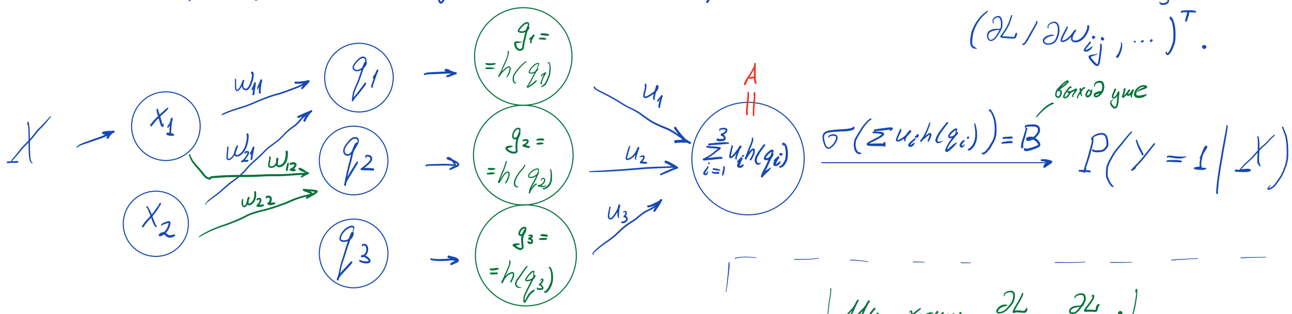
В этом примере $\frac{\partial L}{\partial x} = \begin{pmatrix} \partial L/\partial x_1 \\ \partial L/\partial x_2 \end{pmatrix}$ нам не надо считать, т.к. 1 слой + лог.рег. на 2 класса, не требующий ничего сложного.

$X = (x_1, x_2)$, классификация на 2 класса; Мы хотим: $\left(\frac{\partial L}{\partial u_1}, \frac{\partial L}{\partial u_2}, \frac{\partial L}{\partial u_3}\right)^T$;
$(\partial L / \partial w_{ij}, \dots)^T$.



$$(q_1, q_2, q_3) = \underset{1\times 2}{(x_1, x_2)} \underset{2\times 3}{\begin{pmatrix} w_{11} & w_{12} & w_{13} \\ w_{21} & w_{22} & w_{23} \end{pmatrix}}$$

$q_1 = x_1 \cdot w_{11} + x_2 \cdot w_{21}$, $q_2 = x_1 \cdot w_{12} + x_2 \cdot w_{22}$, $q_3 = \dots$

|Мы хотим: $\frac{\partial L}{\partial u_1}, \frac{\partial L}{\partial w_{11}}$:|

$\frac{\partial L}{\partial B}$ – мы можем посчитать;

$\frac{\partial B}{\partial A}$ – тоже можем посчитать; $\left(\frac{\partial L}{\partial A} = \frac{\partial L}{\partial B} \cdot \frac{\partial B}{\partial A}\right.$ – тоже смогли$\left.\right)$

$\frac{\partial A}{\partial u_1}$ – тоже можем посчитать;

L от u_1 явно не зависит, но! A – зависит, поэтому:

$$\boxed{\frac{\partial L}{\partial u_1} = \underbrace{\frac{\partial L}{\partial B}}_{\text{знаем}} \cdot \underbrace{\frac{\partial B}{\partial A}}_{\text{знаем}} \cdot \underbrace{\frac{\partial A}{\partial u_1}}_{\text{знаем}}}$$

Разберёмся с $\frac{\partial L}{\partial w_{11}}$:

(Уже видна структура графа вычислений – считаем производную соседа справа по соседу слева, потом перемножаем)

A зависит от g_1 => $\frac{\partial A}{\partial g_1}$ – можем посчитать

g_1 зависит от q_1 => $\frac{\partial g_1}{\partial q_1}$ – можем посчитать

q_1 зависит от w_{11} и w_{21} => $\frac{\partial q_1}{\partial w_{11}}, \frac{\partial q_1}{\partial w_{21}}$ можем посчитать

$$\Rightarrow \frac{\partial L}{\partial w_{11}} = \frac{\partial L}{\partial A} \cdot \frac{\partial A}{\partial g_1} \cdot \frac{\partial g_1}{\partial q_1} \cdot \frac{\partial q_1}{\partial w_{11}} = \frac{\partial L}{\partial B} \cdot \frac{\partial B}{\partial A} \cdot \frac{\partial A}{\partial g_1} \cdot \frac{\partial g_1}{\partial q_1} \cdot \frac{\partial q_1}{\partial w_{11}}.$$

No brain страта:

То что справа дифф. по тому что слева – получаем граф вычислений, все промежуточные произв. мы можем посчитать, а все остальные получаем проходом по графу.

