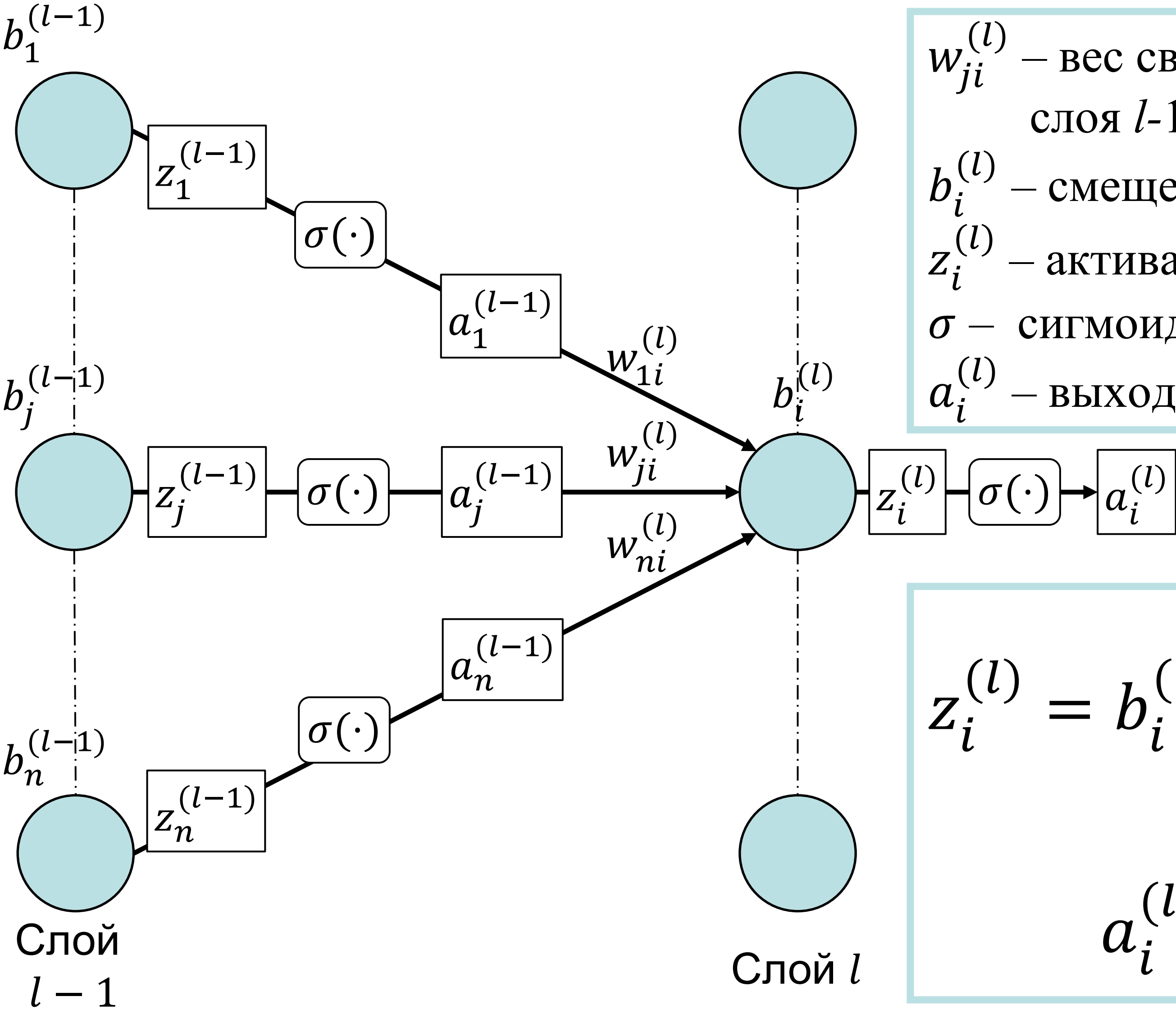
Глубокие нейронные сети

Метод обратного распространения ошибки

(Error Back Propagation)

Лекция 4

Обозначения



$w_{ji}^{(l)}$ – вес связи от j-того нейрона слоя l-1 к i-тому нейрону слоя l

$b_i^{(l)}$ – смещение

$z_i^{(l)}$ – активационный потенциал

σ – сигмоидная функция активации

$a_i^{(l)}$ – выходной сигнал

$$z_i^{(l)} = b_i^{(l)} + \sum_{j=1}^{n} w_{ji}^{(l)} a_j^{(l-1)}$$

$$a_i^{(l)} = \sigma\left(z_i^{(l)}\right)$$

Обозначения

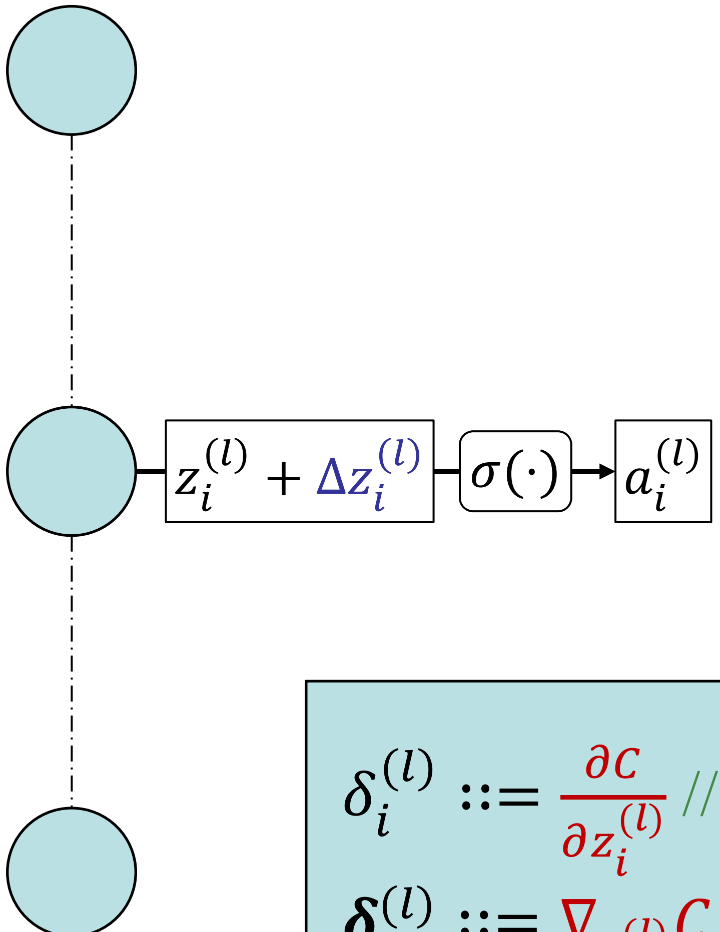
Везде далее

L – количество слоев нейронной сети

$\boldsymbol{a}^{(L)} = \boldsymbol{\alpha}(\boldsymbol{x})$ – выходной сигнал нейронной сети

