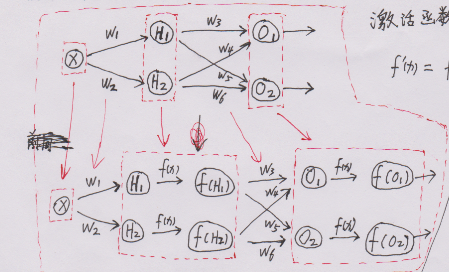
# 反向传播

### 前向传播



### 误差



### 反向传播





### 更新



### 特殊环节的反向传播

* **ReLU**：



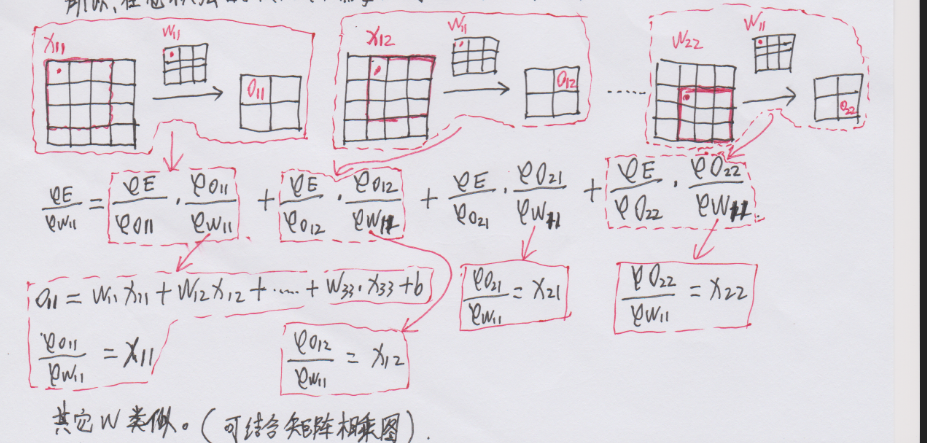
ReLU在x=0处不可微，直接将其在x=0处的导数置为1，



* **Pooling：**

把一个梯度传递给nxn个像素，需要保证梯度传递的总和不变；

* **卷积层**

在反向传播过程中，若第L层的a节点通过权值w对L+1层的b节点有贡献，则在方向过程中，梯度通过权值w从b节点传播回a节点。都遵循这个规律，所以，在卷积层的反向中，需要找到卷积层L中的每个单元和L+1层中哪些单元相关联。

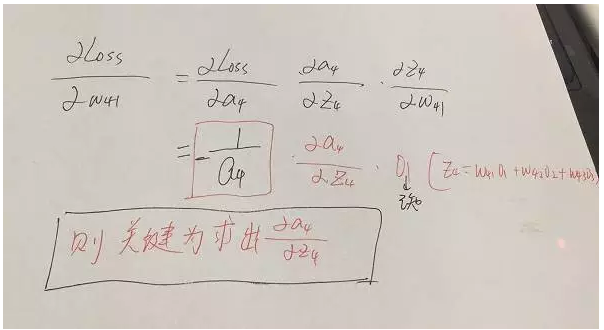
### 核心思想

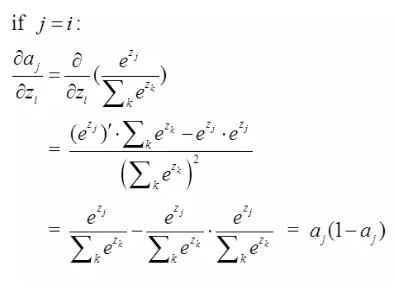
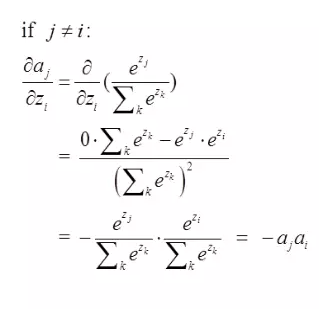
链式求导法则，动态规划

### Softmax反向传播

<https://mp.weixin.qq.com/s/MS8h8BUv1BC3Ql9w2oxmJg>

* 因为只有一个y=1，其他都为0，所以这里导数可以写成下面形式。



将这三个式子组合得到神奇的效果是：

如果输出是：**[0.0903,0.2447,0.665]**

**那么梯度就是[0.0903,0.2447-1,0.665]=[0.0903,-0.7553,0.665]**