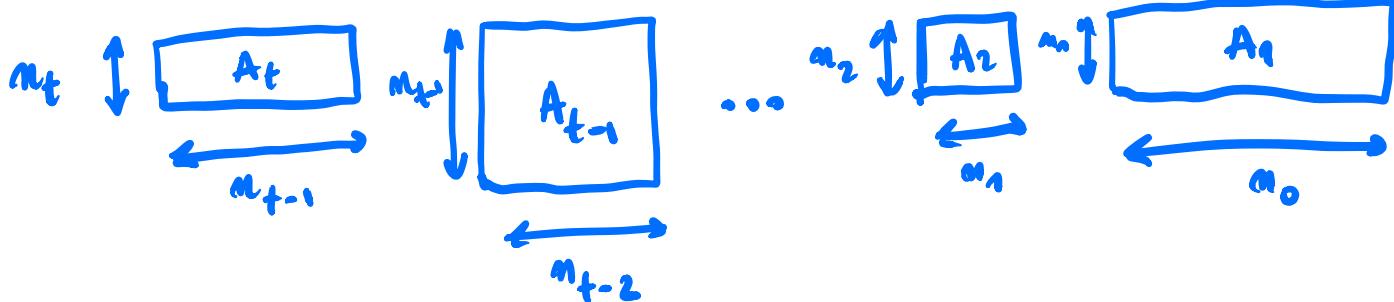


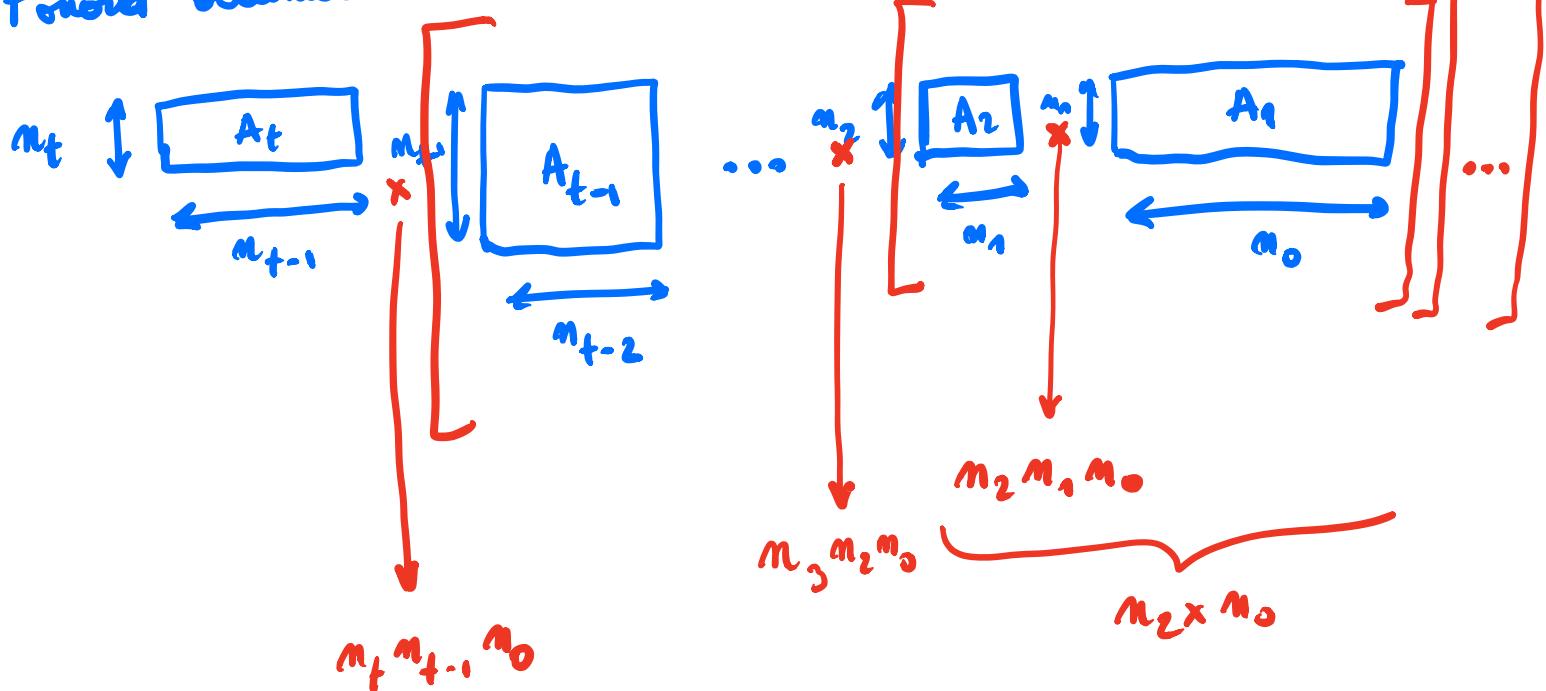
Lecture 3

Complexity (Slide 17)

$$\frac{\partial x_t}{\partial x_0} = \frac{\partial x_t}{\partial x_{t-1}} \cdot \frac{\partial x_{t-1}}{\partial x_{t-2}} \cdots \frac{\partial x_2}{\partial x_1} \cdot \frac{\partial x_1}{\partial x_0}$$

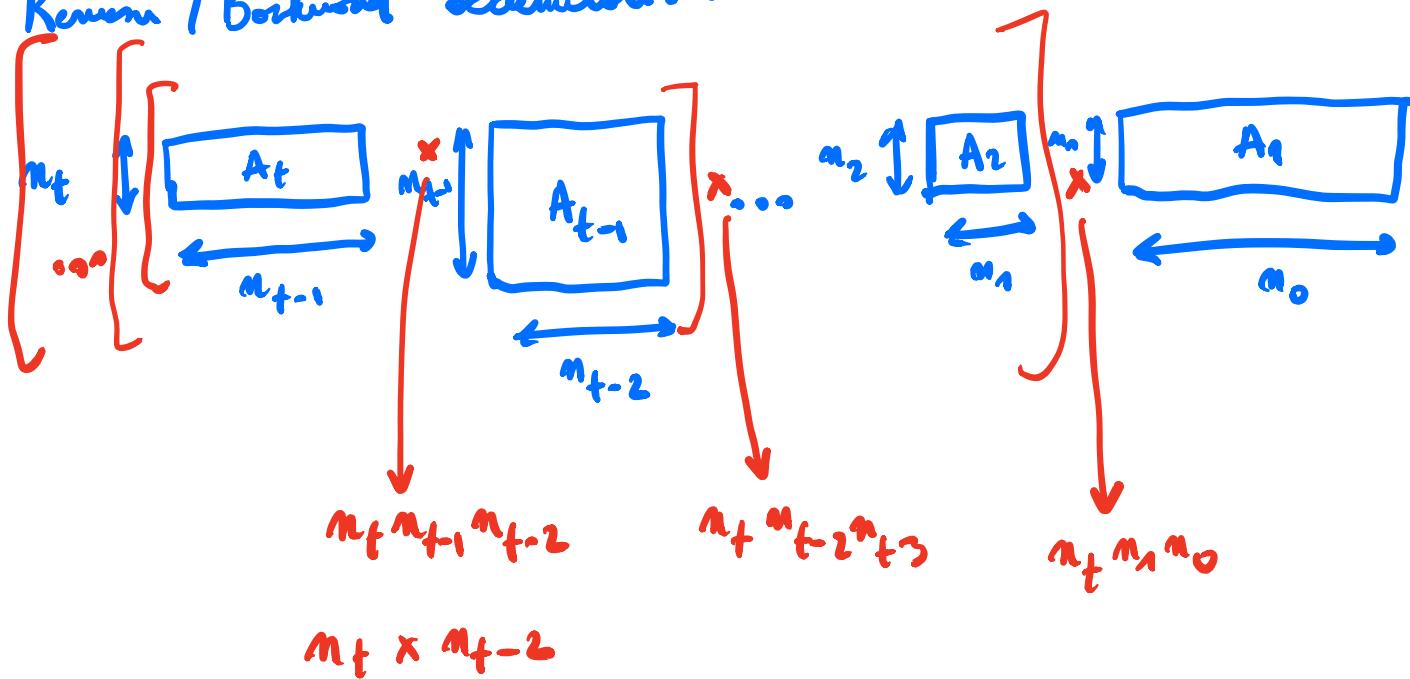


Forward accumulation:

