Dense

$$y = f(x)$$

$$x = \begin{pmatrix} 3 \\ -1 \\ 2 \end{pmatrix} \qquad y = \underbrace{\begin{pmatrix} 1 & 2 & 3 \\ 3 & 2 & 1 \\ -1 & -2 & -1 \end{pmatrix}}_{} \cdot x + \begin{pmatrix} -1 \\ 0 \\ -7 \end{pmatrix}$$

$$W \cdot x + b$$

Offset Bias

$$y_j = \sum_{i=0}^{I-1} w_{ji} \cdot x_i + b_j$$

↳ lower case i: the running index
upper case I: the limit

$$\left(\left(\tanh x = \frac{e^x - e^{-x}}{e^x + e^{-x}} \right)\right)$$

$$y_j = f(x) = \frac{e^{x_j}}{\sum_{i=0}^{I-1} e^{x_i}}$$

$$x = \begin{pmatrix} 0.2 \\ 0.4 \\ 0.1 \end{pmatrix} \qquad \begin{pmatrix} e^{0.2} \\ e^{0.4} \\ e^{0.1} \end{pmatrix} = \begin{pmatrix} 1.22 \\ 1.49 \\ 1.11 \end{pmatrix}$$

$$y = \begin{pmatrix} \frac{1.22}{1.22 + 1.49 + 1.11} \\ \\ \end{pmatrix} = \begin{pmatrix} 0.32 \\ 0.39 \\ 0.29 \end{pmatrix} \begin{matrix} j=0 \\ j=1 \\ j=2 \end{matrix}$$

$$\frac{dy_1}{dx_1} = \frac{d}{dx_1} \frac{e^{x_1}}{e^{x_0} + e^{x_1} + e^{x_2}} = \frac{e^{x_1}(e^{x_0} + e^{x_1} + e^{x_2}) - e^{x_1} \cdot e^{x_1}}{(e^{x_0} + e^{x_1} + e^{x_2})^2}$$

$$= \frac{e^{x_1}}{e^{x_0} + e^{x_1} + e^{x_2}} - \frac{e^{x_1}}{e^{x_0} + e^{x_1} + e^{x_2}} \cdot \frac{e^{x_1}}{e^{x_0} + e^{x_1} + e^{x_2}}$$

$$= y_1 - y_1 \cdot y_1$$

$$= y_1 \quad (1 - y_1)$$

$$\frac{dy_1}{dx_0} = \frac{d}{dx_0} \frac{e^{x_1}}{e^{x_0} + e^{x_1} + e^{x_2}} = e^{x_1} \frac{d}{dx_0} (e^{x_0} + e^{x_1} + e^{x_2})^{-1}$$

$$= -e^{x_1} (e^{x_0} + e^{x_1} + e^{x_2})^{-2} \cdot e^{x_0}$$

$$= \frac{-e^{x_1} e^{x_0}}{(e^{x_0} + e^{x_1} + e^{x_2})^2}$$

$$= - \frac{e^{x_1}}{e^{x_0} + e^{x_1} + e^{x_2}} \cdot \frac{e^{x_0}}{e^{x_0} + e^{x_1} + e^{x_2}}$$

$$= - y_1 \cdot y_0$$

$$= y_1 (0 - y_0)$$

$$\left.\begin{aligned} \frac{dy_1}{dx_1} &= y_1 (1 - y_1) \\ \frac{dy_1}{dx_0} &= y_1 (0 - y_0) \end{aligned}\right\} \frac{dy_j}{dx_i} = y_j \cdot (\delta_{ji} - y_i)$$

$$\delta_{ji} = \begin{cases} 1 & \text{for } i = j \\ 0 & \text{for } i \neq j \end{cases}$$

Delta

$$\frac{dy_1}{dx_2} =$$

backward procedure

$$\frac{dL}{dy_j} \longrightarrow \frac{dL}{dx_i} = \sum_{j=0}^{J-1} \frac{dL}{dy_j} \cdot \frac{dy_j}{dx_i}$$

$$f(x) = \sin(x^2) \quad = \cos(x^2) \cdot \; 2x$$

$$z = \sin(y) \qquad \frac{dz}{dy} = \cos(y)$$

$$y = x^2 \qquad \frac{dy}{dx} = 2x$$

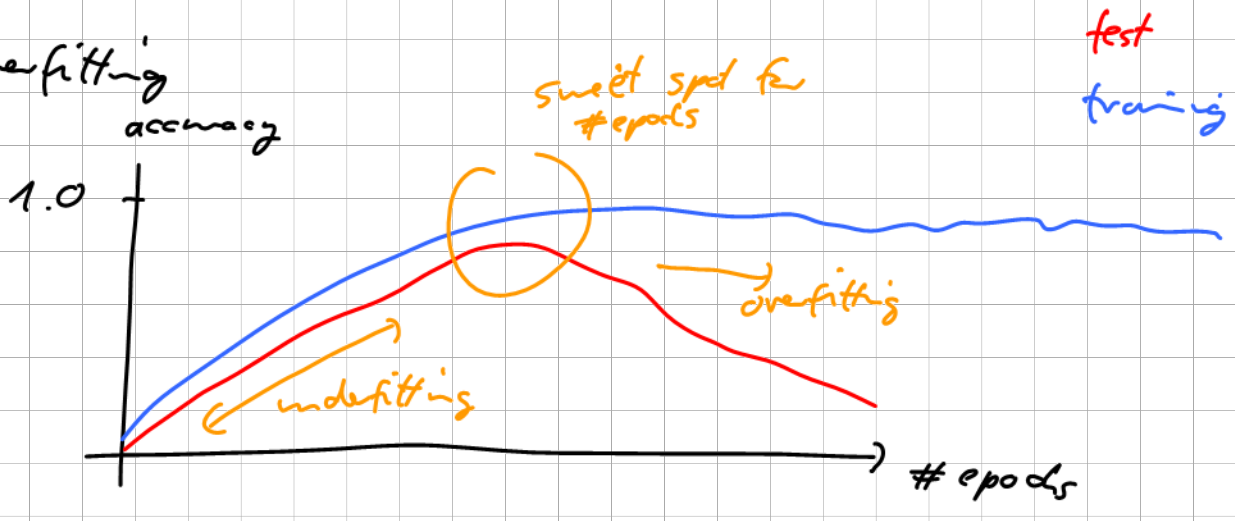
$$\frac{dz}{dx} = \frac{dz}{dx} \cdot \underbrace{\frac{dy}{dy}}_{\cdot 1} = \frac{dz}{dy} \cdot \frac{dy}{dx} = \cos(y) \cdot 2x$$

$$= \cos(x^2) \cdot 2x$$

$$\frac{dL}{dx_i} = \sum_{j=0}^{J-1} \underbrace{\frac{dL}{dy_j}}_{\texttt{d\_L\_d\_out}} \cdot \underbrace{y_j \cdot (\delta_{ji} - y_i)}_{\frac{dy_j}{dx_i}}$$

```
def    Delta(i, j):
    if  i == j:
        return 1
    else
        return 0
```

Overfitting



		prediction	
test_data	shirt	shirt	a)
	boot	shirt	b)
	shirt	shirt	c)
	umbrella	umbrella	d)
	⋮	⋮	

a)

CM	shirt
shirt	+1
boot	+1

b)