Backward Propagation

$\mathbf{R}^{(3)}$

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

$\mathbf{R}^{(2)}$

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

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

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

L

$$\begin{aligned} r_{L_q}^{(1)} = \frac{\partial E(w)}{\partial a_{Lq}^{(1)}} &= \sum_k \left(\frac{\partial E(w)}{\partial a_{L_k}^{(2)}}\frac{\partial a_{L_k}^{(2)}}{\partial a_{L_q}^{(1)}} + \frac{\partial E(w)}{\partial a_{R_k}^{(2)}}\frac{\partial a_{R_k}^{(2)}}{\partial a_{L_q}^{(1)}} + \frac{\partial E(w)}{\partial a_{LR_k}^{(2)}}\frac{\partial a_{LR_k}^{(2)}}{\partial a_{L_q}^{(1)}}\right) \\ &= \sum_k r^{(2)}((w_{L_q}^{(2)})^T \cdot sech(a_{L_q,k}^{(1)}) + (w_{LR,k}^{(2)})^T \cdot sech(a_{L_q}^{(1)})) \end{aligned}$$

R

$$\begin{aligned} r_{R_q}^{(1)} = \frac{\partial E(w)}{\partial a_{R_q}^{(1)}} &= \sum_k \left(\frac{\partial E(w)}{\partial a_{L_k}^{(2)}}\frac{\partial a_{L_k}^{(2)}}{\partial a_{R_q}^{(1)}} + \frac{\partial E(w)}{\partial a_{R_k}^{(2)}}\frac{\partial a_{R_k}^{(2)}}{\partial a_{R_q}^{(1)}} + \frac{\partial E(w)}{\partial a_{LR_k}^{(2)}}\frac{\partial a_{LR_k}^{(2)}}{\partial a_{R_q}^{(1)}}\right) \\ &= \sum_k r^{(2)}((w_{R_q}^{(2)})^T \cdot sech(a_{R_q,k}^{(1)}) + (w_{LR,k}^{(2)})^T \cdot sech(a_{R_q}^{(1)})) \end{aligned}$$