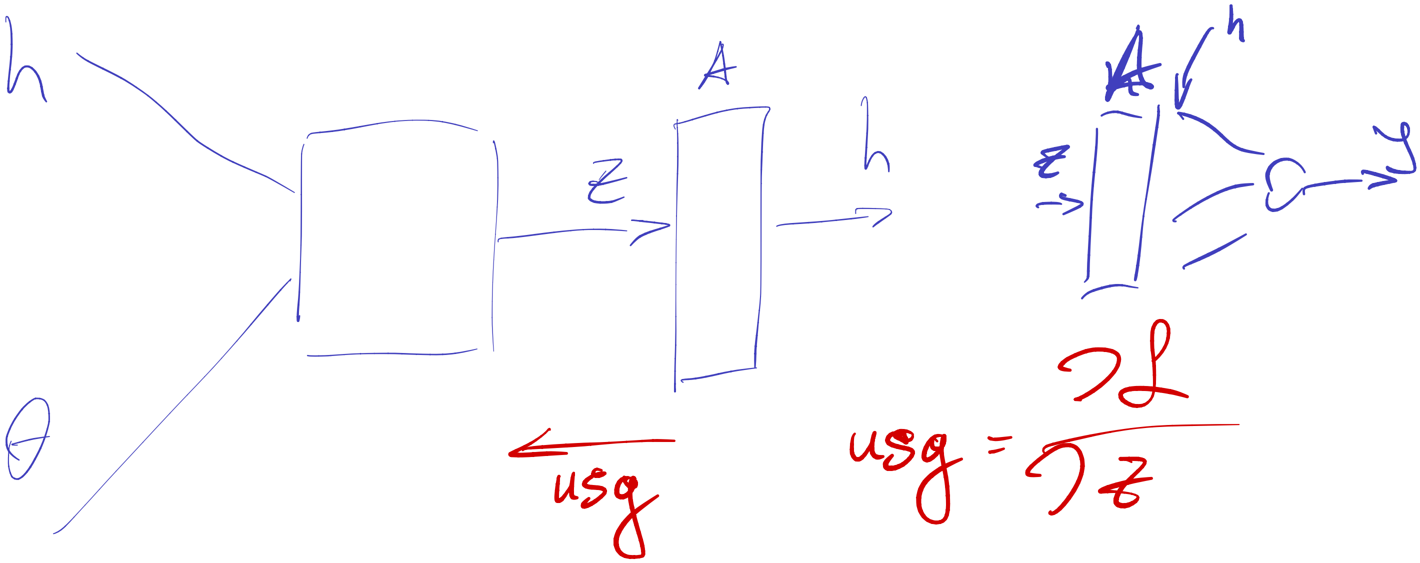


$$usg = \frac{\partial \mathcal{L}}{\partial z}$$

$$z = \theta^T h^{(l-1)}$$

$$\frac{\partial \mathcal{L}}{\partial \theta} = usg \cdot h$$

step 0: $\theta = 0$

$$z^{(0)} = \theta^T x = 0$$

$$h^{(0)} = 0$$

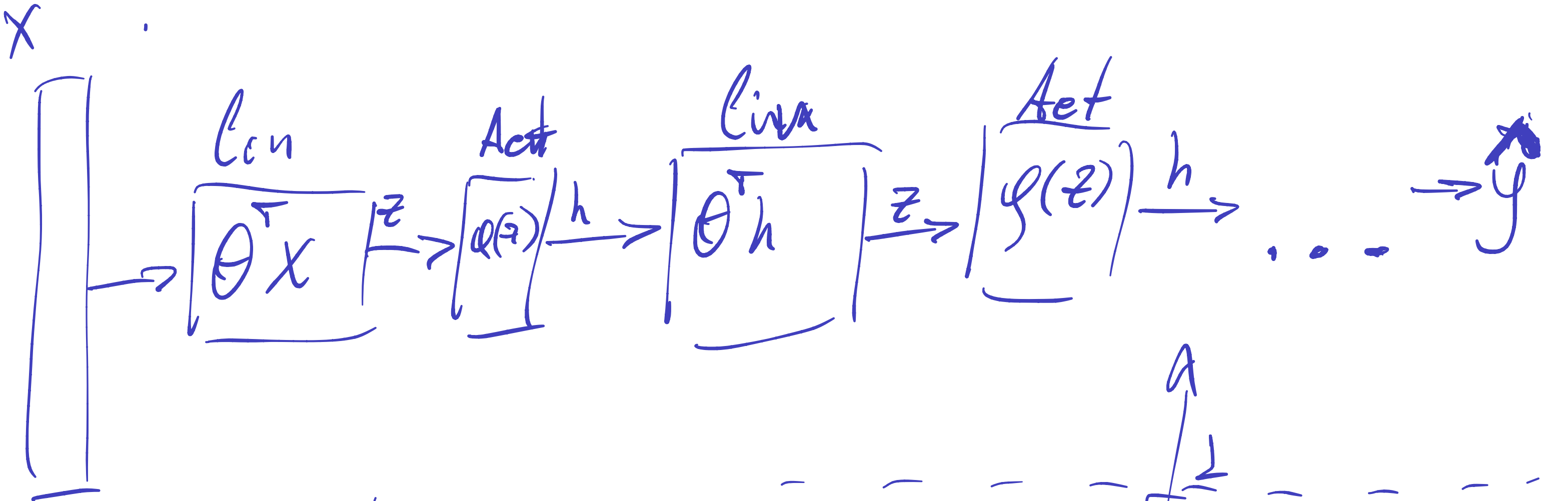
$$\dots$$

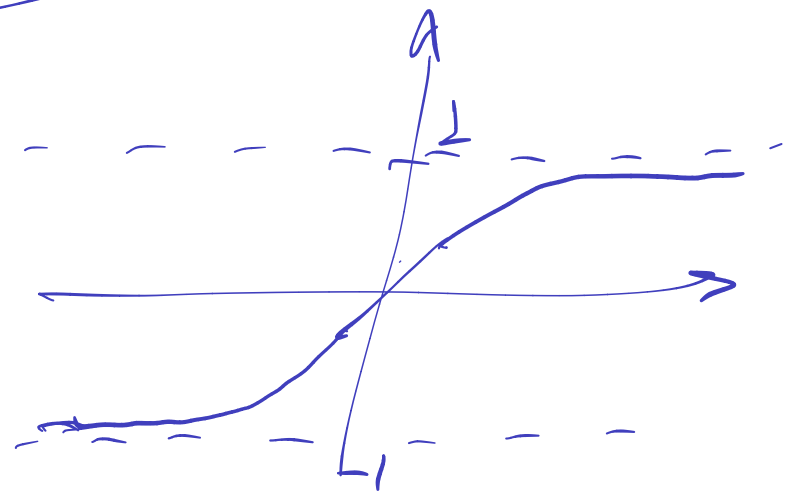
$$z^{(\dots)} = 0$$

$$h^{(\dots)} = 0$$

$$y = \varphi(\theta h^{(-1)}) \to \mathcal{L}$$