последний слой:
(A – это $\sigma(\sum u_i g_i)$, g_i нам уже вычислили при прямом распространении)

$\frac{\partial L}{\partial B}$ | $\frac{\partial L}{\partial w_{ij}} = \frac{\partial L}{\partial B} \cdot \Theta(B)$ зависит от структуры L

$\leftarrow \frac{\partial B}{\partial A} \leftarrow$

$\frac{\partial A}{\partial u_1}$, $\frac{\partial A}{\partial u_2}$, $\frac{\partial A}{\partial u_3}$

первый слой:

$\frac{\partial g_1}{\partial q_1} = \frac{\partial h}{\partial q_1}\Big|_{q_1}$

$\frac{\partial g_2}{\partial q_2} = \frac{\partial h}{\partial q_2}\Big|_{q_2}$

$\frac{\partial g_3}{\partial q_3} = \frac{\partial h}{\partial q_3}\Big|_{q_3}$

$\leftarrow$

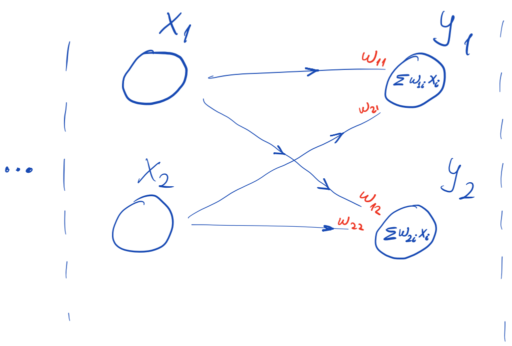
$\partial q_1 / \partial w_{i1}$
$\vdots$
$\partial q_2 / \partial w_{i2}$
$\vdots$
$\partial q_3 / \partial w_{i3}$

$\leftarrow$ Все град. теперь считаются:
$\vec{u} = \vec{u} - \alpha \nabla \vec{u}$; $W = W - \alpha \nabla W$

Вся суть: 1) Прямое распространение – считаем все промежуточные числа на текущих весах

2) Обратное распр.: считаем по-очереди произв, т.к. все числа для их вычисл. уже есть; обновляем веса.

Рассм. линейный слой: $\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} w_{11} & w_{12} \\ w_{21} & w_{22} \end{pmatrix} \begin{pmatrix} y_1 \\ y_2 \end{pmatrix} \Leftrightarrow x = Wy$



... $\Rightarrow$ L считаем

$\frac{dL}{dx_{step-1}}$ надо, чтобы назад распр.

$\frac{dL}{dW_{step-1}}$ — чтобы веса обновлять

Дошли до сюда back prop-ом и $\frac{dL}{dy} = \begin{pmatrix} \partial L / \partial y_1 \\ \partial L / \partial y_2 \end{pmatrix}$ — уже знаем

Хотим: $\frac{dL}{dW} = \begin{pmatrix} \frac{\partial L}{\partial w_{11}} & \frac{\partial L}{\partial w_{21}} \\ \frac{\partial L}{\partial w_{21}} & \frac{\partial L}{\partial w_{22}} \end{pmatrix}$ — ? $+$ $\frac{dL}{dx}$ — ?, чтобы back prop дальше вести;

Очевидно: $\frac{dL}{dx}$ гораздо сложнее, чем $\frac{dL}{dW}$ найти, т.к. и x_1, и x_2 входят и в y_1, и в y_2

$\frac{\partial y_1}{\partial x_1} = w_{11}$; $\quad \frac{\partial y_2}{\partial x_1} = w_{21}$;

$\frac{\partial y_1}{\partial x_2} = w_{12}$; $\quad \frac{\partial y_2}{\partial x_2} = w_{22}$;

$\Rightarrow \frac{dy}{dx} = \begin{pmatrix} w_{11} & w_{21} \\ w_{12} & w_{22} \end{pmatrix}$

$L = L(y_1, y_2)$

$y_1 = y_1(x_1, x_2)$
$y_2 = y_2(x_1, x_2)$

Так вся цепочка и строится

Но!: $y_1 = y_1(w_{11}, w_{12})$
$y_2 = y_2(w_{21}, w_{22})$,

т.е. ничего лишнего в $\frac{\partial y_i}{\partial w_{ij}}$ не будет

поэтому в $\frac{dL}{dW}$ лишь одно слаг., а в $\frac{dL}{dx}$ два!

Запишем $\frac{dL}{dx}$ тот самый страшный:

$$\frac{\partial L}{\partial x_1} = \frac{\partial L}{\partial y_1} \cdot \frac{\partial y_1}{\partial x_1} + \frac{\partial L}{\partial y_2} \cdot \frac{\partial y_2}{\partial x_1} = \frac{\partial L}{\partial y_1} \cdot w_{11} + \frac{\partial L}{\partial y_2} \cdot w_{21},$$

$$\frac{\partial L}{\partial x_2} = \frac{\partial L}{\partial y_1} \cdot \frac{\partial y_1}{\partial x_2} + \frac{\partial L}{\partial y_2} \cdot \frac{\partial y_2}{\partial x_2} = \frac{\partial L}{\partial y_1} \cdot w_{12} + \frac{\partial L}{\partial y_2} \cdot w_{22}$$

Т.е. видно: $\frac{dL}{dx} = W^T \cdot \frac{dL}{dy}$, можем дальше back prop-ить назад.