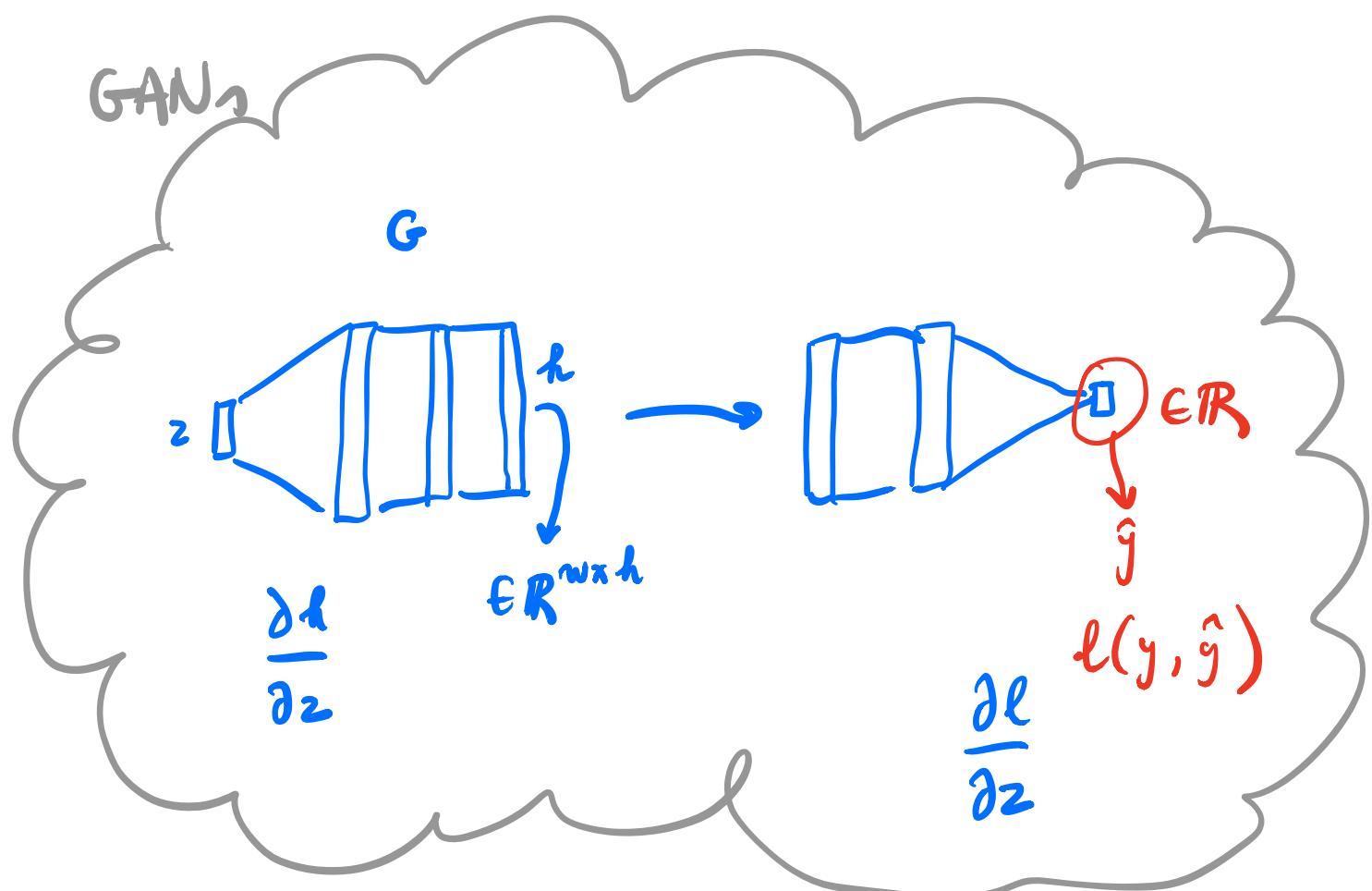


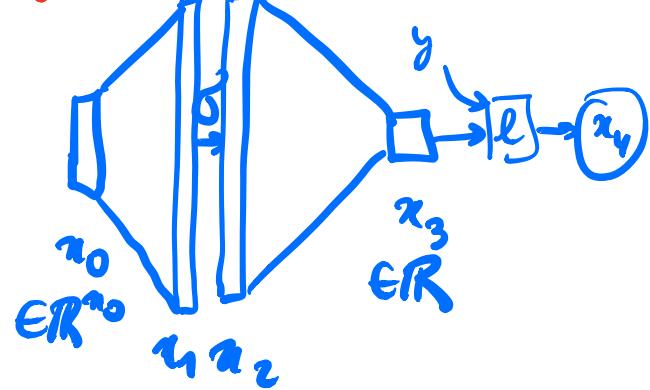
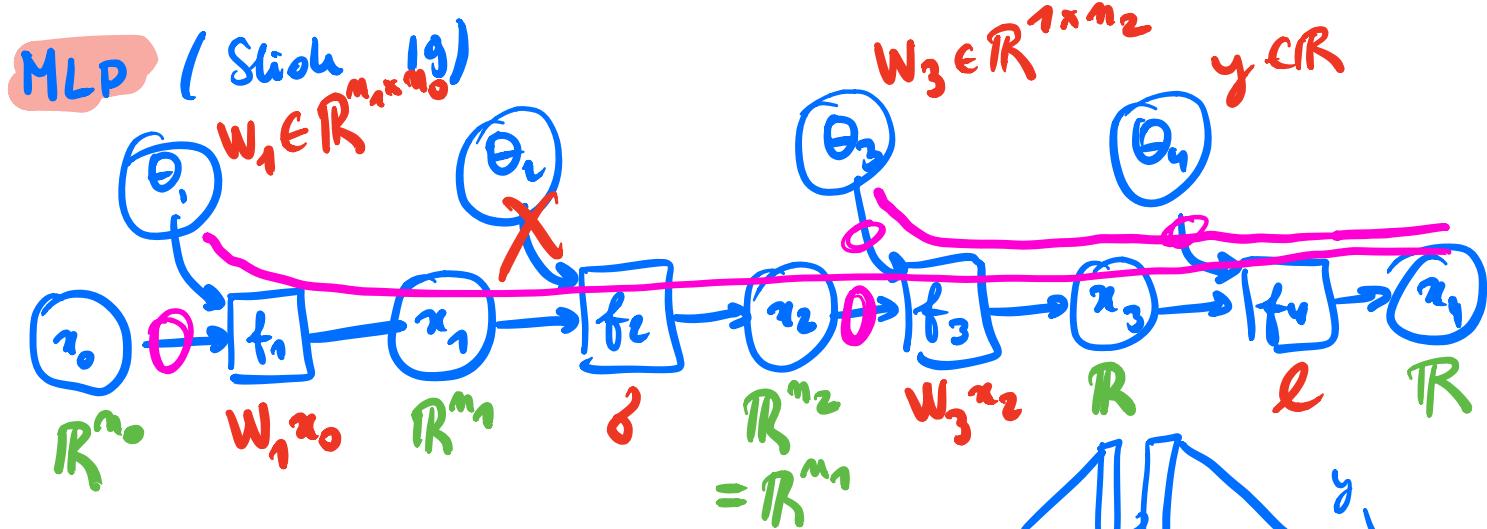
$$\Rightarrow m_0 \sum_{k=1}^{t-1} m_k m_{k+1} = \left(m_0 \sum_{k=1}^{t-2} m_k m_{k+1} \right) + m_0 m_{t-1} m_t$$

