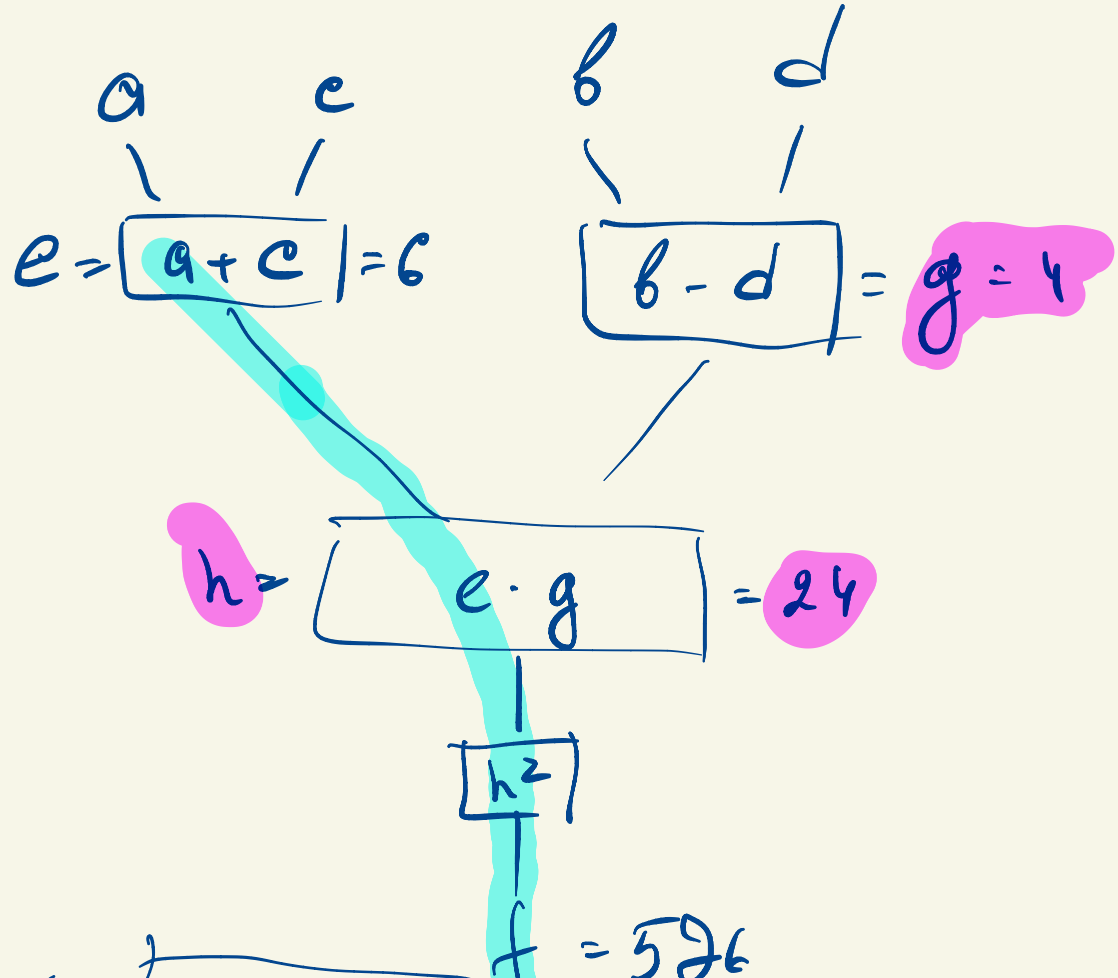
previously on ML4ES-2

$$f = ((\underset{1}{a} + \underset{5}{c}) \cdot (\underset{7}{b} - \underset{3}{d}))^2$$



$$\frac{\partial f}{\partial a} = \boxed{\frac{\partial f}{\partial h}\frac{\partial h}{\partial e}\frac{\partial e}{\partial a}} + \cancel{\frac{\partial f}{\partial h}\frac{\partial h}{\partial g}\frac{\partial g}{\partial a}} = 0$$

$$\frac{\partial f}{\partial h} = 2h = 48 \qquad \frac{\partial h}{\partial e} = g = 4 \qquad \frac{\partial e}{\partial a} = 1 \qquad \frac{\partial f}{\partial a} = 192$$

Выводы:

1. Если есть граф вычислений, то легко можно посчитать $\frac{\partial f}{\partial \cdot}$ по любой переменной этого графа

2. Кэшировать промежуточные результаты вычислений полезно

3. Можно считать $\frac{\partial}{\partial \theta}$, $\frac{\partial}{\partial x}$, ...