$C = C_{(\boldsymbol{x},\boldsymbol{y})} = \dfrac{\left\|\boldsymbol{a}^{(L)} - \boldsymbol{y}\right\|^2}{2}$ – ошибка на прецеденте $(\boldsymbol{x}, \boldsymbol{y})$

Мера влияния нейрона на величину ошибки



Слой l

$$C = \frac{\left\|\boldsymbol{a}^{(L)} - \boldsymbol{y}\right\|^2}{2}$$

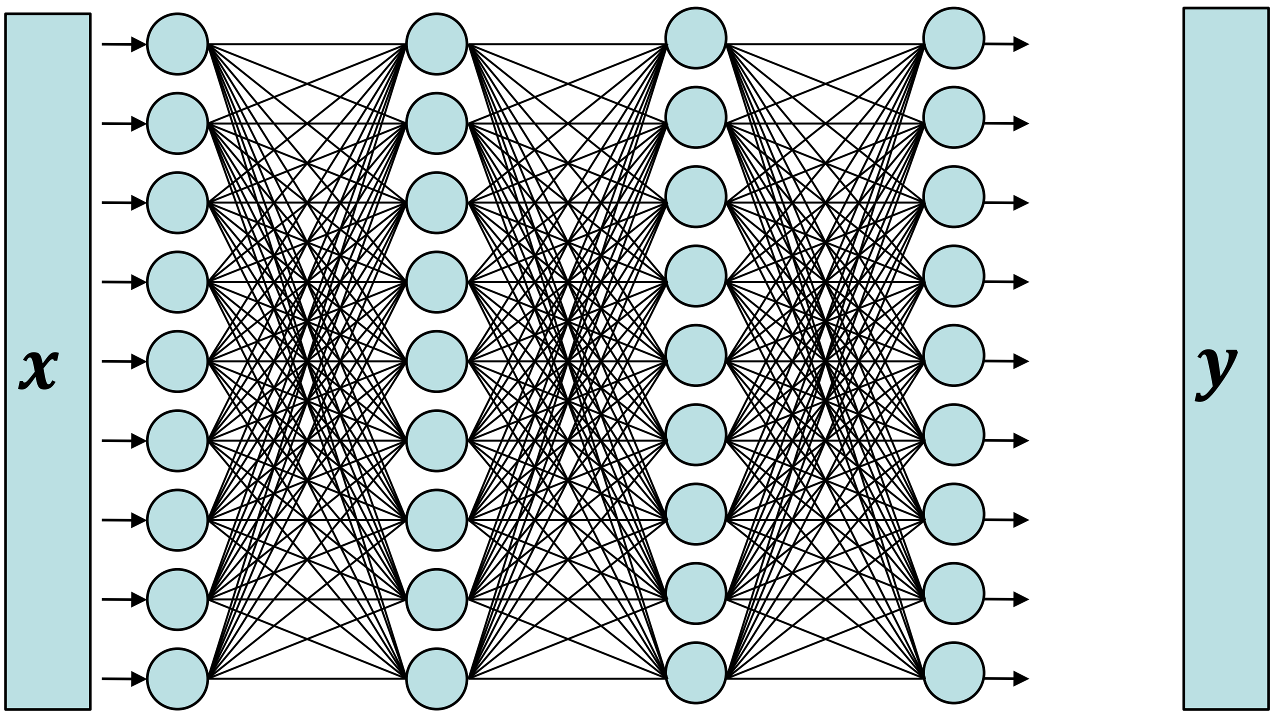
$$\Delta C \approx \frac{\partial C}{\partial z_i^{(l)}} \Delta z_i^{(l)}$$

$$\delta_i^{(l)} ::= \frac{\partial C}{\partial z_i^{(l)}}$$ // для одного нейрона

$$\boldsymbol{\delta}^{(l)} ::= \nabla_{\boldsymbol{z}^{(l)}} C$$ // для слоя

Идея метода обратного распространения ошибки

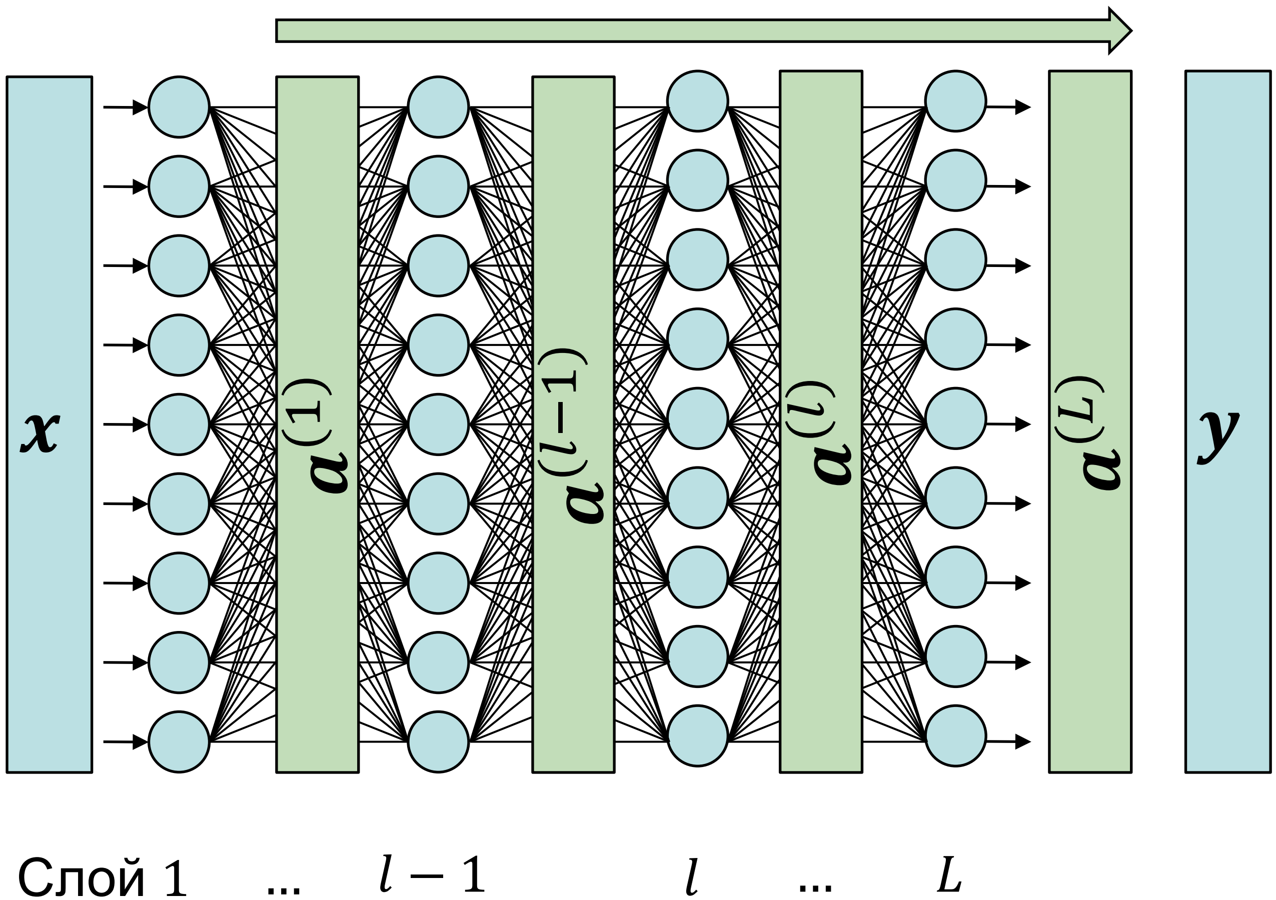
Шаг 1. Взять образец $(\boldsymbol{x}, \boldsymbol{y})$ и подать сигнал x на вход нейронной сети



