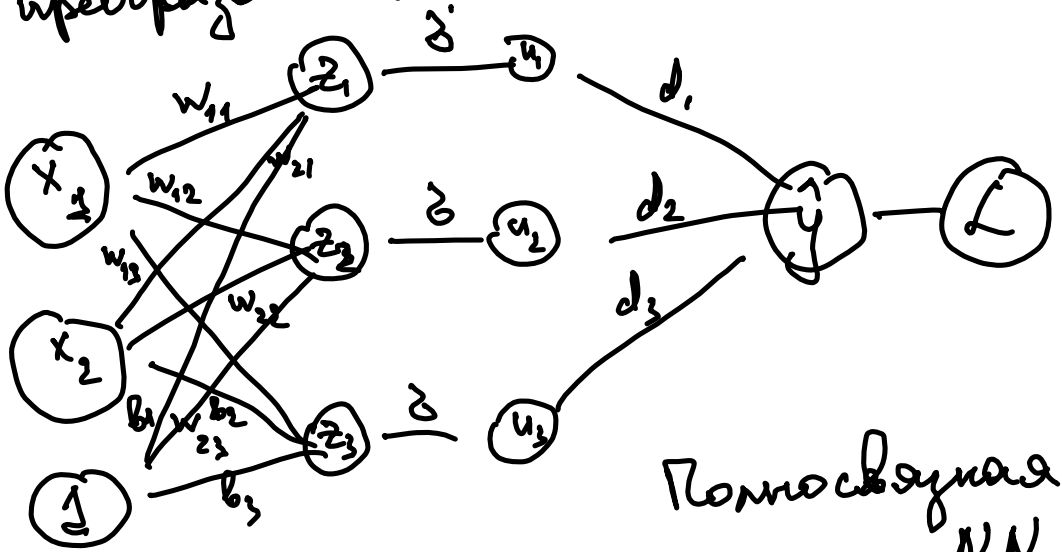
Neural Networks

Ациклический орграф, в кот. вершины – входные данные или нейроны, а рёбра – какие-то преобразования.

