Мультиномиальная логистическая регрессия

$$p(x, \theta_k) = \frac{e^{\theta_k x}}{\sum\limits_{c=1}^{K} e^{\theta_c x}}$$

$$\mathcal{L}(x, \theta_1, \theta_2, \dots \theta_K) = -\sum_{c=1}^{K} y_c \log(p(x, \theta_c))$$

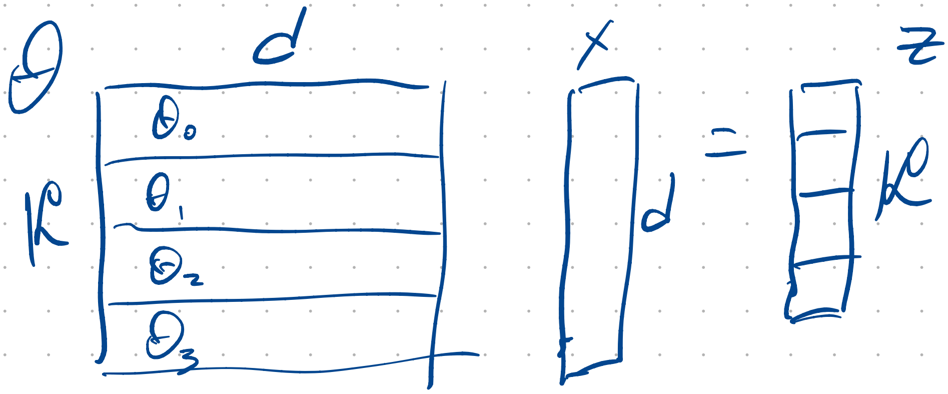
One-hot кодирование

$$\begin{array}{c} \\ A \\ B \\ C \\ D \end{array}
\begin{array}{c} y = B \\ \begin{bmatrix} 0 \\ 1 \\ 0 \\ 0 \end{bmatrix} \end{array}
\qquad
\begin{array}{c} y = C \\ \begin{bmatrix} 0 \\ 0 \\ 1 \\ 0 \end{bmatrix} \end{array}
\qquad
\begin{array}{c} P \\ \begin{bmatrix} 0{,}2 \\ 0{,}37 \\ 0{,}4 \\ 0{,}03 \end{bmatrix} \end{array}
\begin{array}{l} \\ = p(x, \theta_0) \\ = p(x, \theta_1) \\ = p(x, \theta_2) \\ = p(x, \theta_3) \end{array}$$

$$p_c \sim e^{\theta_c x}$$

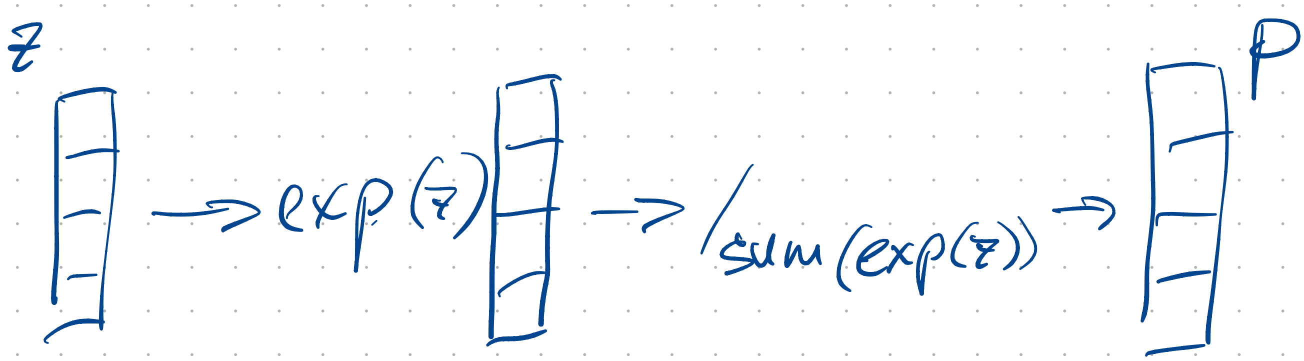
$$\begin{array}{c} \theta_0 \\ \theta_1 \\ \theta_2 \\ \theta_3 \end{array}
\begin{bmatrix} \theta_{c_0} & \theta_c \\ & \\ & \\ & \end{bmatrix}
*
\begin{bmatrix} 1 \\ x \end{bmatrix}
\begin{array}{l} = z_0 \\ = z_1 \\ = z_2 \\ = z_3 \end{array}$$