Слой 0 ... $l-1$ l ... L

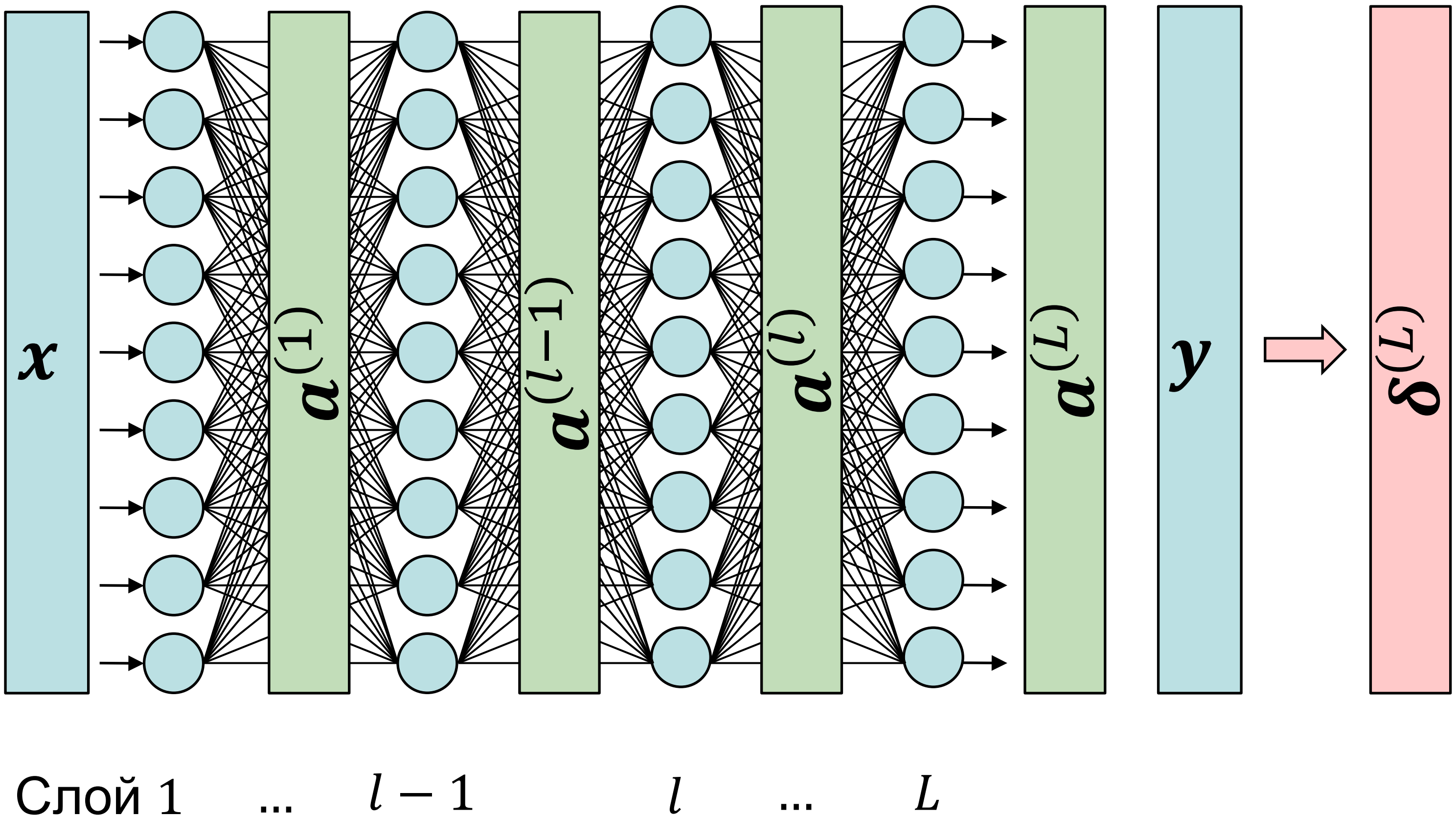
Идея метода обратного распространения ошибки

Шаг 2. Последовательно вычислить выходные сигналы $\boldsymbol{a}^{(l)}$ для каждого слоя



