$$xy + \sin(x-y) = w_1 \varphi_1(x,y) + w_2 \varphi_2(x,y) + \dots + w_n \varphi_n$$

$w^T \varphi$ — лин. модель.

$$f = xy = \frac{1}{2}(x+y)^2 - \frac{1}{2}x^2 - \frac{1}{2}y^2$$

$w_1(a_1 x + a_2 y)^2 + w_2(b_1 x + b_2 y)^2$

$$= w_1 \varphi_1(x,y) + w_2 \varphi_2(x) + w_3 \varphi_3(y)$$

$$w = \begin{pmatrix} 1 \\ \frac{1}{2} \\ -\frac{1}{2} \\ -\frac{1}{2} \end{pmatrix}$$

$\{\varphi = \varphi_1 = \varphi_2 = \varphi_3 = \cdot^2\}$

features

$$= w^T \varphi(y)$$

— лин. мод.

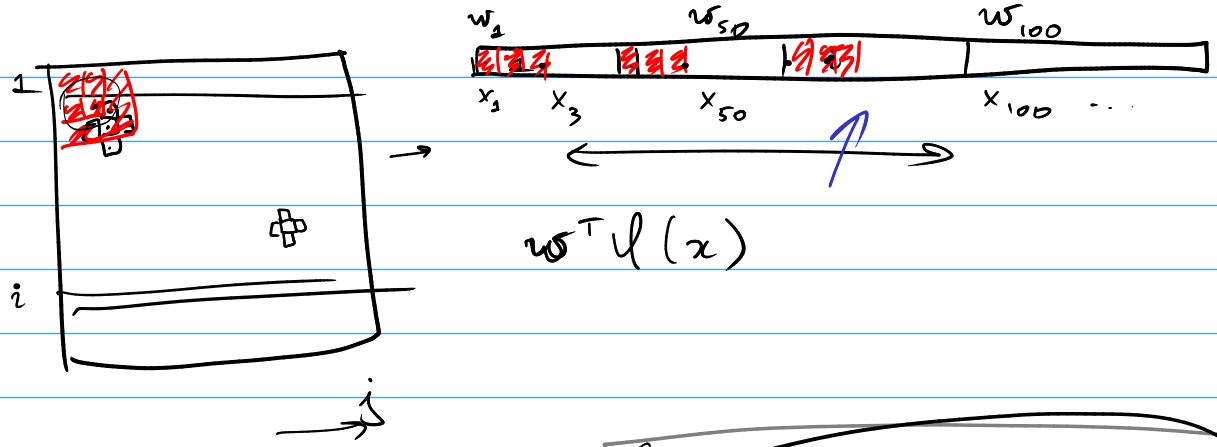
$$\varphi(y) = \begin{pmatrix} y_1 = x+y \\ y_2 = x \\ y_3 = y \end{pmatrix} \quad \begin{pmatrix} 1 & 1 \\ 1 & 0 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} x+y \\ x \\ y \end{pmatrix}$$