2a. Чтобы посчитать $\frac{\partial f}{\partial \cdot}$ хорошо (но не обяз-но) посчитать f

$$\hat{y} = \mu = \theta^T x$$

θ, x – вектор-столбцы

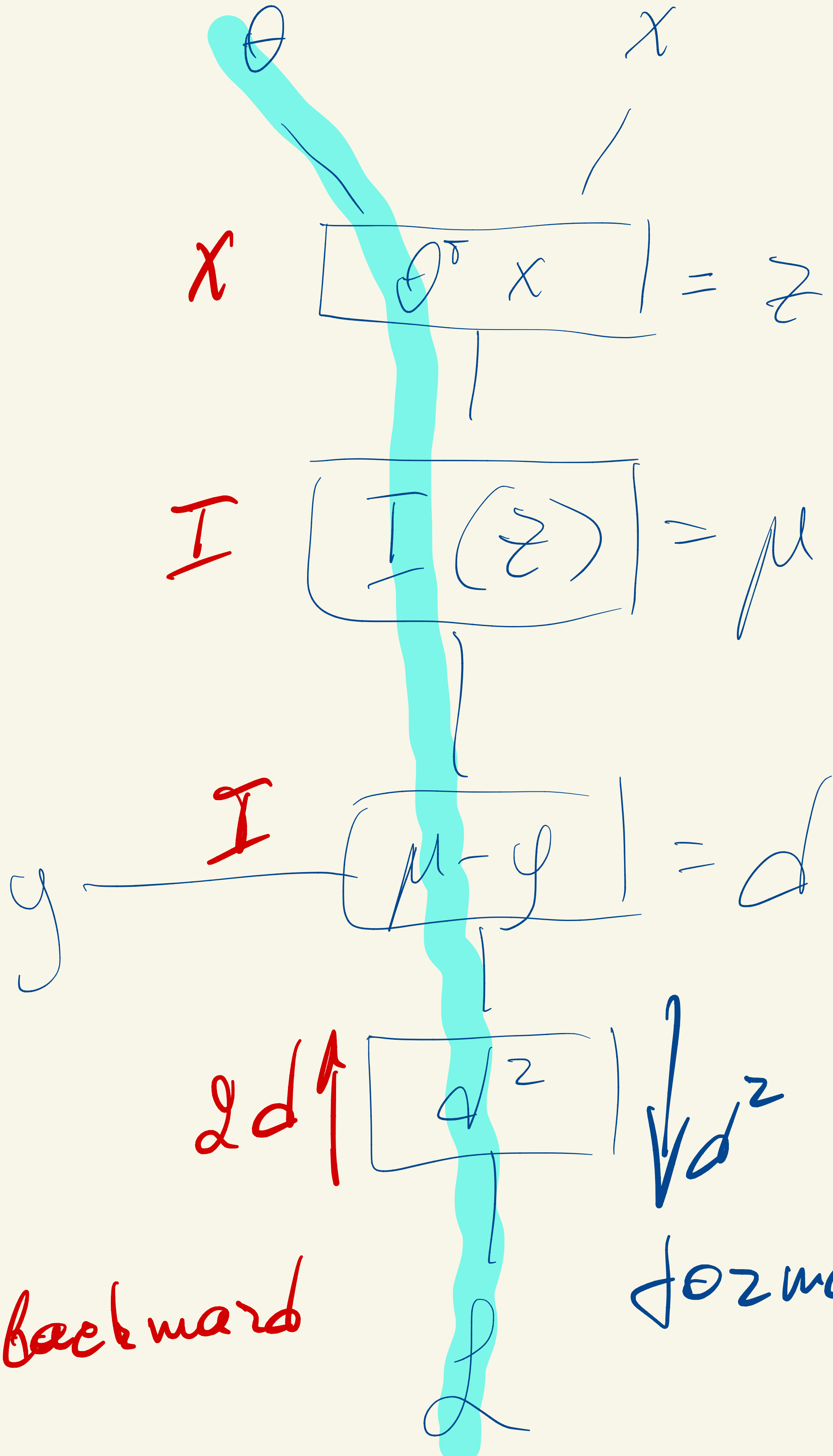
$\underline{\underline{\theta}}$ $\quad x$