$$\frac{\partial \mathcal{L}}{\partial \theta^{(l)}} = usg \cdot h$$





$\varphi = \tanh(\cdot)$

$z \approx 0: \ \varphi(z) \approx z$

$\Theta^{(l)}: \ n \times m$

$h^{(l-1)}: \ B \times m \qquad z^{(l)} = h^{(l-1)} \cdot \Theta^{(l)}$

$\qquad\qquad\qquad\qquad \ \ B \times n \quad B \times m \quad m \times n$

$z^{(l)}: \ B \times n$

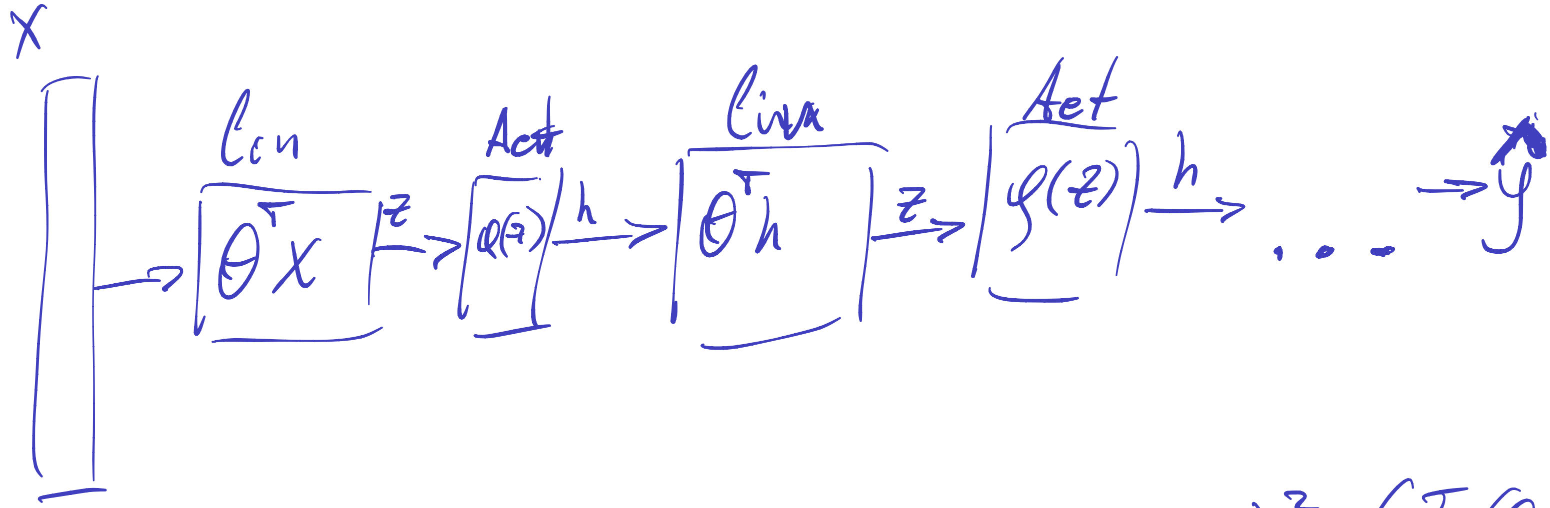
① $x_i \sim$ i.i.d. Independent, identically distributed

