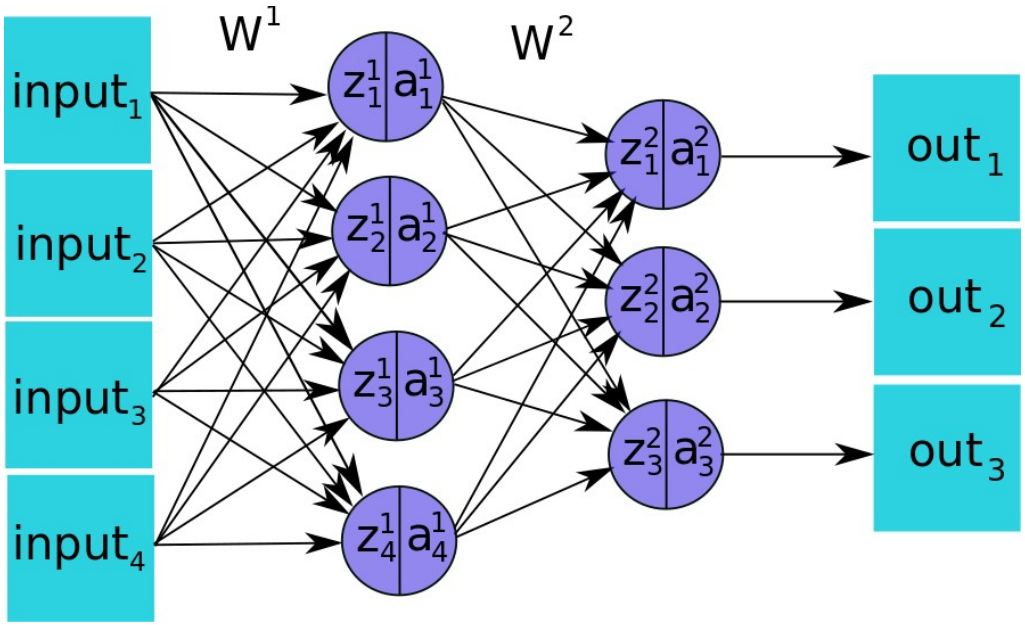
DL: СЕТЬ ПРЯМОГО РАСПОСТРАНЕНИЯ

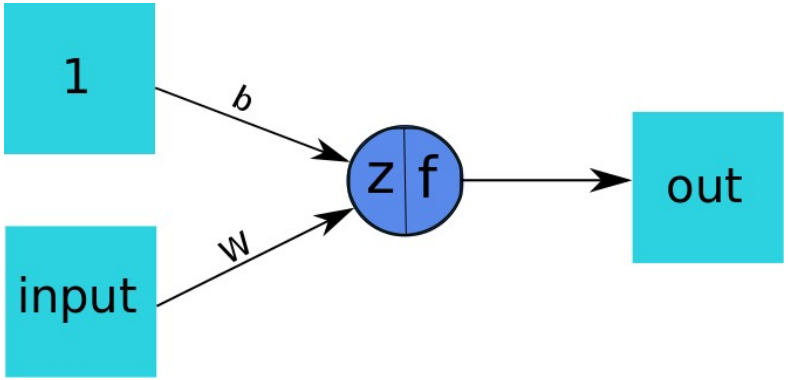
План

- Сеть прямого распостранения
- Функция активации
- Алгоритм обратного распостранения
- Автоматический градиент

