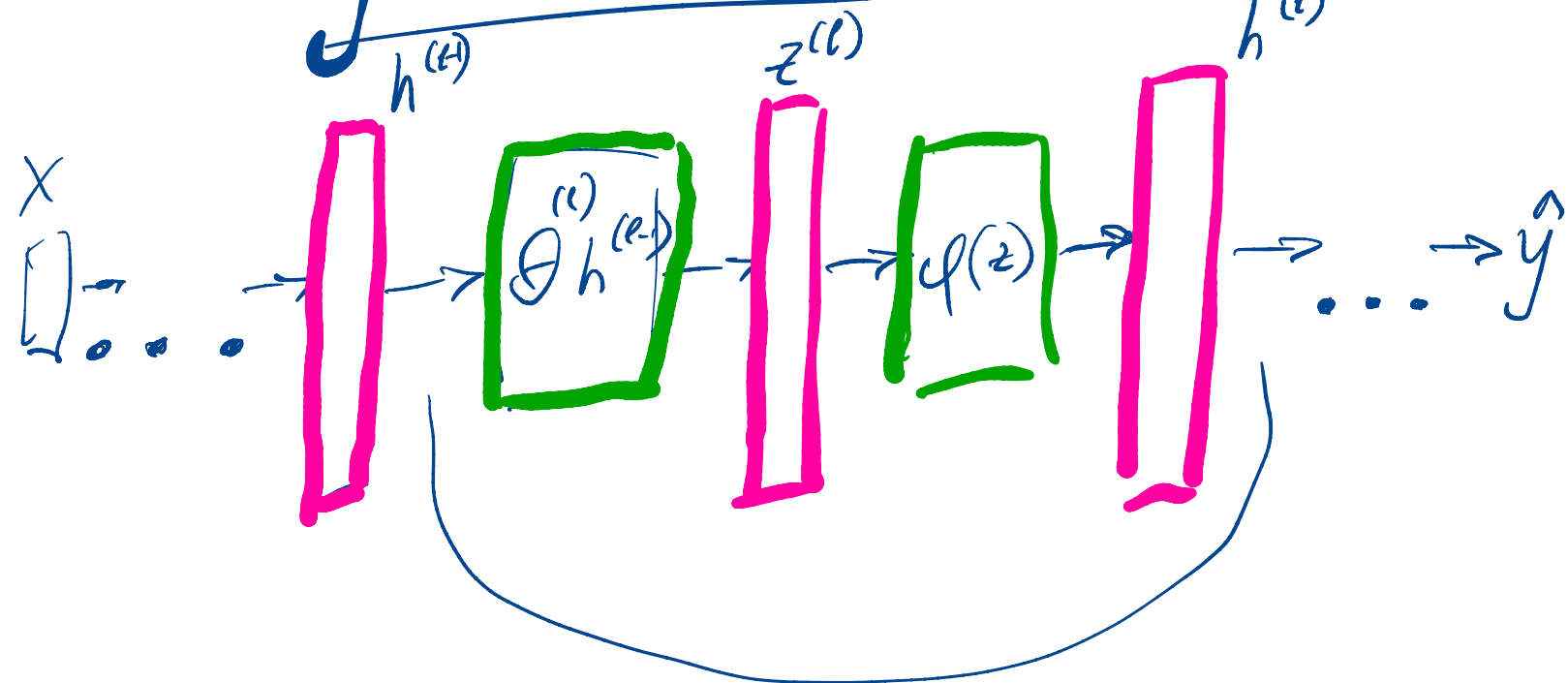
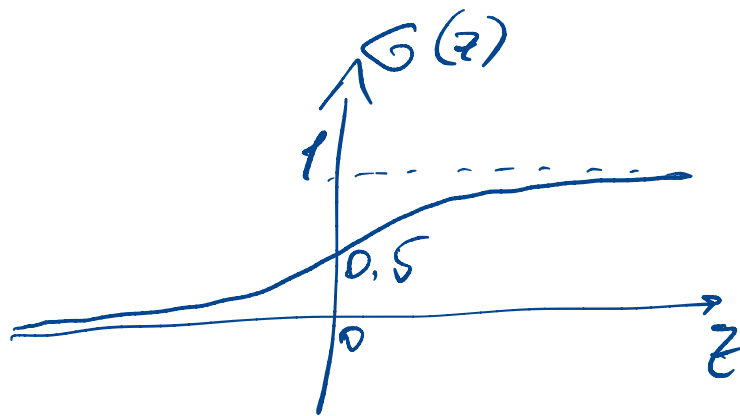