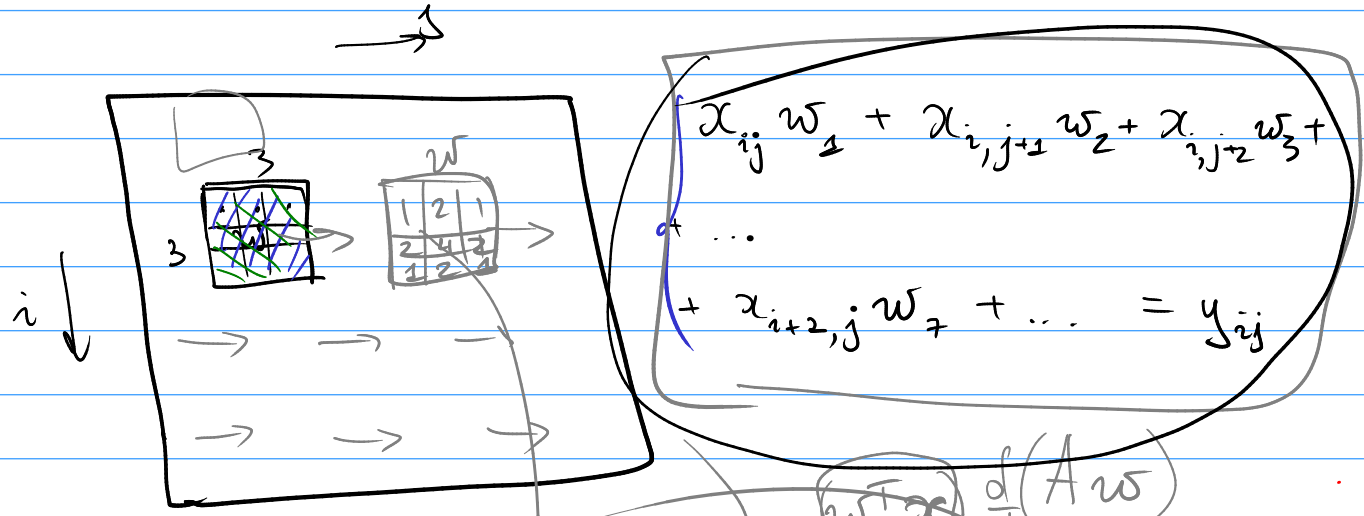
A

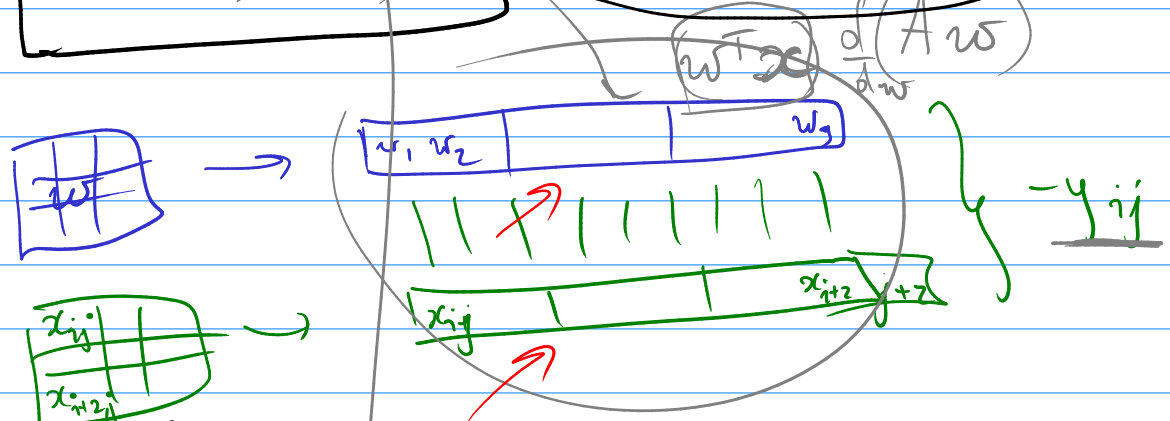
$$y = A \begin{pmatrix} x \\ y \end{pmatrix} = AX, \quad X = \begin{pmatrix} x \\ y \end{pmatrix}$$

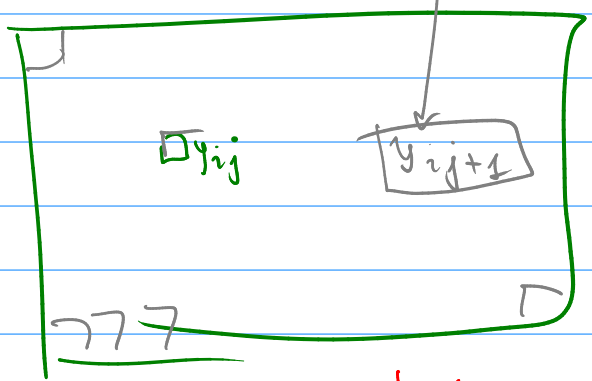
$$f(x,y) = xy = w^T \varphi(AX) \qquad C \subseteq \mathbb{R}^n$$

