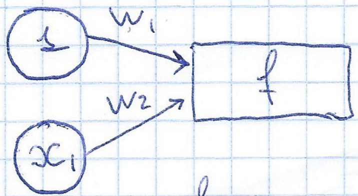
Упражнение



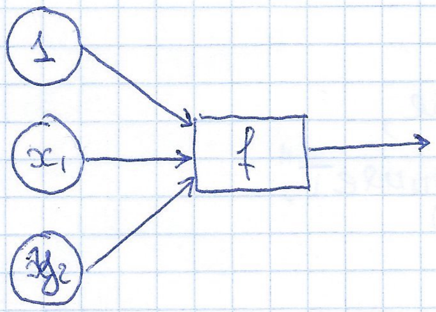
$0 \longrightarrow 1$

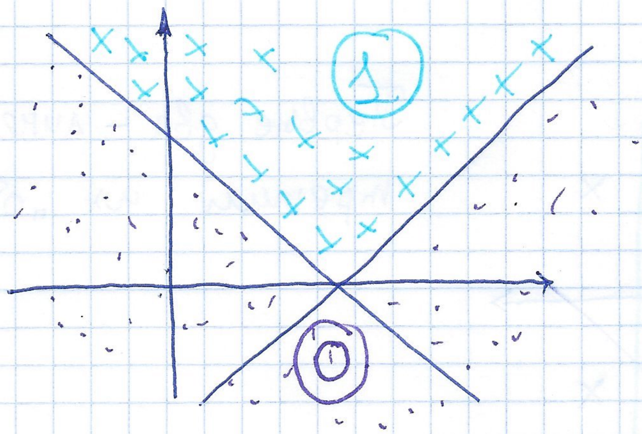
$1 \longrightarrow 0$

$f = \max(0, t)$

$$\max(w_0 + w_1 \cdot 0, 0) = 1 \quad \Rightarrow \quad w_0 = 1$$
$$\max(w_0 + w_1 \cdot 1, 0) = 1 \qquad\qquad w_1 = -2$$

