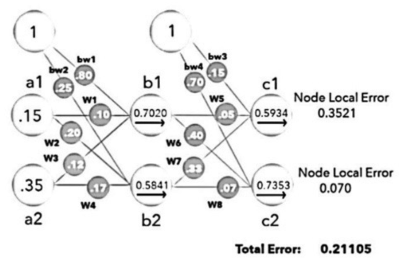
Lab 1 Calculating Gradients



$$\frac{\partial E}{\partial W_5} = (z-t)\, z(1-z)\, out_{b1}$$

① Every "z" in this formula is the actual output for **node c1**.

② This "t" is the target output for **node c1**.

$\frac{\partial E}{\partial W_5} = (0.5934 - 0) * 0.5934(1 - 0.5934) * 0.7020 = 0.1005$

$\frac{\partial E}{W_6} = (0.7353 - 1) * 0.7353(1 - 0.7353) * 0.7020 = -0.0362$

$\frac{\partial E}{W_7} = (0.5934 - 0)\ 0.5934(1 - 0.5934) * 0.5841 = 0.0836$

$\frac{\partial E}{W_8} = (0.7353 - 1)\ 0.7353(1 - 0.7353) * 0.5841 = -0.0301$

Bias Weights

$$\delta_z = (z-t)\, z(1-z)$$

Delta z

$BW_3 = (0.5934 - 0) * .5934(1 - .5934) = 0.1432$

$Bw_4 = (0.7353 - 1) * .7353(1 - .7353) = -0.0515$

Hidden Layer

$$\frac{\partial E}{\partial W_1} = \Big(\sum_c \delta_z W_i\Big)\, out_{b1}(1 - out_{b1})\, out_{a1}$$

① $W_1 \rightarrow$

$0.1432 * \overset{W_5}{0.05} = 0.0072$

$-0.0515 * \overset{W_6}{0.4} = -0.0206$

$0.0072 - 0.0206 = -0.0134$

$\sum \delta_z W_i = -0.0134$

$out_{b_1}(1 - out_{b_1}) \rightarrow 0.7020 * (1 - 0.7020) = 0.2092$

$-0.0134 * .2092 * 0.15 = -0.00042$

$W_2 \rightarrow$

$\delta_3 = 0.1432 * \overset{W_7}{.33} = 0.0473$

$\delta_2 = -0.0515 * \overset{W_8}{.07} = -0.0036$

0.0437

② $out_{b_2}(1 - out_{b_2})$

$0.5841 * (1 - 0.5841) = 0.2429$

③ $0.0437 * 0.2429 * .15 = 0.00159$

$W_3 \rightarrow -0.0134 * 0.2092 * 0.35 = -0.00098$

$W_4 \rightarrow 0.0437 * 0.2429 * 0.35 = 0.00372$

Hidden Layer Bias Weights

$$\delta_b = \left(\sum_c \delta_z W_i\right) out_i (1 - out_i)$$

1 All of the letter "i"'s refer to a unique value. This value depends on the gradient we are calculating.

$$\frac{\partial E}{\partial BW_1} = \left(\sum \delta W_{5+6}\right) out_{b_1} (1 - out_{b_1})$$

$$\frac{\partial E}{BW_2} = \left(\sum \delta W_{7+8}\right) out_{b_2} (1 - out_{b_2})$$

$(-0.0134) * 0.7020 \ (1 - 0.7020) = \underline{-0.0028}$

$0.0437 * 0.5841 \ (1 - 0.5841) = \underline{0.0106}$

Updating Weights

1 The old weight **w8**, which is being updated.

3 The partial derivative of **the total error** with respect to the weight **w8.**

$$W8_{new} = W8 - \eta * \frac{\partial E}{\partial W8}$$

2 The greek letter eta represents the learning rate.

$B_{W_4} = 0.7 - 0.5 * -0.0515 = 0.7258$

$B_{W_3} = 0.15 - 0.5 * 0.1432 = 0.0784$

$B_{W_2} = 0.25 - 0.5 * 0.016 = 0.242$

$B_{W_1} = 0.8 - 0.5 * -0.0028 = 0.814$

$W_8 = 0.07 - 0.5 * (-0.0301) = 0.0851$

$W_7 = 0.33 - 0.5 * .0836 = 0.2882$

$W_6 = .4 - 0.5 * -0.0362 = 0.4181$

$W_5 = 0.05 - 0.5 * 0.1005 = -0.00025$

$W_4 = 0.17 - 0.5 * 0.00372 = 0.1514$

$W_3 = 0.12 - 0.5 * -0.00098 = 0.1205$

$W_2 = 0.2 - 0.5 * 0.00159 = 0.1992$

$W_1 = 0.1 - 0.5 * -0.000042 = 0.1002$