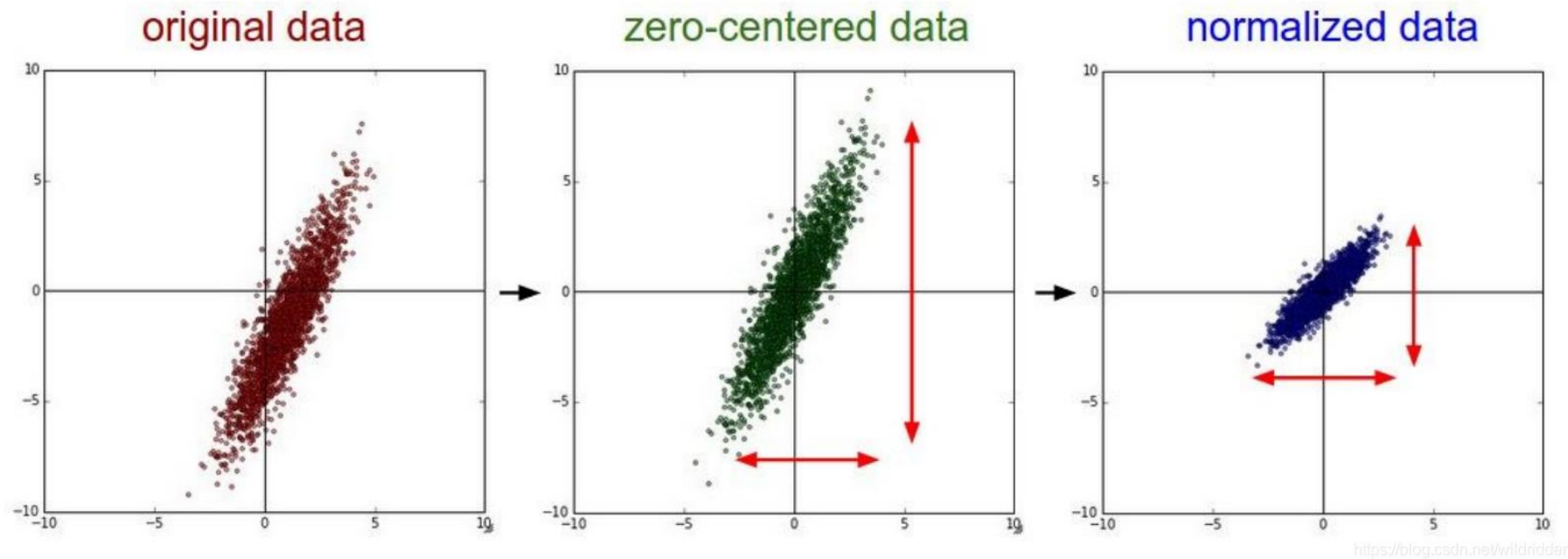




# Элементы нейронных сетей

# Нормализация данных



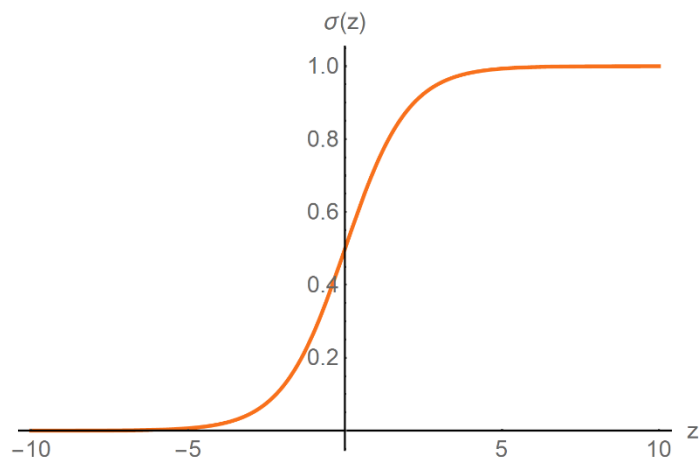
$$\hat{x} = x - \bar{x}$$

$$\hat{x} = \frac{x - \bar{x}}{\sigma}$$

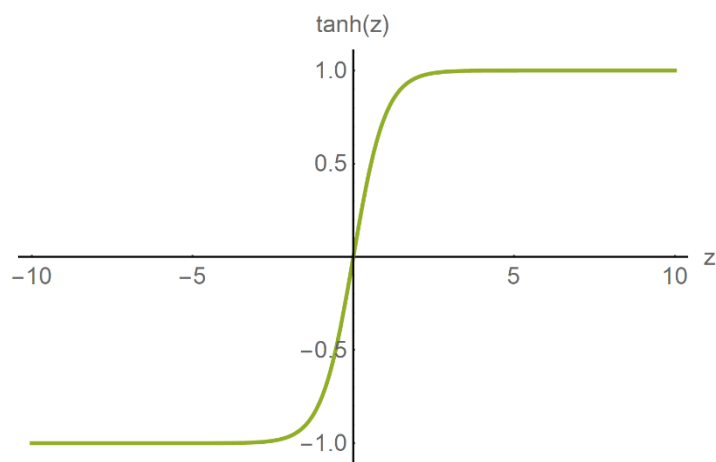
$\bar{x}$  – выборочное среднее;