Функции активации

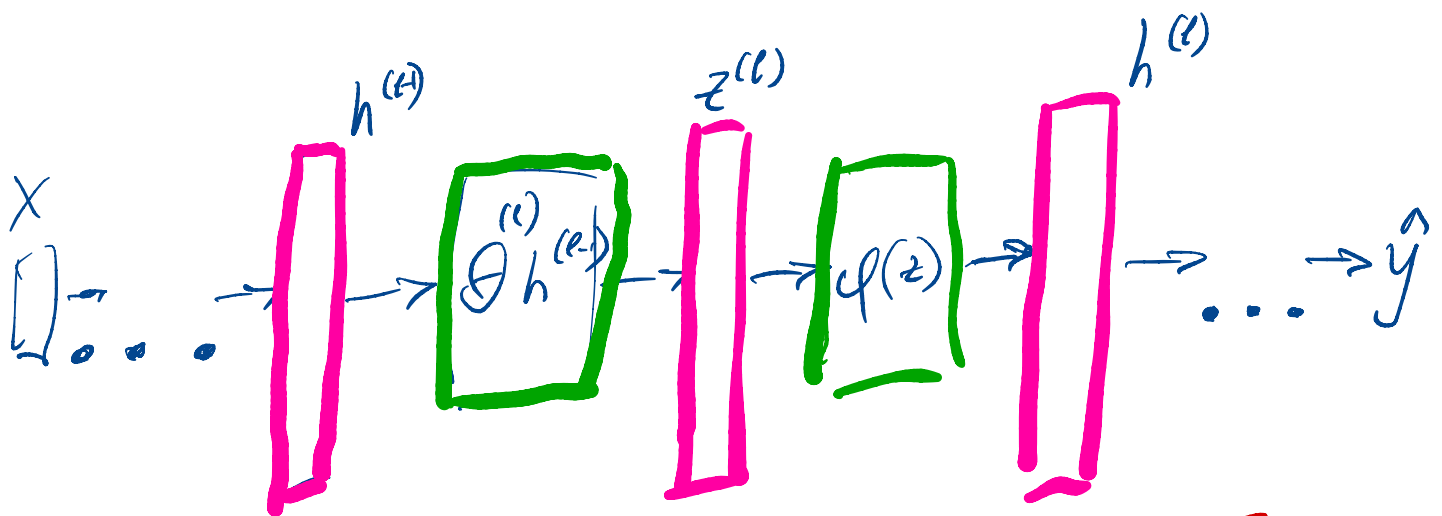


$\sigma(z)$:

$$\sigma(z) = \frac{1}{1 + e^{-z}}$$



$$\frac{d\sigma(z)}{dz} = \sigma(z)(1 - \sigma(z))$$



$$\frac{\partial L}{\partial \theta^{(l)}} = \text{usg} \cdot \frac{\partial \varphi(z)}{\partial z} \cdot h^{(l-1)} \leq 0.25 \leq 1$$

← usg

← ← ←

$$\frac{\partial L}{\partial \theta^{(L)}} \leq 0.25$$

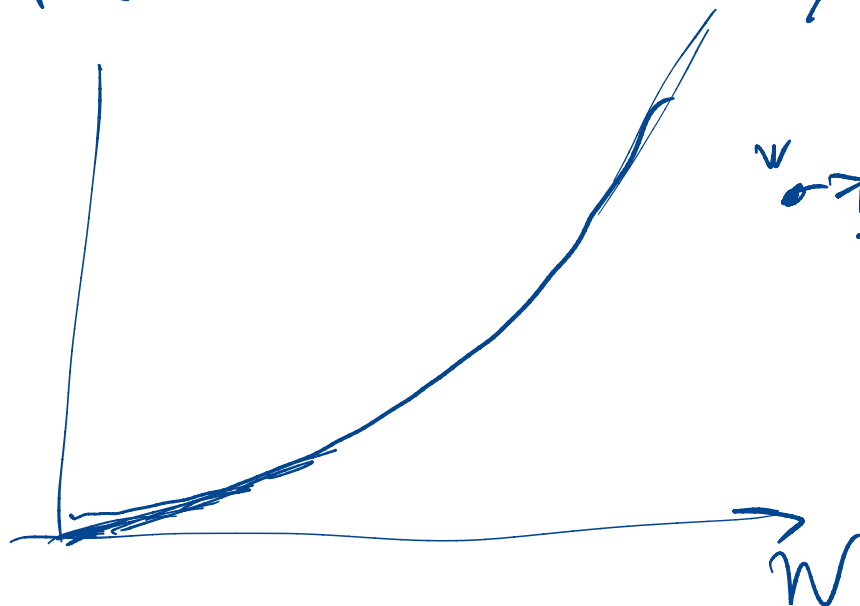
$$\frac{\partial L}{\partial \theta^{(L-1)}} \leq 0.25^2$$