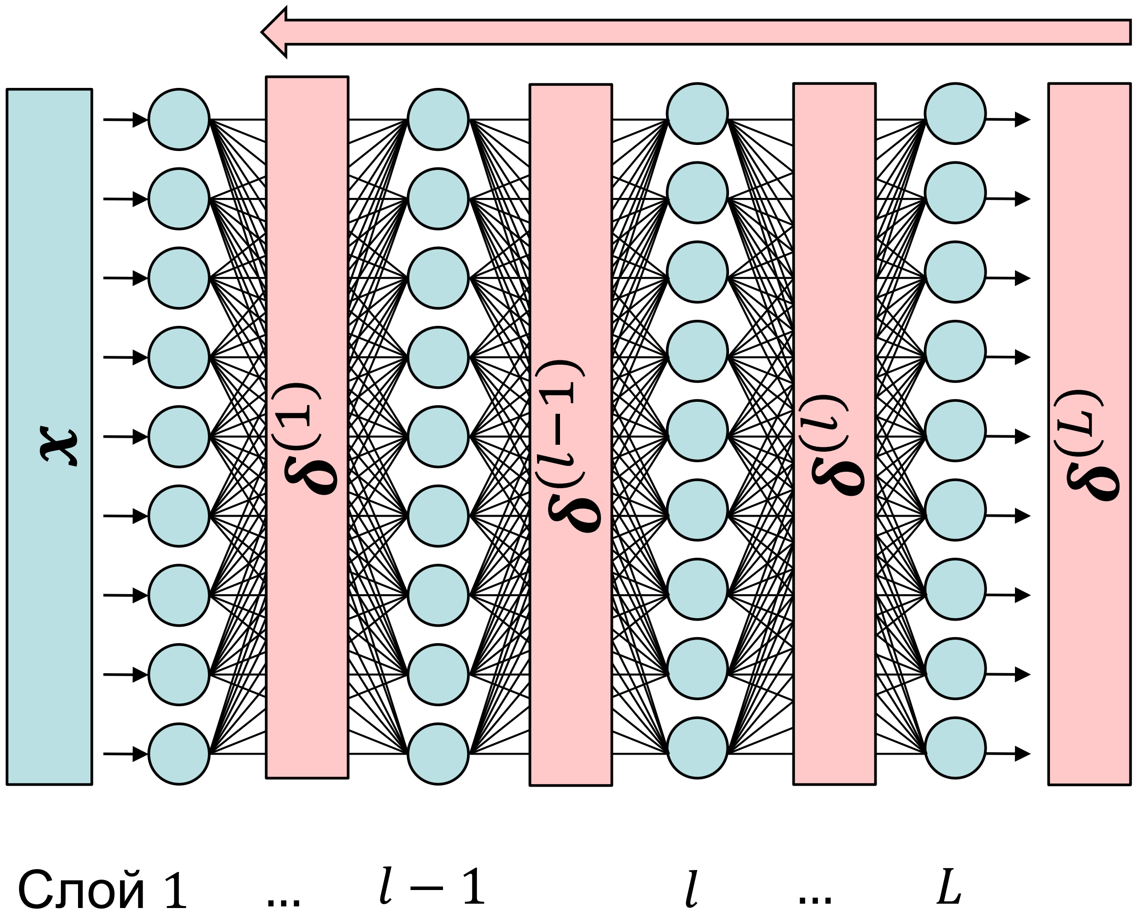
Идея метода обратного распространения ошибки

Шаг 3. Вычислить меру влияния $\boldsymbol{\delta}^{(L)}$ нейронов выходного слоя L на ошибку C



Идея метода обратного распространения ошибки

Шаг 4. Вычислить в обратном порядке меру влияния $\boldsymbol{\delta}^{(l)}$ на ошибку для каждого слоя



Идея метода обратного распространения ошибки

Шаг 5. Используя $\boldsymbol{\delta}^{(l)}$, вычислить $\nabla_{\boldsymbol{w}} C_{(\boldsymbol{x},\boldsymbol{y})}$ и $\nabla_{\boldsymbol{b}} C_{(\boldsymbol{x},\boldsymbol{y})}$ для каждого слоя l

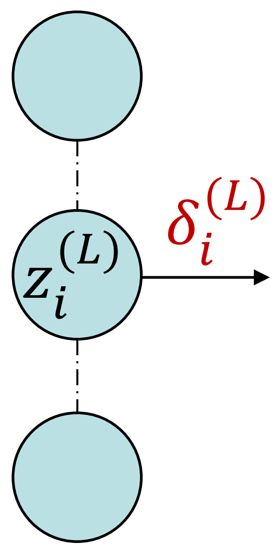
Шаг 6. Выполнив шаги 1-5 для всей подвыборки V_i, вычислить

$$\nabla_{\boldsymbol{w}} \mathbb{C}_{V_i} := \frac{1}{|V_i|} \sum_{(\boldsymbol{x},\boldsymbol{y}) \in V_i} \nabla_{\boldsymbol{w}} C_{(\boldsymbol{x},\boldsymbol{y})}$$

$$\nabla_{\boldsymbol{b}} \mathbb{C}_{V_i} := \frac{1}{|V_i|} \sum_{(\boldsymbol{x},\boldsymbol{y}) \in V_i} \nabla_{\boldsymbol{b}} C_{(\boldsymbol{x},\boldsymbol{y})}$$

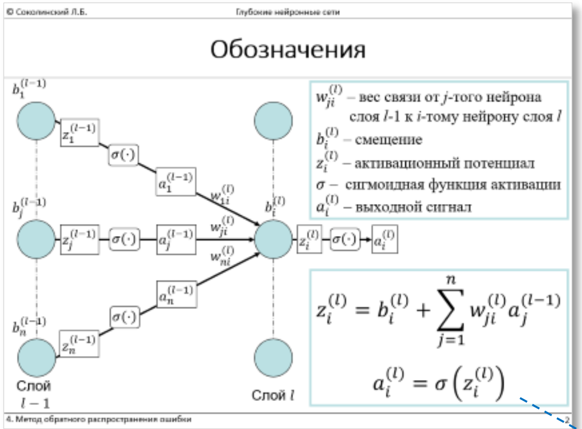
Мера влияния нейронов выходного слоя L на величину ошибки

$$\delta_i^{(L)} = \frac{\partial C}{\partial a_i^{(L)}} \cdot \sigma'\left(z_i^{(L)}\right) \qquad \text{BP1}$$



