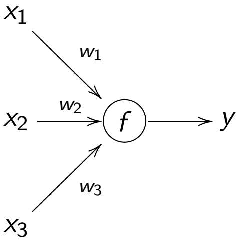
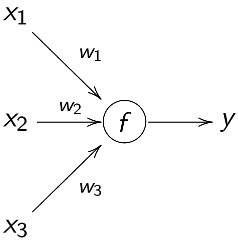
Персептрон

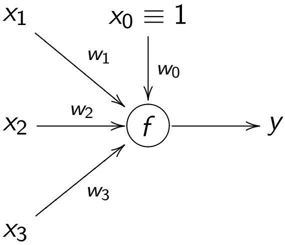


$$y = f\left(\sum_{i=1}^{n} w_i x_i\right)$$

$$f(x) = \mathrm{sign}(x)$$

Вес активации





Реализация конъюнкции

x_1	x_2	$x_1 \wedge x_2$
0	0	0
0	1	0
1	0	0
1	1	1

$$\begin{cases} w_0 & < & 0 \\ w_0 + w_1 & < & 0 \\ w_0 + w_2 & < & 0 \\ w_0 + w_1 + w_2 & > & 0 \end{cases}$$

Реализация конъюнкции

x_1	x_2	$x_1 \wedge x_2$
0	0	0
0	1	0
1	0	0
1	1	1

$$\begin{cases} w_0 & < & 0 \\ w_0 + w_1 & < & 0 \\ w_0 + w_2 & < & 0 \\ w_0 + w_1 + w_2 & > & 0 \end{cases}$$

$$\begin{aligned} w_0 &= 3 \\ w_1 &= 2 \\ w_2 &= 2 \end{aligned}$$

Реализация дизъюнкции

x_1	x_2	$x_1 \vee x_2$
0	0	0
0	1	1
1	0	1
1	1	1

$$\begin{cases} w_0 & < & 0 \\ w_0 + w_1 & > & 0 \\ w_0 + w_2 & > & 0 \\ w_0 + w_1 + w_2 & > & 0 \end{cases}$$

$$\begin{aligned} w_0 &= 1 \\ w_1 &= 2 \\ w_2 &= 2 \end{aligned}$$

Геометрическая интерпретация

в PP

Обучение персептрона

Тут про советчиков, м.б. тоже в РР

Функция XOR
Функция XOR
Многослойный персептрон

Постановка задачи

Дано:

$\mathcal{I} = (I_1, \ldots, I_k)$	входные вектора размерности n
$\mathcal{A} = (A_1, \ldots, A_k)$	правильные выходные вектора размерности m
$(\mathcal{I}, \mathcal{A})$	обучающая выборка
$N(W, I)$	функция, соответствующая нейронной сети
$O_i = N(W, I_i)$	ответ нейронной сети, вектор размерности m
$E(O_i, A_i) = \sum_{j=1}^{m}(O_i[j] - A_i[j])^2$	функция ошибки

Найти: вектор W такой, что $\sum_{i=1}^{k} E(N(W, I_i) - A_i) \to \min$

Обучение онлайн

Решим задачу для одной пары (I, A)

В этом случае $E(N(W_i, I) - A)$ является функцией от вектора весов $E = E(W)$.

Алгоритм градиентного спуска

1. Инициализировать x_1 случайным значением из $\mathbb{R}$
2. $i := 1$
3. $x_{i+1} = x_i + \varepsilon f'(x_i)$
4. $i++$
5. $\text{if} \;\; |x_{i+1} - x_i| > c \;\; \text{goto } 3$

Алгоритм градиентного спуска

1. Инициализировать W_1 случайным значением из $\mathbb{R}^n$
2. $i := 1$
3. $W_{i+1} = W_i + \varepsilon \nabla f(W_i)$
4. $i++$
5. $\text{if } \; ||W_{i+1} - W_i|| > c \; \text{ goto } 3$

Частные производные

Функция n переменных:

$$F : \mathbb{R}^n \to \mathbb{R}$$

$$F(x_1, \dots, x_n)$$

Частная производная по i-й переменной:

$$\frac{\partial F}{\partial x_i}(x_1, \dots, x_n) =$$

$$= \lim_{\varepsilon \to 0} \frac{F(x_1, x_2, \dots, x_i + \varepsilon, \dots, x_n) - F(x_1, x_2, \dots, x_i, \dots, x_n)}{\varepsilon}$$

$$\frac{\partial F}{\partial x_i} : \mathbb{R}^n \to \mathbb{R}$$

Частные производные

$$F(x, y, z, u) = x^3 + y^u + \sin z^2 u^3$$

$$\frac{\partial F}{\partial x} =$$
$$\frac{\partial F}{\partial y} =$$
$$\frac{\partial F}{\partial z} =$$