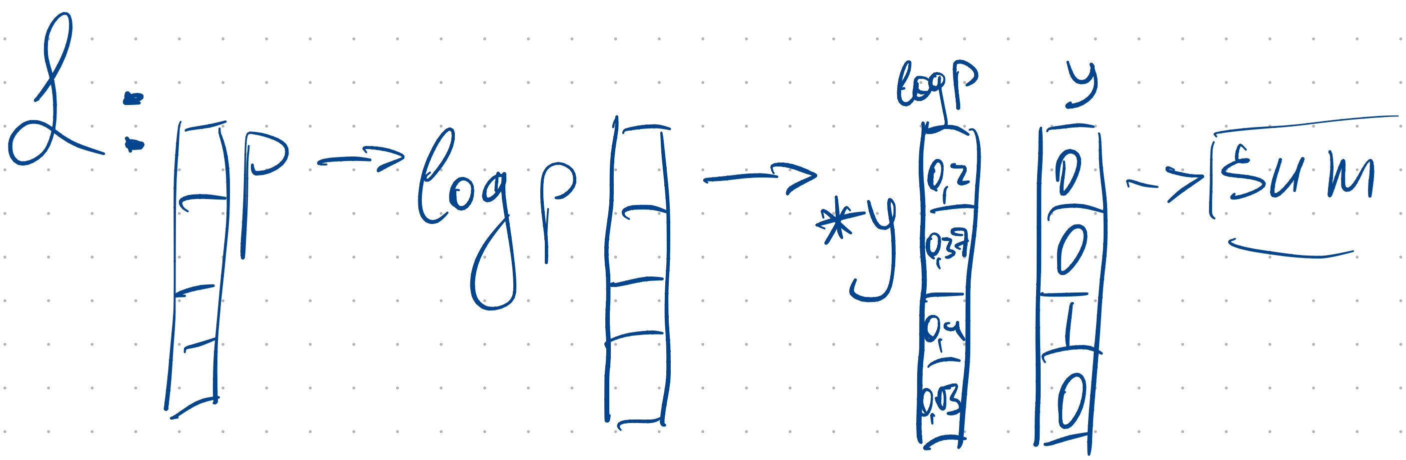
$$p_c \sim e^{z_c}$$



$$\underset{\Theta}{k \times d} \cdot \underset{x}{d \times 1} \longrightarrow \underset{z}{k \times 1}$$

$$p_i = \text{Softmax}(z_i) = \frac{e^{z_i}}{\sum_{j=1}^{k} e^{z_j}}$$





Softmax: стабилизация вычисления

$$\mathrm{Softmax}(z) = \frac{e^{z_i}}{\sum_{j=1}^{K} e^{z_j}}$$

float64: $\max \quad 1{,}7 \cdot 10^{308}$

$$\ln(1{,}7 \cdot 10^{308}) \approx 710$$

z

$\rightarrow z^* = z - \max(z)$

$\qquad\qquad\quad z_m$

$$\mathrm{Sm}(z^*) = \frac{e^{z - z_m}}{\sum_{j=1}^{K} e^{z_j - z_m}} = \frac{e^{z} / \cancel{e^{z_m}}}{\cancel{\frac{1}{e^{z_m}}} \sum_{j=1}^{K} e^{z_j}} = \mathrm{Sm}(z)$$

$$\mathrm{Sm}(z) = \mathrm{Sm}(z^*)$$

$$\sigma(z) = \frac{1}{1+e^{-z}} \qquad z \to \pm 700$$