Слой L

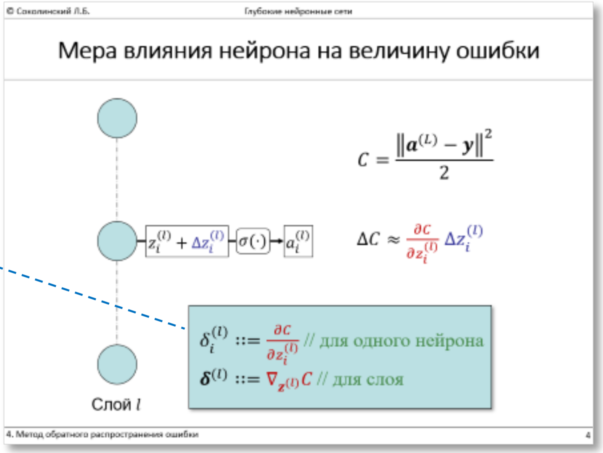
Доказательство BP1



$$\delta_i^{(L)} = \frac{\partial C}{\partial z_i^{(L)}}$$

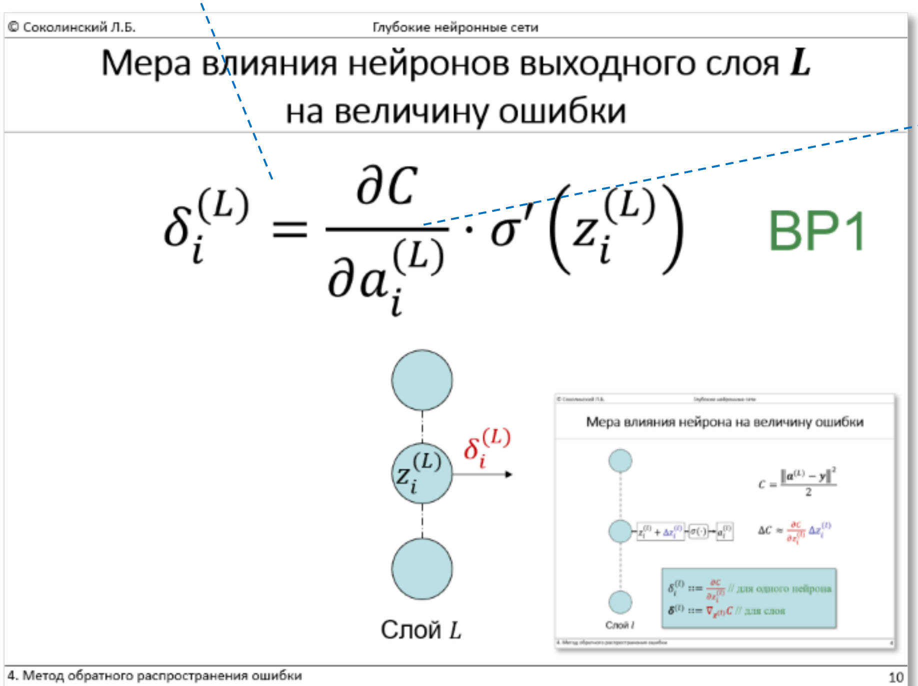
$$= \frac{\partial C}{\partial a_i^{(L)}} \cdot \frac{\partial a_i^{(L)}}{\partial z_i^{(L)}}$$

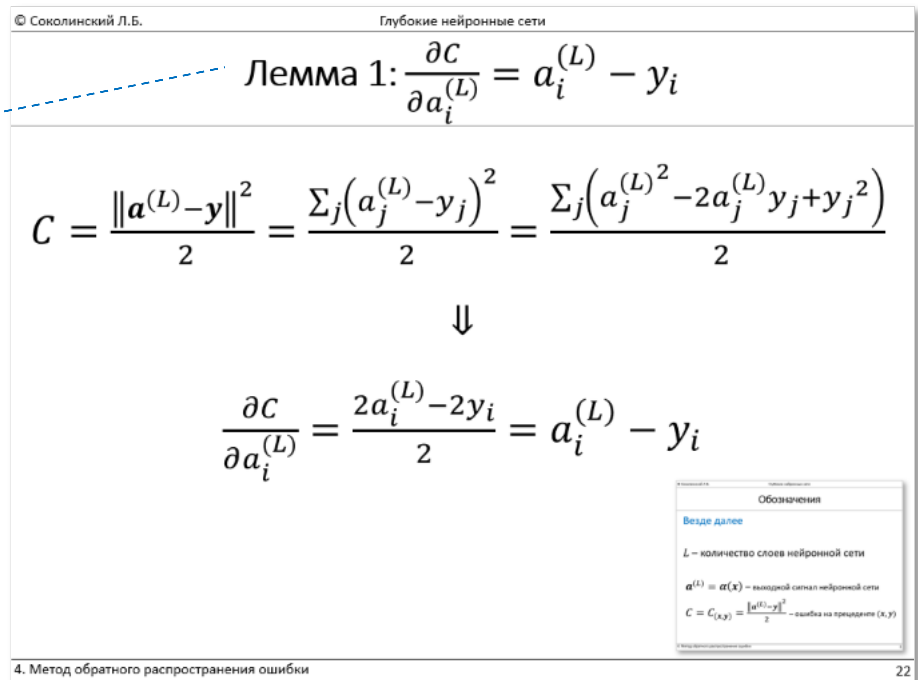
$$= \frac{\partial C}{\partial a_i^{(L)}} \cdot \sigma'\left(z_i^{(L)}\right)$$



Формула BP1 для случая среднеквадратичной ошибки

$$\delta_i^{(L)} = \left(a_i^{(L)} - y_i\right) \cdot \sigma'\left(z_i^{(L)}\right) \quad \text{BP1'}$$

