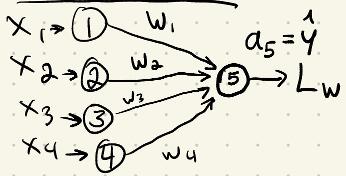
c] • Want gradient of objective function L_w w.r.t to the neural network weights, $\frac{dL_w}{dw\_}$

Neural Network



Loss funct.

$$L_w = \sum_{n=1}^{N} (y - \hat{y})^2, \text{ where}$$

$$\hat{y} = a_5 = \sigma(x_1 w_1 + x_2 w_2 + x_3 W_3 + x_4 W_4 + w_0)$$

$$\sigma(z) = 1/(1 + e^{-z})$$

Apply chain rule and solve for a single sample n and weight w_1

$$\frac{dL_w}{dw_1} = \sum_n \frac{dL_{wn}}{da_{5,n}} \times \frac{da_{5,n}}{dw_1} \rightarrow \frac{dL_{wn}}{da_{5,n}} \times \frac{da_{5,n}}{dw_1}$$

Find $dL_{w,n}/da_{5,n}$

$$\frac{d}{da_{5,n}} (y - a_{5,n})^2 = 2(y - a_{5,n}) \cdot (0 - 1) = -2(y_n - a_{5,n})$$

Find $da_{5,n}/dw_1$, substitute $z = x_1 w_1 + x_2 w_2 + x_3 W_3 + x_4 W_4 + w_0$

$$\frac{d}{dw_1} \sigma(z) = \sigma(z)(1 - \sigma(z)) \frac{dz}{dw_1} \rightarrow \text{chain rule for outer \& inner of } \sigma(z)$$

$$= \sigma(z)(1 - \sigma(z))(x_1^n)$$

So,

$$\frac{dL_{w,n}}{dw_1} = -2(y - a_{5,n}) \cdot \sigma(z)(1 - \sigma(z)) \cdot x_1^n$$

Finally, sum over all N samples in dataset

$$\frac{dL_w}{dw_1} = \sum_{n=1}^{N} -2(y - a_5) \cdot \sigma(z)(1 - \sigma(z)) \cdot x_1$$

(where $z = x_1 w_1 + x_2 w_2 + x_3 w_3 + x_4 W_4 + w_0$,

$\sigma(z) = \frac{1}{1 + e^{-z}}$, and

$a_5 = \sigma(z)$

Repeat for $\underline{w_2}$:

$$\frac{dL_{w,n}}{dw_2} = \frac{dL_{w,n}}{da_{5,n}} \times \frac{da_{5,n}}{dw_2} \rightarrow \text{Only } \frac{da_{5n}}{dw_2} \text{ changes}$$

$$= -2(y^n - a_{5,n})\,\sigma(z)(1-\sigma(z))\frac{dz}{dw_2}$$

$$\frac{dz}{dw_2} = \frac{d}{dw_2}\left[x_1^n w_1 + x_2^n w_2 + x_3^n w_3 + x_4^n w_4 + w_0\right] = x_2^n$$

So $$\frac{dL_w}{dw_2} = \sum_{n=1}^{N} -2(y - a_5)\,\sigma(z)(1-\sigma(z)) \cdot x_2$$

w_3 and w_4 work the same as w_1 and w_2 so repeat the pattern - only need to update dz/dw_k

$$\frac{dL_{w_n}}{dw_3} = \sum_n \left[-2(y - a_5)\,\sigma(z)(1-\sigma(z))(x_3)\right]$$

$$\frac{dL_{w_n}}{dw_4} = \sum_n \left[-2(y - a_5)\,\sigma(z)(1-\sigma(z))(x_4)\right)$$

For w_0, compute $\frac{d}{dw_0}\left[x_1 w_1 + \ldots + x_4 w_4 + w_0\right] = 1$, so

$$\frac{dL_{w,n}}{dw_0} = \sum_n \left[-2(y - a_5)\,\sigma(z)(1-\sigma(z))\right)$$