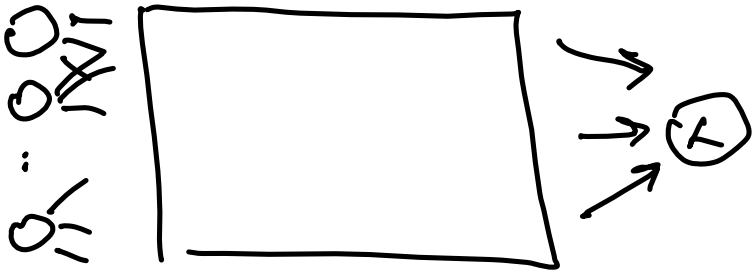


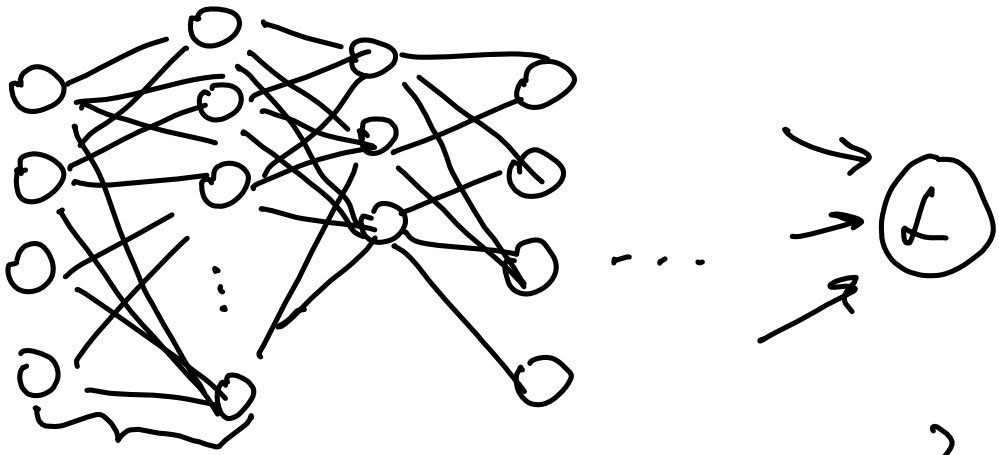
Полносвязная NN

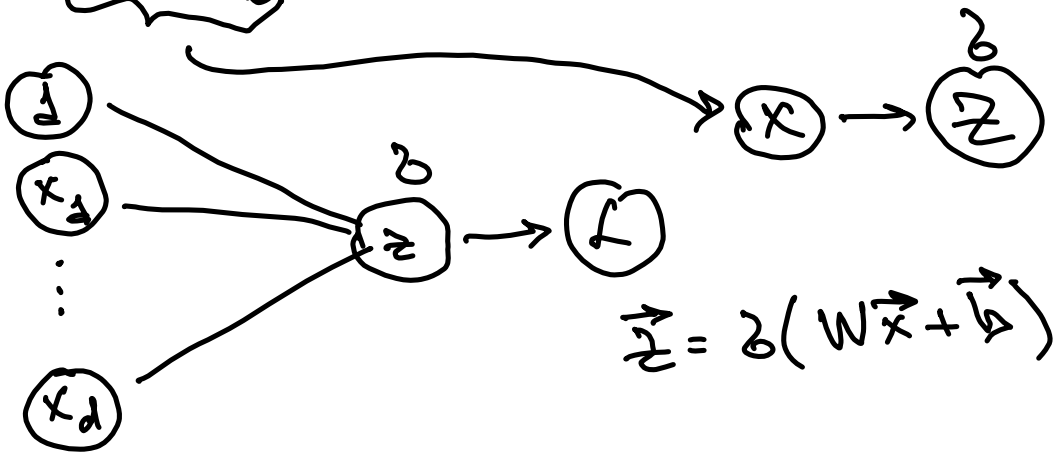
$$z_2 = x_1 \cdot w_{12} + x_2 \cdot w_{22} + b_2$$

$$u_i = g(z_i)$$

$$\hat{y} = d_1 \cdot u_1 + d_2 u_2 + d_3 u_3$$







$$\vec{z} = \sigma(W\vec{x} + \vec{b})$$

$$a(x) = \sigma(\langle w, x \rangle)$$

Нелинейности:

$$\sigma(x) = \frac{1}{1+e^{-x}} \quad (\text{Sigma})$$

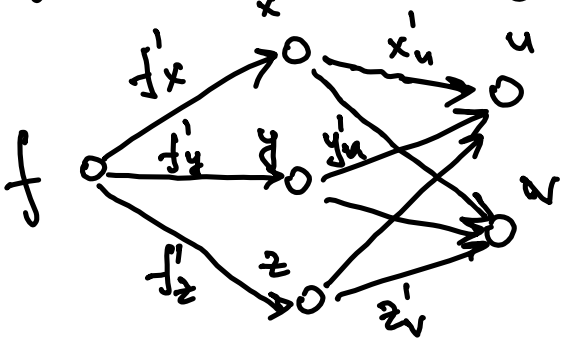
$$ReLU(x) = \max(0, x) \quad (ReLU)$$

$$\tanh(x) = \frac{e^{x} - e^{-x}}{e^{x} + e^{-x}} \quad (\tanh)$$

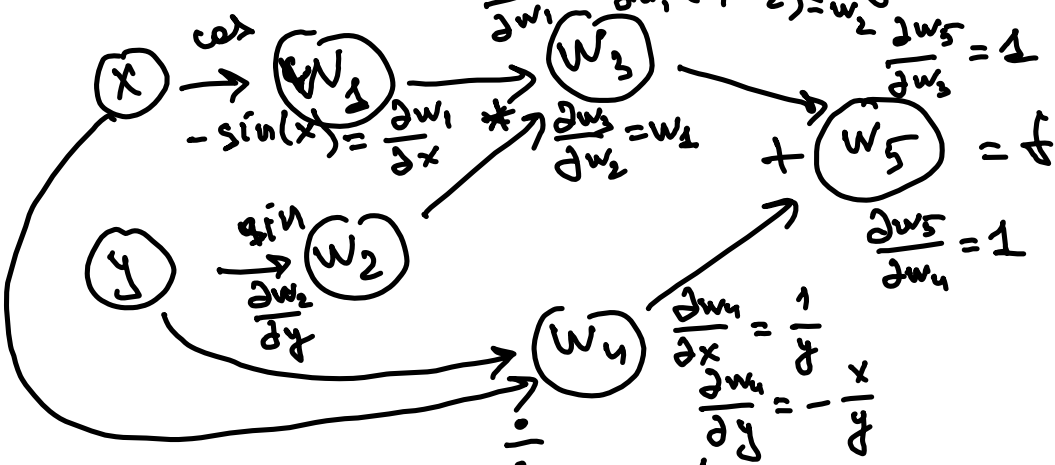
Autodiff.