НС прямого распостранения

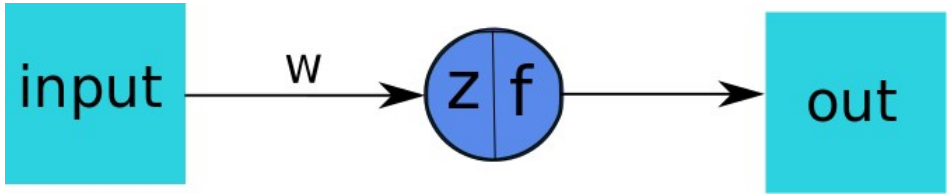


Один вход, один нейрон



$$out = f\left(w * input + 1 * b\right)$$

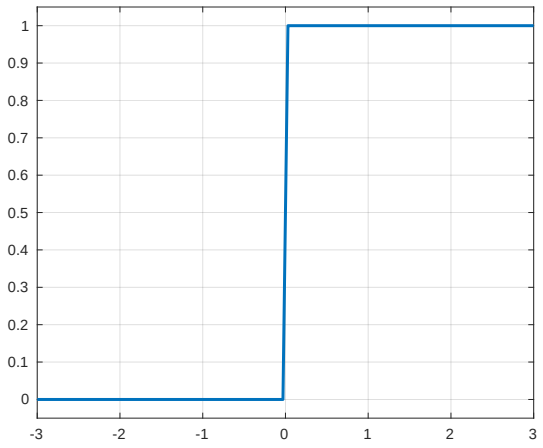
$$z = w * input + 1 * b$$



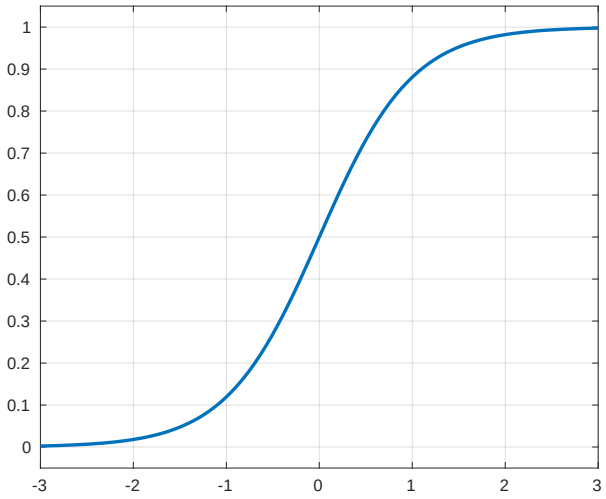
$$y = a * x + b$$

Функция активации

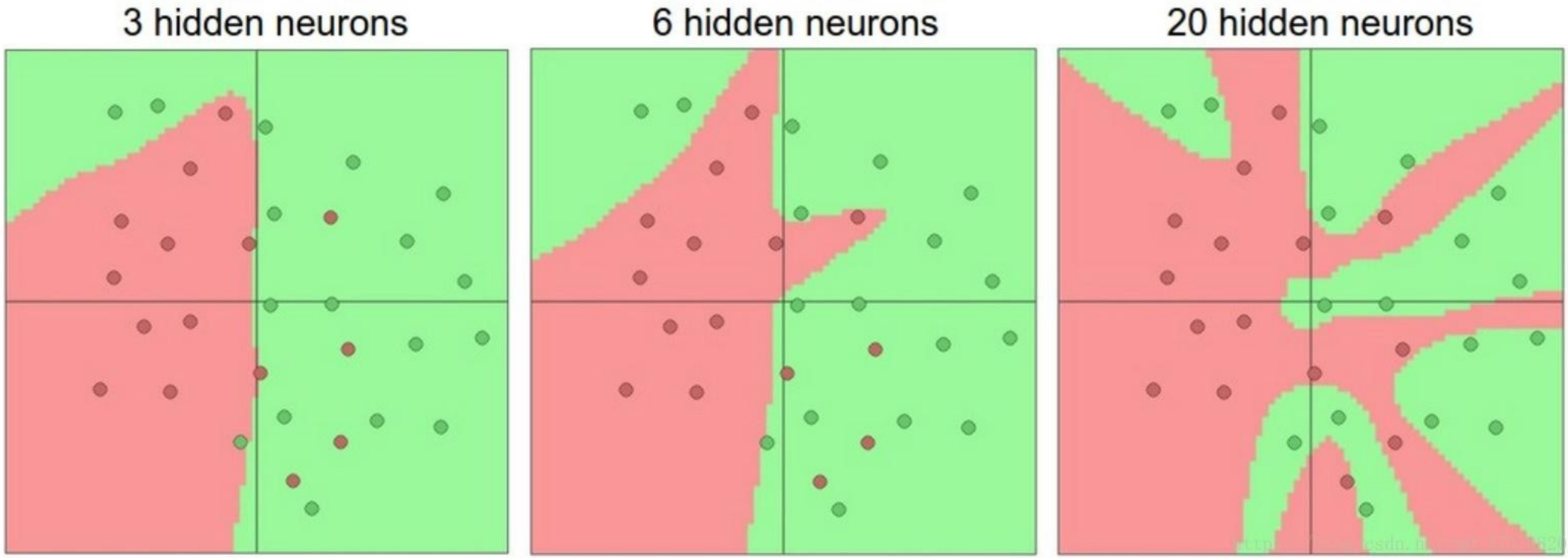
$$f(x)=\begin{cases} 1, & if\ x\geq 0 \\ -1, & if\ x<0 \end{cases}$$