$\sigma$  – выборочное отклонение

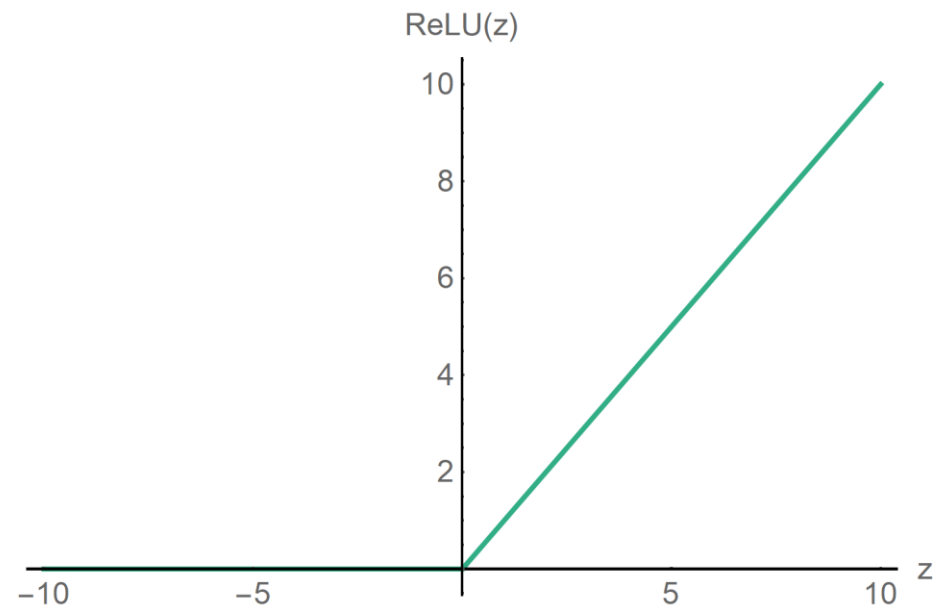
# Выбор функции активации



$$\sigma(z) = \frac{1}{1 + e^{-z}}$$

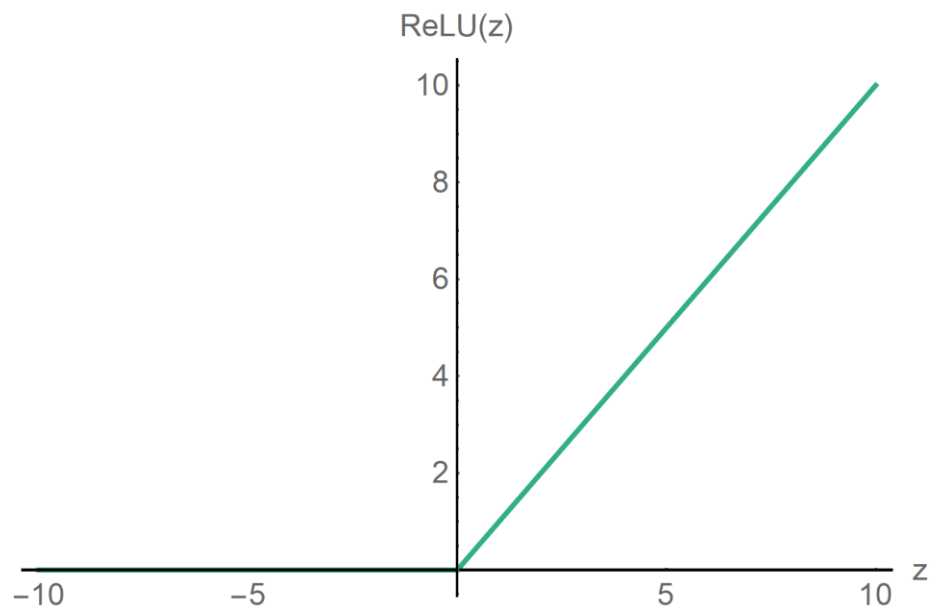


$$\tanh(z) = \frac{e^z - e^{-z}}{e^z + e^{-z}}$$

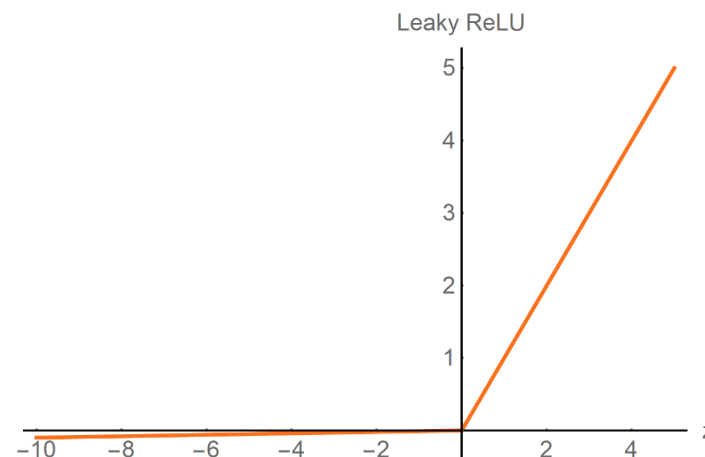


$$\text{ReLU}(z) = \begin{cases} 0, & z < 0, \\ z, & z \geq 0 \end{cases}$$