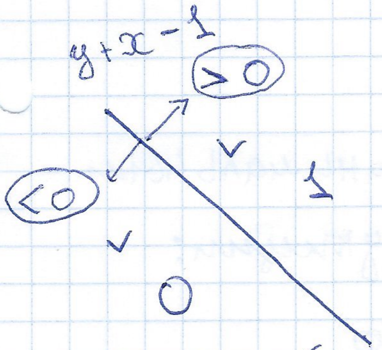
Упражнение



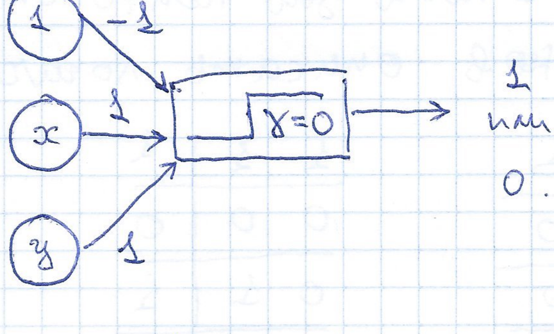


$x - y = 1$

$x + y = 1$

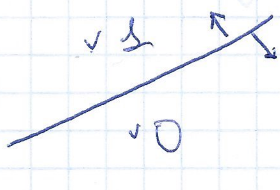


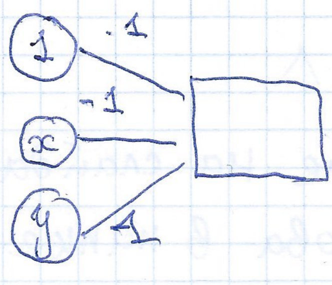
линейный классификатор

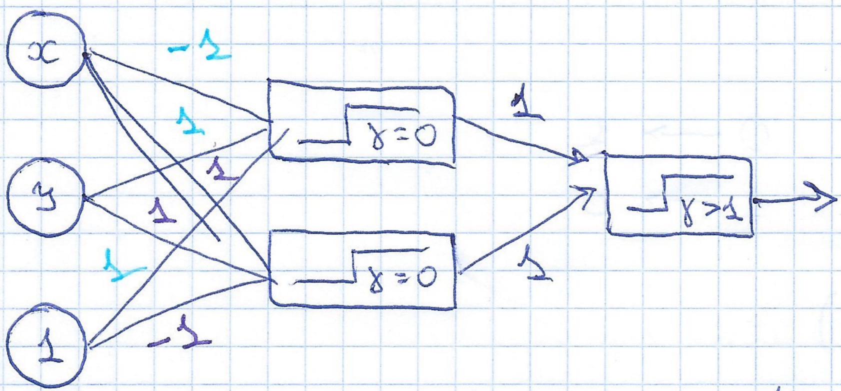


$y - x + 1$

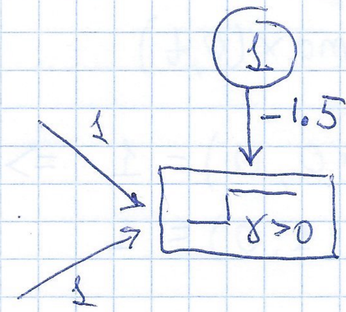
$-y + x - 1$



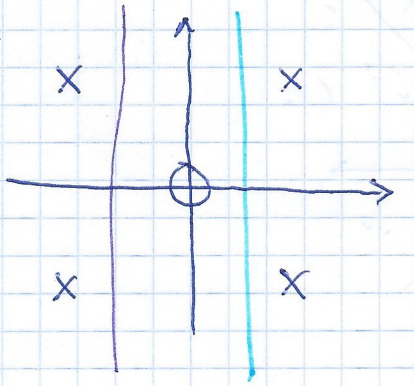




1	1	1
0	1	0
1	0	0
0	0	0



Упражнение



Первые два – линейные, третий их „объединяет".

Упражнение

Попробуйте с помощью нейросетки с минимальным числом нейронов описать логические функции:

1	1	1
0	0	0
0	1	0
1	0	0

$\wedge$

1	1	1
0	0	0
0	1	1
1	0	1

$\vee$

1	1	0
0	0	0
0	1	1
1	0	1

XOR (исключающее или)

[Решение на слайдах Воронцова в папке

Упражнение

Маша

x	1	2
y	2	3

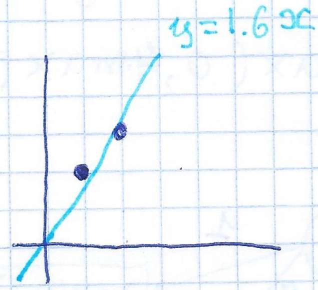
а) Теоретическая оценка

б) GD

в) SGD $1 \to 2$.

$\eta = 0.1$

а) $y = \beta x$ $\qquad x \xrightarrow{\beta} \boxed{f(t)=t} \longrightarrow y$



$$L = \sum_i (y_i - \hat{y}_i)^2 = \sum_i (y_i - x_i\beta)^2 \longrightarrow \min_\beta$$

$$L'_\beta = -2\sum_i (y_i - x_i\hat{\beta})x_i = 0$$

$$\sum y_i x_i - \hat{\beta}\sum x_i^2 = 0$$

$$\Rightarrow \hat{\beta} = \frac{\sum_i y_i x_i}{\sum_i x_i^2} = \frac{1\cdot 2 + 2\cdot 3}{1^2 + 2^2} = \frac{8}{5} = 1.6$$

б) $\nabla L = L'_\beta = -2(y_1 - x_1\beta)x_1 - 2(y_2 - x_2\beta)x_2$.

① Инициализация $\beta_0 = 0$

② $\beta_1 = \beta_0 - 0.1 \cdot \underbrace{\nabla L(\beta_0)}$

$-2(2-0)\cdot 1 - 2(3 - 2\cdot 0)\cdot 2 =$

$= -4 - 4 = -8$

$\beta_1 = 0 + 0.8 = 0.8$

③ $\beta_2 = \beta_1 - 0.1\,\nabla L(\beta_1)$

и т.д.

в) ① $\beta_0 = 0$

② $\beta_1 = 0 - 0.1 \cdot \left(-2(2-0)\cdot 1\right) = 0.4$

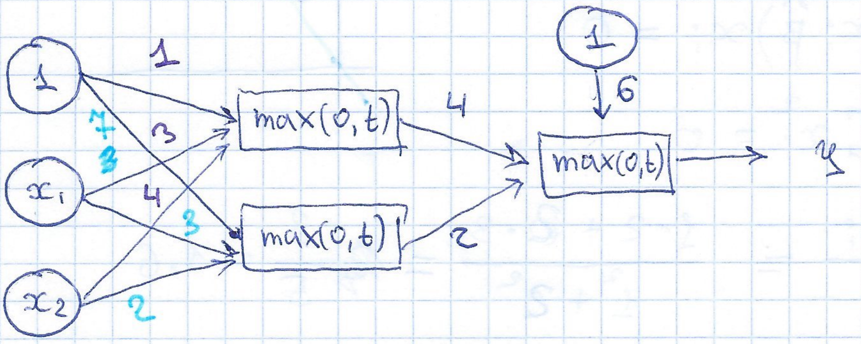
③ $\beta_2 = 0.4 - 0.1 \cdot \left(-2(3 - 0.4\cdot 2)\cdot 2\right) = 0.4 + 0.88 = 1.28$

и т.д.