Частные производные

$$F(x, y, z, u) = x^3 + y^u + \sin z^2 u^3$$

$$\begin{aligned} \frac{\partial F}{\partial x} &= 3x^2 \\ \frac{\partial F}{\partial y} &= \\ \frac{\partial F}{\partial z} &= \end{aligned}$$

Частные производные

$$F(x, y, z, u) = x^3 + y^u + \sin z^2 u^3$$

$$\begin{aligned}
\frac{\partial F}{\partial x} &= 3x^2 \\
\frac{\partial F}{\partial y} &= uy^u - 1 \\
\frac{\partial F}{\partial z} &=
\end{aligned}$$

Частные производные

$$F(x, y, z, u) = x^3 + y^u + \sin z^2 u^3$$

$$\begin{aligned} \frac{\partial F}{\partial x} &= 3x^2 \\ \frac{\partial F}{\partial y} &= uy^u - 1 \\ \frac{\partial F}{\partial z} &= (-\cos z^2 u^3)(u^3 2z) \end{aligned}$$

Производная сложной функции

$$F = F(x_1, \ldots, x_n)$$

$$G_i = G_i(y_1, \ldots, y_m)$$

$$H(y_1, \ldots, y_n) = F(G_1(y_1, \ldots, y_m), \ldots, G_n(y_1, \ldots, y_n))$$

$$\frac{\partial H}{\partial y_i} = \sum_{j=1}^{n} \frac{\partial H}{\partial G_j} \frac{\partial G_j}{\partial y_i}$$

Производная сложной функции

$$F(x_1, \ldots, x_n) = \sum_{k=1} na_i x_i$$

$$G_i(y_1, \ldots, y_m) = \sum_{k=1} my_k^i$$

$$H(y_1, \ldots, y_n) = F(G_1(y_1, \ldots, y_m), \ldots, G_n(y_1, \ldots, y_n))$$

$$\frac{\partial H}{\partial y_j} =$$

Производная сложной функции

$$F(x_1,\dots,x_n) = \sum_{k=1} n a_i x_i$$

$$G_i(y_1,\dots,y_m) = \sum_{k=1} m y_k^i$$

$$H(y_1,\dots,y_n) = F(G_1(y_1,\dots,y_m),\dots,G_n(y_1,\dots,y_n))$$

$$\frac{\partial F}{\partial G_i} = a_i,\ \frac{\partial G_i}{\partial y_j} = i y_j^{i-1}$$

$$\frac{\partial H}{\partial y_j} =$$

Производная сложной функции

$$F(x_1, \ldots, x_n) = \sum_{k=1} n a_i x_i$$

$$G_i(y_1, \ldots, y_m) = \sum_{k=1} m y_k^i$$

$$H(y_1, \ldots, y_n) = F(G_1(y_1, \ldots, y_m), \ldots, G_n(y_1, \ldots, y_n))$$

$$\frac{\partial F}{\partial G_i} = a_i, \ \frac{\partial G_i}{\partial y_j} = i y_j^{i-1}$$

$$\frac{\partial H}{\partial y_j} = \sum_{i=1}^{n} \frac{\partial H}{\partial G_i} \frac{\partial G_i}{\partial y_j} = \sum_{i=1}^{n} a_i i y_j^{i-1}$$

Градиент

$$\nabla F = \left(\frac{\partial F}{\partial x_1}, \frac{\partial F}{\partial x_2}, \ldots, \frac{\partial F}{\partial x_n}\right)$$

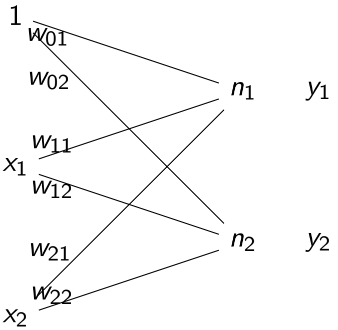
$$\nabla F :$$

Градиент

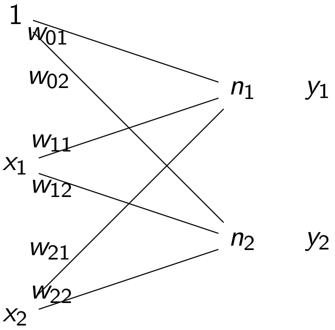
$$\nabla F = \left(\frac{\partial F}{\partial x_1}, \frac{\partial F}{\partial x_2}, \dots, \frac{\partial F}{\partial x_n} \right)$$