$$\|f(x) - w^T \sigma_2(A_1(\sigma_2(A_2 X)))\| < \varepsilon$$

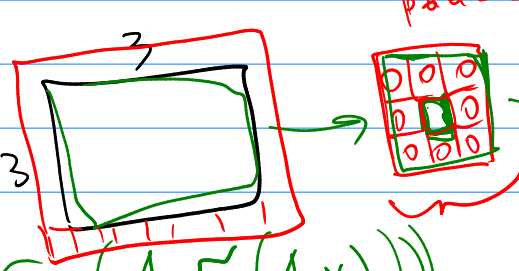








$$W = \begin{pmatrix} 1 & 2 & 1 \\ 2 & 4 & 2 \\ 1 & 2 & 1 \end{pmatrix}$$

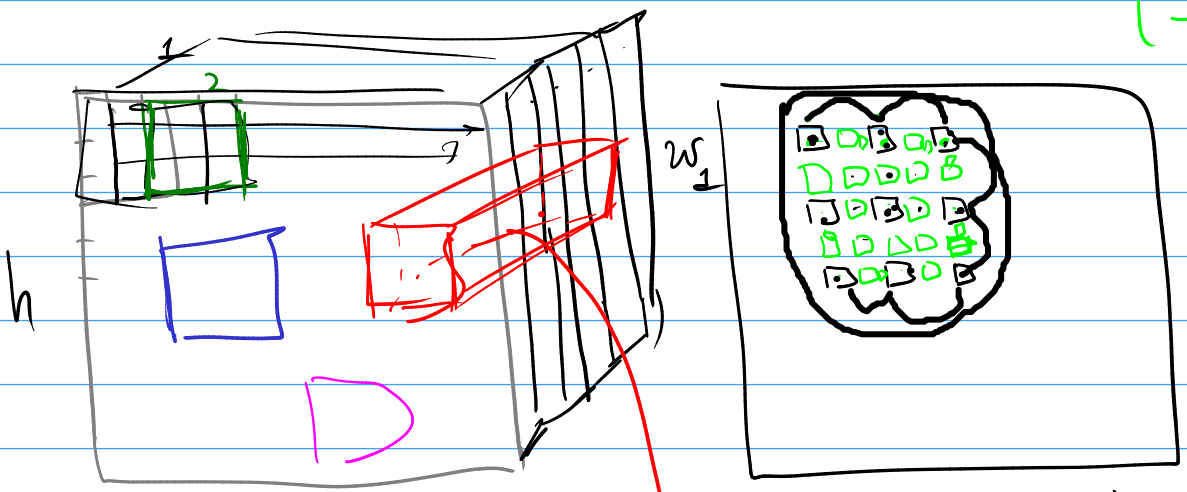


$$w \begin{pmatrix} 1 & 0 & -1 \\ 0 & 0 & 0 \\ 1 & 0 & -1 \end{pmatrix}$$

$$\sigma(A(\sigma(A\sigma(Ax))))$$

stride

dilation = 3 $w = (\#)$



Лин.

