1. K层前馈网络BP算法推导(考虑跨层连接权值)

考虑如下K层前馈网络

$y_1 \quad y_2 \quad \cdots \quad y_l$ $\qquad$ m个

$x_{K-1,1} \quad x_{K-1,2} \quad \cdots \quad x_{K-1,n_{K-1}}$ $\qquad$ n_{K-1}个

$\vdots$ $\qquad$ $\vdots$

$x_{2,1},\ x_{2,2},\ \cdots,\ x_{2,n_2}$ $\qquad$ n_2个

$x_{1,1},\ x_{1,2},\ \cdots,\ x_{1,n_1}$ $\qquad$ $n_1=n$个

其中 $x_{1,i}$ 表示网络的输入 x_i，为了推导方便写为 $x_{1,i}$，其中相邻层的每两个结点一定有边相连，不同层的结点也可以有边相连.

STEP 1 各层广义误差δ的推导(考虑一个输入点)

$$E=\frac{1}{2}\sum_{l=1}^{m}(t_l-y_l)^2$$

$$\delta_l^K=(t_l-y_l)\cdot y_l(1-y_l)$$

由三层BP算法的推导得

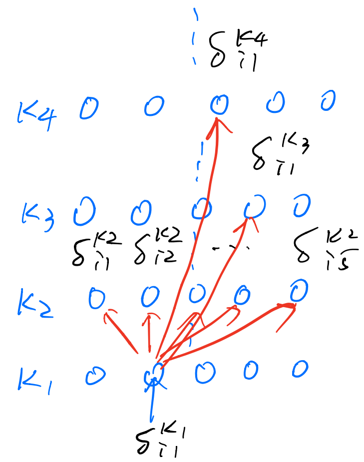
$$\delta_i^{K-1}=\sum_{l=1}^{m}\delta_l^K\cdot w_{K-1,l}^{K,i}$$

其中 δ_i^t 表示t层第i个的广义误差

$w_{a,i}^{b,j}$ 表示a层i个到b层j个的误差

则有 $\delta_i^t = \sum_{p=t+1}^{K} \sum_{q=1}^{n_p} \delta_q^p \cdot W_{t,i}^{p,q}$

如下图所示



则有 $\delta_{i1}^{K_1} = \sum_{K} \sum_{i} \delta_i^K$

STEP 2 对某一样本u，$\frac{\partial E}{\partial W_{K_1,i}^{K_2,j}}$ 的推导

$$\frac{\partial E^u}{\partial W_{K_1,i}^{K_2,j}} = \frac{\partial E^u}{\partial x_j^{K_2}} \cdot \frac{\partial x_j^{K_2}}{\partial u_j^{K_2}} \cdot \frac{\partial u_j^{K_2}}{\partial x_i^{K_1}}$$

$$= \delta_j^{K_2} \cdot x_i^{K_1}$$

则 $\frac{\partial E}{\partial W_{K_1,i}^{K_2,j}} = \sum_{u=1}^{N} \delta_j^{K_2} \cdot x_i^{K_1}$

STEP 3 综上，δ_i^K 的更新公式在考虑跨层连接之后略有不同，其余与三层BP算法一致，可以使用三层BP算法学习，只需将第一步修改为：

对所有的u，计算$(\alpha_i^k)_u$与$(\delta_i^k)_u$，

其中 $\delta_i^k = \sum_{t>k} \sum_{j=1}^{n_k} \delta_j^t \cdot W_{k,i}^{t,j}$

若该条边不存在，则对应的$W=0$