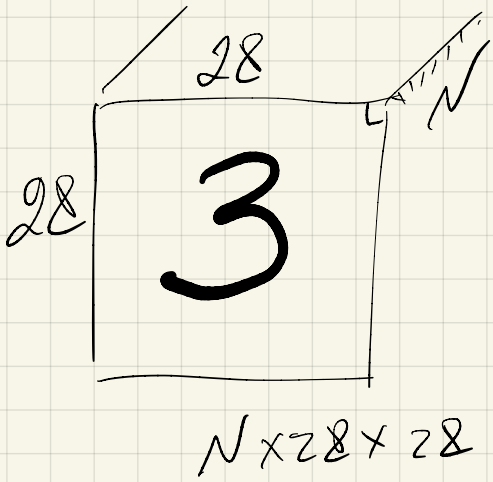


$\rightarrow N \times 784$

3

0	1	2	3	4	5	6	7	8	9
0	0	0	1	0	0	0	0	0	0

$i \rightarrow 0\ 1\ 2\ 3\ 4\ 5\ 6\ 7\ 8\ 9$

$K = 10$

Cross-Entropy

$$L(y, \hat{y}) = -\sum_{i=1}^{K} y_i \ln \hat{y}_i = -\ln \hat{y}_y$$

Accuracy = Доля верных ответов

$$k = \underset{i}{\operatorname{argmax}}(\hat{y})$$

$$\hat{y}_k = k$$

0	1	2	3	4	5	6	7
0.0	0.1	0.1	0.7	0.1	0.0	0.0	0.0

$i \rightarrow 0\ 1\ 2\ 3\ 4\ 5\ 6\ 7$

Sm

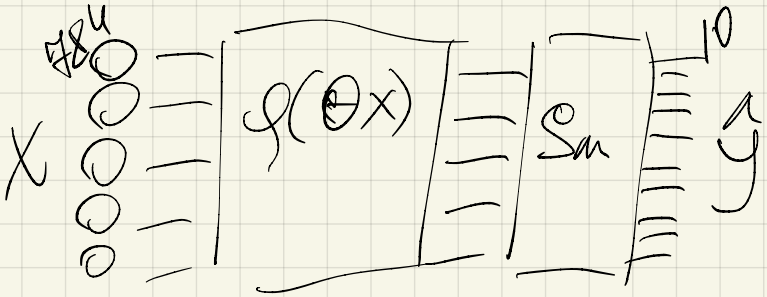
Softmax

$$Accuracy = \frac{\sum_{i=1}^{N} [\hat{y}_i^k == y_i]}{N}$$

$$Sm(z_j) = \frac{e^{z_j}}{\sum_{i=1}^{K} e^{z_i}}$$

$\sim 94\%$

$$J(x, \theta) = CE(Sm(\varphi(\theta x)))$$



self. cache = None

$$l(\theta, x) = \theta x \qquad \nabla_\theta l = x$$

$$\frac{\partial J}{\partial \theta} = \frac{\partial J}{\partial Sm} \frac{\partial Sm}{\partial \varphi} \frac{\partial \varphi}{\partial (\theta x)} \frac{\partial (\theta x)}{\partial \theta}$$

$$\varphi(z) = ReLU = \begin{cases} 0, & \text{if } z < 0 \\ z, & \text{if } z \geq 0 \end{cases}$$

$$\frac{\partial \varphi}{\partial z} = \begin{cases} 0, & \text{if } z < 0 \\ 1, & \text{if } z \geq 0 \end{cases}$$

$$\ln(Sm(h))$$

Derivative of cross-entropy

$$\frac{\partial J}{\partial \theta} = \frac{\partial J}{\partial Sm} \frac{\partial Sm}{\partial \varphi} \frac{\partial \varphi}{\partial \theta} \qquad \left(\frac{\partial \varphi}{\partial \theta} = \frac{\partial \varphi}{\partial (\theta x)} \cdot x\right)$$