$$f(x(u,v), y(u,v), z(u,v))$$

$$f'_u = f'_x \cdot x'_u + f'_y \cdot y'_u + f'_z \cdot z'_u$$



$$f(x,y) = \cos(x)\cdot\sin(y) + \frac{x}{y}$$



$$w_5'_x = -\sin(x)\cdot\sin(y) + \frac{1}{y}$$

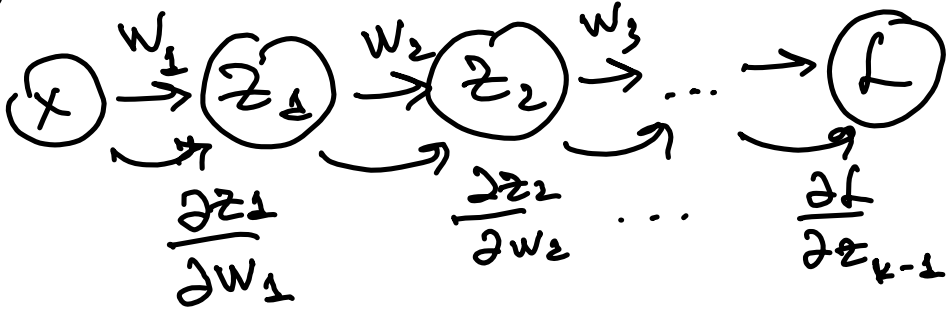
$$w_5'_y = \cos(x)\cdot\cos(y) - \frac{x}{y^2}$$

$$w_5'_x = \frac{\partial w_5}{\partial w_3}\cdot\frac{\partial w_3}{\partial w_1}\cdot\frac{\partial w_1}{\partial x} + \frac{\partial w_5}{\partial w_4}\cdot\frac{\partial w_4}{\partial x} =$$

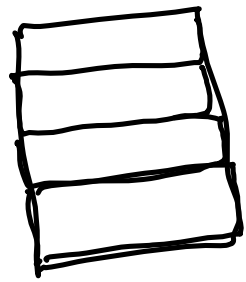
$$= 1\cdot w_2\cdot(-\sin(x)) + 1\cdot\frac{1}{y} = -\sin(x)\cdot\sin(y) + \frac{1}{y}$$

Backward Propagation (Backprop)

