# Домашнее задание №1

# Машинное обучение. 6 курс. Осенний семестр.

Горбунов Егор Алексеевич

3 октября 2016 г.

## 1 Back propagation алгоритм

$$s_i^l = \sum_j \omega_{ij}^l \cdot x_j^{l-1} \qquad \delta_i^L = \frac{\partial L(w)}{\partial s_i^L} \qquad \delta_i^{l-1} = \sum_j \delta_j^l \cdot \omega_{ji}^l \cdot \sigma'(s_i^l)$$

**Задание №1** Выписать в явном виде выражения для $\delta^L$:

- Решающая функция: softmax
$$\sigma^L(s^L)_i = \frac{\exp s_i^L}{\sum_j \exp s_j^L}$$
- Функция активации: сигмоида
$$\sigma(s) = \frac{1}{1+e^{-s}}$$
- Функция потерь: кросс-энтропия
$$C(w) = -\sum_k o_k \cdot \log x_k^L$$

***Решение:***

$$\begin{aligned}
\delta_i^L = \frac{\partial L(w)}{\partial s_i^L} &= \sum_k \frac{\partial C(w)}{\partial x_k^L} \cdot \frac{\partial x_k^L}{\partial s_i^L} \\
&= \left[\text{от } s_i^L \text{ зависит только } x_i^L\right] \\
&= \frac{\partial C(w)}{\partial x_i^L} \cdot \frac{\partial x_i^L}{\partial s_i^L} \\
&= (-o_i \cdot \log x_i^L)' \cdot \frac{\partial \sigma^L(s^L)_i}{\partial s_i^L} = -\frac{o_i}{x_i^L} \cdot \frac{\exp s_i^L \cdot \sum_j \exp s_j^L - \exp 2s_i^L}{\left(\sum_j \exp s_j^L\right)^2} \\
&= -\frac{o_i}{x_i^L} \cdot \frac{\exp s_i^L \cdot \sum_{j\neq i} \exp s_j^L}{\left(\sum_j \exp s_j^L\right)^2}
\end{aligned}$$

**Задание №2** Выписать в явном виде выражения для $\delta^{l-1}$ для следующий функций активации:

(a) Сигмоида: $\sigma(s) = \frac{1}{1+e^{-s}}$

(b) Гиперболический тангенс: $\sigma(s) = \frac{e^{2s}-1}{e^{2s}+1}$

(c) ReLU: $\sigma(s) = \max(0, s)$

***Решение:***

(a) $\sigma'(s) = \frac{e^{-s}}{(1+e^{-s})^2} \Longrightarrow \delta_i^{l-1} = \sum_j \delta_j^l \cdot \omega_{ji}^l \cdot \frac{e^{-s_i^l}}{\left(1+e^{-s_i^l}\right)^2}$

(b) $\sigma'(s) = \frac{2e^{2s}\left(e^{2s}+1\right)-2e^{2s}\left(e^{2s}-1\right)}{\left(e^{2s}+1\right)^2} = \frac{4e^{2s}}{\left(e^{2s}+1\right)^2} \Longrightarrow \delta_i^{l-1} = \sum_j \delta_j^l \cdot \omega_{ji}^l \cdot \frac{4e^{2s_i^l}}{\left(e^{2s_i^l}+1\right)^2}$

(c) $\sigma'(s) = \begin{cases} 0, & s \le 0 \\ 1, & s > 0 \end{cases} \Longrightarrow \delta_i^{l-1} = \begin{cases} 0, & s_i^l \le 0 \\ \sum_j \delta_j^l \cdot \omega_{ji}^l, & s_i^l > 0 \end{cases}$