Упражнение

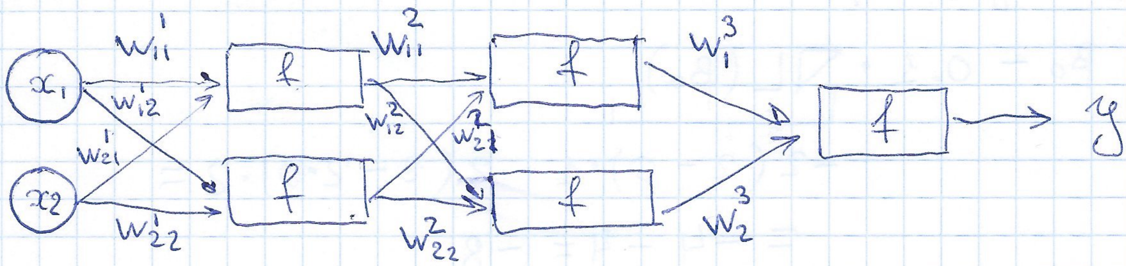
Записать функцию в виде нейросетки.

$$\max(0, 4\cdot\max(0, 3x_1 + 4x_2 + 1) + 2\cdot\max(0, 3x_1 + 2x_2 + 7) + 6)$$



Упражнение

а) Записать нейросеть как сложную функцию.



$y =$

$$= f\Big(w^3_1 \cdot f\big(w^2_{11}\cdot f(w^1_{11}x_1 + w^1_{21}x_2) + w^2_{21}\cdot f(w^1_{12}x_1 + w^1_{22}x_2)\big) +$$

$$+ w^3_2 \cdot f\big(w^2_{12}\cdot f(w^1_{11}x_1 + w^1_{21}x_2) + w^2_{22}\cdot f(w^1_{12}x_1 + w^1_{22}x_2)\big)\Big)$$

Сложная функция !!

б) Запишем её же в матричном виде!

$$X = \begin{pmatrix} x_{11} & x_{21} \\ \vdots & \vdots \\ x_{1n} & x_{2n} \end{pmatrix}_{n\times 2} \qquad h_1 = XW_1 = \begin{pmatrix} h_{11} & h_{21} \\ \vdots & \vdots \\ h_{1n} & h_{2n} \end{pmatrix}_{n\times 2}$$

$$W_1 = \begin{pmatrix} w_{11}^1 & w_{12}^1 \\ w_{21}^1 & w_{22}^1 \end{pmatrix}_{2\times 2}$$

$$(x_1\ x_2) \cdot \begin{pmatrix} w_{11}^1 & w_{12}^1 \\ w_{21}^1 & w_{22}^1 \end{pmatrix} = \Big(\underbrace{w_{11}^1 x_1 + w_{21}^1 x_2}_{h_1};\ \underbrace{w_{12}^1 x_1 + w_{22}^1 x_2}_{h_2}\Big)$$

отдельное наблюдение

Первый слой выплёвывает: $f(XW_1)$.

По аналогии получаем, что:

$$y = f(f(f(XW_1)W_2)W_3)$$

Сетка – сложная функция!

в) $\hat{y} = f(f(f(XW_1)W_2)W_3)$

$$L(W_1, W_2, W_3) = \frac{1}{2} \cdot (y - f(f(f(XW_1)W_2)W_3))^2$$

функция потерь

Секрет успеха в умении брать производную.

$$f(g(x))' = f'(g(x)) \cdot g'(x)$$

$$\frac{\partial L}{\partial w_3} = -(y - f(f(f(xw_1)w_2)w_3)) \cdot f'_{w_3}(\ldots) \cdot f(f(xw_1)w_2)$$