$$f(x)=\frac{1}{1+e^{-x}}$$



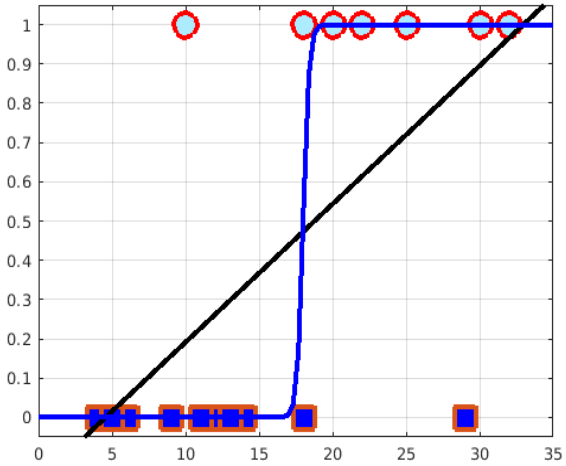
Функция активации

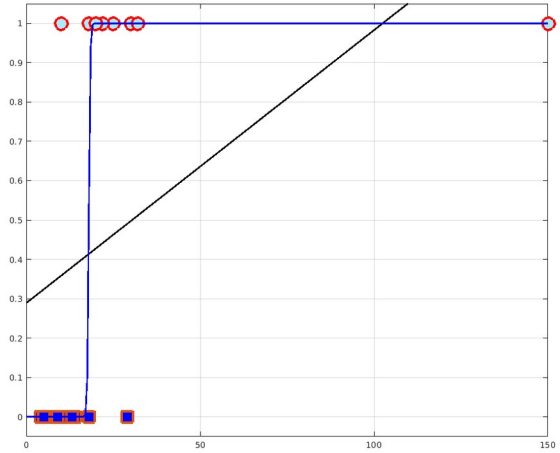


Функция активации: сигмоид

$$z = \beta_0 + \beta_1 x$$

$$z \in (-\infty; +\infty) \Rightarrow P \in [0; 1]$$





$$OR = \frac{P}{1-P} = e^{\ln \frac{P}{1-P}}$$

$$OR \in [0; \infty)$$

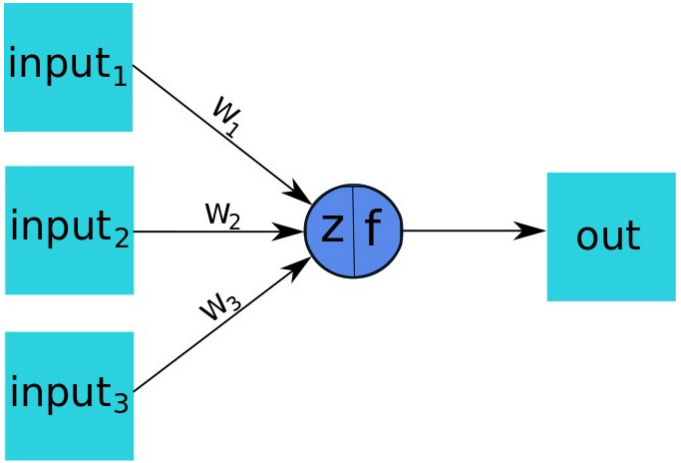
$$\beta_0 + \beta_1 x = \ln \frac{P}{1-P} = z$$

$$OR = \frac{P}{1-P} \Rightarrow P = \frac{OR}{1+OR} = \frac{e^z}{1+e^z} = \frac{1}{1+e^{-z}}$$