Forward / Backward accumulation



$$\Rightarrow m_t \sum_{k=0}^{t-2} m_k a_{k+1} = \left(m_t \sum_{k=1}^{t-2} m_k a_{k+1} \right) + m_t m_0 a_1$$





We want $\frac{\partial z_4}{\partial \theta_3}$ and $\frac{\partial z_4}{\partial \theta_1}$

$$\frac{\partial z_4}{\partial \theta_3} = \frac{\partial z_4}{\partial z_3} \cdot \frac{\partial z_3}{\partial \theta_3}$$

$\theta_3 := W_3 \in \mathbb{R}^{1 \times m_2}$
 $| \theta_3 | = m_2$

$$= \left[(z_3 - y) \right] \left[\begin{matrix} & \\ & \end{matrix} \right]$$

$z_3 = W_3 z_2$
 $= \boxed{W_3} \boxed{z_2}$

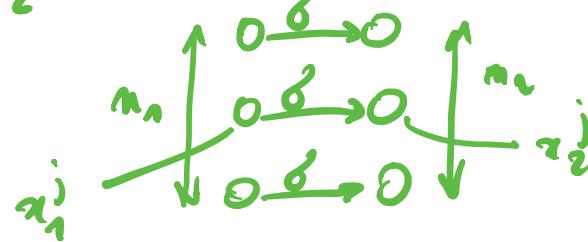
$$= \left[x_2^{(1)} \dots x_2^{(m_2)} \right]$$

$\frac{\partial z_3}{\partial W_3^{1,j}} = x_2^{(j)}$
 $= \sum_j W_3^{1,j} x_2^{(j)}$

$$\frac{\partial z_4}{\partial \theta_1} = \frac{\partial z_4}{\partial z_3} \frac{\partial z_3}{\partial z_2} \frac{\partial z_2}{\partial z_1} \frac{\partial z_1}{\partial \theta_1} \quad \theta_1 := W_1 \in \mathbb{R}^{n_1 \times n_0} \quad |W_1| = n_1 n_0$$

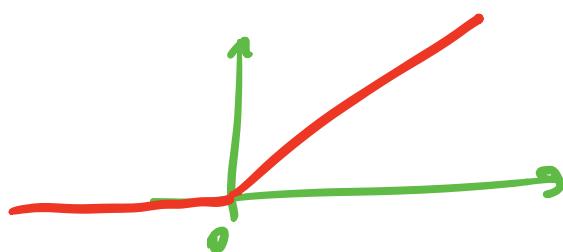
$$= [(z_3 - g)] \begin{bmatrix} 1 \times n_2 \\ ; \end{bmatrix} \begin{bmatrix} n_2 \times n_1 \\ [1] \end{bmatrix} \begin{bmatrix} n_1 \times |W_1| \\ \begin{array}{c} \text{---} \\ \text{---} \end{array} \end{bmatrix}^i$$

$\frac{\partial z_3}{\partial z_2^{(j)}} = W_3^{(j)}$



$$\delta := \text{ReLU}$$

$$z_1 > 0$$



$$z_1 = W_1 z_0$$

$$\hat{z}_1 = \begin{bmatrix} n_0 \\ \vdots \\ n_0 \end{bmatrix} \quad \hat{z}_0 = \begin{bmatrix} 1 \\ \vdots \\ 1 \end{bmatrix} n_0$$