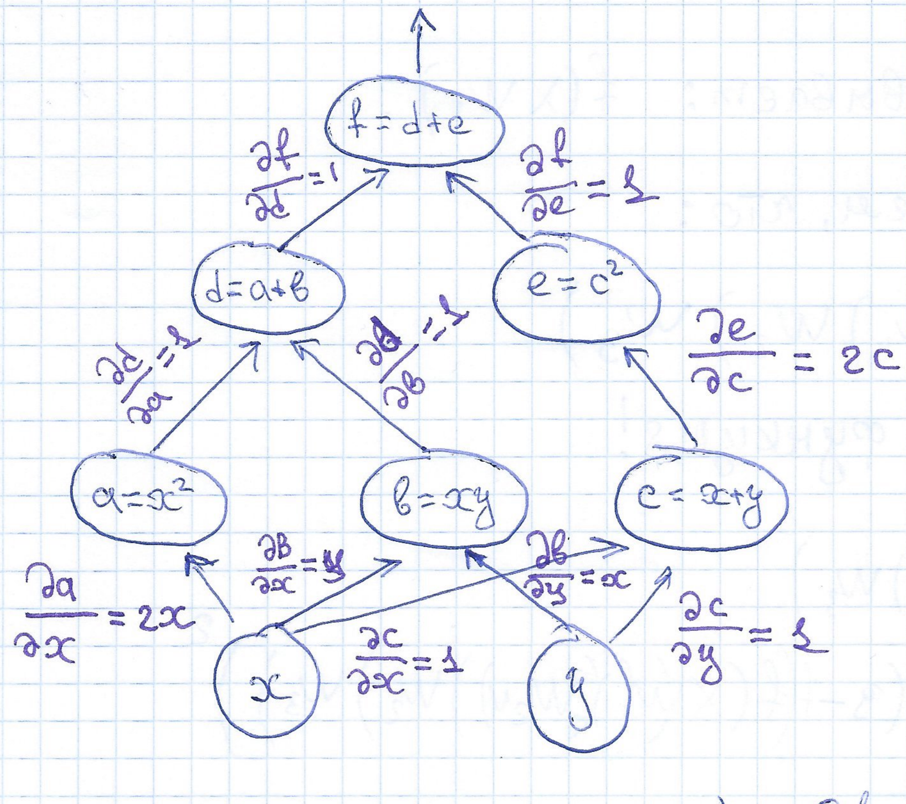
$$\frac{\partial L}{\partial w_2} = -(y - f(f(f(xw_1)w_2)w_3)) \cdot f'_{w_2}(\ldots) \cdot f'_{w_2}(\ldots) \cdot w_3 \cdot f(xw_1)$$

$$\frac{\partial L}{\partial w_1} = -(y - f(f(f(xw_1)w_2)w_3)) \cdot f'_{w_1}(\ldots) \cdot f'_{w_1}(\ldots) \cdot w_3 \cdot f'_{w_1}(xw_1) \cdot X$$

Упражнение

Изобразить граф вычислений и найти производные

$$f(x, y) = x^2 + xy + (x+y)^2$$



$$\frac{\partial f}{\partial x} = \underset{1}{\frac{\partial f}{\partial d}} \cdot \left(\underset{1}{\frac{\partial d}{\partial a}} \cdot \underset{2x}{\frac{\partial a}{\partial x}} + \underset{1}{\frac{\partial d}{\partial b}} \cdot \underset{y}{\frac{\partial b}{\partial x}} \right) + \underset{1}{\frac{\partial f}{\partial e}} \cdot \underset{\underbrace{2c}_{2(x+y)}}{\frac{\partial e}{\partial c}} \cdot \underset{1}{\frac{\partial c}{\partial x}}$$

$$\frac{\partial f}{\partial x} = 2x + y + 2(x+y)$$

б) Запишем её же в матричном виде!

$$X = \begin{pmatrix} x_{11} & x_{21} \\ \vdots & \vdots \\ x_{1n} & x_{2n} \end{pmatrix}_{n\times 2} \qquad h_1 = XW_1 = \begin{pmatrix} h_{11} & h_{21} \\ \vdots & \vdots \\ h_{1n} & h_{2n} \end{pmatrix}_{n\times 2}$$

$$W_1 = \begin{pmatrix} w^1_{11} & w^1_{12} \\ w^1_{21} & w^1_{22} \end{pmatrix}_{2\times 2}$$

$$\underset{\text{отдельное наблюдение}}{(x_1 \ x_2)} \cdot \begin{pmatrix} w^1_{11} & w^1_{12} \\ w^1_{21} & w^1_{22} \end{pmatrix} = (\underbrace{w^1_{11}x_1 + w^1_{21}x_2}_{h_1} ; \underbrace{w^1_{12}x_1 + w^1_{22}x_2}_{h_2})$$

Первый слой выплёвывает: $f(XW_1)$.

По аналогии получаем, что:

$$y = f(f(f(XW_1)W_2)W_3)$$

Сетка - сложная функция!

