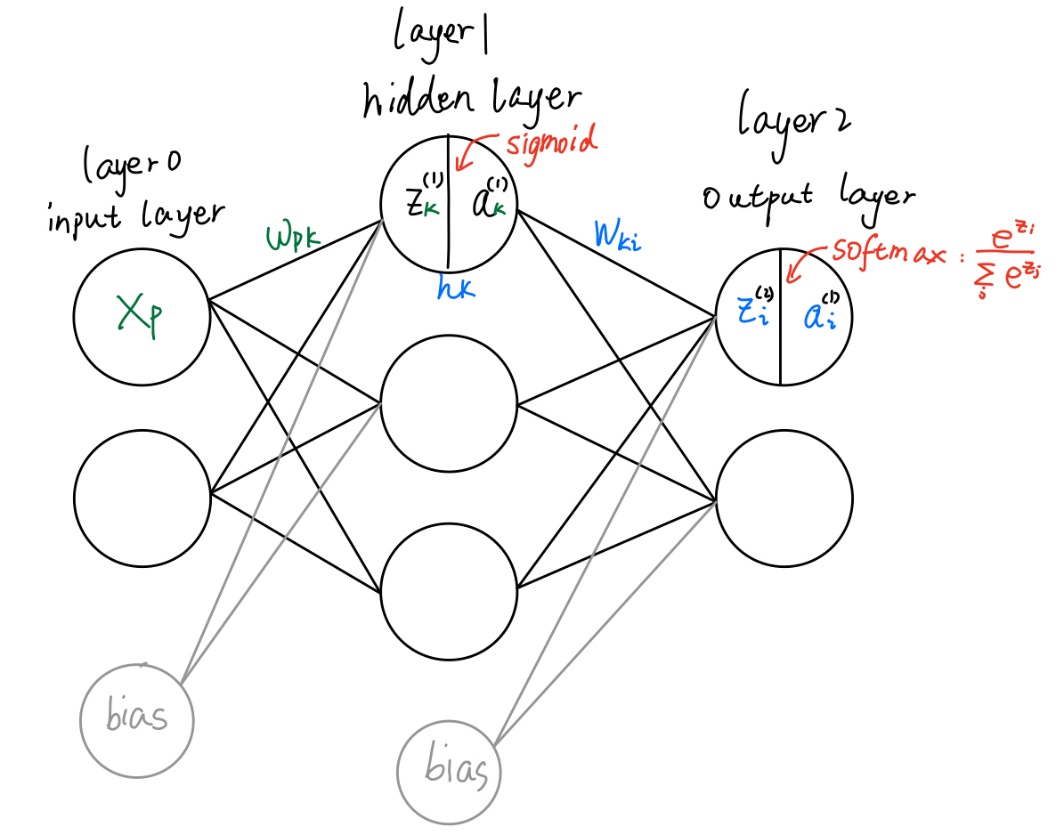
# softmax的推导

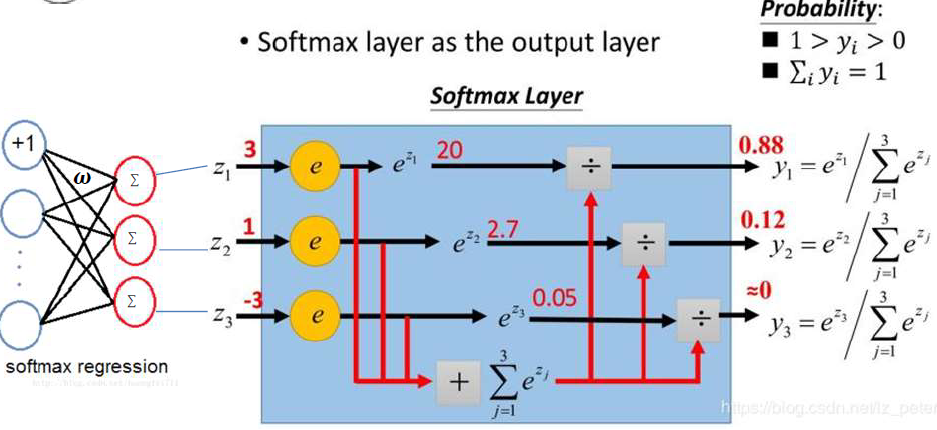
## 定义

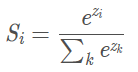
设神经元输出为，权重为ω，偏置为b，真实样本标签为y（onehot编码），

神经网络如图所示

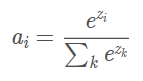


softmax层展开如下：

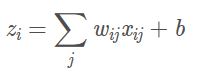


（1）softmax函数的公式：

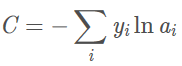
神经元的输出为：



其中z为：

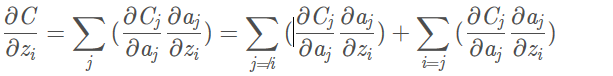