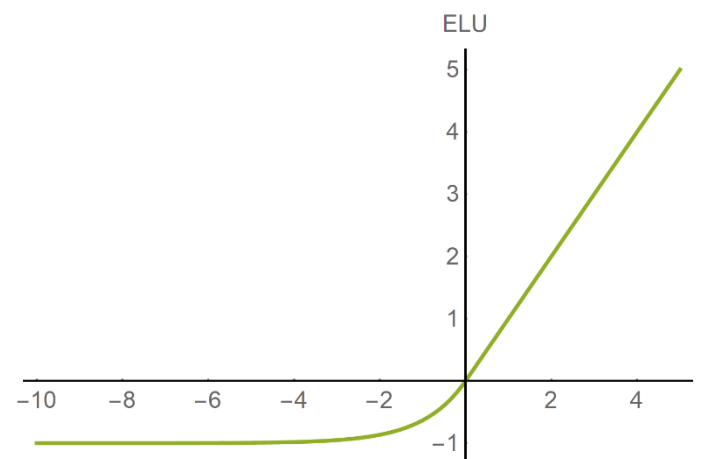
# Выбор функции активации



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



$$\text{Leaky ReLU}(z) = \begin{cases} 0.01 z, & z < 0, \\ z, & z \geq 0 \end{cases}$$



$$\text{ELU}(\alpha, z) = \begin{cases} \alpha(e^z - 1), & z < 0, \\ z, & z \geq 0 \end{cases}$$

