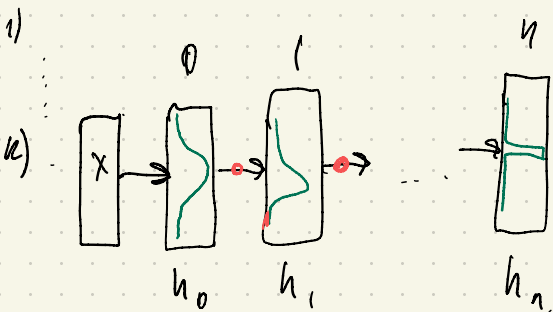
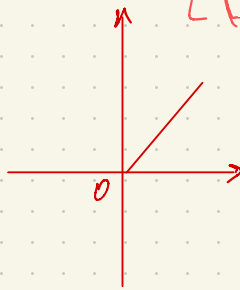
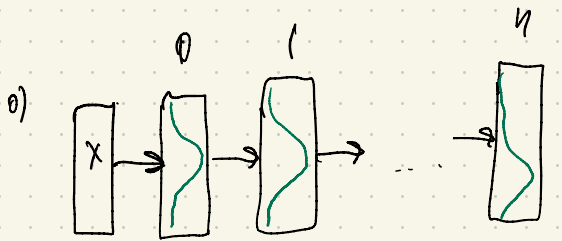
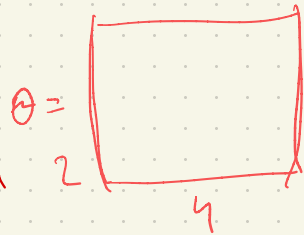


$$\theta_i \sim N(0, 1)$$

$$t = \theta_n(\theta_{n-1}(\dots(\theta_0 x)\dots))$$



batch normalization

$$h_i = \frac{h_i - \mathbb{E}_{x \sim b} h_i}{\text{std}(h_i) + \epsilon} \cdot \gamma + \beta$$

$$h_i = \frac{h_i - \mu(h_i)}{\sigma(h_i) + \epsilon} \cdot \gamma + \beta$$

$$\mu_i = \mu_{i-1} \cdot \beta + \mathbb{E} h_i$$

$$\sigma_i = \sigma_{i-1} \cdot \alpha + \sigma(h_i)$$

1) μ, σ

+

4d

2) регуляризаторы

3) скалярные изображения

1) 4 d neurons

2) noxo pasowala ge narekataro batch size.

Dropout.

$$\hat{h}_i = h_i \cdot \boxed{b / \alpha} \quad b \sim \text{Ber}(1 - \alpha)$$

$$\alpha = 0.7.$$

$$\mu \cdot \mu_b$$
