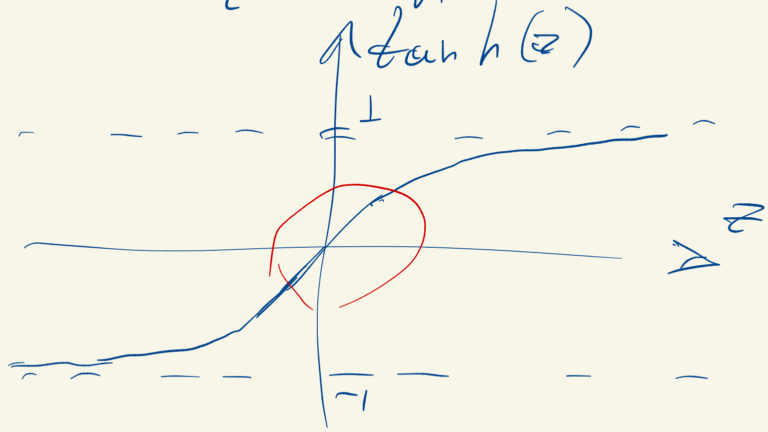
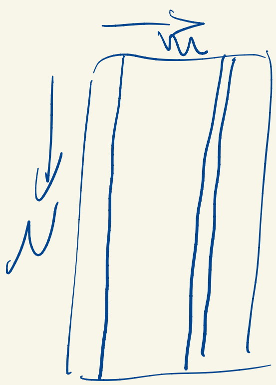
$$x \to \underbrace{\boxed{\theta^T x}}_{=z^{(1)}} - \underbrace{\boxed{g(z)}}_{=h^{(1)}} - \underbrace{\boxed{\theta^T h}}_{=z^{(2)}} \to \underbrace{\boxed{g(z)}}_{=h^{(2)}} - \dots - \underbrace{\boxed{\theta^T h^{(\ell-1)}}}_{=z^{(\ell)}} - \underbrace{\boxed{g(z^{(\ell)})}}_{=h^{(\ell)}} - \dots \hat{y}$$

$$g(z) = \tanh(z)$$



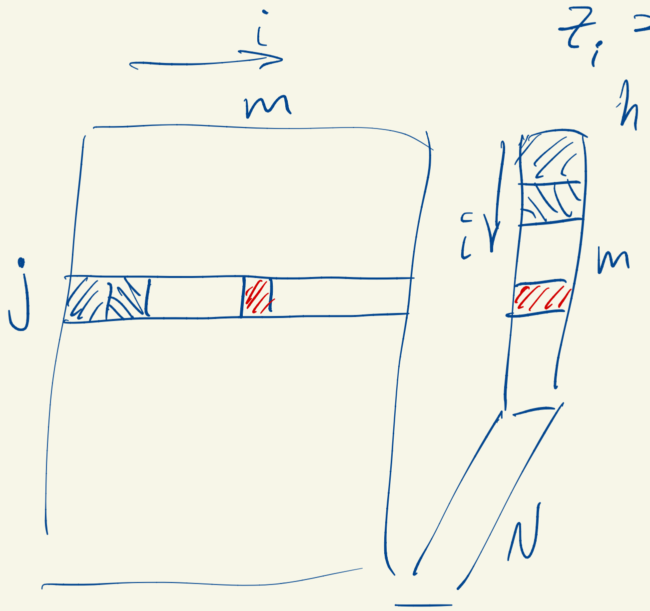
$h \approx 0$ но $h \neq 0$

1) $x_i \sim$ i.i.d.
 $h_i \sim$ i.i.d.
2) $\theta_i \sim$ i.i.d.
3) $x \perp \theta$
 $h \perp \theta$



$$z^{(\ell)} = \theta^{T(\ell)} h^{(\ell-1)} \qquad z_{(j)} = \theta^T h = \sum_{i=1}^{m} \theta_i h_i$$

$$z_i = \theta_i h_i$$



z_i

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

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

$$= \Big(\underbrace{\mathbb{E}\theta_i^2 - (\mathbb{E}\theta_i)^2}_{\operatorname{Var}\theta_i} + (\mathbb{E}\theta_i)^2\Big)\Big(\overbrace{\mathbb{E} h_i^2 - (\mathbb{E} h_i)^2}^{\operatorname{Var} h_i} + (\mathbb{E} h_i)^2\Big) - (\mathbb{E}\theta_i)^2 (\mathbb{E} h_i)^2 =$$

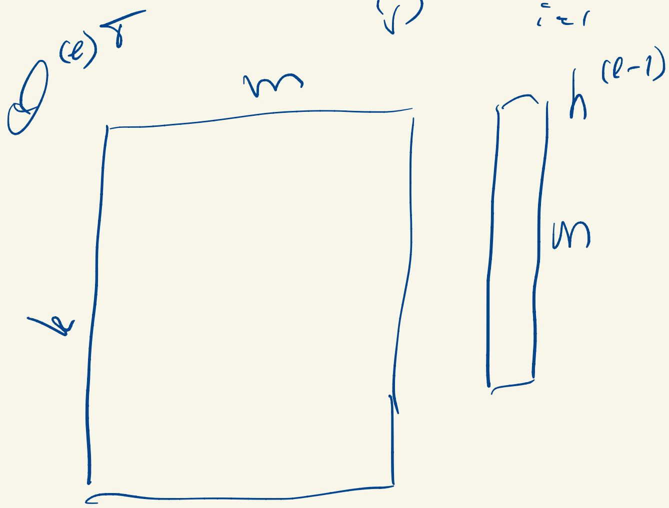
$$= \Big(\operatorname{Var}\theta_i + (\underbrace{\mathbb{E}\theta_i}_{0})^2\Big)\Big(\operatorname{Var} h_i + (\underbrace{\mathbb{E} h_i}_{0})^2\Big) - (\underbrace{\mathbb{E}\theta_i}_{0})^2 (\underbrace{\mathbb{E} h_i}_{0})^2 =$$