$$\theta^T x$$

$$\frac{\partial(\theta^T x)}{\partial \theta} = \ldots = x$$

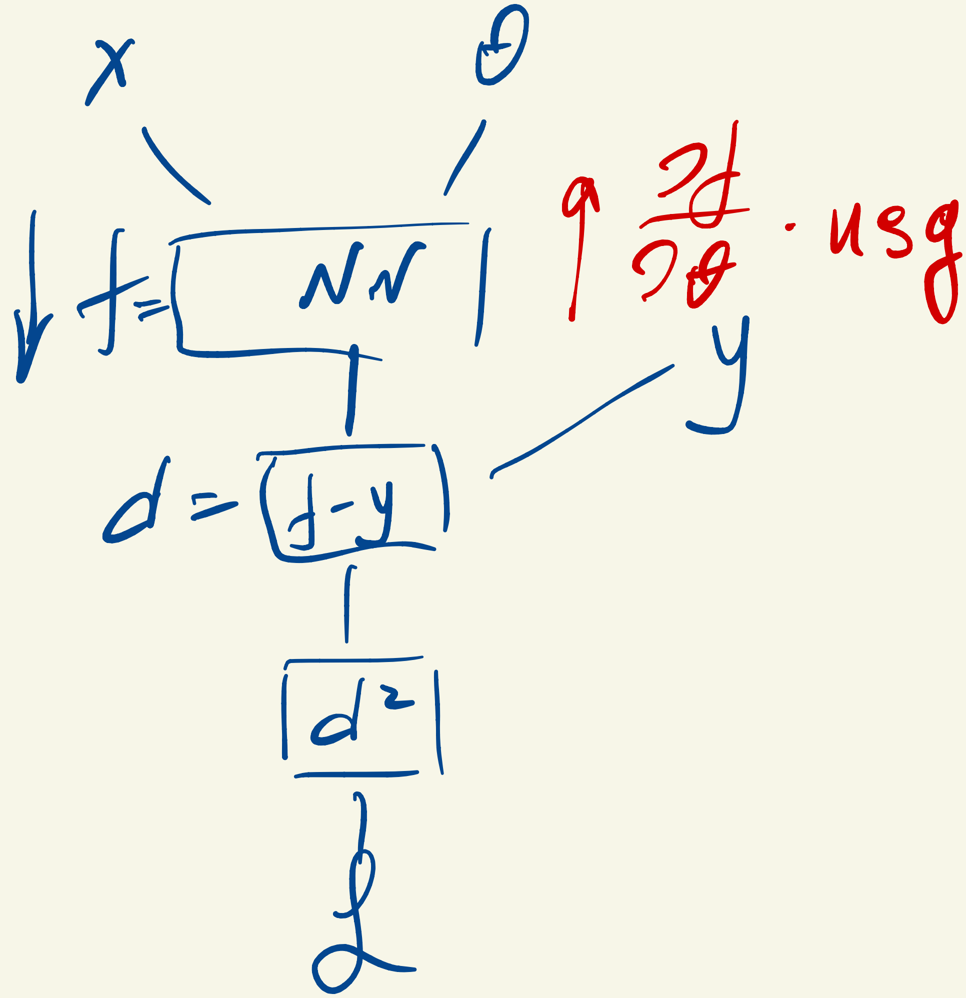
$$\mu = I(\theta^T x)$$

$$\mathcal{L} = \frac{1}{N}\sum_{i=1}^{N}(\mu_i - y_i)^2$$



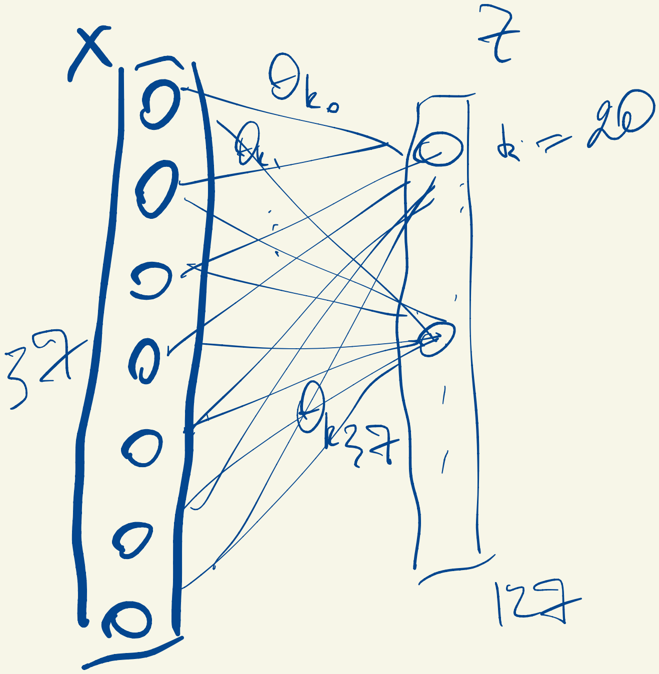
$$\frac{\partial \ell}{\partial \theta} = \frac{\partial \ell}{\partial d} \cdot \frac{\partial d}{\partial \mu} \cdot \frac{\partial \mu}{\partial z} \cdot \frac{\partial z}{\partial \theta}$$

