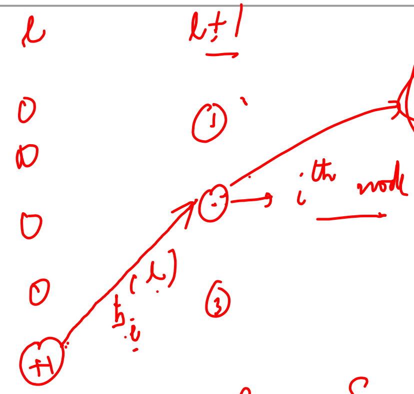
$$\frac{\partial J(W,b;x,y)}{\partial b_i^{(l)}} = \underbrace{\frac{\partial J}{\partial z_i^{(l+1)}}}_{\delta_i^{(l+1)}} \underbrace{\frac{\partial z_i^{(l+1)}}{\partial b_i^{(l)}}}_{=1}$$



$$\delta_i^{(l+1)} = \frac{\partial J}{\partial z_i^{(l+1)}}$$

a , δ

$$z^{(l+1)} = W^{(l)} a^{(l)} + b^{(l)}$$

$$z_i^{(l+1)} = W_{i1}^{(l)} a_1^{(l)} + W_{i2} a_2^{(l)} + \dots + b_i^{(l)}$$

$$\frac{\partial z_i^{(l+1)}}{\partial b_i^{(l)}} = 1$$

$$\boxed{\frac{\partial J}{\partial b_i^{(l)}} = \delta_i^{(l+1)}} \qquad \boxed{\frac{\partial J}{\partial b^{(l)}} = \delta^{(l+1)}}$$

$$b^{(l)} = \begin{pmatrix} b_1^{(l)} \\ \vdots \end{pmatrix}_{N_{l+1}\times 1}$$

$$\frac{\partial J}{\partial b^{(l)}} = \begin{pmatrix} \frac{\partial J}{\partial b_1^{(l)}} \\ \frac{\partial J}{\partial b_2^{(l)}} \\ \vdots \end{pmatrix}_{N_{l+1}\times 1} = \begin{pmatrix} \delta_1^{(l+1)} \\ \delta_2^{(l+1)} \\ \vdots \end{pmatrix}_{N_{l+1}\times 1} = \delta^{(l+1)}$$

$$\frac{\partial J}{\partial W_{ij}^{(l)}} = \frac{\partial J}{\partial z_i^{(l+1)}} \frac{\partial z_i^{(l+1)}}{\partial W_{ij}^{(l)}}$$

$\delta_i^{(l+1)}$ $\quad a_j^{(l)}$

$(l) \quad (l+1)$

j^{th} node $\quad W_{ij}^{(l)} \quad i^{th}$ node $\quad z_i^{(l+1)}$

$$z^{(l+1)} = W^{(l)} a^{(l)} + b^{l}$$

$$z_i^{(l+1)} = W_{i1}^{(l)} a_1^{(l)} + W_{i2}^{(l)} a_2^{(l)} + \dots + W_{ij}^{(l)} a_j^{(l)} + b_i^{(l)}$$

$$\frac{\partial z_i^{(l+1)}}{\partial W_{ij}^{(l)}} = a_j^{(l)}$$

$$\frac{\partial J}{\partial W_{ij}^{(l)}} = a_j^{(l)} \delta_i^{(l+1)}$$

$$z^{(l+1)} = W a^{(l)} + b^{(l)}$$

$N_{(l+1)} \times 1 \quad (N_{l+1} \times N_l) \quad N_l \times 1$

$$\boxed{\frac{\partial J}{\partial W_{ij}^{(l)}} = \delta_i^{(l+1)} a_j^{(l)}}$$

$$\boxed{\frac{\partial J}{\partial W} = \delta^{(l+1)} (a^{(l)})^T}$$

$$\frac{\partial J}{\partial W} = \begin{pmatrix} \frac{\partial J}{\partial W_{11}} & \frac{\partial J}{\partial W_{12}} & \dots \\ \frac{\partial J}{\partial W_{21}} & \frac{\partial J}{\partial W_{22}} & \dots \end{pmatrix} = \delta^{(l+1)} a^{(l)T}$$