# Регуляризация

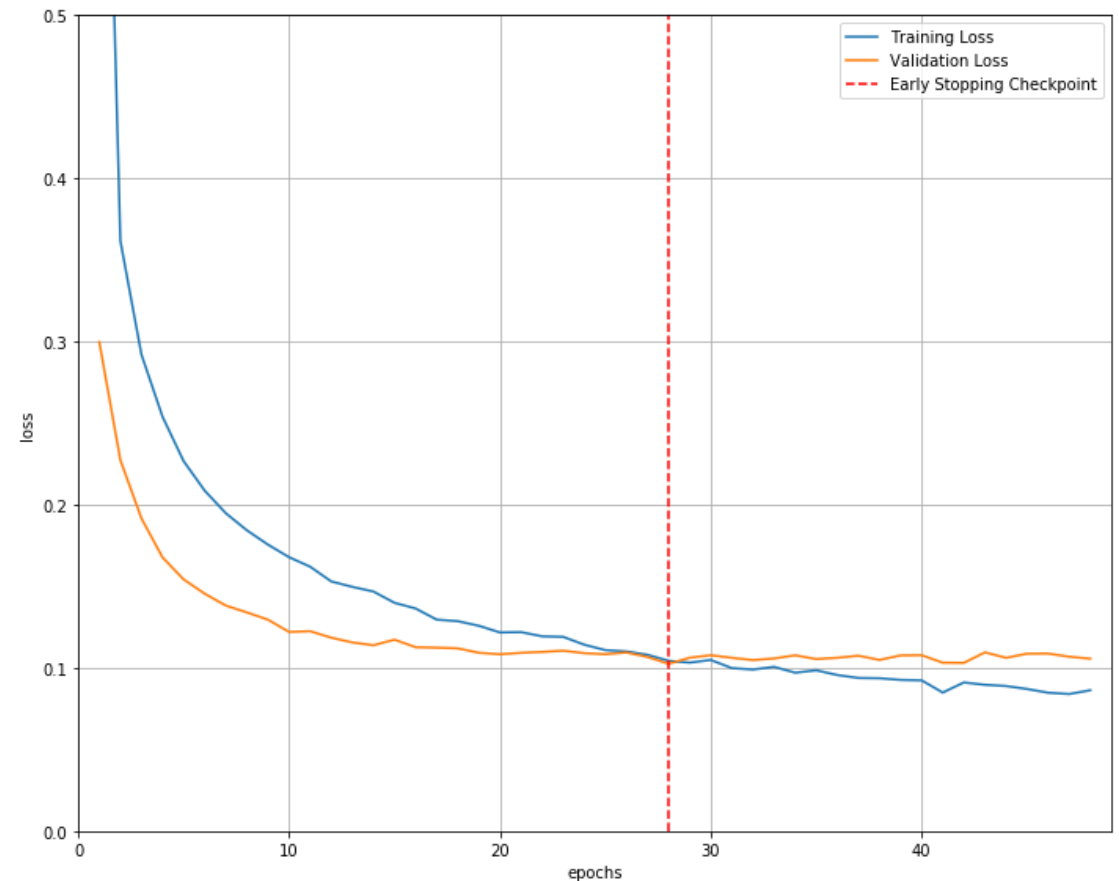
- $L_2$ -регуляризация:

