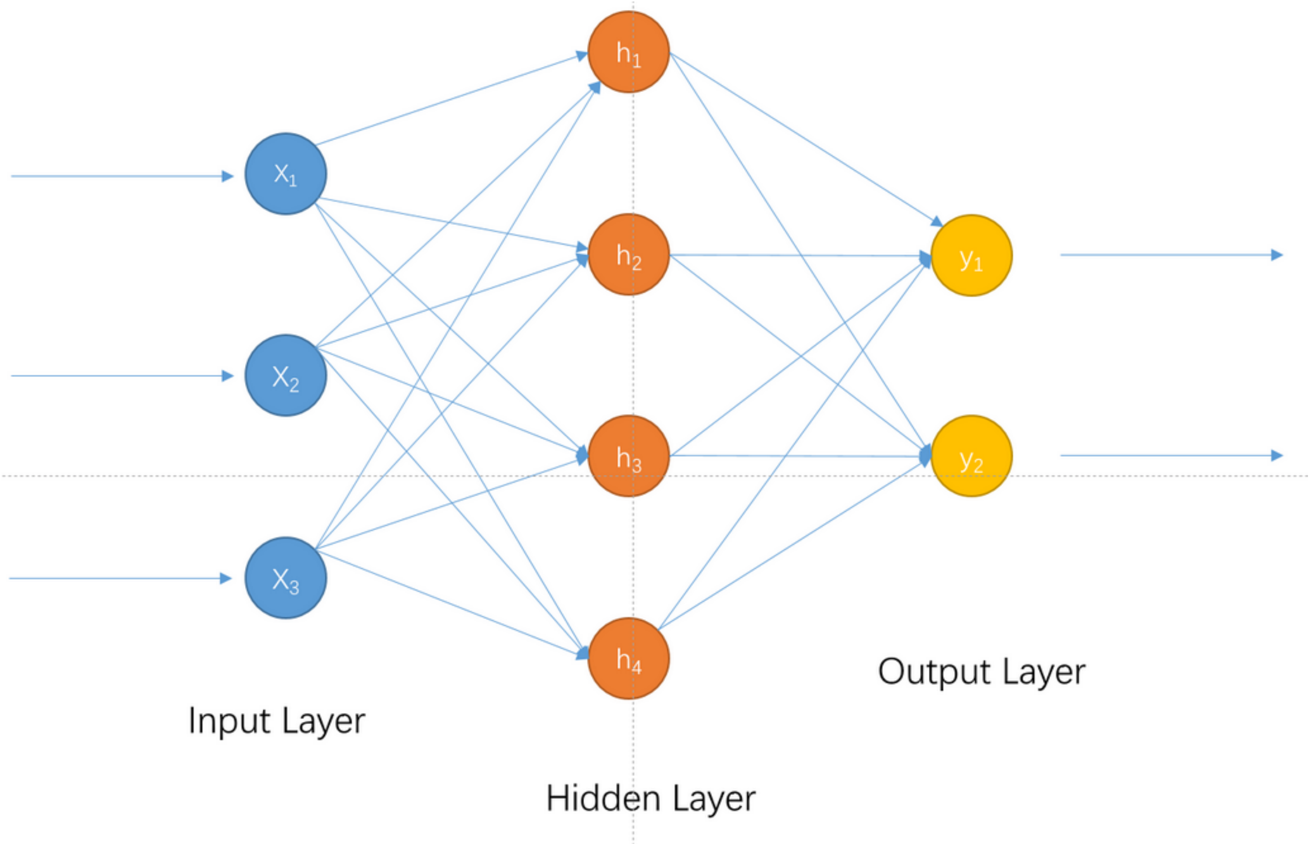
Author：Martinlwx

E-mail：Martinlwx@163.com

推导过程



如图中所示我们规定W_{ij}^k表示第k层与第k+1层之间神经元的权重值，z^k表示对应输入第k层的值，o^k表示对应第k层输出的值

定义**损失函数**为：

$$l = \frac{1}{2}(y - \hat{y})^2 (y\text{表示预期输出})$$

比如图中**第2层和第3层**的**误差分配**如下：

$$l_{23} = \begin{pmatrix} w_{11} & w_{12} \\ w_{21} & w_{22} \\ w_{31} & w_{32} \\ w_{41} & w_{42} \end{pmatrix} \begin{pmatrix} l_1 \\ l_2 \end{pmatrix} \text{（省去}w\text{的上角标）}$$

我们可以发现：**此处的权重矩阵就是前向传播的时候第2层所乘的矩阵的转置矩阵，也就是**$(w^k)^T$

引出**记号**：

$$\frac{\partial L}{\partial W^k} \text{（表示损失函数的值是如何根据权重矩阵变化的）}$$

则**损失函数**改写为：

$$L = \sum_{i=1}^{n} \frac{1}{2}(y - \hat{y})^2$$

暂时把$\sum$忽略可得：

$$\frac{\partial L}{\partial W^k} = \frac{\partial}{\partial W^k} \frac{1}{2}(y - \hat{y})^2$$

又根据**链式法则**：

$$\frac{\partial L}{\partial W^k} = \frac{\partial L}{\partial \hat{y}} \cdot \frac{\partial \hat{y}}{\partial z^{k+1}} \cdot \frac{\partial z^{k+1}}{\partial W^k}$$（y_i为激活函数）

如果**取激活函数为sigmoid函数**可得其导函数为：

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

进一步**改写**得：

$$\frac{\partial L}{\partial W^k} = -(y - \hat{y}) \cdot \frac{\partial}{\partial W^k} \sigma(z^{k+1})$$

$$\frac{\partial L}{\partial W^k} = -(y - \hat{y}) \cdot \sigma(z^{k+1})(1 - \sigma(z^{k+1})) \cdot \frac{\partial}{\partial W^k} \sigma(z^{k+1})$$

$$\frac{\partial L}{\partial W^k} = -(y - \hat{y}) \cdot \sigma(z^{k+1})(1 - \sigma(z^{k+1})) \cdot o^k$$

$$\frac{\partial L}{\partial W^k} = (o^k)^T \cdot (-(y - \hat{y}) \cdot \sigma(z^{k+1})(1 - \sigma(z^{k+1})))$$

依次**倒推回去**即可

注：以上图的网络为例，$z^{k+1} = o^k W^k$对W^k来说o^k就是斜率