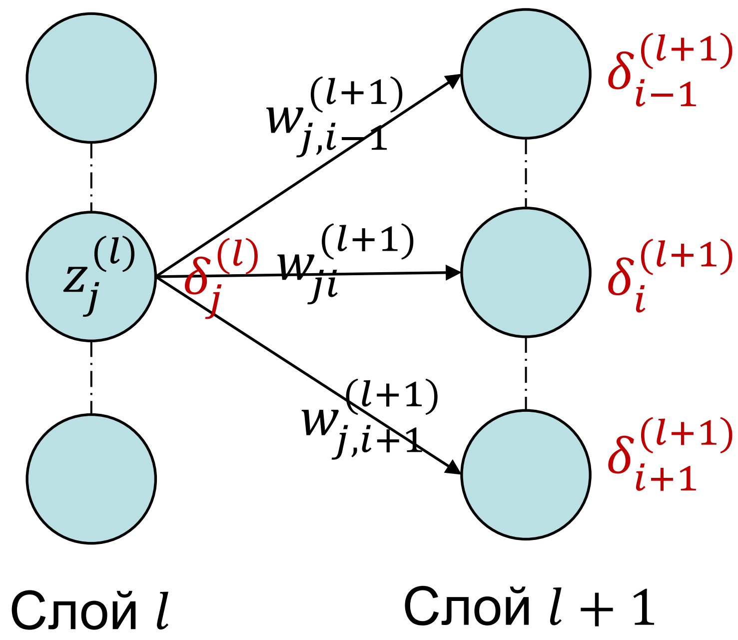


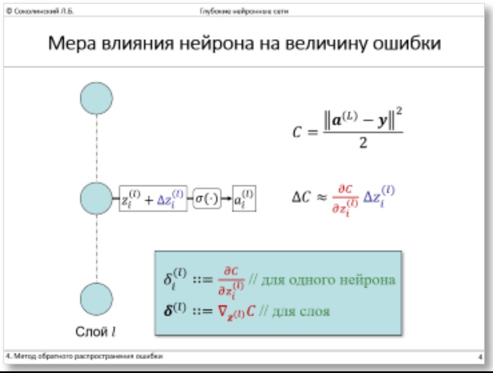


Мера влияния нейрона слоя l на величину ошибки

$$\delta_j^{(l)} = \sigma'\left(z_j^{(l)}\right)\sum_i w_{ji}^{(l+1)}\delta_i^{(l+1)}$$

BP2





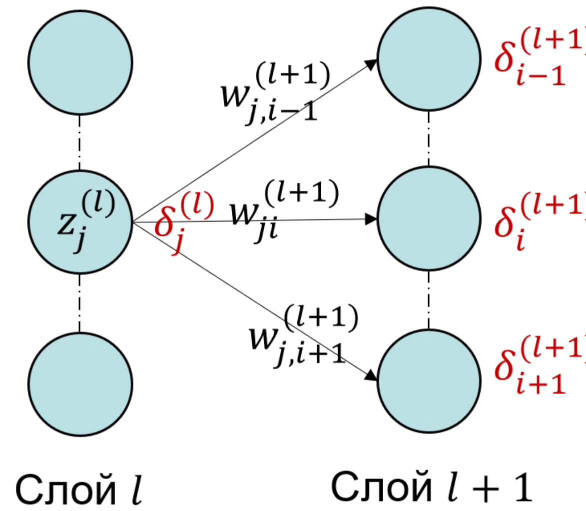
Доказательство BP2

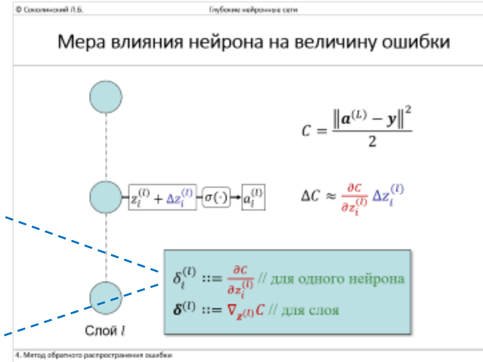
$$\delta_j^{(l)} = \frac{\partial C}{\partial z_j^{(l)}}$$

$$= \sum_i \frac{\partial C}{\partial z_i^{(l+1)}} \cdot \frac{\partial z_i^{(l+1)}}{\partial z_j^{(l)}}$$

$$= \sum_i \delta_i^{(l+1)} w_{ji}^{(l+1)} \sigma'\left(z_j^{(l)}\right)$$

$$= \sigma'\left(z_j^{(l)}\right) \sum_i \delta_i^{(l+1)} w_{ji}^{(l+1)}$$




© Соколинский Л.Б. Глубокие нейронные сети

Мера влияния нейрона на величину ошибки

$C = \frac{\left\|\boldsymbol{a}^{(L)} - \boldsymbol{y}\right\|^2}{2}$

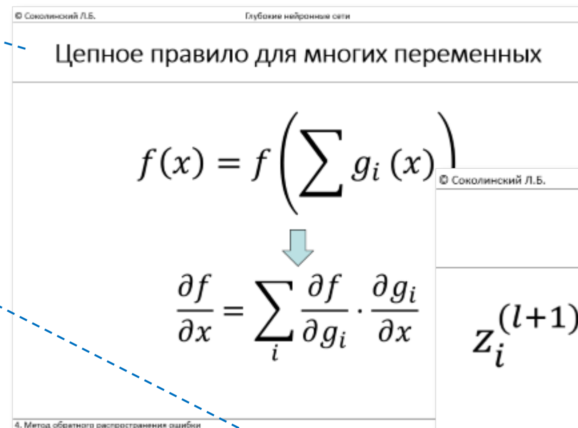
$\Delta C \approx \frac{\partial C}{\partial z_i^{(l)}} \Delta z_i^{(l)}$

$\delta_i^{(l)} ::= \frac{\partial C}{\partial z_i^{(l)}}$ // для одного нейрона

$\boldsymbol{\delta}^{(l)} ::= \nabla_{\boldsymbol{z}^{(l)}} C$ // для слоя

Слой l

4. Метод обратного распространения ошибки 4

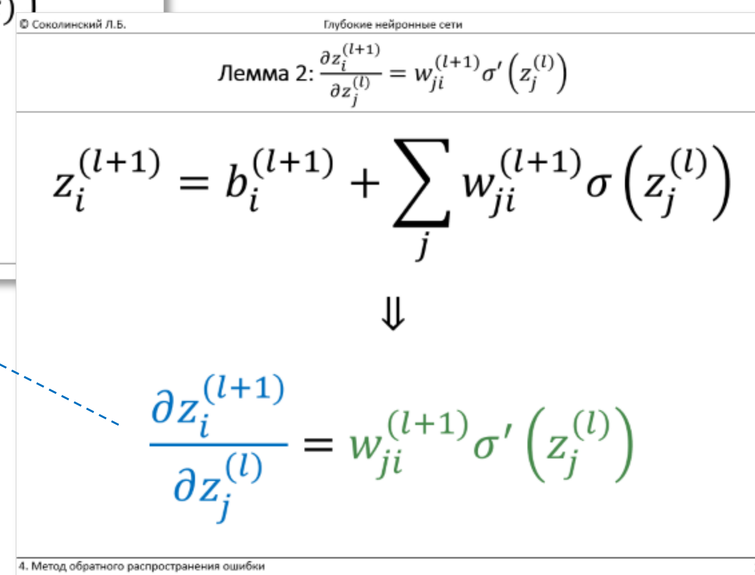

© Соколинский Л.Б. Глубокие нейронные сети

Цепное правило для многих переменных

$$f(x) = f\left(\sum g_i(x)\right)$$

$$\frac{\partial f}{\partial x} = \sum_i \frac{\partial f}{\partial g_i} \cdot \frac{\partial g_i}{\partial x}$$