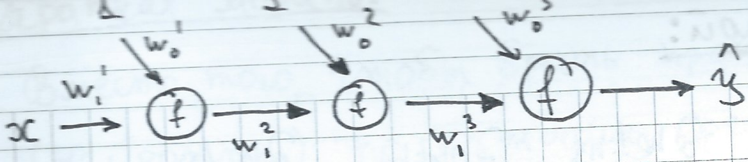
в) $\hat{y} = f(f(f(XW_1)W_2)W_3)$

$$L(W_1, W_2, W_3) = \frac{1}{2}\cdot(y - f(f(f(XW_1)W_2)W_3))^2$$

функция потерь

Секрет успеха в умении брать производную.

$$f(g(x))' = f'(g(x))\cdot g'(x)$$



$$f(t) = \frac{e^t}{1+e^t} \qquad f'(t) = f(t)(1-f(t))$$

$$\mathcal{E}\left(\hat{y}_i - f\left(w_1^3 \cdot f\left(w_1^2 \cdot f\left(w_1^1 x + w_0^1\right) + w_0^2\right) + w_0^3\right)\right)^2$$

Надо минимизировать такую штуку. Влоб сложно.

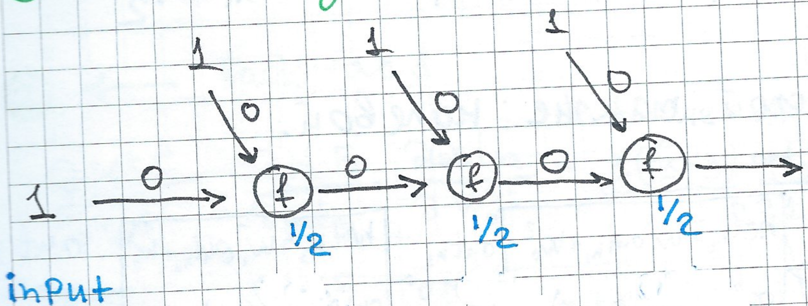
$$w_0 = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}$$

x	y
5	0
1	1

Сделать 2 шага SGD:
- сначала 2 набл. потом 1-ое
- $\gamma = 1$

I шаг

① → проход (forward)



$\hat{y} = \frac{1}{2}$

$y = 1$

$MSE = (1 - 1/2)^2 = 1/4$

$net_{h_1} = w_1^1 \cdot input + w_0^1 \cdot 1 = 0$

$net_{h_2} = w_1^2 \cdot out_{h_1} + w_0^2 \cdot 1 = 0$

$net_{h_3} = w_1^3 \cdot out_{h_2} + w_0^3 \cdot 1 = 0$

$out_{h_1} = f(net_{h_1}) = 1/2$

$out_{h_2} = f(net_{h_2}) = 1/2$

$out_{h_3} = f(net_{h_3}) = \hat{y} = 1/2$

② ← проход (backward)

$$\frac{\partial MSE}{\partial w_0^3} = \underbrace{\frac{\partial MSE}{\partial out_{h_3}}}_{2(y - out_{h_3})} \cdot \underbrace{\frac{\partial out_{h_3}}{\partial net_{h_3}}}_{out_{h_3}(1-out_{h_3})} \cdot \underbrace{\frac{\partial net_{h_3}}{\partial w_0^3}}_{1} = 2 \cdot (1 - 1/2) \cdot \frac{1}{2} \cdot \frac{1}{2} \cdot 1 = \frac{1}{4}$$

$$\frac{\partial MSE}{\partial w_1^3} = \frac{\partial MSE}{\partial out_{h_3}} \cdot \frac{\partial out_{h_3}}{\partial net_{h_3}} \cdot \frac{\partial net_{h_3}}{\partial w_1^3} = 2(1 - 1/2) \cdot \frac{1}{2} \cdot \frac{1}{2} \cdot \overbrace{out_{h_2}}^{1/2} = \frac{1}{8}$$

Переходим в скрытый слой:

$$\frac{\partial MSE}{\partial w_0^2} = \underbrace{\frac{\partial MSE}{\partial out_{h_3}} \cdot \frac{\partial out_{h_3}}{\partial net_{h_3}}}_{\substack{\text{уже нашли} \\ 1/4}} \cdot \underbrace{\frac{\partial net_{h_3}}{\partial out_{h_2}}}_{w_1^3 = 0} \cdot \underbrace{\frac{\partial out_{h_2}}{\partial net_{h_2}}}_{\substack{1/2 \cdot 1/2 \\ out_{h_2}(1-out_{h_2})}} \cdot \underbrace{\frac{\partial net_{h_2}}{\partial w_0^2}}_{1} = 0$$