1) FP

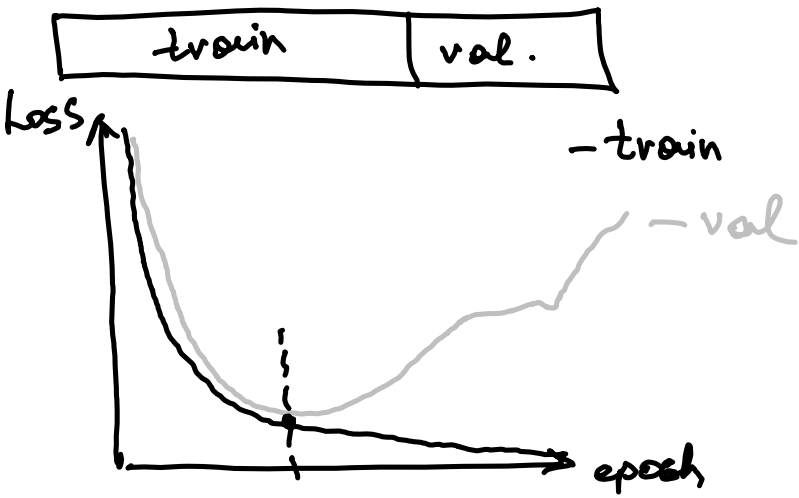


$$z_i = f(W_k \cdot z_{i-1} + b_k)$$

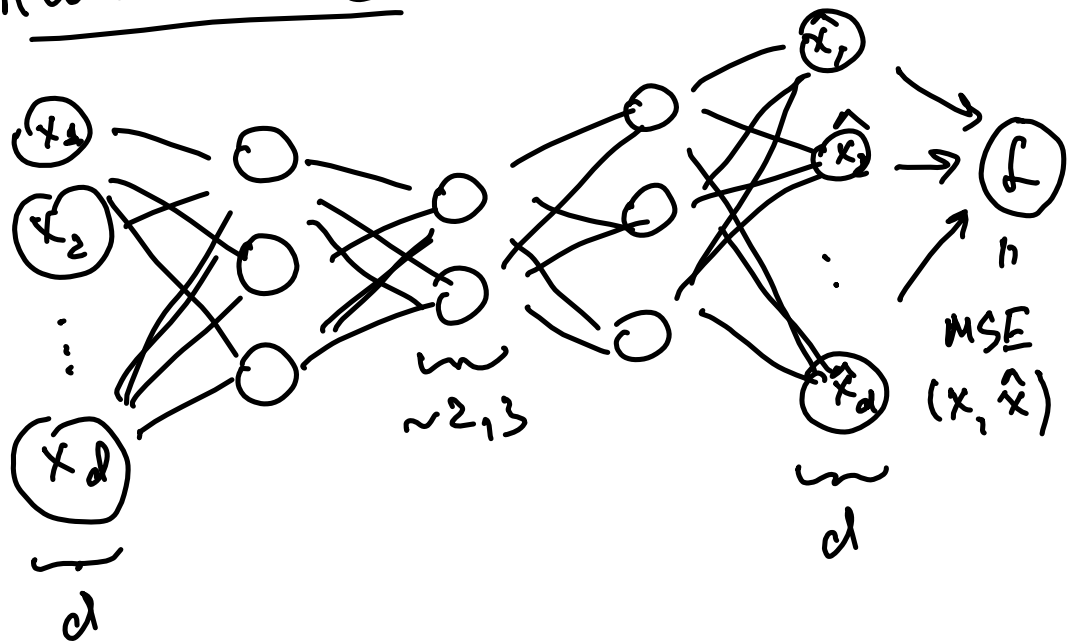


2) BP

$$W_1^{new} = W_1^{old} - \alpha \cdot \frac{\partial L}{\partial W_1^{old}}$$



Autoencoder



Семинар

1) EM - алгоритм.

$$\log \mathbb{P}(X|\theta) = KL(q_f(z) \| p(z|x,\theta)) +$$

$$+ \underbrace{ELBO}$$

$$\int q(z) \cdot \log \frac{P(x,z|\theta)}{q(z)} dz = \mathcal{L}(q,\theta)$$

E - шаг:

$$q(z) = P(z \mid x, \theta^{old})$$

M - шаг:

$$L(q, \theta) \to \max$$

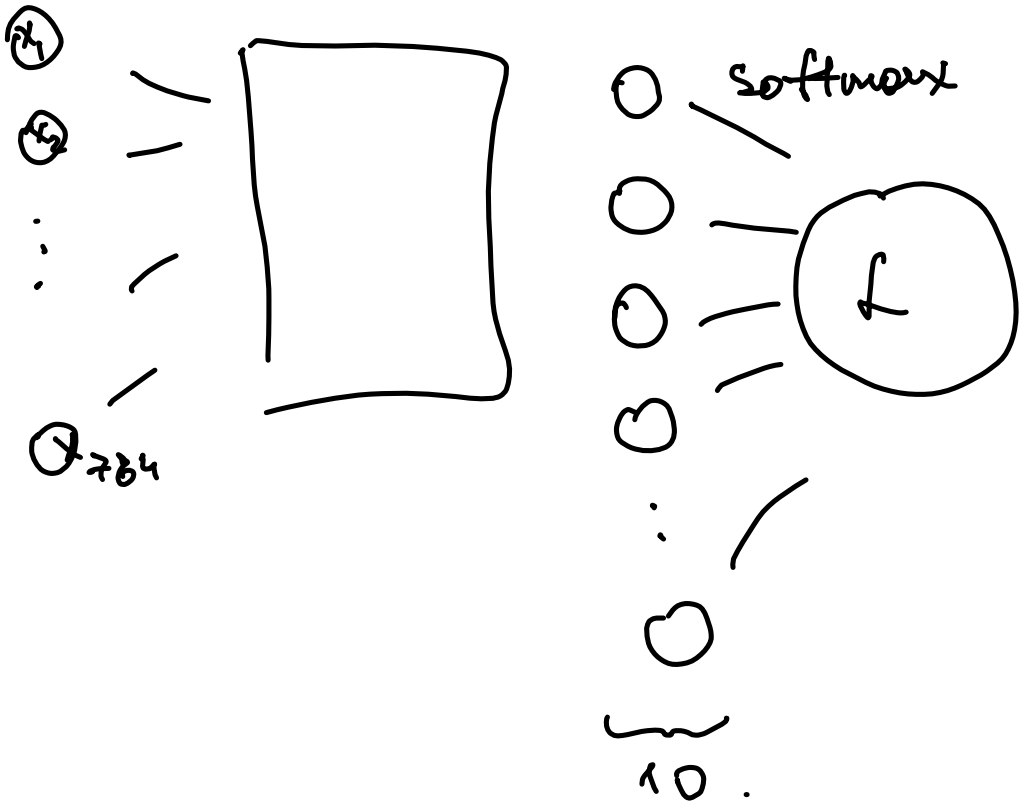
$$\theta^{new} = \arg\max_{\theta} \mathbb{E}_{q(z)} \log P(x, z \mid \theta)$$

2) bvd:

$$\text{noise} - \mathbb{E}_{x,y}\left[(y - \mathbb{E}(y \mid x))^2\right]$$

$$\text{bias} - \mathbb{E}_{x,y}\left[\left(\mathbb{E}_{x^n}(\mu(x^n)(x)) - \mathbb{E}(y \mid x)\right)^2\right]$$

$$\text{var} - \mathbb{E}_{x,y}\left[\mathbb{E}_{x^n}\left[(\mu(x^n)(x) - \mathbb{E}_{x^n}(\mu))^2\right]\right]$$



$$L = \log \text{loss}_i = -\sum_{k=1}^{K} [y_i = k] \log p_{ik}$$