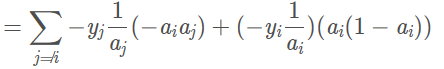
（2）交叉熵损失函数：

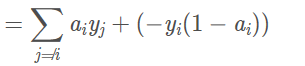


## 求解过程：

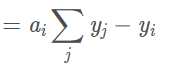
（1）求解：





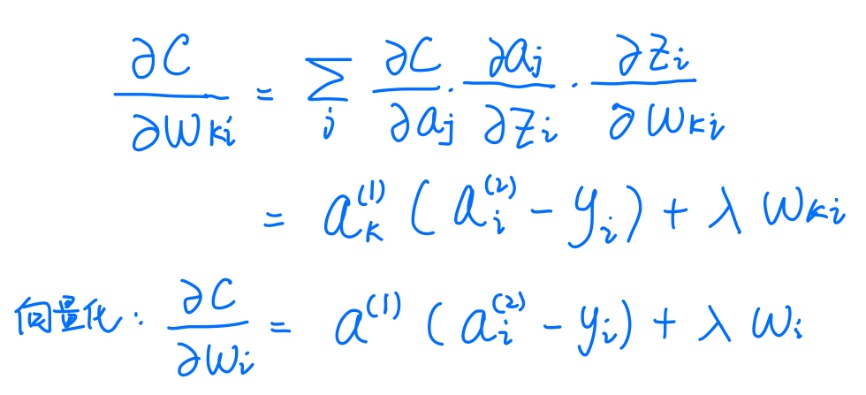




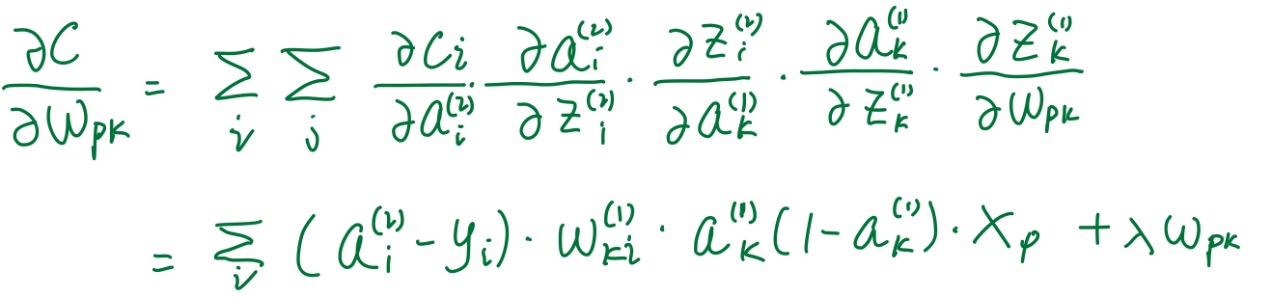




（2）求C对隐层到输出层的权重的偏导



（3）求C对输入层到隐层的权重的偏导



（4）根据 ω = ω + Δω

Δω = η \* ∂*C/*∂ω​

更新参数

参考

简单易懂的softmax交叉熵损失函数求导https://blog.csdn.net/qian99/article/details/78046329