases} x & x \geq 0 \\ \alpha(e^x - 1) & x < 0 \end{cases}$$



- c. 命名：x层神经网络，x-1层隐藏层神经网络

络



全连接层的每一个结点都与上一层的所有结点相连，用来把前边提取到的特征综合起来。