$$\nabla F : \mathbb{R}^n \to \mathbb{R}^n$$

Обратное распространение ошибки



Обратное распространение ошибки



$$E(y_1, y_2) = (y_1 - a_1)^2 + (y_2 - a_2)^2$$

$$\frac{\partial E}{\partial y_1} = 2(y_1 - a_1)$$

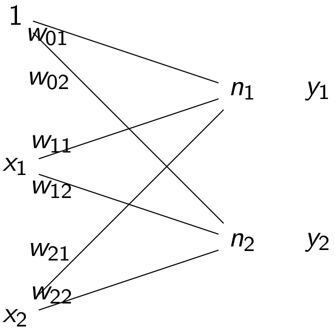
$$y_1 = f\underbrace{\left(w_{01} + x_1 w_{11} + x_2 w_{21}\right)}_{S_1}$$

$$\frac{\partial y_1}{\partial w_{21}} = f'(S_1)x_2$$

$$\frac{\partial E}{\partial w_{21}} = \frac{\partial E}{\partial y_1}\frac{\partial y_1}{\partial w_{21}}$$

$$= 2(y_1 - a_1)f'(S_1)x_2$$

Обратное распространение ошибки



$$E(y_1, \dots, y_n) = (y_i - a_i)^2 + \dots + (y_n - a_n)^2$$

$$\frac{\partial E}{\partial y_i} = 2(y_i - a_i)$$

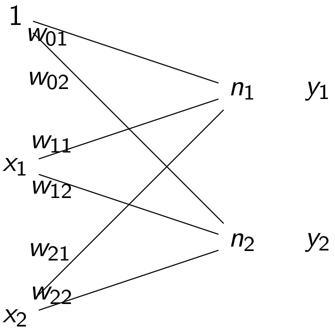
$$y_i = f\,\underbrace{\left(w_{0i} + x_1 w_{1i} + \dots + x_n w_{ni}\right)}_{S_i}$$

$$\frac{\partial y_i}{\partial w_{ji}} = f'(S_i)x_j$$

$$\frac{\partial E}{\partial w_{ji}} = \frac{\partial E}{\partial y_i}\frac{\partial y_i}{\partial w_{ji}}$$

$$= 2(y_i - a_i)f'(S_i)x_j$$

Обратное распространение ошибки



$$E(y_1, y_2) = (y_1 - a_1)^2 + (y_2 - a_2)^2$$

$$\frac{\partial E}{\partial y_1} = 2(y_1 - a_1)$$

$$y_1 = f\,\underbrace{(w_{01} + x_1 w_{11} + x_2 w_{21})}_{S_1}$$

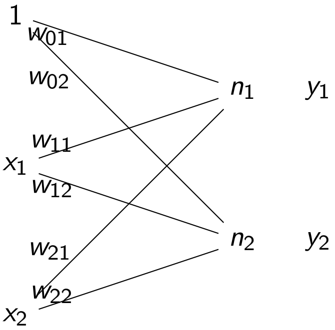
$$\frac{\partial y_1}{\partial x_2} = f'(S_1) w_{12}$$

$$\frac{\partial y_2}{\partial x_2} = f'(S_2) w_{22}$$

$$\frac{\partial E}{\partial x_2} = \frac{\partial E}{\partial y_1}\frac{\partial y_1}{\partial x_2} + \frac{\partial E}{\partial y_1}\frac{\partial y_1}{\partial x_2} =$$

$$= 2(y_1 - a_1) f'(S_1) w_{12} + 2(y_2 - a_2) f'(S_2) w_{22}$$

Обратное распространение ошибки



$$E(y_1, ..., y_n) = (y_i - a_i)^2 + ... + (y_n - a_n)^2$$

$$\frac{\partial E}{\partial y_i} = 2(y_i - a_i)$$

$$y_i = f\underbrace{\left(w_{0i} + x_1 w_{1i} + ... + x_n w_{ni}\right)}_{S_i}$$

$$\frac{\partial y_i}{\partial x_j} = f'(S_i) w_{ji}$$

$$\frac{\partial E}{\partial x_j} = \sum_{i=1}^{n} \frac{\partial E}{\partial y_i} \frac{\partial y_i}{\partial x_j} = \sum_{i=1}^{n} 2(y_i - a_i) f'(S_i) w_{ji}$$