Reminder. The gradient $\nabla\mathcal{L} = \frac{\partial \mathcal{L}}{\partial x} = \left[\frac{\partial \mathcal{L}}{\partial x_1}, \frac{\partial \mathcal{L}}{\partial x_2}, \ldots, \frac{\partial \mathcal{L}}{\partial x_n}\right]$ and how we need to change to decrease or increase it.

We now use it by applying the chain rule $\frac{\partial \mathcal{L}}{\partial W_{ij}^{(l)}} = \frac{\partial \mathcal{L}}{\partial z_i^{(l)}} \frac{\partial z_i^{(l)}}{\partial W_{ij}^{(l)}}$

$$\frac{\partial z_i^{(l)}}{\partial W_{ij}^{(l)}} = \frac{\partial}{\partial W_{ij}^{(l)}} \sum_{k=1}^{n} W_{ik}^{(l)} a_k^{(l-1)} + b_k^{(l-1)} = a_j^{(l-1)}$$

By taking all derivatives and putting everything together, you can compute all changes. From this analysis, we can propagate the errors, how? Just computing the gradients and update the values of the parameters as:

$$W \quad \Longrightarrow \quad W - \xi \frac{\partial \mathcal{L}}{\partial W}$$

$$\text{b} \quad \Longrightarrow \quad b - \xi \frac{\partial \mathcal{L}}{\partial b}$$