4. Метод обратного распространения ошибки


© Соколинский Л.Б. Глубокие нейронные сети

Лемма 2: $\frac{\partial z_i^{(l+1)}}{\partial z_j^{(l)}} = w_{ji}^{(l+1)} \sigma'\left(z_j^{(l)}\right)$

$$z_i^{(l+1)} = b_i^{(l+1)} + \sum_j w_{ji}^{(l+1)} \sigma\left(z_j^{(l)}\right)$$

$$\Downarrow$$

$$\frac{\partial z_i^{(l+1)}}{\partial z_j^{(l)}} = w_{ji}^{(l+1)} \sigma'\left(z_j^{(l)}\right)$$

4. Метод обратного распространения ошибки 23

Формула для вычисления градиента по смещению

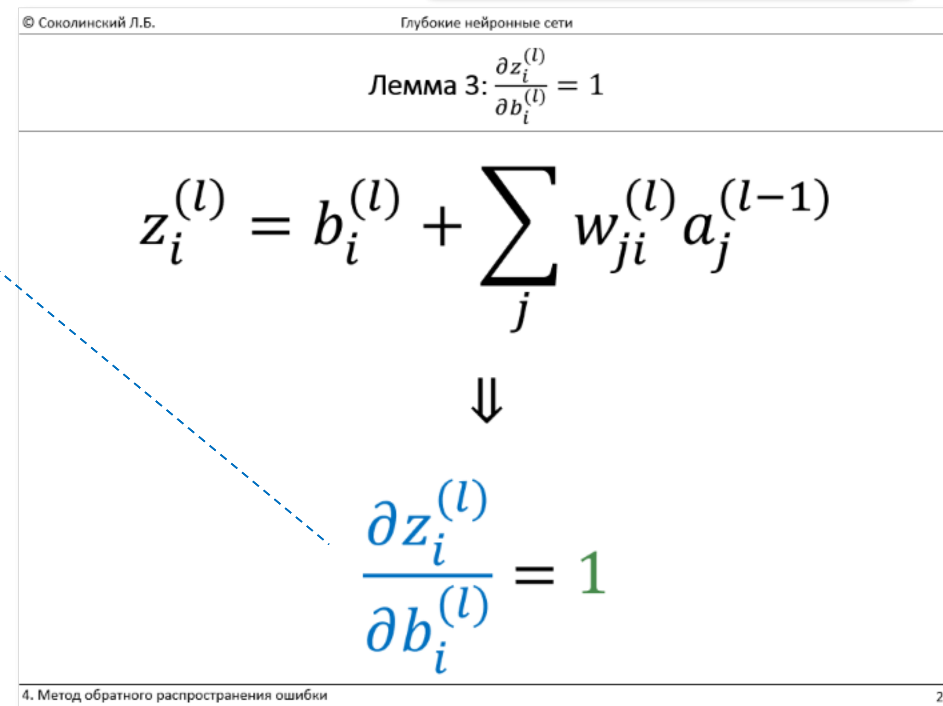
$$\frac{\partial C}{\partial b_i^{(l)}} = \delta_i^{(l)}$$

BP3

Доказательство BP3

$$\frac{\partial C}{\partial b_i^{(l)}} = \frac{\partial C}{\partial z_i^{(l)}} \cdot \frac{\partial z_i^{(l)}}{\partial b_i^{(l)}}$$

$$= \delta_i^{(l)} \cdot \frac{\partial z_i^{(l)}}{\partial b_i^{(l)}}$$



$$= \delta_i^{(l)}$$

Формула для вычисления градиента по весам

$$\frac{\partial C}{\partial w_{ji}^{(l)}} = \delta_i^{(l)} \cdot a_j^{(l-1)}$$

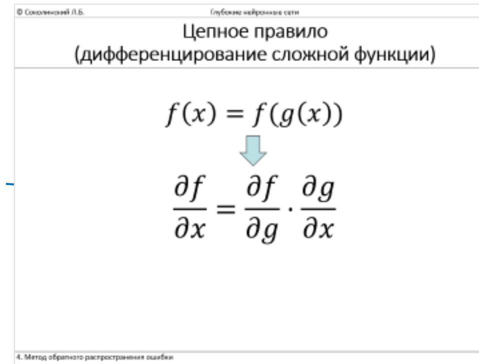
BP4

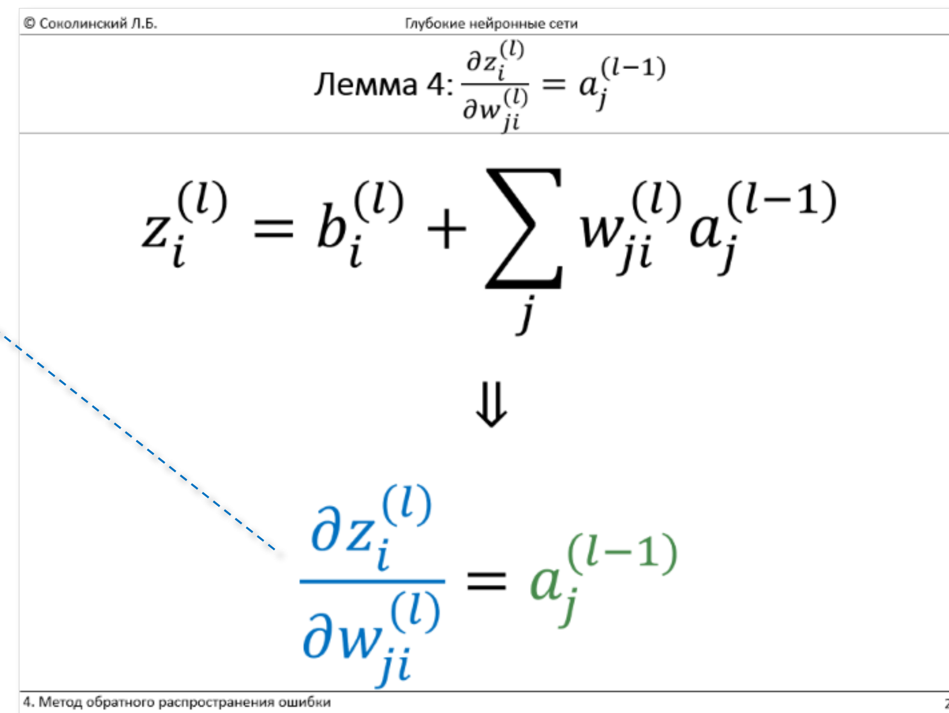
Доказательство BP4

$$\frac{\partial C}{\partial w_{ij}^{(l)}} = \frac{\partial C}{\partial z_i^{(l)}} \cdot \frac{\partial z_i^{(l)}}{\partial w_{ij}^{(l)}}$$