$|w| = 28 \times 28$

$\approx 10^3$

Свёрт

$|w| = 3 \times 3$

$\approx 10^2$

w

$\begin{pmatrix} -1 & 0 \\ 0 & 1 \end{pmatrix}$

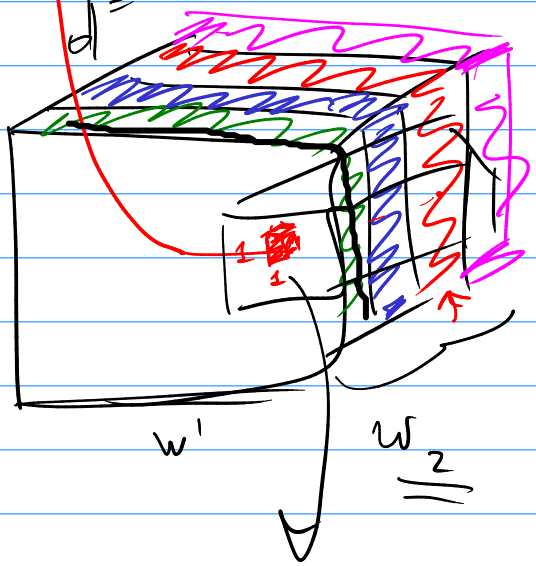
$\begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$

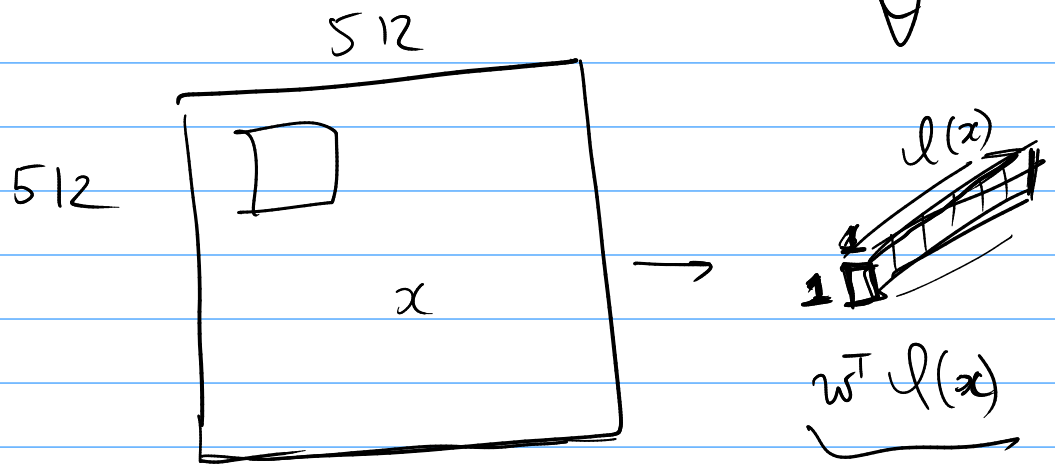
h'

d = out_channels = 512

1000

32





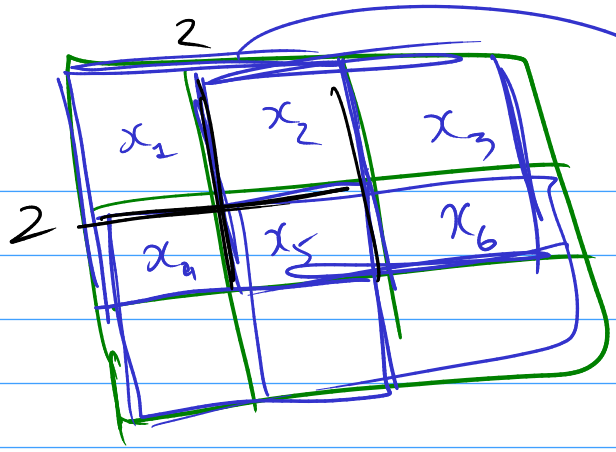
511

Max Pooling

$\log_2 512 = 2^9$

x → conv → maxpool → conv → maxp → ...

