

Let  $\mathbf{f}'_j(\mathbf{a}_j) = (f_j(a_{j1}), f_j(a_{j2}), f_j(a_{j3}), \dots, f_j(a_{jN}))$  and  $\mathbf{d}_j = \frac{\partial E_{\mathbf{o}}}{\partial \mathbf{a}_j}$   
for  $j = 1, 2, 3, \dots, N$ .

$$\frac{\partial E_{\mathbf{o}}}{\partial W_j} = \frac{\partial E_{\mathbf{o}}}{\partial \mathbf{a}_j} \frac{\partial \mathbf{a}_j}{\partial W_j} = \frac{\partial E_{\mathbf{o}}}{\partial \mathbf{a}_j} \mathbf{o}_{j-1}^T = \mathbf{d}_j \mathbf{o}_{j-1}^T \text{ for } j = 1, 2, 3, \dots, N$$

$$\frac{\partial E_{\mathbf{o}}}{\partial \mathbf{b}_j} = \frac{\partial E_{\mathbf{o}}}{\partial \mathbf{a}_j} \frac{\partial \mathbf{a}_j}{\partial \mathbf{b}_j} = \frac{\partial E_{\mathbf{o}}}{\partial \mathbf{a}_j} = \mathbf{d}_j \text{ for } j = 1, 2, 3, \dots, N$$

$$\frac{\partial E_{\mathbf{o}}}{\partial \mathbf{a}_N} = (\mathbf{o} - \mathbf{o}_N) \circ (-\mathbf{f}'_N(\mathbf{a}_N))$$

$$\mathbf{d}_N = (\mathbf{o}_N - \mathbf{o}) \circ \mathbf{f}'_N(\mathbf{a}_N)$$

For  $j < N$ :

$$\frac{\partial E_{\mathbf{o}}}{\partial \mathbf{a}_j} = \frac{\partial E_{\mathbf{o}}}{\partial \mathbf{a}_{j+1}} \cdot \frac{\partial \mathbf{a}_{j+1}}{\partial \mathbf{a}_j} = \mathbf{d}_{j+1} \cdot \frac{\partial \mathbf{a}_{j+1}}{\partial \mathbf{a}_j}$$

$$\frac{\partial E_{\mathbf{o}}}{\partial \mathbf{a}_j} = \sum_k d_{j+1k} \left( \frac{\partial}{\partial \mathbf{a}_j} \sum_l W_{j+1lk} o_{jl} \right)$$

$$\mathbf{d}_j = (W_{j+1}^T \mathbf{d}_{j+1}) \circ \mathbf{f}'_j(\mathbf{a}_j)$$