$$\ln \frac{P}{1-P} \in (-\infty; \infty)$$

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

НС: 3 входа, 1 скрытый нейрон

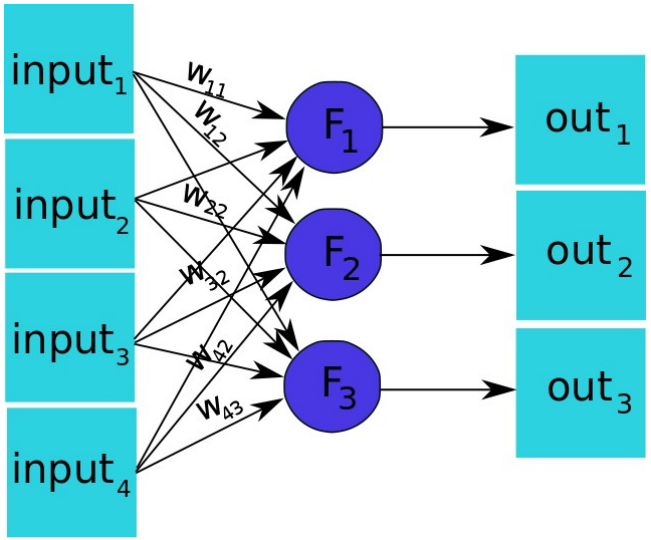


$$out = f(z)$$

$$z = w_1 * input_1 + w_2 * input_2 + w_3 * input_3 + b$$

$$f(z) = \frac{1}{1+e^{-z}} = \frac{1}{1+e^{-(w_1 * input_1 + w_2 * input_2 + w_3 * input_3 + b)}}$$

НС: 4 входа, 3 выхода

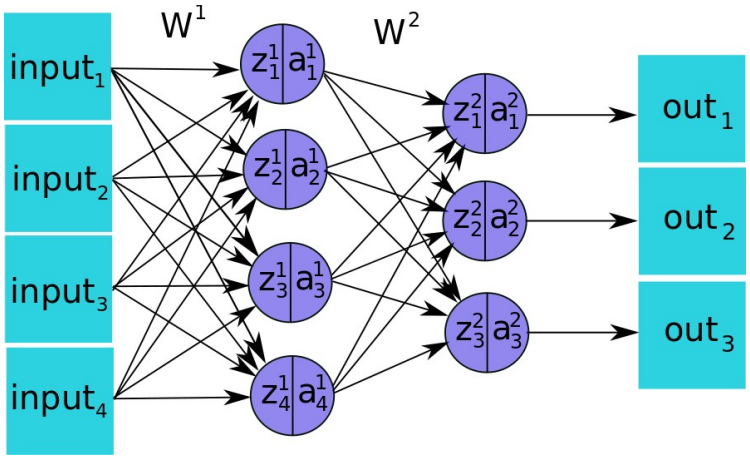


$$INPUT = \begin{pmatrix} input_1 \\ input_2 \\ input_3 \\ input_4 \end{pmatrix} \qquad B = \begin{pmatrix} b_1 \\ b_2 \\ b_3 \end{pmatrix} \qquad W = \begin{pmatrix} w_{11} & w_{12} & w_{13} \\ w_{21} & w_{22} & w_{23} \\ w_{31} & w_{32} & w_{33} \\ w_{41} & w_{42} & w_{43} \end{pmatrix}$$

$$OUT = F\left(W^T \cdot INPUT + B\right)$$

$$\begin{cases} out_1 = F\left(w_{11} \cdot input_1 + w_{21} \cdot input_2 + w_{31} \cdot input_3 + w_{41} \cdot input_4 + b_1\right) \\ out_2 = F\left(w_{12} \cdot input_1 + w_{22} \cdot input_2 + w_{32} \cdot input_3 + w_{42} \cdot input_4 + b_2\right) \\ out_3 = F\left(w_{13} \cdot input_1 + w_{23} \cdot input_2 + w_{33} \cdot input_3 + w_{43} \cdot input_4 + b_3\right) \end{cases}$$

Многослойная сеть



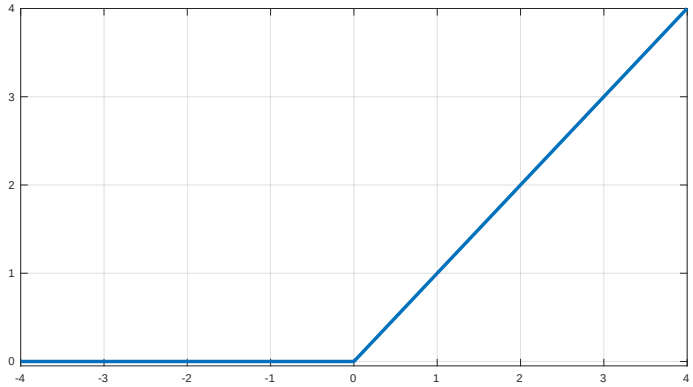
$$OUT^1 = A\left(W^{1T} \cdot INPUT + B^1\right)$$

$$OUT = A\left(W^{2T} \cdot INPUT^2 + B^2\right) = A\left(W^{2T} \cdot OUT^1 + B^2\right)$$