512 512 256 256 128

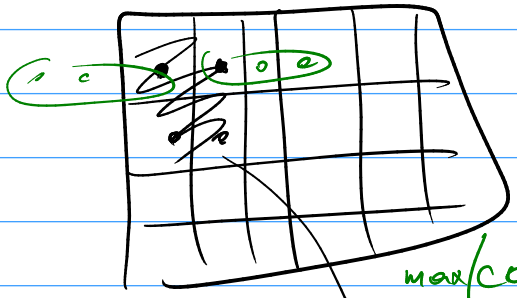


$y_1 = \max(x_1, x_2, x_4, x_5)$

y_1	y_2
y_3	y_4

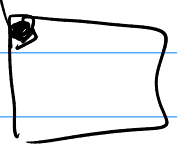
stride = 2

$h \rightarrow$	$h/2$
$w \rightarrow$	$w/2$

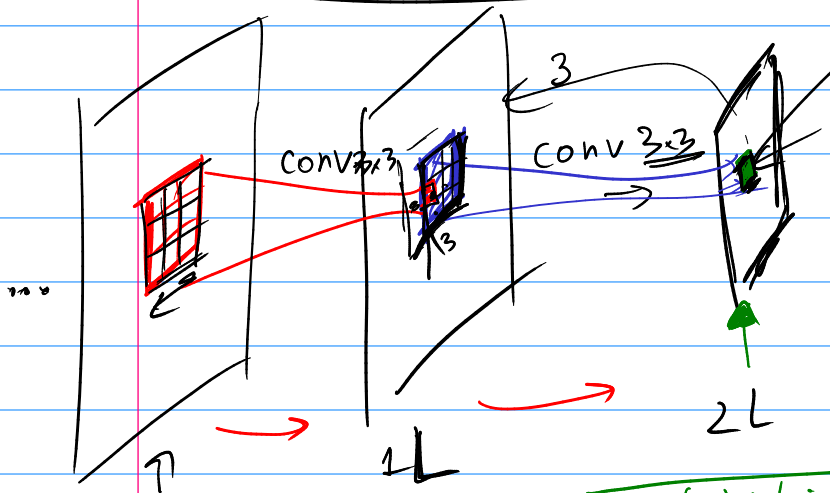


max(conv)

