

**5.18** For a pattern  $n$ , the error function becomes:

$$E_n = \frac{1}{2} \sum_{k=1}^K (y_k - t_k)^2$$

The forward propagation is done using:

$$a_j = \sum_{i=0}^D w_{ji}^{(1)} x_i$$

$$z_j = \tanh(a_j)$$

$$y_k = \left( \sum_{j=0}^M w_{kj}^{(2)} z_j \right) + \left( \sum_{i=0}^D w_{ki}^{(3)} x_i \right)$$

Note: Skip level connections are represented by  $w_{ki}$ .

Computing the  $\delta$ 's for all the output units:

$$\delta_k = y_k - t_k$$

Computing the  $\delta$ 's for all the hidden units:

$$\delta_j = h'(a_j) \sum_k w_{kj} \delta_k = (1 - z_j^2) \sum_k w_{kj} \delta_k$$

Finally, the derivatives with respect to the second-layer, skip-layer weights and first-layer are given by:

$$\frac{\partial E_n}{\partial w_{kj}^{(2)}} = \delta_k z_j = (y_k - t_k) z_j$$

$$\frac{\partial E_n}{\partial w_{ki}^{(3)}} = \delta_k x_i = (y_k - t_k) x_i$$

$$\frac{\partial E_n}{\partial w_{ji}^{(1)}} = \delta_j x_i = \left( (1 - z_j^2) \sum_k w_{kj} \delta_k \right) x_i$$