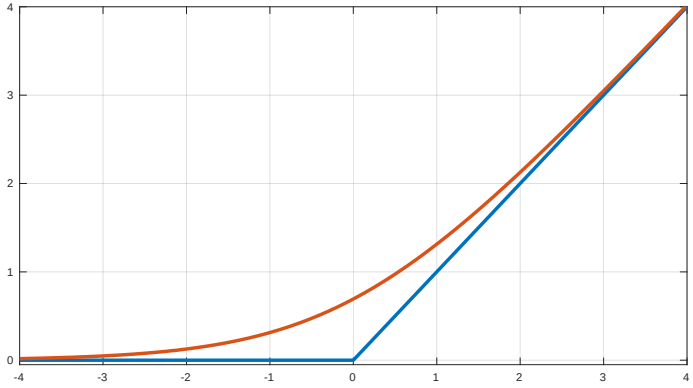
$$OUT = A\left(W^{2T} \cdot A\left(W^{1T} \cdot INPUT + B^1\right) + B^2\right)$$

Ректификация (relu)

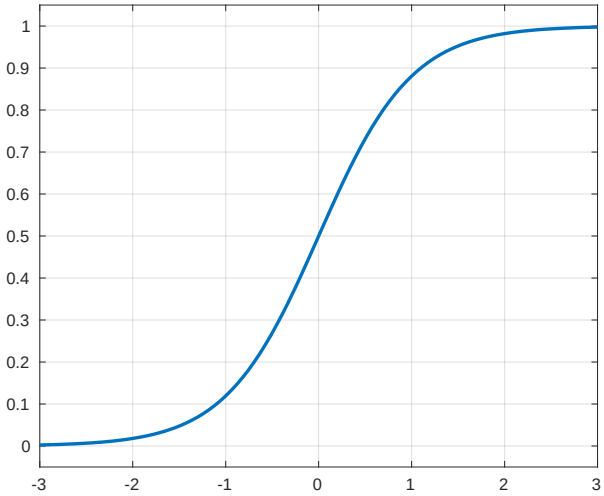
$$y = max(x, 0)$$



$$y = \ln(1 + e^{x})$$



Ректификация (relu)

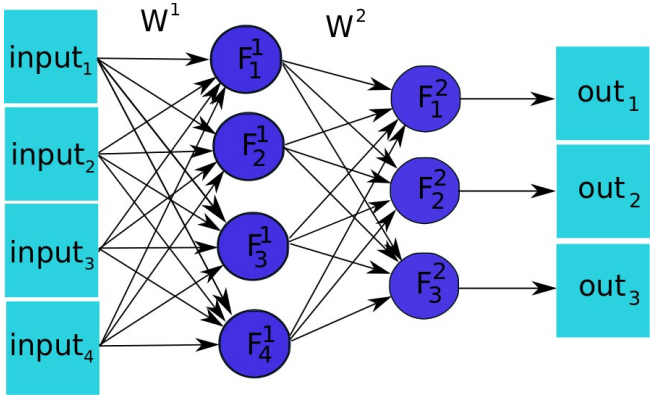


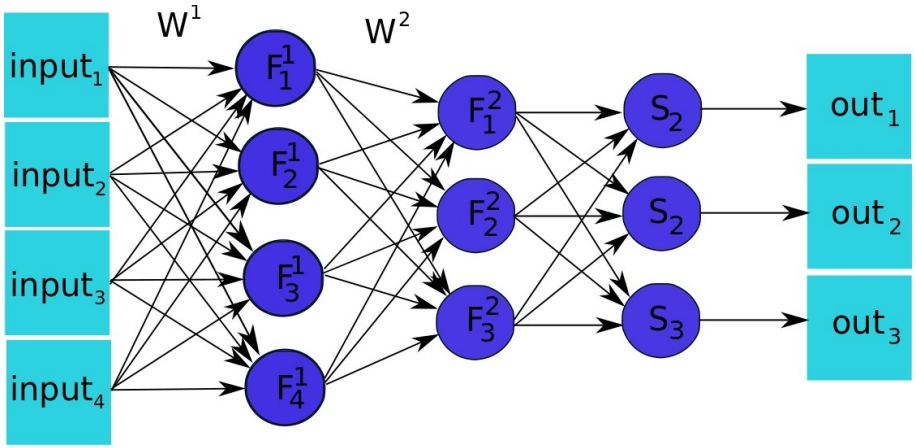
$$f(x)=\frac{1}{1+e^{-x}} \qquad f(5)\approx f(10)$$

