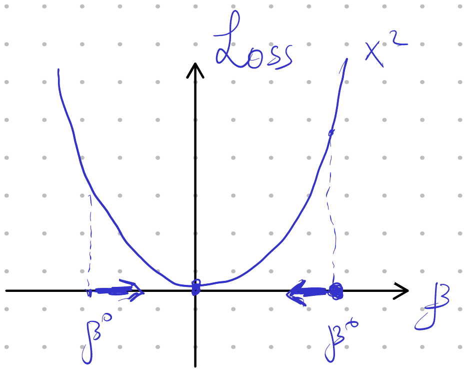
① Backpropagation idea.



$$\beta^{t+1} = \beta^t - \gamma \cdot L'(\beta^t)$$

$$\Delta L \approx L'(\beta)\cdot \Delta\beta$$

$2x$

$$L(\hat{y}) = (\hat{y}-y)^2, \quad \hat{y} = \sigma(a),$$

$$a = \varphi_0 + \varphi_1 h^1 + \varphi_2 h^2,$$

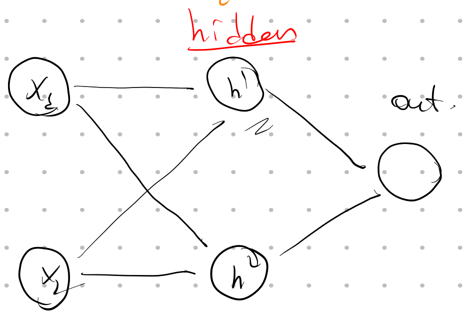
$$h^1 = \sigma(t^1), \quad t^1 = \omega_0^1 + \omega_1^1 x_1 + \omega_2^1 x_2$$

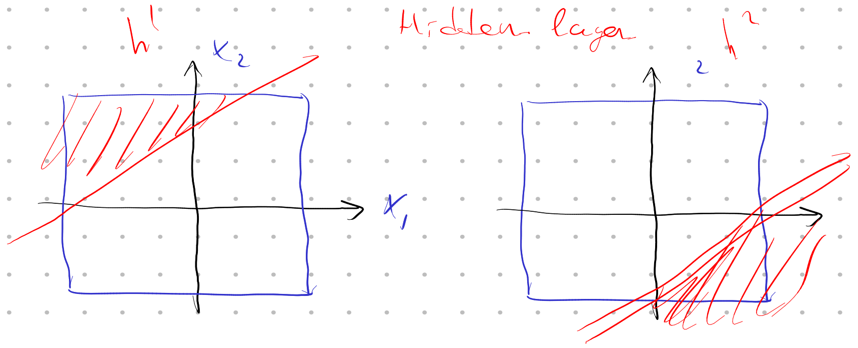
$$f(x) \longmapsto \Delta f \approx f'(x)\cdot\Delta x$$

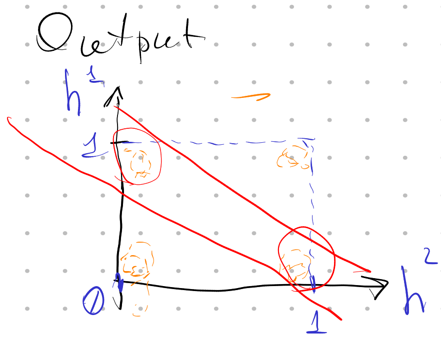
$$f(g(x)) \longmapsto \Delta f \approx f'(g)\cdot\Delta g \approx f'(g)\cdot g'(x)\cdot\Delta x$$

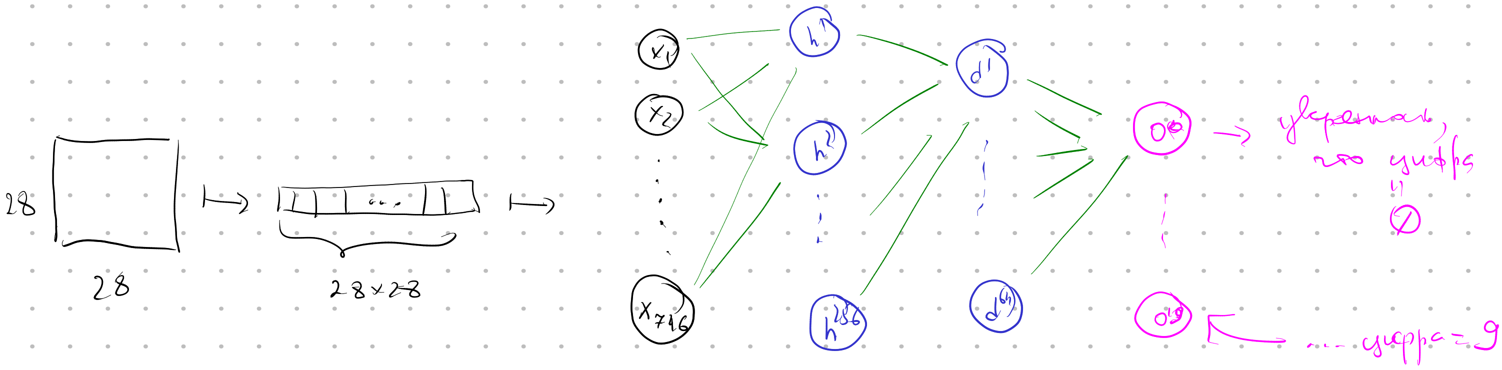
$$\underline{\omega_1^1}: \; L'_{\omega_1^1} = 2(\hat{y}-y)\cdot \underbrace{\hat{y}'(a)}_{\sigma'(a)} \cdot \underbrace{a'(h^1)}_{\varphi_1} \cdot \underbrace{[h^1]'(t)}_{\sigma'(t)} \cdot \underbrace{t'(\omega_1^1)}_{x_1}$$

$$[\omega_1^1]^{t+1} = [\omega_1^1]^{t} - \gamma \cdot L'_{\omega_1^1}$$








28
28
28×28
x_1
x_2
x_{784}
h^1
h^2
h^{256}
d^1
d^{64}
o^0
o^9
уверенность, что цифра = 0
... цифра = 9