$$= \delta_i^{(l)} \cdot \frac{\partial z_i^{(l)}}{\partial w_{ij}^{(l)}}$$

$$= \delta_i^{(l)} \cdot a_j^{(l-1)}$$





Формулы обратного распространения ошибки (для случая среднеквадратичной ошибки)

$$\delta_i^{(L)} = \left(a_i^{(L)} - y_i\right) \cdot \sigma'\left(z_i^{(L)}\right) \qquad \text{(BP1')}$$

$$\delta_j^{(l)} = \sigma'\left(z_j^{(l)}\right) \sum_i w_{ji}^{(l+1)} \delta_i^{(l+1)} \qquad \text{(BP2)}$$

$$\frac{\partial C}{\partial b_i^{(l)}} = \delta_i^{(l)} \qquad \text{(BP3)}$$

$$\frac{\partial C}{\partial w_{ji}^{(l)}} = \delta_i^{(l)} \cdot a_j^{(l-1)} \qquad \text{(BP4)}$$

Алгоритм обратного распространения ошибки

1. **Вход** x: Установить соответствующие значения активации $\boldsymbol{a}^{(1)}$ для входного слоя

2. **Прямое распространение**: Для $l = 2, \dots, L$ последовательно вычислить $\boldsymbol{z}^{(l)} = W^{(l)}\left(\boldsymbol{a}^{(l-1)}\right) + \boldsymbol{b}^{(l)}$ и $\boldsymbol{a}^{(l)} = \boldsymbol{\sigma}\left(\boldsymbol{z}^{(\boldsymbol{l})}\right)$

3. **Вычислить**: $\boldsymbol{\delta}^{(L)} = \nabla_{\boldsymbol{a}} C \circ \boldsymbol{\sigma}'\left(\boldsymbol{z}^{(L)}\right)$

4. **Обратное распространение**: Для $l = L - 1, \dots, 2$ последовательно вычислить $\boldsymbol{\delta}^{(l)} = \left(W^{(l+1)}\boldsymbol{\delta}^{(l+1)}\right) \circ \boldsymbol{\sigma}'\left(\boldsymbol{z}^{(l)}\right)$

5. **Выход**: Для $l = 2, \dots, L$ вычислить

$$\nabla_{\boldsymbol{w}^{(l)}} C_{(\boldsymbol{x},\boldsymbol{y})} = {\boldsymbol{\delta}^{(l)}}^{T} \circ \boldsymbol{a}^{(l-1)}$$

$$\nabla_{\boldsymbol{b}^{(l)}} C_{(\boldsymbol{x},\boldsymbol{y})} = \boldsymbol{\delta}^{(l)}$$

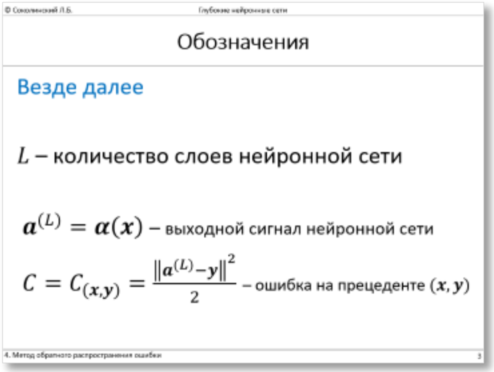
Конец лекции 4

Лемма 1: $\frac{\partial C}{\partial a_i^{(L)}} = a_i^{(L)} - y_i$

$$C = \frac{\left\|\boldsymbol{a}^{(L)} - \boldsymbol{y}\right\|^2}{2} = \frac{\sum_j \left(a_j^{(L)} - y_j\right)^2}{2} = \frac{\sum_j \left({a_j^{(L)}}^2 - 2a_j^{(L)} y_j + {y_j}^2\right)}{2}$$

$$\Downarrow$$

$$\frac{\partial C}{\partial a_i^{(L)}} = \frac{2a_i^{(L)} - 2y_i}{2} = a_i^{(L)} - y_i$$



Обозначения

Везде далее

L – количество слоев нейронной сети

$\boldsymbol{a}^{(L)} = \boldsymbol{\alpha}(\boldsymbol{x})$ – выходной сигнал нейронной сети

$C = C_{(\boldsymbol{x},\boldsymbol{y})} = \frac{\left\|\boldsymbol{a}^{(L)} - \boldsymbol{y}\right\|^2}{2}$ – ошибка на прецеденте $(\boldsymbol{x}, \boldsymbol{y})$

Лемма 2: $\frac{\partial z_i^{(l+1)}}{\partial z_j^{(l)}} = w_{ji}^{(l+1)} \sigma'\left(z_j^{(l)}\right)$

$$z_i^{(l+1)} = b_i^{(l+1)} + \sum_j w_{ji}^{(l+1)} \sigma\left(z_j^{(l)}\right)$$

$$\Downarrow$$

$$\frac{\partial z_i^{(l+1)}}{\partial z_j^{(l)}} = w_{ji}^{(l+1)} \sigma'\left(z_j^{(l)}\right)$$

Лемма 3: $\frac{\partial z_i^{(l)}}{\partial b_i^{(l)}} = 1$

$$z_i^{(l)} = b_i^{(l)} + \sum_j w_{ji}^{(l)} a_j^{(l-1)}$$

$$\Downarrow$$

$$\frac{\partial z_i^{(l)}}{\partial b_i^{(l)}} = 1$$

Лемма 4: $\dfrac{\partial z_i^{(l)}}{\partial w_{ji}^{(l)}} = a_j^{(l-1)}$

$$z_i^{(l)} = b_i^{(l)} + \sum_j w_{ji}^{(l)} a_j^{(l-1)}$$

$$\Downarrow$$

$$\frac{\partial z_i^{(l)}}{\partial w_{ji}^{(l)}} = a_j^{(l-1)}$$

Цепное правило (дифференцирование сложной функции)

$$f(x) = f(g(x))$$

$$\Downarrow$$

$$\frac{\partial f}{\partial x} = \frac{\partial f}{\partial g} \cdot \frac{\partial g}{\partial x}$$

Цепное правило для многих переменных

$$f(x) = f\left(\sum g_i(x)\right)$$

$$\Downarrow$$

$$\frac{\partial f}{\partial x} = \sum_i \frac{\partial f}{\partial g_i} \cdot \frac{\partial g_i}{\partial x}$$