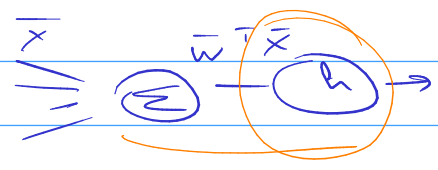
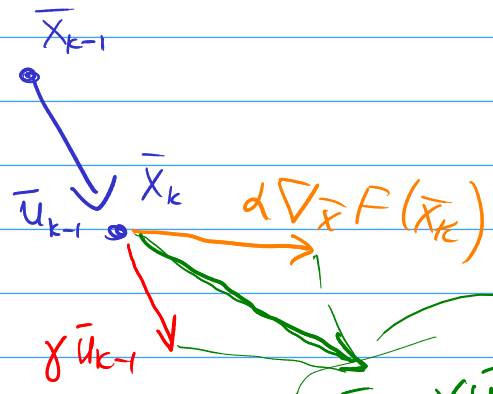


Momentum

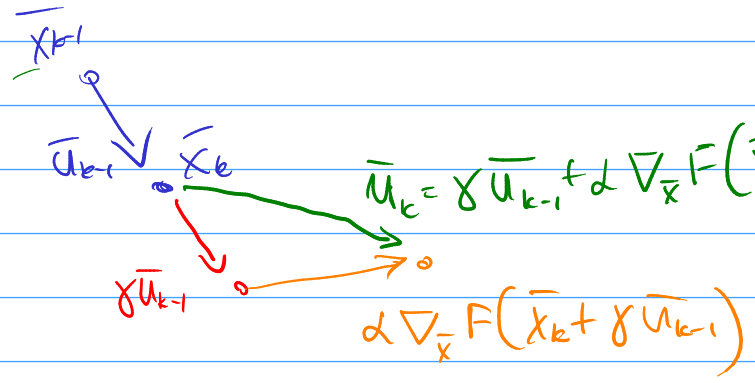


ReLU = max(0, x)

NAG Nesterov accelerated gradient

$$\bar{u}_k = \gamma \bar{u}_{k-1} + \alpha \nabla_{\bar{x}} F(\bar{x}_k)$$

NAS neural arch. search



$$\bar{u}_k = \gamma \bar{u}_{k-1} + \alpha \nabla_{\bar{x}} F(\bar{x}_k + \gamma \bar{u}_{k-1})$$

$$S_{\beta}(g, h) = (g - h) \sigma(\beta(g - h)) + h$$

$$\frac{\max(x, 0)}{x \sigma(\beta \cdot x)}$$

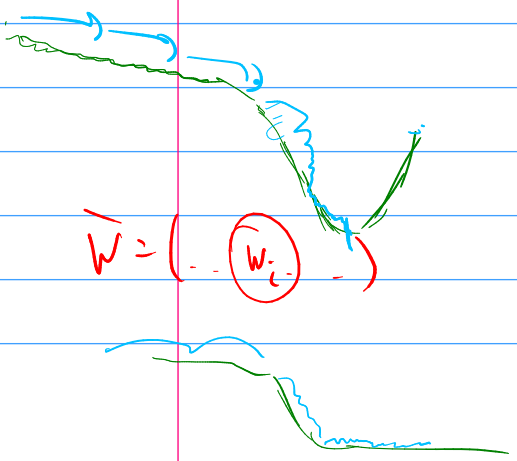
Adagrad

$$\bar{w}_{t+1} = \bar{w}_t - \frac{\alpha}{\sqrt{G_t + \epsilon}} g_t$$

$$w_{t+1, i} = w_{t, i} - \frac{\alpha}{\sqrt{G_{t, i} + \epsilon}} g_{t, i} \quad (-g_{t, i})$$

$$G_{t, i} = G_{t-1, i} + g_{t, i}^2$$

$$\bar{w} = (\dots, w_i, \dots)$$



RMS' prop

$F(\bar{w}) - \text{mean}$

$$\bar{w}_{t+1} = \bar{w}_t - \alpha_t \bar{g}_t$$

$$\bar{w}_{t+1} = \bar{w}_t - \frac{\alpha}{\sqrt{G_t + \epsilon}} \bar{g}_t$$

$$\bar{w}_{t+1} = \bar{w}_t - \frac{1}{2} H_t \bar{g}_t$$

$$G_{t+1} = \rho \cdot G_{t+1} + (1-\rho) g_{t+1}$$

Adadelta

$$\bar{w}_{t+1} = \bar{w}_t - \frac{\sqrt{E_t[\Delta \bar{w}]^2 + \epsilon}}{\sqrt{G_t^2 + \epsilon}} \bar{g}_t$$

$$E_t[\Delta \bar{w}]^2 = \rho E_{t-1}[\Delta \bar{w}] + (1-\rho) \Delta \bar{w}^2$$

