D In scalar form, take only $\frac{dL_{\vec{w}}}{dw_i}$ for one sample, n

$$\frac{dL_{w,n}}{dw_1} = -2(y - a_5^n) \cdot \sigma(z)(1 - \sigma(z)) \cdot x_1^n,$$

$$\left(\begin{aligned} &\text{where } z = x_1^n w_1 + x_2^n w_2 + x_3^n w_3 + x_4^n w_4 + w_0, \\ &\sigma(z) = \frac{1}{1+e^{-z}}, \text{ and} \\ &a_{5,n} = \sigma(z) \end{aligned} \right)$$

$$\frac{dL_{w,n}}{dw_2} = -2(y - a_5^n) \cdot \sigma(z)(1 - \sigma(z)) \cdot x_2^n,$$

$$\frac{dL_{w,n}}{dw_3} = -2(y - a_5^n) \cdot \sigma(z)(1 - \sigma(z)) \cdot x_3^n,$$

$$\frac{dL_{w,n}}{dw_4} = -2(y - a_5^n) \cdot \sigma(z)(1 - \sigma(z)) \cdot x_4^n,$$

$$\frac{dL_{w,n}}{dw_0} = -2(y - a_5^n) \cdot \sigma(z)(1 - \sigma(z))$$

Now in <u>Vector form</u>, for all N samples n

$$\frac{dLw}{dw_1} = \begin{bmatrix} -2(y-a_5^1)\cdot\sigma(z)(1-\sigma(z))\cdot x_1^1 \\ -2(y-a_5^2)\cdot\sigma(z)(1-\sigma(z))\cdot x_1^2 \\ \cdots \\ -2(y-a_5^N)\cdot\sigma(z)(1-\sigma(z))\cdot x_1^N \end{bmatrix} \begin{array}{l} \leftarrow n=1 \\ \leftarrow n=2 \\ \ddots \\ \leftarrow n=N \end{array}$$

$$\frac{dLw}{dw_2} = \begin{bmatrix} -2(y-a_5^1)\cdot\sigma(z)(1-\sigma(z))\cdot x_2^1 \\ -2(y-a_5^2)\cdot\sigma(z)(1-\sigma(z))\cdot x_2^2 \\ \cdots \\ -2(y-a_5^N)\cdot\sigma(z)(1-\sigma(z))\cdot x_2^N \end{bmatrix}$$

$$\frac{dLw}{dw_3} = \begin{bmatrix} -2(y-a_5^1)\cdot\sigma(z)(1-\sigma(z))\cdot x_3^1 \\ -2(y-a_5^2)\cdot\sigma(z)(1-\sigma(z))\cdot x_3^2 \\ \cdots \\ -2(y-a_5^N)\cdot\sigma(z)(1-\sigma(z))\cdot x_3^N \end{bmatrix}$$

$$\frac{dLw}{dw_4} = \begin{bmatrix} -2(y-a_5^1)\cdot\sigma(z)(1-\sigma(z))\cdot x_4^1 \\ -2(y-a_5^2)\cdot\sigma(z)(1-\sigma(z))\cdot x_4^2 \\ \cdots \\ -2(y-a_5^N)\cdot\sigma(z)(1-\sigma(z))\cdot x_4^N \end{bmatrix}$$

$$\frac{dLw}{dw_0} = \begin{bmatrix} -2(y-a_5^1)\cdot\sigma(z)(1-\sigma(z)) \\ -2(y-a_5^2)\cdot\sigma(z)(1-\sigma(z)) \\ \cdots \\ -2(y-a_5^N)\cdot\sigma(z)(1-\sigma(z)) \end{bmatrix} \begin{array}{l} \leftarrow n=1 \\ \leftarrow n=2 \\ \\ \leftarrow n=N \end{array}$$