$$f(x)=\sigma(x+\frac{1}{2})+\sigma(x-\frac{1}{2})+\sigma(x-\frac{3}{2})+\sigma(x-\frac{5}{2})+\ldots$$

$$f(x)=\sum_{i=0}^{\infty}\sigma(x+\frac{1}{2}-i)\approx\int_{1/2}^{\infty}\sigma(x+\frac{1}{2}-y)\,dy=[-\log(1+e^{x+\frac{1}{2}-y})]_{y=1/2}^{y=\infty}=\log(1+e^{x})$$

Софтмакс





$$P(C_j|data) = \frac{P(data|C_j)P(C_j)}{\sum_{k=1}^{K} P(data|C_k)P(C_k)}$$

$$z_k = \ln P(data|C_k)P(C_k)$$

$$\sigma(z_j) = \frac{e^{z_j}}{\sum_{k=1}^{K} e^{z_k}}$$

Функция ошибки

Регрессия

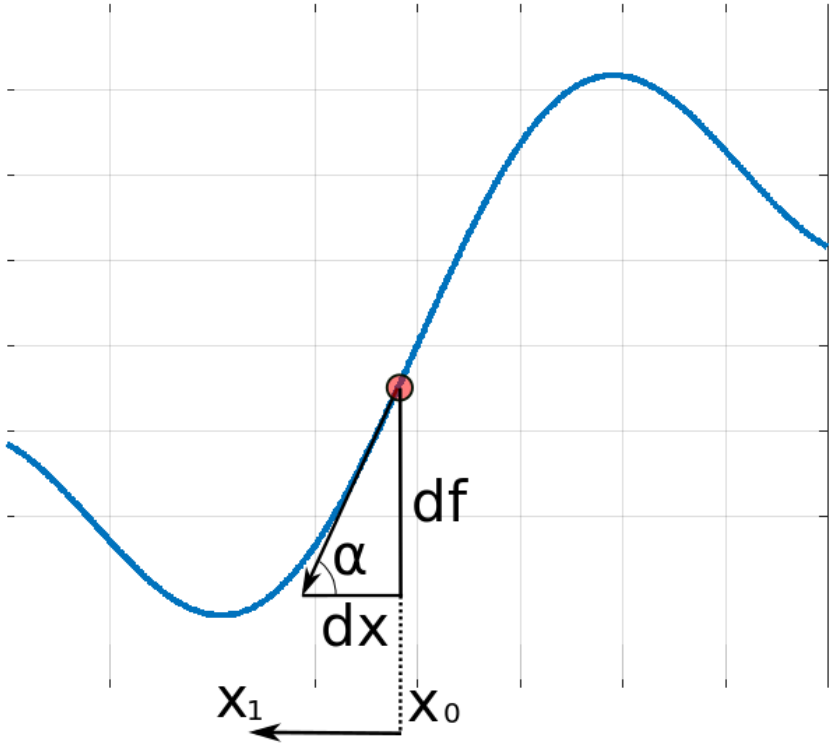
$$E = \sqrt{Y^2 - T^2}$$

Классификация

$$H(p,q) = H(p) + D_{KL}(p,q)$$

$$H(p,q) = -\sum_x p(x) \log(q(x))$$

Оптимизация



$$x_1 = x_0 - \eta \left. \frac{df}{dx} \right|_{x=x_0}$$

$$\Delta x_0 = -\eta \left. \frac{df}{dx} \right|_{x=x_0}$$

$$x_1 = x_0 + \Delta x_0$$

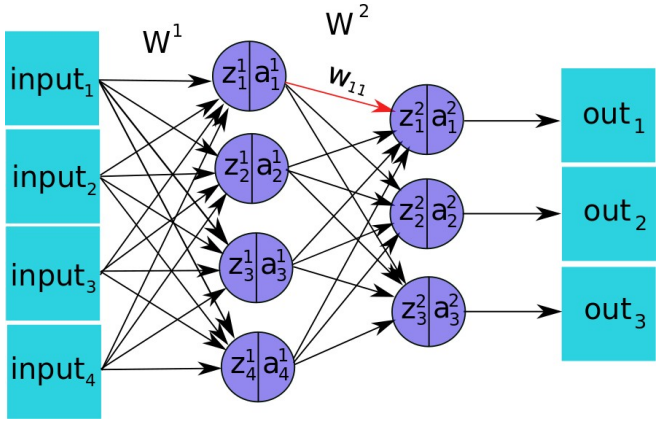
Алгоритм обратного распостранения ошибки: последний слой