...

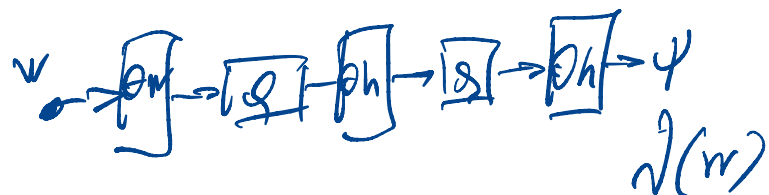
$$\frac{\partial L}{\partial \theta^{(L-k)}} \leq 0.25^{k+1}$$

Эффект затухающих градиентов

$J(w)$



$$J(w) = NN^{(3)}(w)$$

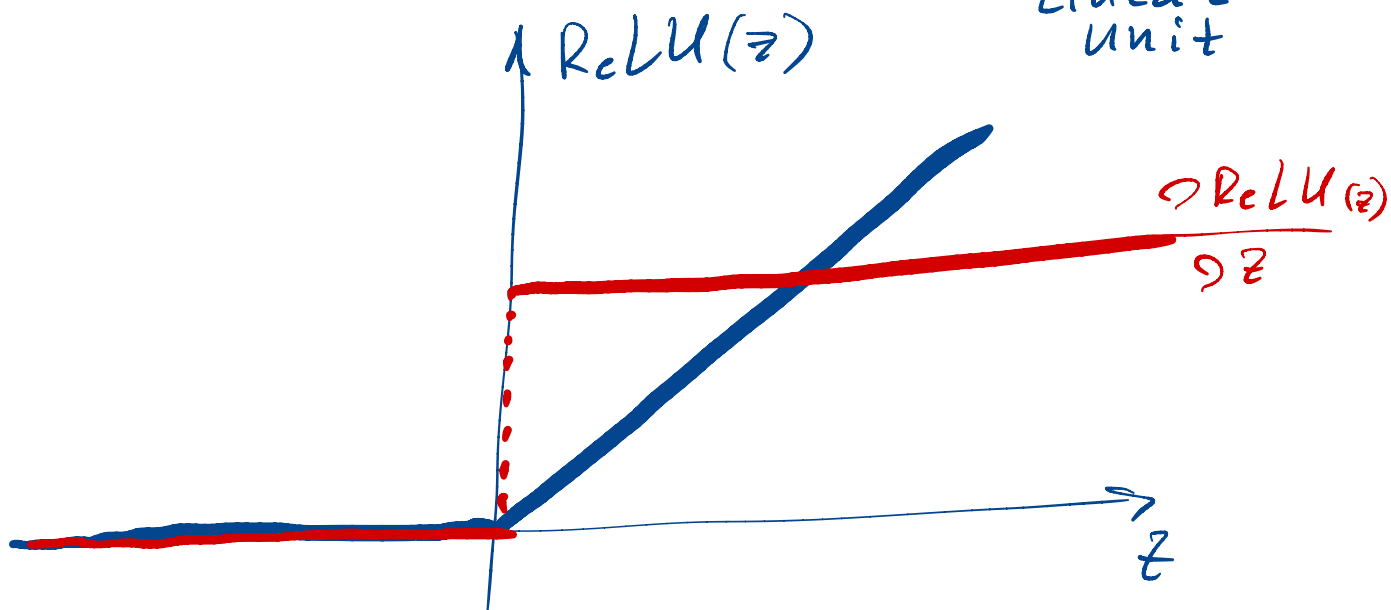


$$\tanh(z) = \frac{e^z - e^{-z}}{e^z + e^{-z}} = 2G(2z) - 1$$

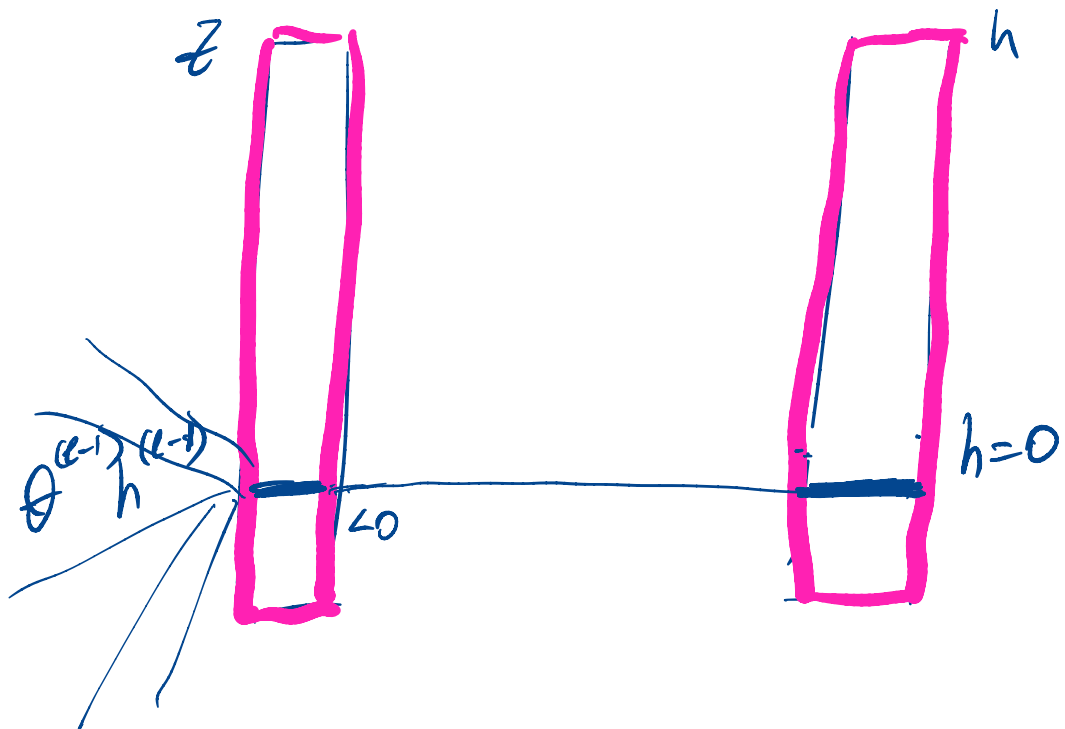
$$\frac{\partial \tanh(z)}{\partial z} = 4G(2z)(1 - G(2z))$$

$$\text{ReLU}(z) = \max(0, z)$$

ReLU =
rectified
linear
unit



$$\frac{\partial \text{ReLU}(z)}{\partial z} = \begin{cases} 1, & \text{if } z \geq 0 \\ 0, & \text{if } z < 0 \end{cases}$$



ELU

$$\text{ELU}(z) = \begin{cases} z, & z \geq 0 \\ 2(e^z - 1), & \text{if } z < 0 \end{cases}$$

$$\text{Если } z \sim \mathcal{N}(0; 1)$$

$$\theta \sim \mathcal{N}\left(0; \frac{\sqrt{6}}{\sqrt{n+1}}\right)$$

$$\text{тогда } \mathbb{E} g(z) = \mathbb{E}(z)$$

Scaled ELU (SELU)

$$\psi(z) = \lambda \begin{cases} z, & z \geq 0 \\ 2(e^z - 1), & z < 0 \end{cases} \quad \begin{aligned} \lambda &= 1.0507 \\ \lambda &= 1.6752 \end{aligned}$$

$$z \sim \mathcal{N}(0, 1)$$

$$\theta \sim \mathcal{N}\left(0, \frac{1}{\sqrt{n}}\right)$$

$$\begin{aligned} \mathbb{E}(\psi(z)) &= \mathbb{E}(z) \\ \Rightarrow \text{Var}(\psi(z)) &= \text{Var}(z) \end{aligned}$$

↳ размерность

входных данных

$$\text{сход} = \sqrt{\ln(n^{(p-1)})}$$