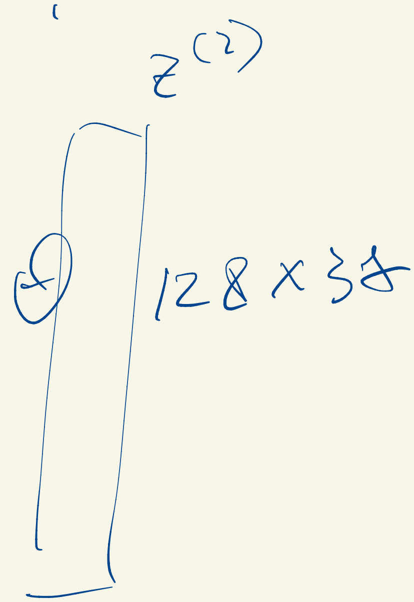
$2d \quad I \quad I \quad X$

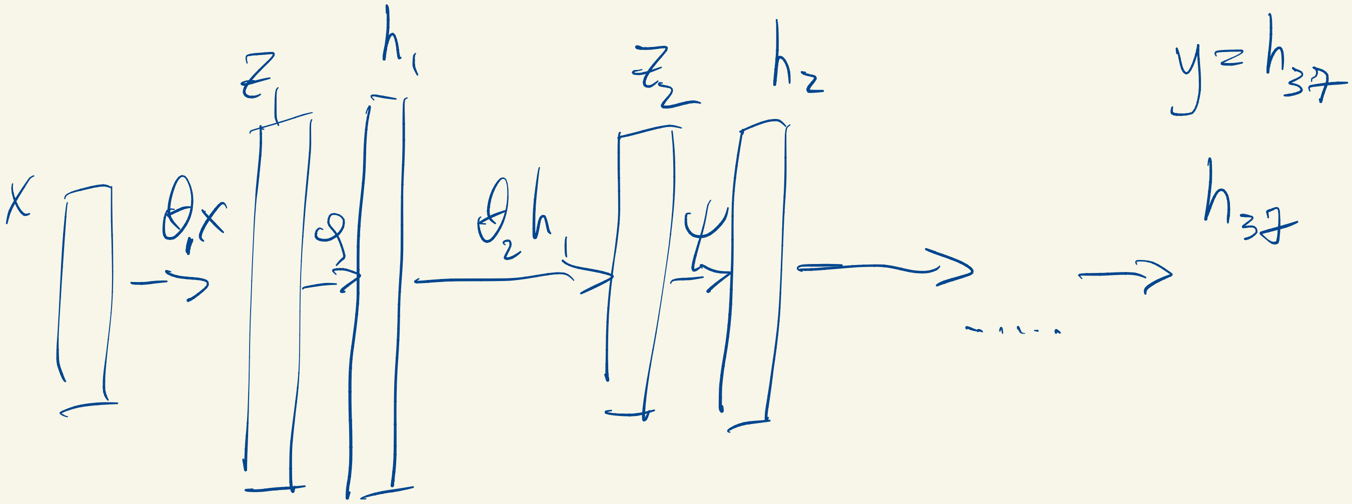

x
θ
f = NN
y
d = f - y
d²
ℒ
∂f/∂θ · usg

$$Z = \Theta^T X$$

$$Z_k = \sum_i \Theta_{ki} X_i$$







$$y = \varphi(\ldots \varphi(\theta_3 \varphi(\theta_2 \varphi(\theta_1 x))))$$

$$y = kx + b \qquad \theta = \begin{pmatrix} k \\ b \end{pmatrix} \qquad 128 \times 37$$

$$\mathcal{L} = \sum (y - y_t)^2 = \sum_{i=1}^{N} (NN(x_i) - y_i)^2$$

$$\theta_1 \; \theta_2 \; \theta_3$$

$\theta \quad \varphi$

$\downarrow \xi$

$z = \boxed{f(\xi)}$

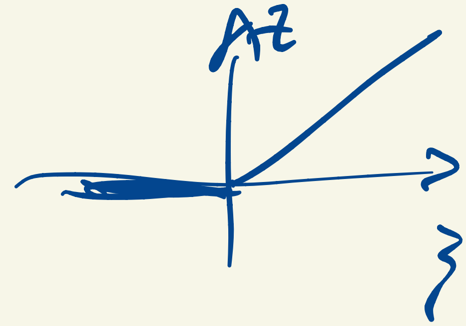
$\downarrow \qquad \uparrow \mathrm{usg}$

$\downarrow$

$\mathcal{L}$

$$\frac{\partial \mathcal{L}}{\partial \xi} = \mathrm{usg} \cdot \frac{\partial z}{\partial \xi}$$

$$z = f(\xi) = e^{\xi}$$



$$\frac{\partial z}{\partial \xi} = e^{\xi}$$

$$z = f(\xi) = \mathrm{ReLU}(\xi)$$

$$\frac{\partial z}{\partial \xi} = \begin{cases} 0, & \xi < 0 \\ 1, & \xi \geq 0 \end{cases}$$