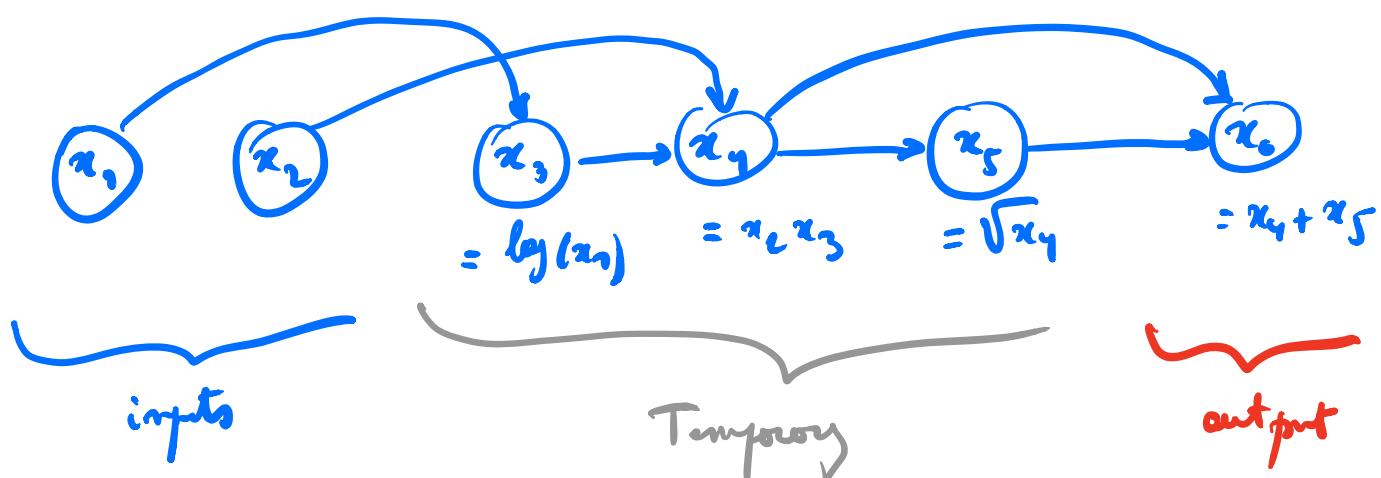
$$z_1^{(i)} = \sum_j W_1^{i,j} z_2^{(j)} \quad \frac{\partial z_1}{\partial W_1^{i,j}}$$

$$\begin{bmatrix} z_0 & 0 & 0 \\ 0 & z_0 & 0 \\ 0 & 0 & z_0 \end{bmatrix} \quad \overbrace{\quad}^{\mathbf{z}_0^T \otimes \mathbf{1}}$$

Forward mode AD

(Slide 25)

$$f(x_1, x_2) = x_2 \log(x_1) + \sqrt{x_2 \log(x_1)}$$



$$\frac{\partial x_6}{\partial x_1} :$$

$$\frac{\partial x_6}{\partial x_1}$$

$$\frac{\partial x_6}{\partial x_1} = 11$$

$$\frac{\partial x_6}{\partial x_2} = 0$$

$$\frac{\partial x_6}{\partial x_1} = \left[\frac{\partial x_6}{\partial x_3} \right] \frac{\partial x_3}{\partial x_1} = \frac{1}{x_1} \cdot 1$$

$$\begin{aligned} \frac{\partial x_6}{\partial x_1} &= \left[\frac{\partial x_6}{\partial x_4} \right] \frac{\partial x_4}{\partial x_1} + \left[\frac{\partial x_6}{\partial x_5} \right] \frac{\partial x_5}{\partial x_1} \\ &= x_2 \frac{\partial x_3}{\partial x_1} + 0 \end{aligned}$$

$$\boxed{\frac{\partial x_k}{\partial x_i} = \sum_{l \in \text{points}(k)} \left[\frac{\partial x_k}{\partial x_l} \right] \frac{\partial x_l}{\partial x_i}}$$

$$\frac{\partial x_5}{\partial x_1} = \left[\frac{\partial x_5}{\partial x_4} \right] \frac{\partial x_4}{\partial x_1}$$

$$= \frac{1}{2} \cdot \frac{1}{\sqrt{x_4}} \cdot \frac{\partial x_4}{\partial x_1}$$

