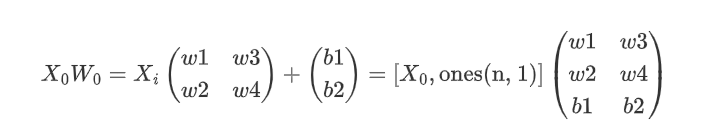
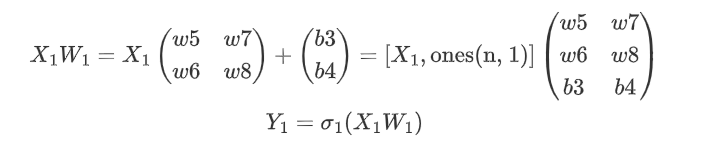
为简化描述，我们用一个三元组来描述第层layer，分表表示in\_feature、out\_feature、bias，使用表示每层的输入inputs[i]，表示每层的输出outputs[i]。

## 前向传播

### 原理

正向传播过程十分简单，假设为样本数，输入为的输入矩阵，那么正向传播过程为：

1. 首先经过第0层的权重矩阵：
2. 再经过激活函数得到第0层的输出也就是第1层的输入：
3. 第1层再经过同样的步骤得到最终网络的输出：
4. 以此类推。

### 实现细节

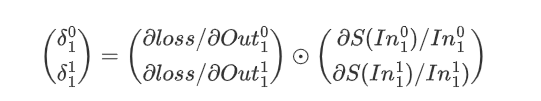
设样本数为，则开始有为的网络输入，则有：

显然后面的层也是一样的：

## 反向传播

### 原理

我们用表示第层的反向传播中间变量delta向量（向量里面也从0开始数，表示第一个元素），用表示激活函数，表示网络输入向量，网络输入输出表示为，即 ：

1. 首先计算输出层（我们这里就是第1层）delta向量：
2. 然后计算隐藏层（我们这里就是第0层）delta向量：
3. 有了delta向量后就能轻而易举地算出梯度进而算出权重变化值：
4. 显然，对于超过2层的网络，中间层数对应的delta计算和权重更新如下：