$z_i^i \sim$ i.i.d.

② $\Theta \sim$ i.i.d.

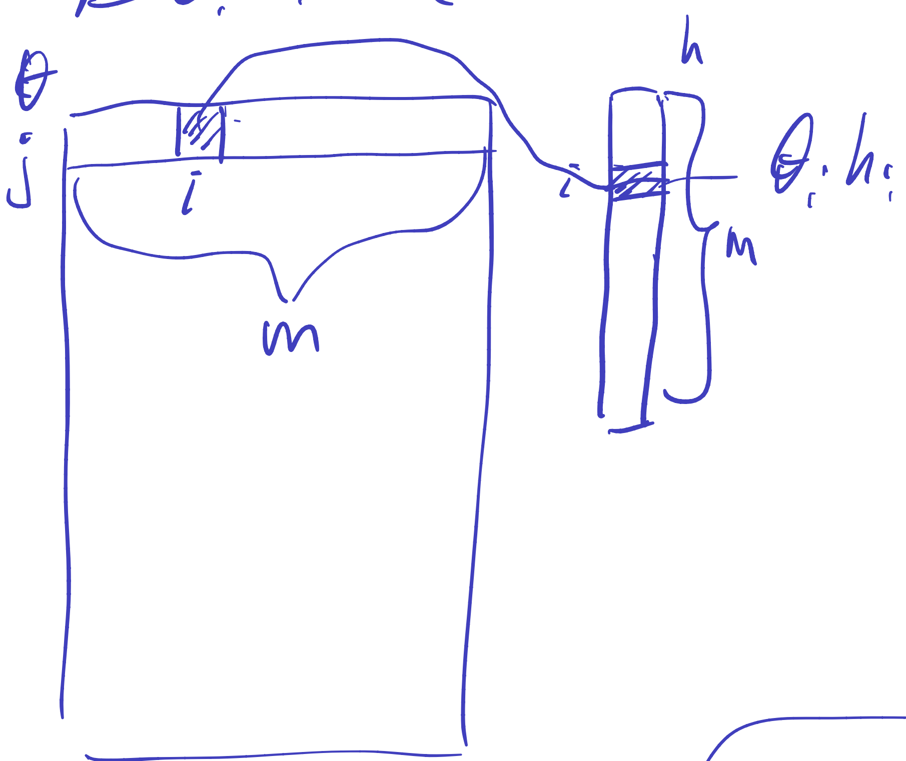
③ $x \perp \Theta$

$h \not\perp \Theta$



$$Var(z_i) = Var(\theta_i h_i) = \mathbb{E}(\theta_i h_i)^2 - (\mathbb{E}(\theta_i h_i))^2$$

$$= \mathbb{E}\theta_i^2 h_i^2 - (\mathbb{E}\theta_i)^2 (\mathbb{E}h_i)^2 \ominus$$



$$\ominus\ \mathbb{E}\theta_i^2\, \mathbb{E}h_i^2 - (\mathbb{E}\theta_i)^2 (\mathbb{E}h_i)^2 =$$

$$= \left(\mathbb{E}\theta_i^2 - (\mathbb{E}\theta_i)^2 + (\mathbb{E}\theta_i)^2\right) \cdot \left(\mathbb{E}h_i^2 - (\mathbb{E}h_i)^2 + (\mathbb{E}h_i)^2\right) - \circledcirc =$$

$$= \left(\mathrm{Var}\,\theta_i + \cancel{(\mathbb{E}\theta_i)^2}\right)\left(\mathrm{Var}\,h_i + \cancel{(\mathbb{E}h_i)^2}\right) -$$

$$- \cancel{(\mathbb{E}\theta_i)^2}\cancel{(\mathbb{E}h_i)^2} = \mathrm{Var}\,\theta_i \, \mathrm{Var}\,h_i$$

$$z_j = \sum_{i=1}^{m} \theta_i h_i$$

$$\mathrm{Var}\, z_j = \sum_{i=1}^{m} \mathrm{Var}\,\theta_i \, \mathrm{Var}\,h_i = \underline{m \, \mathrm{Var}\,\theta \, \mathrm{Var}\,h}$$

