

$$d = a @ b + c$$

broadcasted  
↓

$$\begin{bmatrix} d_{11} & d_{12} \\ d_{21} & d_{22} \end{bmatrix} = \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix} \begin{bmatrix} b_{11} & b_{12} \\ b_{21} & b_{22} \end{bmatrix} + \begin{bmatrix} c_1 & c_2 \\ c_1 & c_2 \end{bmatrix}$$

$$\Rightarrow \begin{aligned} d_{11} &= a_{11} b_{11} + a_{12} b_{21} + c_1 \\ d_{12} &= a_{11} b_{12} + a_{12} b_{22} + c_2 \\ d_{21} &= a_{21} b_{11} + a_{22} b_{21} + c_1 \\ d_{22} &= a_{21} b_{12} + a_{22} b_{22} + c_2 \end{aligned}$$

Loss depends on all d's

$$L = L(d_{11}, d_{12}, d_{21}, d_{22})$$

Chain Rule:-

$$\frac{\partial L}{\partial a_{11}} = \sum_{ij} \frac{\partial L}{\partial d_{ij}} \frac{\partial d_{ij}}{\partial a_{11}}$$

Now which  $d_{ij}$  actually contains  $a_{11}$

Chain rule = multiply along a path, sum across paths.

$$d_{11} = a_{11} b_{11} + a_{12} b_{21} + c_1$$

$$d_{12} = a_{11} b_{12} + a_{12} b_{22} + c_2$$

only row 1 depends on  $a_{11}$

So sum reduces:

$$\frac{\partial L}{\partial a_{11}} = \frac{\partial L}{\partial d_{11}} \cdot \frac{\partial d_{11}}{\partial a_{11}} + \frac{\partial L}{\partial d_{12}} \cdot \frac{\partial d_{12}}{\partial a_{11}}$$

Now differentiate

$$\frac{\partial d_{11}}{\partial a_{11}} = b_{11}$$

$$\frac{\partial d_{12}}{\partial a_{11}} = b_{12}$$


Similarly,

$$\begin{aligned} \Rightarrow \frac{\partial L}{\partial a_{11}} &= \frac{\partial L}{\partial d_{11}} \cdot b_{11} + \frac{\partial L}{\partial d_{12}} \cdot b_{12} \\ \frac{\partial L}{\partial a_{12}} &= \frac{\partial L}{\partial d_{11}} \cdot b_{21} + \frac{\partial L}{\partial d_{12}} \cdot b_{22} \\ \frac{\partial L}{\partial a_{21}} &= \frac{\partial L}{\partial d_{21}} \cdot b_{11} + \frac{\partial L}{\partial d_{22}} \cdot b_{12} \\ \frac{\partial L}{\partial a_{22}} &= \frac{\partial L}{\partial d_{21}} \cdot b_{21} + \frac{\partial L}{\partial d_{22}} \cdot b_{22} \end{aligned} \Rightarrow \begin{bmatrix} \frac{\partial L}{\partial a_{11}} & \frac{\partial L}{\partial a_{12}} \\ \frac{\partial L}{\partial a_{21}} & \frac{\partial L}{\partial a_{22}} \end{bmatrix} = \frac{\partial L}{\partial \mathbf{a}}$$

$$\begin{bmatrix} \frac{\partial L}{\partial a_{11}} & \frac{\partial L}{\partial a_{12}} \\ \frac{\partial L}{\partial a_{21}} & \frac{\partial L}{\partial a_{22}} \end{bmatrix} = \frac{\partial L}{\partial a}$$

$$\Rightarrow \begin{bmatrix} \frac{\partial L}{\partial d_{11}} & \frac{\partial L}{\partial d_{12}} \\ \frac{\partial L}{\partial d_{21}} & \frac{\partial L}{\partial d_{22}} \end{bmatrix} \begin{bmatrix} b_{11} & b_{21} \\ b_{12} & b_{22} \end{bmatrix}$$

Transpose of 'b'



$$\Rightarrow \frac{\partial L}{\partial d} @ b^T$$

Similarly, if we do w.r.t 'b' then we get

$$\frac{\partial L}{\partial b} = a^T @ \frac{\partial L}{\partial d}$$

$$d_{11} = a_{11}b_{11} + a_{12}b_{21} + c_1$$

$$d_{21} = a_{21}b_{11} + a_{22}b_{21} + c_1$$

$$\frac{\partial L}{\partial c_1} = \frac{\partial L}{\partial d_{11}} \cdot \frac{\partial d_{11}}{\partial c_1} + \frac{\partial L}{\partial d_{21}} \cdot \frac{\partial d_{21}}{\partial c_1}$$

$$\frac{\partial d_{11}}{\partial c_1} = 1 \quad \frac{\partial d_{21}}{\partial c_1} = 1$$

w.r.t 'c<sub>1</sub>' and 'c<sub>2</sub>'

$$\frac{\partial L}{\partial c_1} = \frac{\partial L}{\partial d_{11}} \cdot 1 + \frac{\partial L}{\partial d_{21}} \cdot 1$$

$$\frac{\partial L}{\partial c_2} = \frac{\partial L}{\partial d_{12}} \cdot 1 + \frac{\partial L}{\partial d_{22}} \cdot 1$$

$$\Rightarrow \frac{\partial L}{\partial c} = \frac{\partial L}{\partial d} \cdot \text{sum}(\text{dim} = 0)$$