$$\Delta w_{11}^2 = -\eta \frac{\partial E}{\partial w_{11}^2}, \qquad \frac{\partial E}{\partial w_{11}^2} = \frac{\partial E}{\partial a_1^2}\frac{\partial a_1^2}{\partial z_1^2}\frac{\partial z_1^2}{\partial w_{11}^2}$$

$$E = \frac{1}{2}\sum_{i=1}^{N}\left(a_{1i}^2 - t_{1i}\right)^2 \Rightarrow \frac{\partial E}{\partial a_{1i}^2} = \sum_{i=1}^{N} a_1^2 - t_{1i}$$

$$a_1^2 = \frac{1}{1+e^{-z_1^2}} = \sigma\left(z_1^2\right) \Rightarrow \frac{\partial a_1^2}{\partial z_1^2} = \sigma\left(z_1^2\right)\left(1-\sigma\left(z_1^2\right)\right) = a_1^2\left(1-a_1^2\right)$$

$$z_1^2 = a_1^1 w_{11}^2 + a_2^1 w_{21}^2 + a_3^1 w_{31}^2 a_4^1 w_{41}^2 + b_1^2 \Rightarrow \frac{\partial z_1^2}{\partial w_{11}^2} = a_1^1$$



$$\frac{\partial E}{\partial w_{11}^2} = a_1^2\left(1-a_1^2\right)a_1^1\sum_{i=1}^{N} a_1^2 - t_i$$

$$\frac{\partial E}{\partial w_{ij}^l} = a_j^l\left(1-a_j^l\right)a_i^{l-1}\sum_{k=1}^{N} a_{jk}^l - t_{jk}$$

$$\frac{\partial E}{\partial a_j^l}\frac{\partial a_j^l}{\partial z_j^l} = \delta_j^l \qquad \frac{\partial E}{\partial w_{ij}^l} = \delta_j^l a_i^{l-1}$$

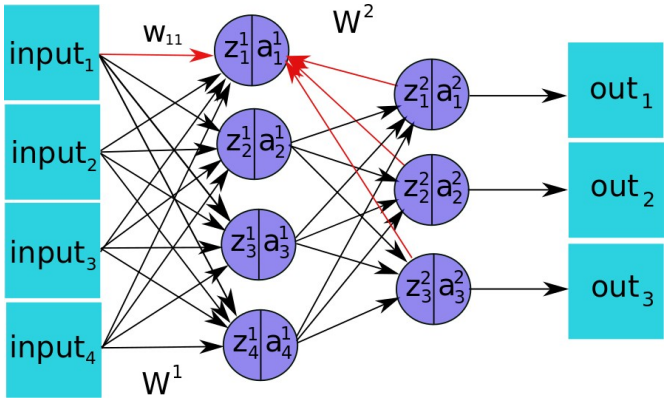
Алгоритм обратного распостранения ошибки: внутренний слой

$$\Delta w_{11}^1 = -\eta \frac{\partial E}{\partial w_{11}^1},$$

$$\frac{\partial E}{\partial w_{11}^1} = \frac{\partial E}{\partial a_1^1}\frac{\partial a_1^1}{\partial z_1^1}\frac{\partial z_1^1}{\partial w_{11}^1} = \frac{\partial E}{\partial z_1^1}\frac{\partial z_1^1}{\partial w_{11}^1}$$

$$\frac{\partial E}{\partial z_1^1} = \sum_{k=1}^{3}\frac{\partial E}{\partial z_k^2}\frac{\partial z_k^2}{\partial z_1^1} = \sum_{k=1}^{3}\frac{\partial E}{\partial z_k^2}\frac{\partial z_k^2}{\partial a_1^1}\frac{\partial a_1^1}{\partial z_1^1}$$

$$z_k^2 = a_1^1 w_{1k}^2 + a_2^1 w_{2k}^2 + a_3^1 w_{3k}^2 + a_4^1 w_{4k}^2 + b_k^2 \Rightarrow \frac{\partial z_k^2}{\partial a_1^1} = w_{1k}^2$$