$$\frac{\partial MSE}{\partial w_1^2} = \underbrace{\frac{\partial MSE}{\partial out_{h_3}} \cdot \frac{\partial out_{h_3}}{\partial net_{h_3}} \cdot \frac{\partial net_{h_3}}{\partial out_{h_2}}}_{\frac{\partial MSE}{\partial out_{h_2}}} \cdot \frac{\partial out_{h_2}}{\partial net_{h_2}} \cdot \frac{\partial net_{h_2}}{\partial w_1^2} = 0$$

$$\underbrace{\frac{\partial MSE}{\partial out_{h_2}}}_{1/4} \quad \underbrace{\quad}_{w_1^3 = 0} \quad \underbrace{\quad}_{1/4} \quad \underbrace{\quad}_{out_{h_1} = 1/2}$$

по аналогии следущий слой также нулевой.

$$\text{input} \to \boxed{\begin{array}{l} net_{h_1} = w_1^1 \, input + w_0^1 \\ out_{h_1} = f(net_{h_1}) \end{array}} \xrightarrow{out_{h_1}} \boxed{\begin{array}{l} net_{h_2} = w_1^2 out_{h_1} + w_0^2 \\ out_{h_2} = f(net_{h_2}) \end{array}} \xrightarrow{out_{h_2}} \boxed{\begin{array}{l} net_{h_3} = w_1^3 out_{h_3} + w_0^3 \\ out_{h_3} = f(net_{h_3}) \end{array}} \to out_{h_3}$$

$$\leftarrow \boxed{\underset{1}{\frac{\partial out_{h_1}}{\partial net_{h_1}}} \cdot \underset{2}{\frac{\partial net_{h_1}}{\partial input}}} \leftarrow \boxed{\underset{1}{\frac{\partial out_{h_2}}{\partial net_{h_2}}} \cdot \underset{2}{\frac{\partial net_{h_2}}{\partial out_{h_1}}}} \leftarrow \boxed{\underset{1}{\frac{\partial MSE}{\partial net_{h_3}}} \cdot \underset{2}{\frac{\partial net_{h_3}}{\partial out_{h_2}}}} \leftarrow$$

1 $out_{h_1}(1-out_{h_1})$ 1 $out_{h_2}(1-out_{h_2})$ 1 $2(\hat{y}-y) \cdot out_{h_3}(1-out_{h_3})$

2 w_1^1 2 w_1^2 2 w_1^3

Делаем первый шаг:

$$\begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} - 1 \cdot \begin{pmatrix} 0 & 0 & 1/4 \\ 0 & 0 & 1/8 \end{pmatrix} = \begin{pmatrix} 0 & 0 & -1/4 \\ 0 & 0 & -1/8 \end{pmatrix}$$

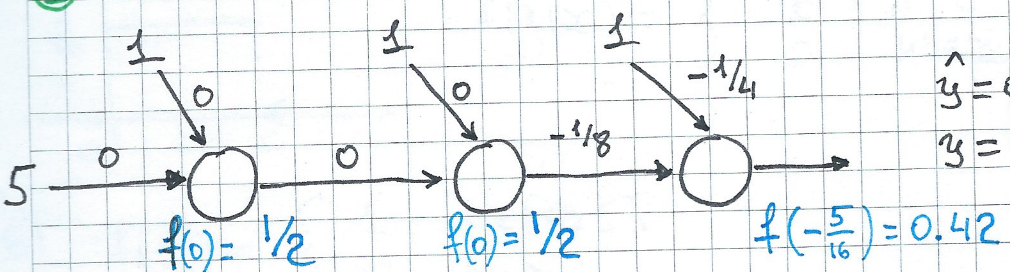
$$w = w - 1 \cdot \frac{\partial MSE}{\partial w}$$

Краткая мораль:

Вместо того, чтобы брать производную большой функции (нейросети) по каждому аргументу, мы заметили, что внутренние части этой производной можно выражать друг через друга и последовательно искать, начиная с конца сетки.