$$= \operatorname{Var}\theta_i \operatorname{Var} h_i$$

$$\mathrm{Var}\, z_i = \mathrm{Var}\, \theta_i \, \mathrm{Var}\, h_i$$

$$z_{(j)} = \sum_{i=1}^{m} \theta_i h_i$$

$$\mathrm{Var}\, z_{(j)} = \sum_{i=1}^{m} \mathrm{Var}\, \theta_i \, \mathrm{Var}\, h_i = \underline{m \, \mathrm{Var}\, \theta \, \mathrm{Var}\, h}$$



$$z = \theta^T h$$
$$h \in \mathbb{R}^m$$
$$z \in \mathbb{R}^k$$

$$\mathrm{Var}\, z = m \, \mathrm{Var}\, \theta \, \mathrm{Var}\, h$$

$$\mathrm{Var}\, g(z) = \mathrm{Var}\, z \qquad \mathrm{Var}\, g(z) = m \, \mathrm{Var}\, \theta \, \mathrm{Var}\, h$$

$$\mathrm{Var}\, h^{(l)} = m \, \mathrm{Var}\, \theta^{(l)} \, \mathrm{Var}\, h^{(l-1)}$$

$$\mathrm{Var}\, h^{(l)} \simeq \mathrm{Var}\, h^{(l-1)}$$

$$m \, \mathrm{Var}\, \theta^{(l)} = 1 \qquad \underline{\mathrm{Var}\, \theta = \frac{1}{m}}$$

$$\boxed{\theta \sim \mathcal{N}\left(0, \sigma^2 = \frac{1}{m}\right)}$$

$$\theta \sim U\left[\underbrace{-\frac{1}{\sqrt{m}}}_{a} ; \underbrace{\frac{1}{\sqrt{m}}}_{b}\right]$$

$$\mathbb{E}\theta = \int_a^b p(\theta)\,\theta\, d\theta = \int_a^b c\,\theta\, d\theta = c\,\frac{\theta^2}{2}\Big|_a^b = c\left(\frac{b^2}{2} - \frac{a^2}{2}\right)$$

$$\frac{b^2}{2} = \frac{1}{2m}, \quad \frac{a^2}{2} = \frac{1}{2m}, \quad \frac{1}{2m} - \frac{1}{2m} = 0$$

$$\boxed{\mathbb{E}\theta = 0}$$

$p(\theta)$: $$\int_{-\infty}^{\infty} p(\theta)\, d\theta = 1$$

$$p(\theta) = p$$

$$\int_{-\frac{1}{\sqrt{m}}}^{\frac{1}{\sqrt{m}}} p\, d\theta = 1 \qquad p\left(\theta\Big|_{-\frac{1}{\sqrt{m}}}^{\frac{1}{\sqrt{m}}}\right) = 1$$

$$\boxed{p = \frac{\sqrt{m}}{2}}$$

$$\mathrm{Var}\,\theta = \int_{-\infty}^{\infty} (\theta - \mathbb{E}\theta)^2 p(\theta)\, d\theta =$$

$$= p\int_a^b \theta^2\, d\theta = \frac{\sqrt{m}}{2}\,\frac{\theta^3}{3}\Big|_{-\frac{1}{\sqrt{m}}}^{\frac{1}{\sqrt{m}}} = \frac{\sqrt{m}}{6}\left(\frac{1}{\sqrt{m}^3} + \frac{1}{\sqrt{m}^3}\right)$$

$$= \frac{2\sqrt{m}}{6\sqrt{m}^3} = \frac{1}{3m} \qquad \mathrm{Var}\,\theta = \frac{1}{3m}$$

$$\mathrm{Var}\, g(z^{(l)}) = m\, \mathrm{Var}\,\theta^{(l)}\, \mathrm{Var}\, h^{(l-1)}$$

$$\mathrm{Var}\, h^{(l)} = \cancel{m} \frac{1}{3\cancel{m}} \mathrm{Var}\, h^{(l-1)}$$

$$\mathrm{Var}\, h^{(l+N)} = \left(\frac{1}{3}\right)^N \mathrm{Var}\, h^{(l-1)}$$

Xavier Glorot, Yoshua Bengio
2010.

$$\theta \sim \mathcal{N}\left(0, \sigma^2 = \frac{1}{m}\right)$$

$$\sigma^2 = \frac{2}{m+k}$$

$$\theta \sim U\left(\frac{-\sqrt{6}}{\sqrt{m+k+1}}, \frac{\sqrt{6}}{\sqrt{m+k+1}}\right)$$

Kaiming He **2015**

$$\mathrm{Var}\,\theta = \frac{2}{k}$$