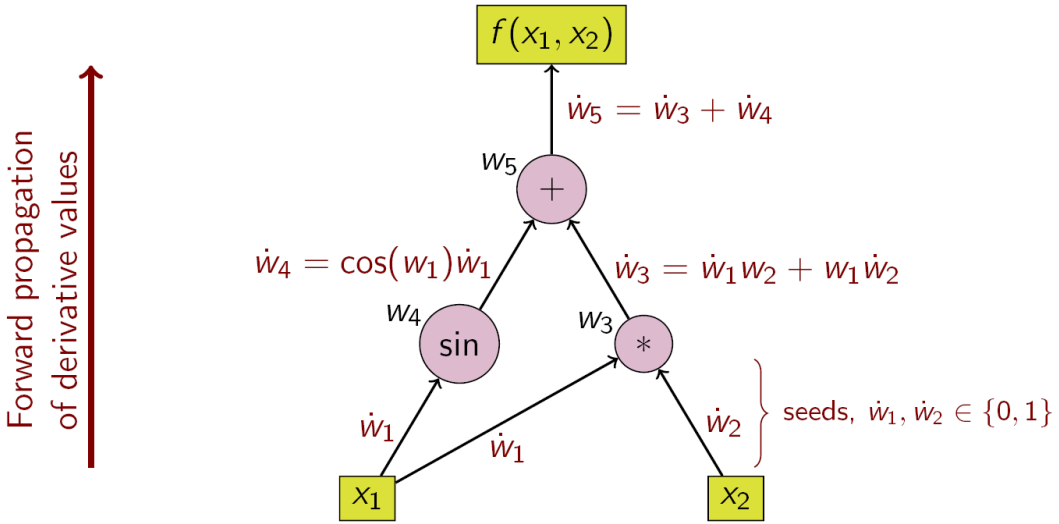


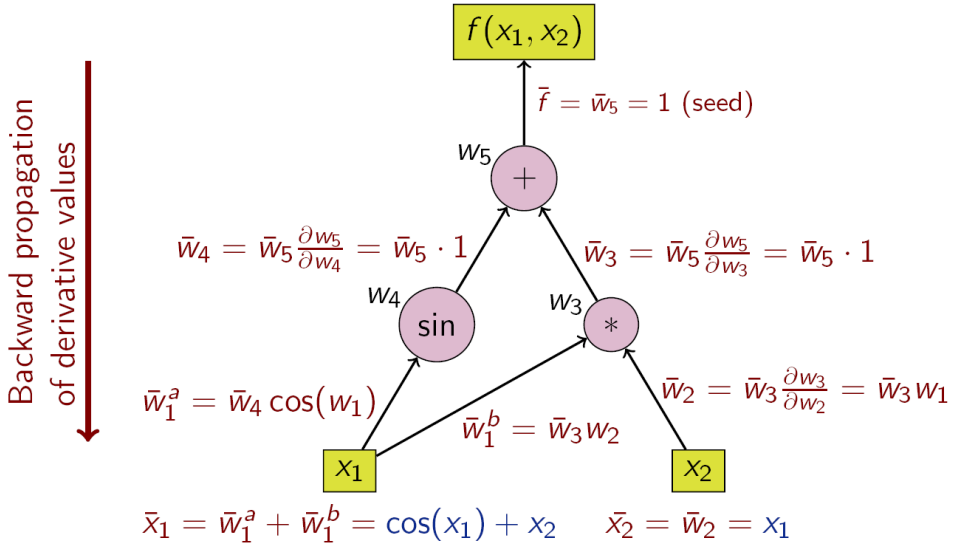
$$\frac{\partial E}{\partial a_j^l}\frac{\partial a_j^l}{\partial z_j^l} = \delta_j^l \Rightarrow \frac{\partial E}{\partial z_k^2} = \delta_k^2$$

$$\frac{\partial E}{\partial w_{11}^1} = input_1 a_1^1(1-a_1^1)\sum_{k=1}^{3}\delta_k^2 w_{1k}^2$$

$$\frac{\partial E}{\partial w_{ij}^{(l)}} = a_i^{(l-1)} a_j^{(l)}\left(1-a_j^{(l)}\right)\sum_{k=1}^{n}\delta_k^{(l+1)} w_{jk}^{(l+1)}$$

Автоматический градиент





$$f(x_1, x_2) = x_1 x_2 + \sin(x_1)$$

Colab

https://colab.research.google.com/notebooks/intro.ipynb