$$L(y; x, \theta) + \frac{\lambda}{2} \|\theta\|^2;$$

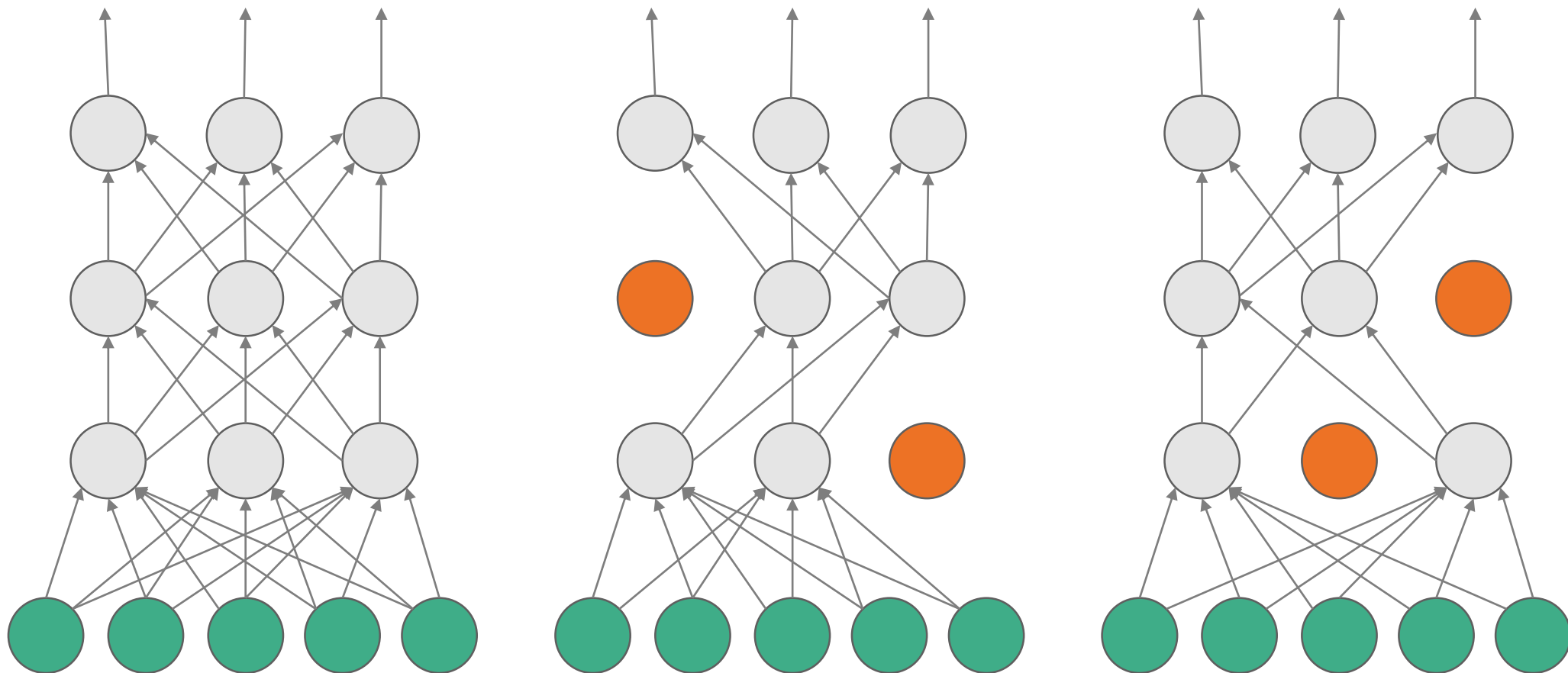
- $L_1$ -регуляризация:

$$L(y; x, \theta) + \lambda |\theta|;$$

- метод ранней остановки (early stopping):

