Connected layer from $i \to j$

$$\frac{\partial E}{\partial w_{jk}} = \underbrace{\frac{\partial E}{\partial O_k}}_{1} \; \underbrace{\frac{\partial O_k}{\partial S_k}}_{2} \; \underbrace{\frac{\partial S_k}{w_{jk}}}_{3} \qquad (9)$$

1 $$\frac{\partial E}{\partial O_k} = -(y_k - O_k) \qquad [5] \qquad (10)$$

2 $$\frac{\partial O_k}{\partial S_k} = \underbrace{O_k(1-O_k) = 1}_{\text{the output is a sum, i.e. there is no threshold}} \qquad (11)$$

3 $$S_k = \sum w_{jk} \times O_j - \frac{\partial S_k}{\partial w_{jk}} = \frac{\partial (w_{jk} \times O_j)}{\partial w_{jk}} = O_j \qquad (12)$$

In Conclusion : $$\frac{\partial E}{\partial w_{jk}} = \underbrace{\underbrace{-(y_k - O_k)}_{\text{Error}} \times 1 \times \underbrace{O_j}_{\text{inputs}}}_{\text{Jacobian}} \qquad (13)$$

$$\text{As } g = \begin{bmatrix} \frac{\partial E}{\partial w_1} \\ \vdots \\ \frac{\partial E}{\partial w_m} \end{bmatrix} \text{ then } g_{ik} = \text{error}_k \begin{bmatrix} O_1 k \\ O_2 k \\ O_3 k \end{bmatrix} \qquad (14)$$

Connected layer from $i \to j$

$$\frac{\partial E}{\partial w_{ij}} = \underbrace{\frac{\partial E}{\partial S_j}}_{①} O_i \qquad (15)$$

① $$\frac{\partial E}{\partial S_j} = -\left[O_j(1-O_j)\sum_k \left(\underbrace{-\beta_k}_{②}\right) w_{jk}\right] \qquad [5]\ (16)$$

② $$\beta_k = \underbrace{O_k(1-O_k)}_{\substack{=1 \\ \text{as unthresholded}}}\left[y_k - O_k\right] = 1 \times (y_k - O_k) \qquad (17)$$

In conclusion: $$\frac{\partial E}{\partial w_{ij}} = -\left[O_j(1-O_j)\sum_k \left[-(y_k - O_k)\right] w_{jk}\right] O_i \qquad (18)$$

$$\frac{\partial E}{\partial w_{ij}} = \left[O_j(1-O_j)\sum_k \underbrace{(y_k - O_k)}_{\text{Error}} \underbrace{w_{jk}}_{\text{weights}}\right] \underbrace{O_i}_{\text{inputs}} \qquad (19)$$

Consequently: $$g_{ij} = \begin{bmatrix} \frac{\partial E}{w_{1j}} \\ \vdots \\ \frac{\partial E}{\partial w_{nj}} \end{bmatrix} \qquad (20)$$