

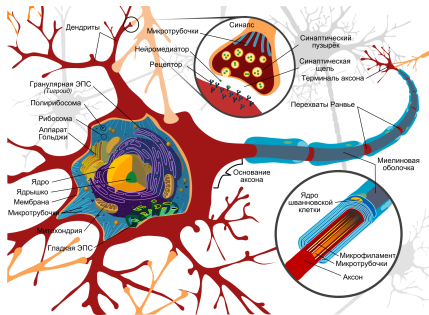
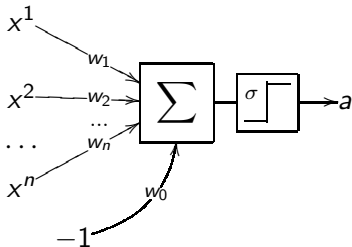
# Линейная модель нейрона МакКаллока-Питтса

$f_j: X \rightarrow \mathbb{R}$ ,  $j = 1, \dots, n$  — числовые признаки,  $x^j = f_j(x)$ ;

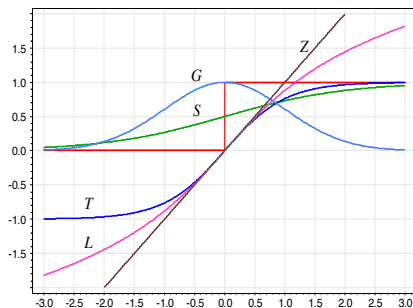
$$a(x, w) = \sigma(\langle w, x \rangle) = \sigma\left(\sum_{j=1}^n w_j f_j(x) - w_0\right),$$

где  $w_0, w_1, \dots, w_n \in \mathbb{R}$  — веса признаков;

$\sigma(s)$  — функция активации (в частности, sign).



## Часто используемые функции активации $\sigma(z)$



$$\theta(z) = [z \geq 0]$$

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

$$\text{th}(z) = 2\sigma(2z) - 1$$

$$\ln(z + \sqrt{z^2 + 1})$$

$$\exp(-z^2/2)$$

$$z$$

— пороговая функция Хевисайда;

— сигмоидная функция (S);

— гиперболический тангенс (T);

— логарифмическая функция (L);

— гауссовская функция (G);

— линейная функция (Z);

**Задача классификации:**  $Y = \{\pm 1\}$ ,  $a(x, w) = \text{sign}\langle w, x_i \rangle$ ;

$$Q(w; X^\ell) = \sum_{i=1}^{\ell} \mathcal{L}(\underbrace{\langle w, x_i \rangle}_{M_i(w)} y_i) \rightarrow \min_w;$$

**Задача регрессии:**  $Y = \mathbb{R}$ ,  $a(x, w) = \sigma(\langle w, x_i \rangle)$ ;

$$Q(w; X^\ell) = \sum_{i=1}^{\ell} (\sigma(\langle w, x_i \rangle) - y_i)^2 \rightarrow \min_w;$$

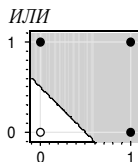
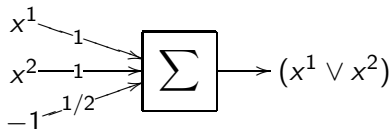
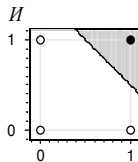
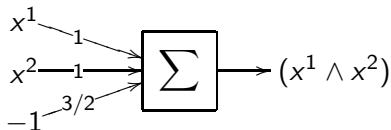
**Насколько богатый класс функций реализуется нейроном?  
А сетью (суперпозицией) нейронов?**

Функции И, ИЛИ, НЕ от бинарных переменных  $x^1$  и  $x^2$ :

$$x^1 \wedge x^2 = [x^1 + x^2 - \frac{3}{2} > 0];$$

$$x^1 \vee x^2 = [x^1 + x^2 - \frac{1}{2} > 0];$$

$$\neg x^1 = [-x^1 + \frac{1}{2} > 0];$$

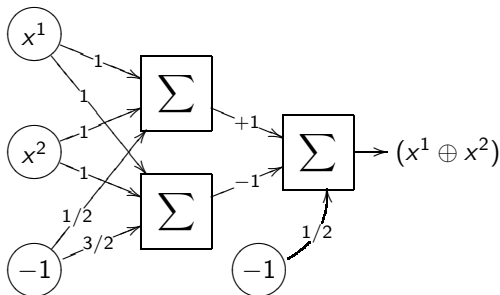


# Логическая функция XOR (исключающее ИЛИ)

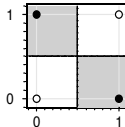
Функция  $x^1 \oplus x^2 = [x^1 \neq x^2]$  не реализуема одним нейроном.

Два способа реализации:

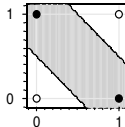
- Добавлением нелинейного признака:  
$$x^1 \oplus x^2 = [x^1 + x^2 - 2x^1x^2 - \frac{1}{2} > 0];$$
- Сетью** (двухслойной суперпозицией) функций И, ИЛИ, НЕ:  
$$x^1 \oplus x^2 = [(x^1 \vee x^2) - (x^1 \wedge x^2) - \frac{1}{2} > 0].$$



1-й способ



2-й способ



# Любую ли функцию можно представить нейросетью?

- Двухслойная сеть в  $\{0, 1\}^n$  позволяет реализовать произвольную булеву функцию (ДНФ).
- Двухслойная сеть в  $\mathbb{R}^n$  позволяет отделить произвольный выпуклый многогранник.
- Трёхслойная сеть  $\mathbb{R}^n$  позволяет отделить произвольную многогранную область, не обязательно выпуклую, и даже не обязательно связную.
- С помощью линейных операций и одной нелинейной функции активации  $\varphi$  можно приблизить любую непрерывную функцию с любой желаемой точностью.

## Практические рекомендации:

- Двух-трёх слоёв обычно достаточно.
- Можно достраивать нейроны в произвольных местах сети по необходимости, вообще не заботясь о числе слоёв.