II шаг

① → forward



$\hat{y} = 0.42$

$y = 0$

② ← backward

$$\frac{\partial MSE}{\partial w_0^3} = \frac{\partial MSE}{\partial neth_3} \cdot \frac{\partial neth_3}{\partial w_0^3} = \underbrace{\overset{2 \cdot 0.42 \cdot 0.42 \quad 0.58}{2(\hat{y}-y)\, out_3 (1-out_3)}}_{0.21} \cdot \underbrace{1} = 0.21$$

$$\frac{\partial MSE}{\partial w_1^3} = \frac{\partial MSE}{\partial neth_3} \cdot \frac{\partial neth_3}{\partial w_1^3} = 0.21 \cdot \underbrace{\frac{1}{2}}_{outh_2} = 0.10$$

$$\frac{\partial MSE}{\partial w_0^2} = \underbrace{\frac{\partial MSE}{\partial outh_2}} \cdot \frac{\partial outh_2}{\partial neth_2} \cdot \frac{\partial neth_2}{\partial w_0^2} = 0.21 \cdot w_1^3 \cdot \frac{1}{2} \cdot \frac{1}{2} \cdot 1 = -0.013$$

$$\frac{\partial MSE}{\partial w_1^2} = \underbrace{\frac{\partial MSE}{\partial outh_2}} \cdot \frac{\partial outh_2}{\partial neth_2} \cdot \frac{\partial neth_2}{\partial w_1^2} = 0.21 \cdot w_1^3 \cdot \frac{1}{2} \cdot \frac{1}{2} \cdot \underbrace{\frac{1}{2}}_{outh_1} = -0.007$$

$$\frac{\partial MSE}{\partial w_0^1} = \underbrace{\frac{\partial MSE}{\partial outh_2}} \cdot \underbrace{\frac{\partial outh_2}{\partial outh_1}} \cdot \frac{\partial outh_1}{\partial neth_1} \cdot \frac{\partial neth_1}{\partial w_0^1} = 0.21 \cdot \cancel{w_1^3} \cdot 1/4 \cdot \overset{0}{w_1^2} \cdot \frac{1}{2} \cdot \frac{1}{2} \cdot 1 = 0$$

$$\frac{\partial MSE}{\partial w_1^1} = \underbrace{\frac{\partial MSE}{\partial outh_2}} \cdot \underbrace{\frac{\partial outh_2}{\partial outh_1}} \cdot \frac{\partial outh_1}{\partial neth_1} \cdot \frac{\partial neth_1}{\partial w_1^1} = 0.21 \cdot w_1^3 \cdot 1/4 \cdot \underset{0}{w_1^2} \cdot \frac{1}{2} \cdot \frac{1}{2} \cdot \frac{5}{\text{input}} = 0$$

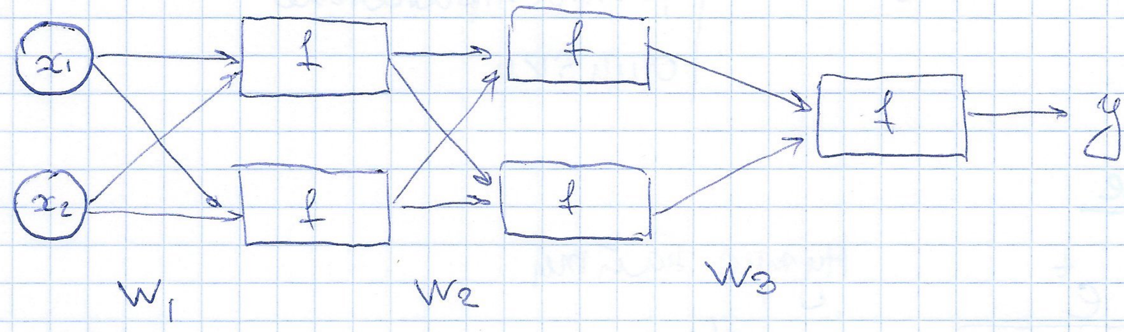
НА каждом шаге копится часть с предыдущего, котору не надо считать заново. Нужно найти только недостающий хвостик и перейти к следующему слою, дозаписав новую накопленную частичку!

Делаем второй шаг:

$$\begin{pmatrix} 0 & 0 & -1/4 \\ 0 & 0 & -1/8 \end{pmatrix} - \begin{pmatrix} 0 & -0.013 & 0.21 \\ 0 & -0.007 & 0.10 \end{pmatrix} = \begin{pmatrix} 0 & 0.013 & -0.46 \\ 0 & 0.007 & -0.225 \end{pmatrix}$$

И так далее !!