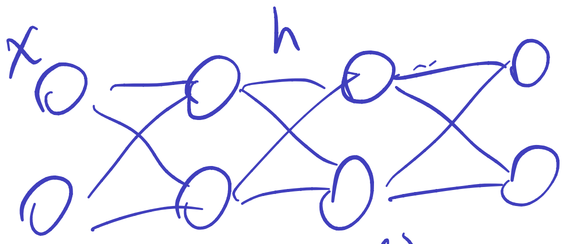
$$h^{(l+1)} = \varphi(z^{(l+1)}) \qquad \mathrm{Var}(h) \approx \mathrm{Var}(z)$$

$$\cancel{\mathrm{Var}\, z} \approx m \, \mathrm{Var}\,\theta \, \cancel{\mathrm{Var}\,h}$$

$$1 = m \, \mathrm{Var}\,\theta$$

$$\mathrm{Var}\,\theta \approx \frac{1}{m}$$

$$\mathrm{Var}(h^{(l+1)}) \approx \mathrm{Var}(h^{(l)})$$



$h^{(1)} = \theta^{(1)} x$

$\theta = 1/2$

$h^{(1)} \approx \frac{1}{2} x$

$h^{(2)} \approx \frac{1}{4} x$

$h^{(n)} \approx \left(\frac{1}{2}\right)^{n} x$

$y \approx \left(\frac{1}{2}\right)^{L} x$

$\theta = 2$

$h^{(1)} \approx 2x$

$h^{(2)} \approx 4x$

$h^{(n)} = 2^{n} x$

$y \approx 2^{L} x$

① $\operatorname{Var} y \approx 0 \Rightarrow \hat{y}$ не зависит от X

② $\operatorname{Var} y \to \infty \Rightarrow \hat{y}$ не зависит от X

$$\theta \sim \mathcal{N}(\mu, \sigma^2)$$

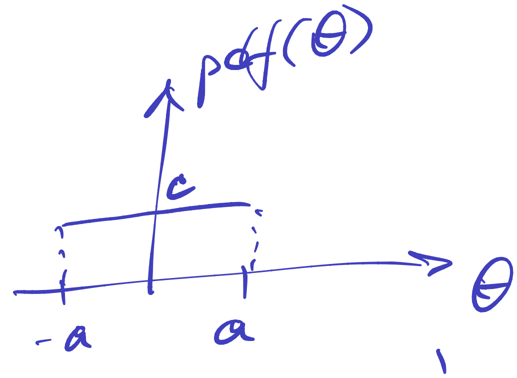
$$\mu = 0$$

$$\sigma^2 = \frac{1}{m}$$

$$Var\,\theta = \frac{1}{m}$$

$$E\,\theta = 0$$

$$\theta \sim U(-a, a)$$



$$c = \frac{1}{2a}$$

$$E\theta = 0$$

$$E\theta = \int_{-\infty}^{+\infty} \theta \, pdf(\theta) \, d\theta = \left. \frac{\theta^2}{2} c \right|_{-a}^{a} =$$

$$= c\left(\frac{a^2}{2} - \frac{a^2}{2}\right) = 0$$

$$Var\,\theta = \frac{1}{m} = \int_{-\infty}^{+\infty} (\theta - \underbrace{\cancel{\mu}}_{0})^2 \, \overset{c}{\cancel{pdf(\theta)}} \, d\theta =$$

$$= c \int_{-a}^{a} \theta^2 \, d\theta = c \left. \frac{\theta^3}{3} \right|_{-a}^{a} = \frac{1}{\cancel{2}\cancel{a}} \frac{\cancel{2} a^{\cancel{3}\,2}}{3} = \frac{a^2}{3}$$

$$\frac{a^2}{3} = \frac{1}{m} \qquad a = \sqrt{\frac{3}{m}}$$

$$Var\,\theta = \frac{1}{m} \Rightarrow \theta \sim U\left(-\sqrt{\frac{3}{m}}; \sqrt{\frac{3}{m}}\right)$$