$2f = 5$

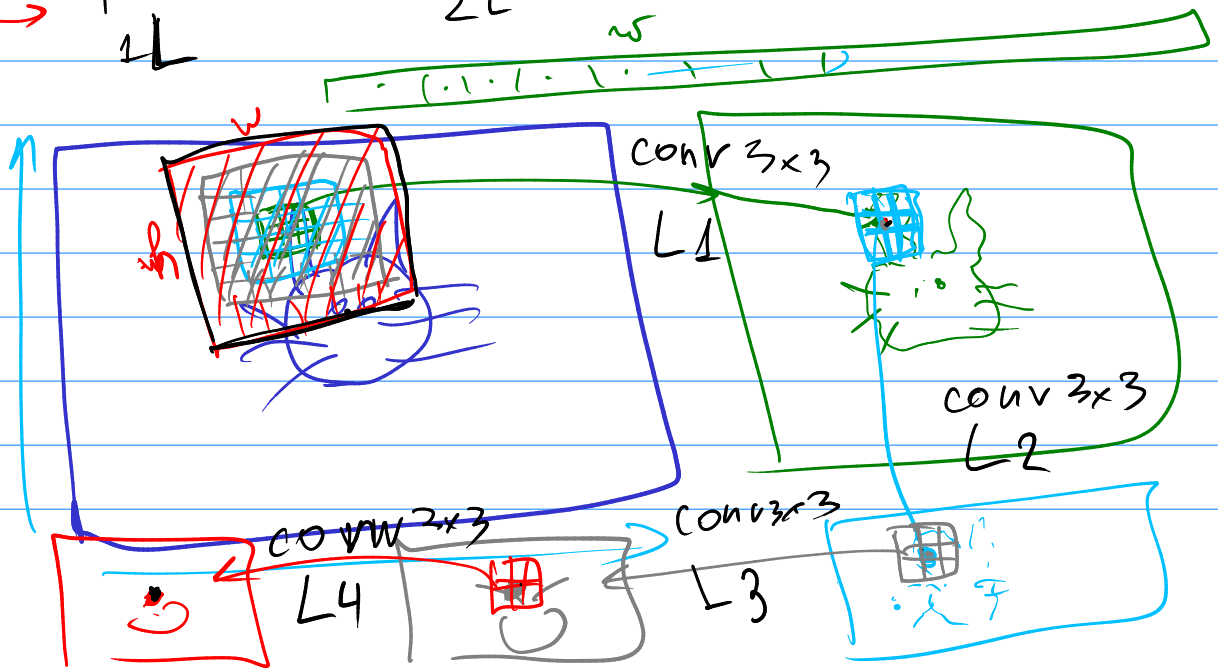


receptive field

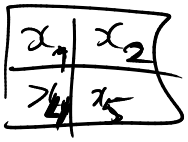


$X = w_1 x_1 + w_2 x_2 + w_3 x_3$

$X = w_1 x_1 \quad w_2 x_2 + x_3$



Conv



$$(\text{maxpool})' = \left(\max(\vec{x})\right)'_{x}$$

$$\left(\max\{x_1, x_2, x_4, x_5\}\right)' = (x_4)' = 1$$

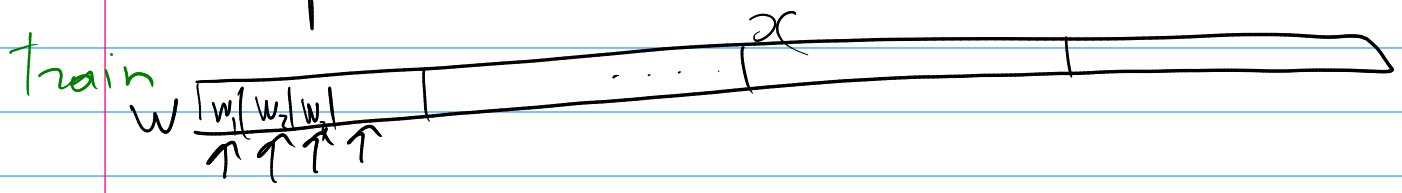
$\frac{\partial \max}{\partial x} = 1$

$\frac{\partial \max}{\partial x} = 0$

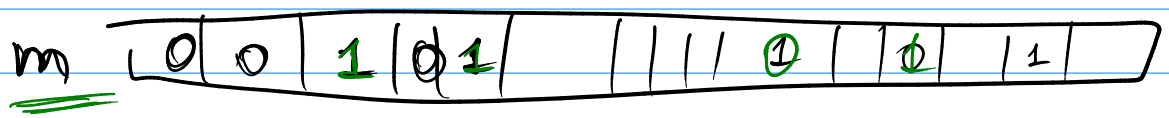
$$(\,)' = \begin{pmatrix} 0 \\ 0 \\ 1 \\ 0 \end{pmatrix}$$

Dropout

800



$m \sim B(p)$ $\quad p = 0.5$

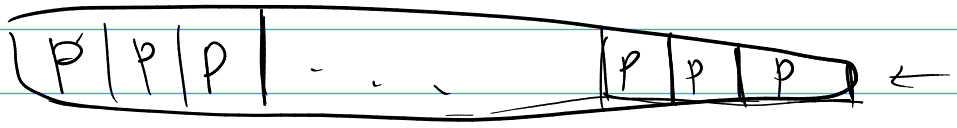


$(w \odot m)^T x$

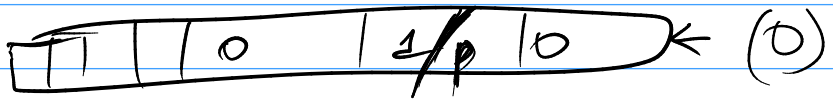
Test?

ex → 1

→ 0

(1) m = 

Inverse Dropout

Train m = 

Test m=1 (1)

- train() ← (0)
- eval() ← (1)