Упражнение



$$X \xrightarrow{W_1} net_{h_1} \xrightarrow{f} out_{h_1} \xrightarrow{W_2} net_{h_2} \xrightarrow{f} out_{h_2} \xrightarrow{W_3} net_{h_3} \xrightarrow{f} out_{h_3} = y$$

$$\frac{\partial MSE}{\partial W_3} = \frac{\partial MSE}{\partial out_{h_3}} \cdot \frac{\partial out_{h_3}}{\partial net_{h_3}} \cdot \frac{\partial net_{h_3}}{\partial W_3}$$

$$\frac{\partial MSE}{\partial W_2} = \frac{\partial MSE}{\partial out_{h_3}} \cdot \frac{\partial out_{h_3}}{\partial net_{h_3}} \cdot \frac{\partial net_{h_3}}{\partial out_{h_2}} \cdot \frac{\partial out_{h_2}}{\partial net_{h_2}} \cdot \frac{\partial net_{h_2}}{\partial W_2}$$

$$\frac{\partial MSE}{\partial W_1} = \frac{\partial MSE}{\partial out_{h_3}} \cdot \frac{\partial out_{h_3}}{\partial net_{h_3}} \cdot \frac{\partial net_{h_3}}{\partial out_{h_2}} \cdot \frac{\partial out_{h_2}}{\partial net_{h_2}} \cdot \frac{\partial net_{h_2}}{\partial out_{h_1}} \cdot \frac{\partial out_{h_1}}{\partial net_{h_1}} \cdot \frac{\partial net_{h_1}}{\partial W_1}$$

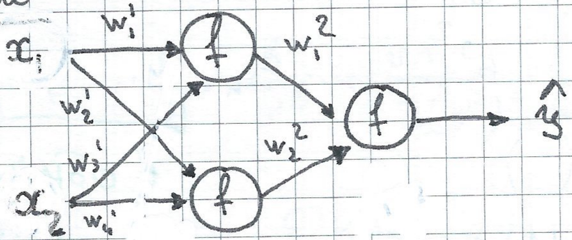
$$\frac{\partial net_{h_2}}{\partial out_{h_1}} = (out_{h_1} \cdot W_2)'_{out_{h_1}} = W_2.$$

$$\frac{\partial out_{h_1}}{\partial net_{h_1}} = \left[f(net_{h_1})\right]'_{net_{h_1}} = f'_{net_{h_1}}$$

$$\frac{\partial net_{h_2}}{\partial W_2} = (out_{h_2} W_2)'_{W_2} = out_{h_2} = f(f(XW_1)W_2)$$

Backpropagation своими руками

Упражнение

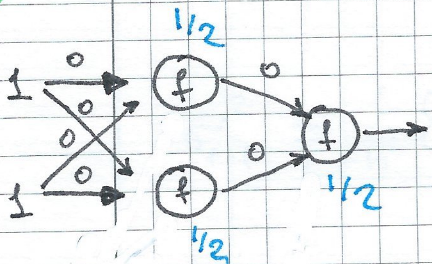


x_1	x_2	y
1	1	1
2	2	0

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

I шаг

① → forward



$\hat{y} = \frac{1}{2}$

$y = 1$

MSE = 1/4

$$neth_{11} = w_1^1 x_1 + w_3^1 x_2 \qquad out_{11} = f(neth_{11})$$

$$neth_{12} = w_2^1 x_1 + w_4^1 x_2 \qquad out_{12} = f(neth_{12})$$

$$neth_2 = w_1^2 out_{11} + w_2^2 out_{12} \qquad out_2 = f(neth_2) = \hat{y}$$

② ← backward

$$\frac{\partial MSE}{\partial w_1^2} = \underbrace{\frac{\partial MSE}{\partial out_2} \cdot \frac{\partial out_2}{\partial neth_2}} \cdot \frac{\partial neth_2}{\partial w_1^2} = 2(y - out_2)\, out_2(1 - out_2) \cdot out_{11} =$$

$$\frac{\partial MSE}{\partial w_2^2} = \underbrace{\frac{\partial MSE}{\partial out_2} \cdot \frac{\partial out_2}{\partial neth_2}} \cdot \frac{\partial neth_2}{\partial w_2^2} = 2(y - out_2)\, out_2(1 - out_2)\, out_{12} =$$

$$\frac{\partial MSE}{\partial w_1^1} = \underbrace{\frac{\partial MSE}{\partial out_2} \cdot \frac{\partial out_2}{\partial neth_2}}_{\text{первый накопленный кусочек}} \cdot \left(\frac{\partial neth_2}{\partial out_{11}} \cdot \frac{\partial out_{11}}{\partial neth_{11}} \cdot \frac{\partial neth_{11}}{\partial w_1^1} + \underbrace{\frac{\partial neth_2}{\partial out_{12}} \cdot \frac{\partial out_{12}}{\partial neth_{12}} \frac{\partial neth_{12}}{\partial w_1^1}}_{} \right)$$

О но если бы был бы ещё слой, было бы не ноль !!

$\boxed{\text{и т.д.}}$