

Backward Propagation

R⁽³⁾

$$r^{(3)} = \frac{\partial E(w)}{\partial a^{(3)}} = \frac{-t}{e^{ta^{(3)}} + 1}$$

R⁽²⁾

$$r_q^{(2)} = \frac{\partial E(w)}{\partial a_q^{(2)}} = \frac{\partial E(w)}{\partial a^{(3)}} \frac{\partial a^{(3)}}{\partial a_q^{(2)}}$$

L

$$r_L^{(2)} = r^{(3)} \frac{(w^{(3)})^T \cdot a_{LR}^{(2)} \cdot e^{a_R^{(2)} + a_L^{(2)}}}{(e^{a_R^{(2)}} + 1)(e^{a_L^{(2)}} + 1)^2}$$

R

$$r_R^{(2)} = r^{(3)} \frac{(w^{(3)})^T \cdot a_{LR}^{(2)} \cdot e^{a_R^{(2)} + a_L^{(2)}}}{(e^{a_L^{(2)}} + 1)(e^{a_R^{(2)}} + 1)^2}$$

LR

$$r_{LR}^{(2)} = r^{(3)} \frac{(w^{(3)})^T}{(e^{-a_L^{(2)}} + 1)(e^{-a_R^{(2)}} + 1)}$$

R⁽¹⁾

$$r_q^{(1)} = \frac{\partial E(w)}{\partial a_q^{(1)}} = \sum_{p \in \{L, R, LR\}} \frac{\partial E(w)}{\partial a_p^{(2)}} \frac{\partial a_p^{(2)}}{\partial a_q^{(1)}}$$

L

$$r_L^{(1)} = \sum_{i=1}^{H2} r^{(2)} ((w_L^{(2)})^T \cdot \text{sech}(a_{L,i}^{(1)}) + (w_{LR,i}^{(2)})^T \cdot \text{sech}(a_L^{(1)}))$$

R

$$r_R^{(1)} = \sum_{i=1}^{H2} r^{(2)} ((w_R^{(2)})^T \cdot \text{sech}(a_{R,i}^{(1)}) + (w_{LR,i}^{(2)})^T \cdot \text{sech}(a_R^{(1)}))$$

L

$$r_{L_q}^{(1)} = \frac{\partial E(w)}{\partial a_{L_q}^{(1)}} = \sum_k \left(\frac{\partial E(w)}{\partial a_{L_k}^{(2)}} \frac{\partial a_{L_k}^{(2)}}{\partial a_{L_q}^{(1)}} + \frac{\partial E(w)}{\partial a_{R_k}^{(2)}} \frac{\partial a_{R_k}^{(2)}}{\partial a_{L_q}^{(1)}} + \frac{\partial E(w)}{\partial a_{LR_k}^{(2)}} \frac{\partial a_{LR_k}^{(2)}}{\partial a_{L_q}^{(1)}} \right)$$

R

$$r_{R_q}^{(1)} = \frac{\partial E(w)}{\partial a_{R_q}^{(1)}} = \sum_k \left(\frac{\partial E(w)}{\partial a_{L_k}^{(2)}} \frac{\partial a_{L_k}^{(2)}}{\partial a_{R_q}^{(1)}} + \frac{\partial E(w)}{\partial a_{R_k}^{(2)}} \frac{\partial a_{R_k}^{(2)}}{\partial a_{R_q}^{(1)}} + \frac{\partial E(w)}{\partial a_{LR_k}^{(2)}} \frac{\partial a_{LR_k}^{(2)}}{\partial a_{R_q}^{(1)}} \right)$$