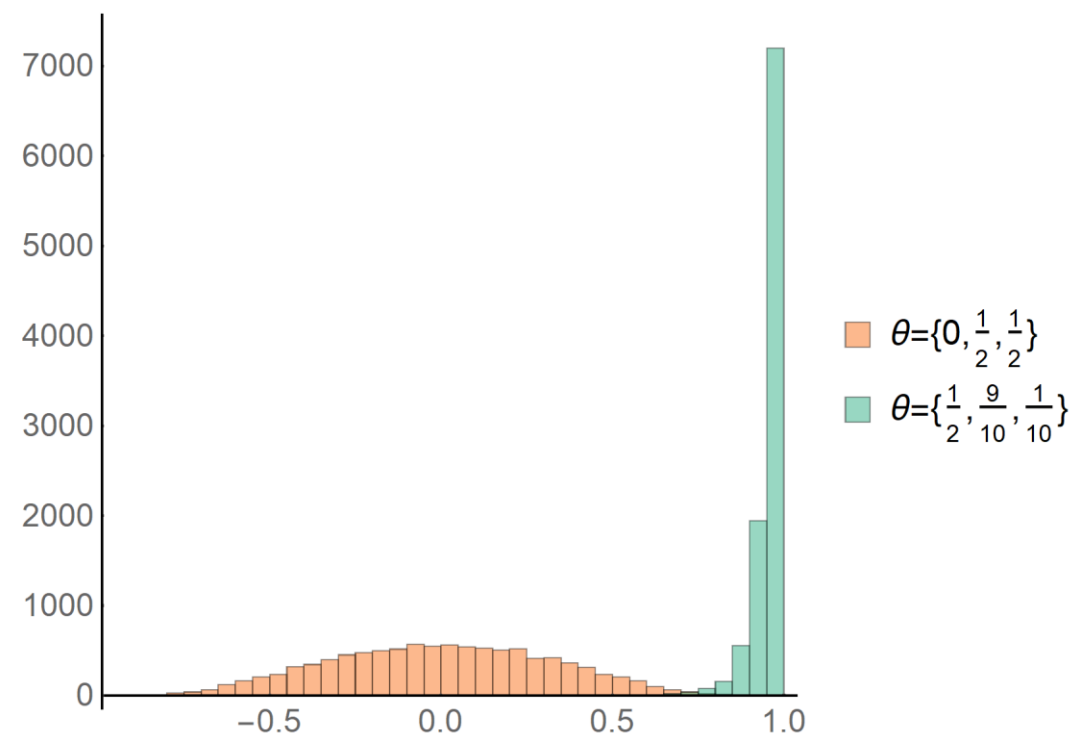
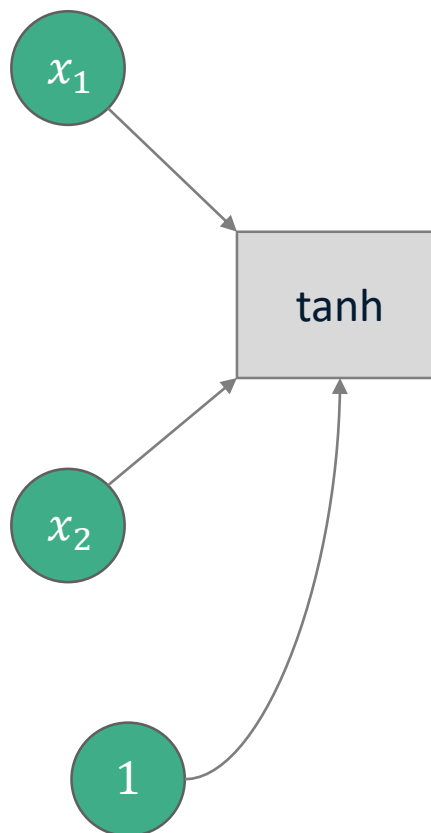
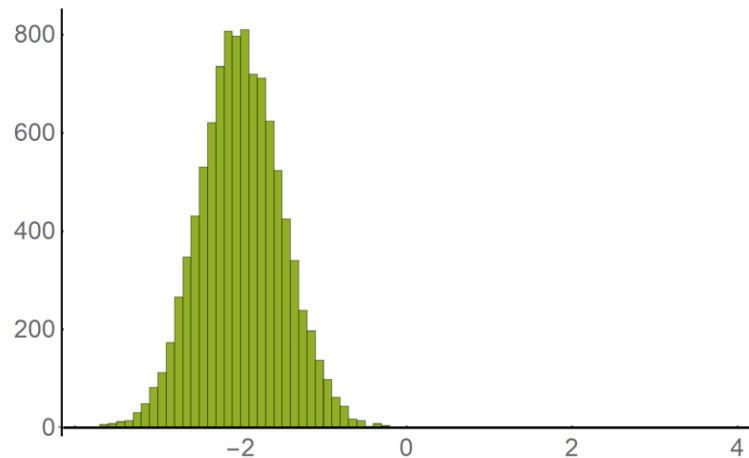
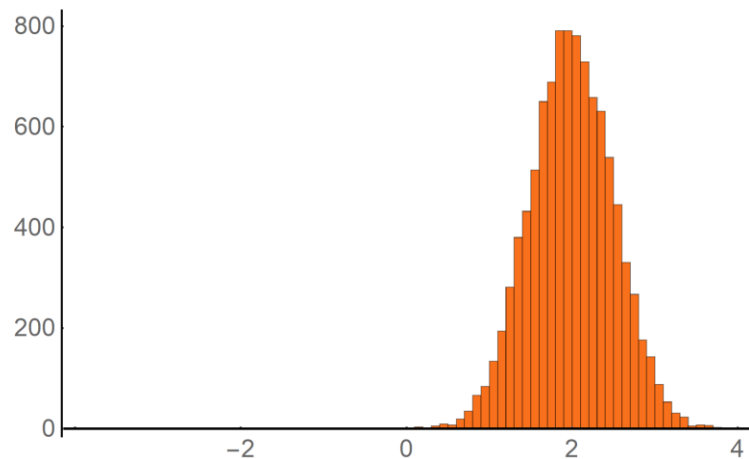