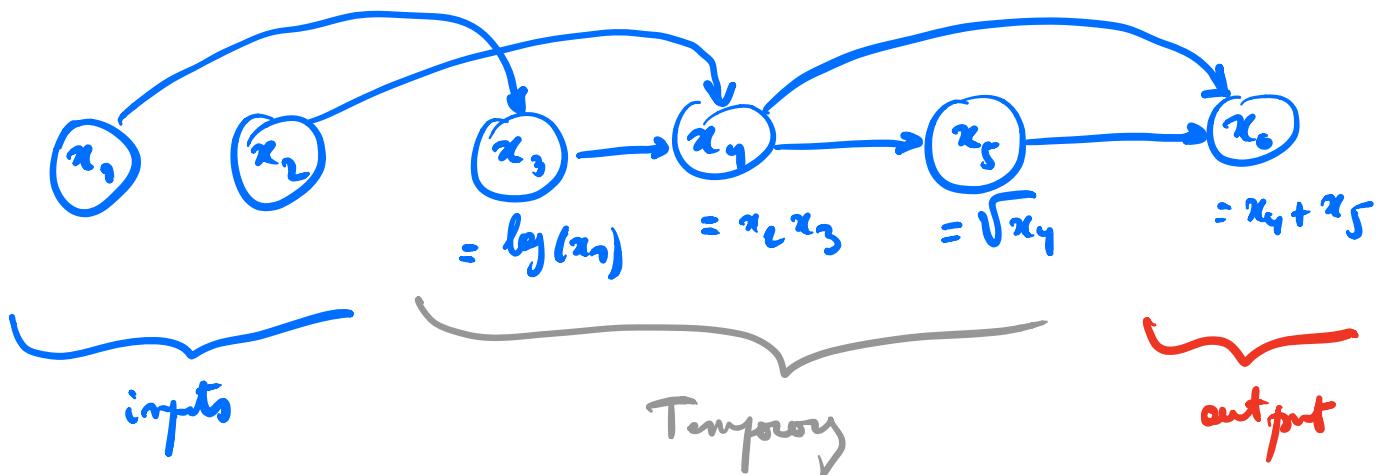
$$\frac{\partial x_6}{\partial x_1} = \left[\frac{\partial x_6}{\partial x_4} \right] \frac{\partial x_4}{\partial x_1} + \left[\frac{\partial x_6}{\partial x_5} \right] \frac{\partial x_5}{\partial x_1}$$

$$= 1 \cdot \frac{\partial x_4}{\partial x_1} + 1 \cdot \frac{\partial x_5}{\partial x_1}$$

$$\frac{\partial x_6}{\partial x_2} : \quad (\dots)$$

Backward mode of AD (Slide 29)

$$f(x_1, x_2) = x_2 \log(x_1) + \sqrt{x_2 \log(x_1)}$$



$$\frac{\partial x_6}{\partial x_6} = 1$$

$$\frac{\partial x_6}{\partial x_5} = \frac{\partial x_6}{\partial x_6} \left[\frac{\partial x_6}{\partial x_5} \right]$$

$$= \frac{\partial x_6}{\partial x_6} \cdot 1$$

$$\frac{\partial x_6}{\partial x_4} = \frac{\partial x_6}{\partial x_5} \cdot \left[\frac{\partial x_5}{\partial x_4} \right] + \frac{\partial x_6}{\partial x_6} \left[\frac{\partial x_6}{\partial x_4} \right]$$

$$= \frac{\partial x_6}{\partial x_5} \cdot \frac{1}{2} \frac{1}{\sqrt{x_4}} + \frac{\partial x_6}{\partial x_6} \cdot 1$$

$$\frac{\partial x_6}{\partial x_3} = \frac{\partial x_6}{\partial x_4} \left[\frac{\partial x_4}{\partial x_3} \right]$$

$$= \frac{\partial x_6}{\partial x_4} \cdot x_2$$

$$\frac{\partial x_6}{\partial x_2} = \frac{\partial x_6}{\partial x_4} \left[\frac{\partial x_4}{\partial x_2} \right]$$

$$= \frac{\partial x_6}{\partial x_4} \cdot x_3$$

$$\frac{\partial x_6}{\partial x_1} = \frac{\partial x_6}{\partial x_4} \left[\frac{\partial x_3}{\partial x_1} \right] = \frac{\partial x_6}{\partial x_3} \cdot \frac{1}{x_1}$$

$$\frac{\partial x_t}{\partial x_k} = \sum_m^m \frac{\partial x_t}{\partial x_m} \left[\frac{\partial x_m}{\partial x_k} \right]_{\in \text{child}(k)}$$

We get both at once!