$\frac{dL}{dW}$ нам нужна, чтобы обновлять веса, а $\frac{dL}{dx}$, чтобы back prop дальше заводить, но dL/dx нам, очевидно, для $\frac{dL}{dW}$ не нужно.

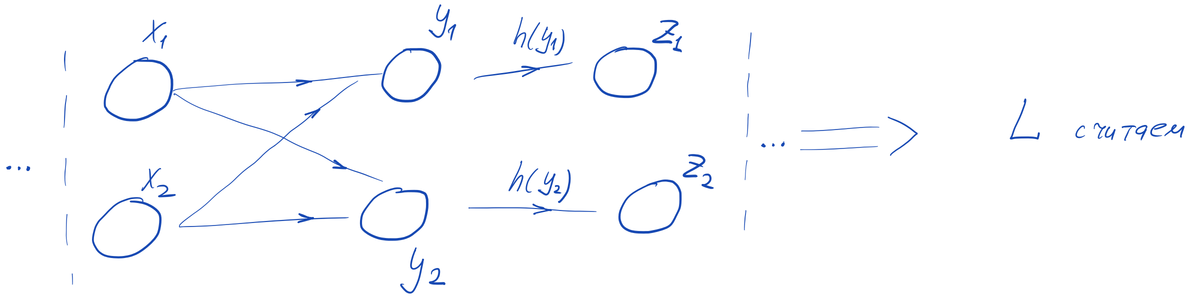
$$\frac{\partial L}{\partial w_{11}} = \frac{\partial L}{\partial y_1}\cdot\frac{\partial y_1}{\partial w_{11}} + \overset{0,\ \text{т.к. в } y_2 \text{ нет } w_{1j}}{\cancel{\frac{\partial L}{\partial y_2}\cdot\frac{\partial y_2}{\partial w_{11}}}} = \frac{\partial L}{\partial y_1}\cdot x_1 \Rightarrow \quad \frac{\partial L}{\partial w_{ij}} = \frac{\partial L}{\partial y_i}\cdot x_j \Rightarrow$$

$$\Rightarrow \frac{dL}{dW} = \begin{pmatrix} \partial L/\partial y_1 \\ \partial L/\partial y_2 \end{pmatrix}\cdot\begin{pmatrix} x_1 & x_2 \end{pmatrix} = \frac{dL}{dy}\cdot\begin{pmatrix} x_1 & x_2 \end{pmatrix}$$

$\Rightarrow$ умеем обновлять веса и проходить полностью back prop.

$\frac{dL}{dw}$ $\frac{dL}{dx}$

Рассмотрю back prop для ф-ии активации:



Дошли до сюда и $\frac{\partial L}{\partial z_1}$, $\frac{\partial L}{\partial z_2}$ можем посчитать

$\frac{dL}{dz} = \begin{pmatrix} \frac{\partial L}{\partial z_1} \\ \frac{\partial L}{\partial z_2} \end{pmatrix}$, Хотим $\frac{dL}{dy}$ получить – дальше ошибку распространять;

$\frac{\partial z_1}{\partial y_1}$ {это мы так $z_1 = h(y_1)$ и $\left.\frac{\partial h}{\partial y_1}\right|_{y_1}$ обозначили} – это мы знаем, т.к. сами h задавали

$\frac{\partial z_2}{\partial y_2} = \frac{\partial z_2}{\partial y_2}$ – ниче не появляется по понятным причинам

$$\left.\begin{aligned} \frac{\partial L}{\partial y_1} &= \frac{\partial L}{\partial z_1}\cdot\frac{\partial z_1}{\partial y_1} + \overset{0 \text{ очевидно}}{\cancel{\frac{\partial L}{\partial z_2}\cdot\frac{\partial z_2}{\partial y_1}}} \\ \frac{\partial L}{\partial y_2} &= \overset{0}{\cancel{\frac{\partial L}{\partial z_1}\cdot\frac{\partial z_1}{\partial y_2}}} + \frac{\partial L}{\partial z_2}\cdot\frac{\partial z_2}{\partial y_2} \end{aligned}\right\} \Rightarrow \frac{dL}{dy} = \begin{pmatrix} \frac{\partial z_1}{\partial y_1} & \frac{\partial z_2}{\partial y_2} \end{pmatrix}\cdot\begin{pmatrix} \frac{\partial L}{\partial z_1} \\ \frac{\partial L}{\partial z_2} \end{pmatrix}$$

$$\Rightarrow \frac{dL}{dy} = \left(\frac{dz}{dy}\right)^T\cdot\left(\frac{dL}{dz}\right)$$ – дальше распространили