Adam

$$\bar{w}_{t+1} = \bar{w}_t - \frac{\alpha}{\sqrt{v_t + \epsilon}} \bar{m}_t$$

$$\bar{m}_t = \beta_1 \bar{m}_{t-1} + (1-\beta_1) \bar{g}_t$$

$$v_t = \beta_2 v_{t-1} + (1-\beta_2) \bar{g}_t^2$$

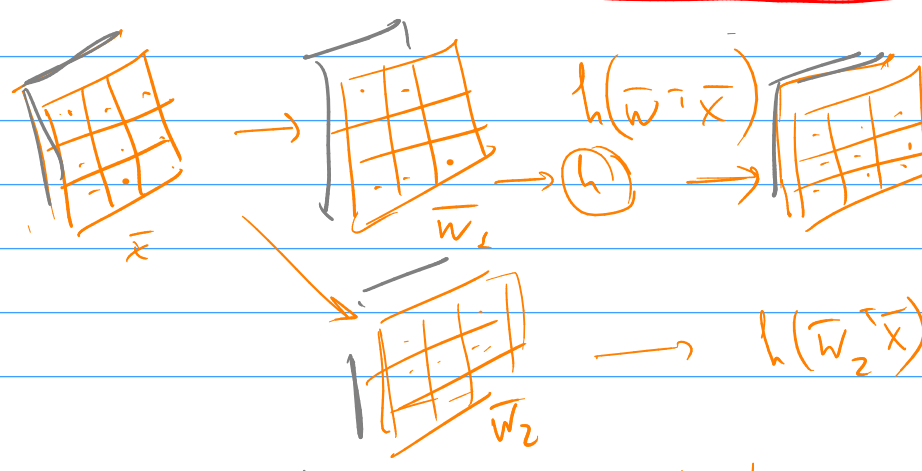
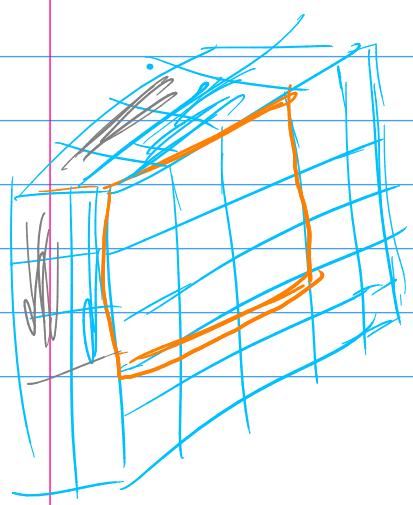
$$F(\bar{w}) + \frac{\alpha}{2} \|\bar{w}\|^2 \rightarrow \min$$

Weight decay

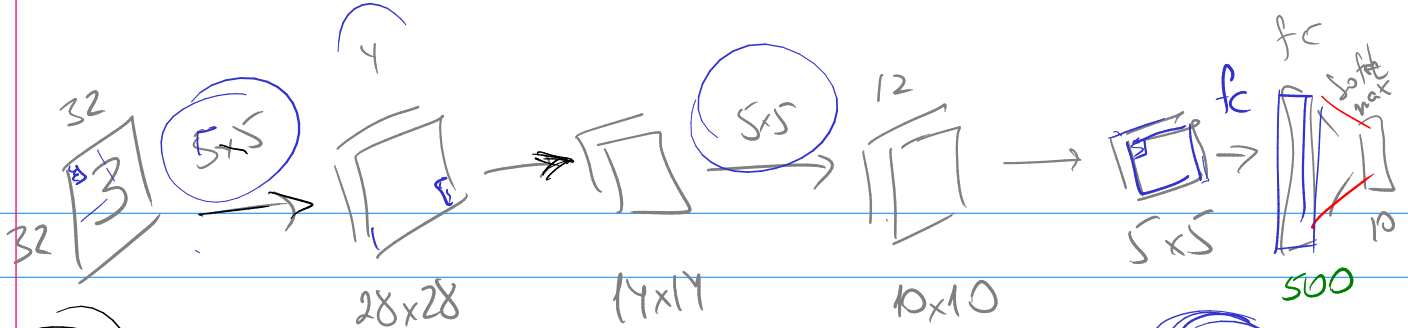
$$\bar{w}_{t+1} = \beta \bar{w}_t - \alpha \nabla_{\bar{w}} F(\bar{w}_t) \quad \text{SGD}$$

$$= \bar{w}_t - \alpha \left( \nabla_{\bar{w}} F(\bar{w}_t) - \frac{(1-\beta) \bar{w}_t}{\alpha} \right)$$

$$= \bar{w}_t - \alpha \nabla_{\bar{w}} \left[ F(\bar{w}_t) - \frac{1-\beta}{2\alpha} \bar{w}^T \bar{w} \right]$$

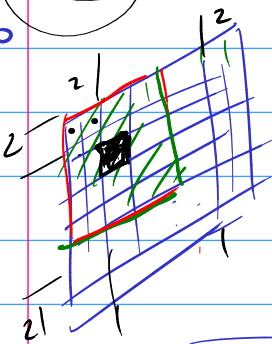


$i$ -channels  $\times$  width  $\times$  height  $\times$  out-channels

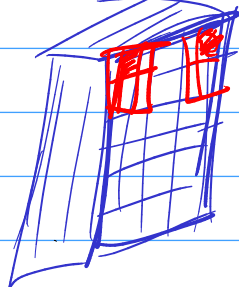


Рассчитать  
теггеров  
и пуллинг

$32 \times 32 \times 1$



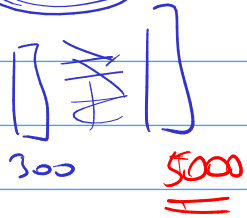
$28 \times 28 \times 4$



$14 \times 14 \times 4$

$10 \times 10 \times 12$

$5 \times 5 \times 12$   
 $300$



Скелето  
бесов

$1 \times 5 \times 5 \times 4 = 100$

0

$4 \times 5 \times 5 \times 12 = 1200$

$300 \times 300$   
 $15000$

$32 \times 32 \times 28 + 28 \times 4 \approx 3.2M$

$4 \times 14^2 \times 10^2 \times 12 \approx 940K$

