

# Мультиклассовая логистическая регрессия

$$p(\theta_k, k, x_i) \propto \exp(\theta_k x_i)$$

$$p(\theta_k, x_i, k) = \frac{e^{\theta_k x_i}}{\sum_{j=1}^K e^{\theta_j x_i}}$$

$$L(x_i, y_i) = - \sum_{k=1}^K \underbrace{y_{ik}}_{\text{one-hot encoding}} \log p_{ik}$$

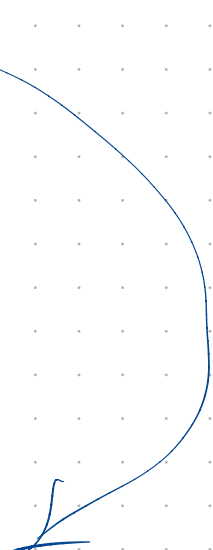
$y_{oh}$  1 0 0 1 0 0 one-hot encoding

$y = 3$

$y_i$ 

--	--	--	--	--	--

$\log \{ p_{i0} \mid p_{i1} \mid \mid \mid p_{i5} \}$



$$L_{\text{bin}} = -(y_i \log p_i + (1-y_i) \log(1-p_i))$$

$$L_{\text{bin}} = -(y_{i1} \log p_1 + y_{i0} \log p_0) = -\sum_{k=0}^1 y_{ik} \log p_k$$

$$L(x_i, y_i) = -\sum_{k=1}^K \underbrace{y_{ik}} \log p_{ik}$$

$$p(x_i, \theta_k, k) = \frac{e^{\theta_k x_i}}{\sum_{j=1}^K e^{\theta_j x_i}}$$

$$p(x_i, \theta) = \text{Softmax}(\theta^T x_i)$$

$$\begin{matrix} \xrightarrow{F} & & M \\ \begin{matrix} \begin{matrix} \vdots \\ \vdots \\ \vdots \\ \vdots \\ \vdots \end{matrix} \\ N \end{matrix} & \begin{bmatrix} \begin{matrix} 1 \\ 1 \\ 1 \\ 1 \\ 1 \end{matrix} \\ \vdots \\ \vdots \\ \vdots \\ \vdots \end{bmatrix} & X & \begin{bmatrix} F & \theta \end{bmatrix} & = & \begin{bmatrix} M \\ N \\ Z \end{bmatrix} \end{matrix}$$

$$N \times F \cdot \underbrace{F \times M}_{\theta} \rightarrow N \times M$$

$$L = -\sum_k \log P \odot Y$$

$$P = \text{Softmax}(Z)$$

$$\begin{array}{|c|} \hline M \\ \hline N \quad Y \\ \hline \end{array} \odot_N \begin{array}{|c|} \hline M \\ \hline \log P \\ \hline \end{array} \rightarrow \sum_{i=1}^N \begin{array}{|c|} \hline \leftarrow \\ \hline z \\ \hline Y \log P \\ \hline \end{array} = L$$

$L \rightarrow \nabla_{\theta} L \rightarrow$  градиентная оптимизация  
 функции потерь  $L(X, Y, \theta)$