$N_{l+1} \times N_l$

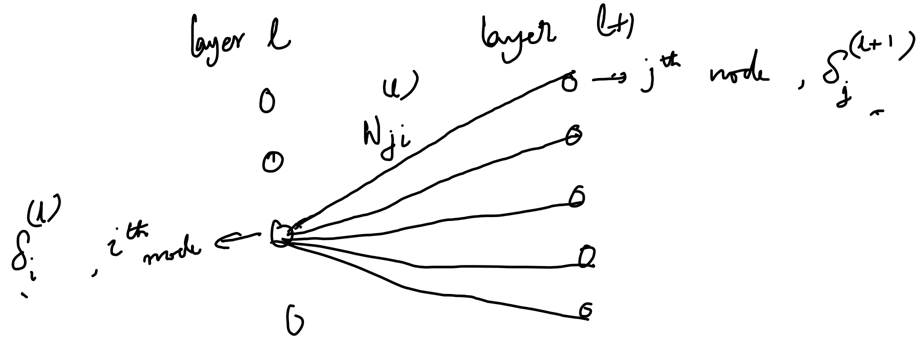
$\delta^{(l+1)}$: $N_{l+1} \times 1$

$a^{(l)}$: $N_l \times 1$

$$i^{th} \quad \begin{pmatrix} \delta_1^{(l+1)} \\ \delta_2^{(l+1)} \\ \vdots \end{pmatrix} (a_1^{(l)} \; a_2^{(l)} \dots)$$

$$\boxed{\delta_i^{(l)} = \left(\sum_{j=1}^{s_{l+1}} W_{ji}^{(l)} \delta_j^{(l+1)}\right) f'\left(z_i^{(l)}\right)}$$

s_{l+1} : # nodes in layer $(l+1)$



$$\delta_i^{(l)} = \frac{\partial J}{\partial z_i^{(l)}} = \sum_{j=1}^{s_{l+1}} \left(\underbrace{\frac{\partial J}{\partial z_j^{(l+1)}}}_{\delta_j^{(l+1)}} \frac{\partial z_j^{(l+1)}}{\partial a_i^{(l)}} \right) \underbrace{\frac{\partial a_i^{(l)}}{\partial z_i^{(l)}}}_{f'\left(z_i^{(l)}\right)}$$

$$z_j^{(l+1)} = W_{ji}^{(l)} a_i^{(l)} + b_j^{(l)}$$

$$\frac{\partial z_j^{(l+1)}}{\partial a_i^{(l)}} = W_{ji}^{(l)}$$

$$\frac{\partial J}{\partial z_j^{(l+1)}} = \delta_j^{(l+1)}$$

$$\boxed{\delta_i^{(l)} = \left(\sum_{j=1}^{s_{l+1}} W_{ji}^{(l)} \delta_j^{(l+1)}\right) f'\left(z_i^{(l)}\right)}$$

forward prop

$$z^{(l+1)} = W^{(l)} a^{(l)} + b^{(l)}$$

$$a^{(l)} = f\left(z^{(l)}\right)$$

$$\frac{\partial a_i^{(l)}}{\partial z_i^{(l)}} = f'\left(z_i^{(l)}\right)$$

$$\delta_i^{(l)} = \underbrace{\left(\sum_{j=1}^{s_{l+1}} W_{ji}^{(l)} \delta_j^{(l+1)} \right)} f'(z_i^{(l)})$$

$$\underline{\delta}^{(l)} = \begin{pmatrix} \delta_1^{(l)} \\ \delta_2^{(l)} \\ \vdots \\ \delta_{s_l}^{(l)} \end{pmatrix} \qquad \underline{\delta}^{(l+1)} = \begin{pmatrix} \delta_1^{(l+1)} \\ \delta_2^{(l+1)} \\ \vdots \\ \delta_{s_{l+1}}^{(l+1)} \end{pmatrix}$$

$$\delta_i^{(l)} = \begin{pmatrix} W_{1i} & W_{2i} & \cdots \end{pmatrix} \begin{pmatrix} \delta_1^{(l+1)} \\ \delta_2^{(l+1)} \\ \vdots \\ \vdots \end{pmatrix}$$

$$\begin{matrix} W_{11} & W_{21} & W_{31} \\ W_{12} & W_{22} & W_{32} \end{matrix} \qquad \begin{pmatrix} W_{11} & W_{12} & W_{13} \\ W_{21} & W_{22} & W_{23} \end{pmatrix}$$

vector-matrix notation for backward recursion

$$\boxed{\underline{\delta}^{(l)}_{N\times 1} = \left(W^{(l)^T} \underline{\delta}^{(l+1)}_{N\times 1} \right) \circ f'(\underline{z}^{(l)})_{N\times 1}}$$

$$\underline{z}^{(l)} = \begin{pmatrix} z_1^{(l)} \\ z_2^{(l)} \\ \vdots \end{pmatrix}_{N\times 1}$$

$$f'(\underline{z}^{(l)}) = \begin{pmatrix} f'(z_1^{(l)}) \\ f'(z_2^{(l)}) \\ \vdots \end{pmatrix}$$