# Регуляризация. Dropout



Вероятность исключения нейрона из сети:  $\delta \sim \text{Bernoulli}(p)$

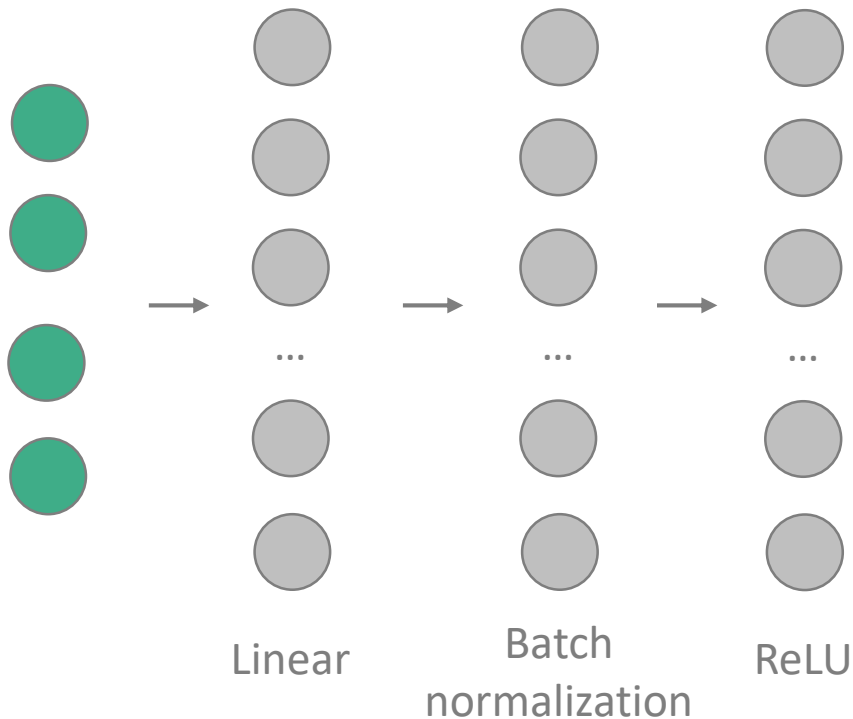
# Нормализация по мини-батчам



# Нормализация по мини-батчам

$$\hat{x}^{(k)} = \frac{x^{(k)} - \mathbb{E}[x^{(k)}]}{\sqrt{\text{Var}[x^{(k)}]}},$$

$$y^{(k)} = \gamma^{(k)} \hat{x}^{(k)} + \beta^{(k)}$$



- стабилизирует обучение;
- регуляризует;
- обучение меньше зависит от начальных значений параметров.



# Элементы нейронных сетей

Для улучшения сходимости обучения нейронных сетей:

- нормализация исходных данных;
- нормализация по мини-батчам.

Методы регуляризации:

- $L_1$  и  $L_2$ -регуляризаторы;
- early stopping (метод ранней остановки);
- dropout (метод отключения случайных нейронов).