



## Model #2: Multinomial logistic regression /

Multinomial logistic regression, also known as softmax regression, models the probability of belonging to any

class, given a feature vector  $\vec{x}_i \in \mathbb{R}^{784}$ .

Think of it as a generalization of logistic regression.

$$P(y=1|X_i)=\sigma(\bar{w}\cdot A$$

$$P(y_i = (3) | \vec{x}_i) =$$

$$e^{i\omega_0 \cdot Aug(x_i)} + e^{i\omega_1 \cdot Aug(x_i)} + \dots + e^{i\omega_q \cdot Aug(x_i)} + \dots + e^{i\omega_q \cdot Aug(x_i)} = 1$$





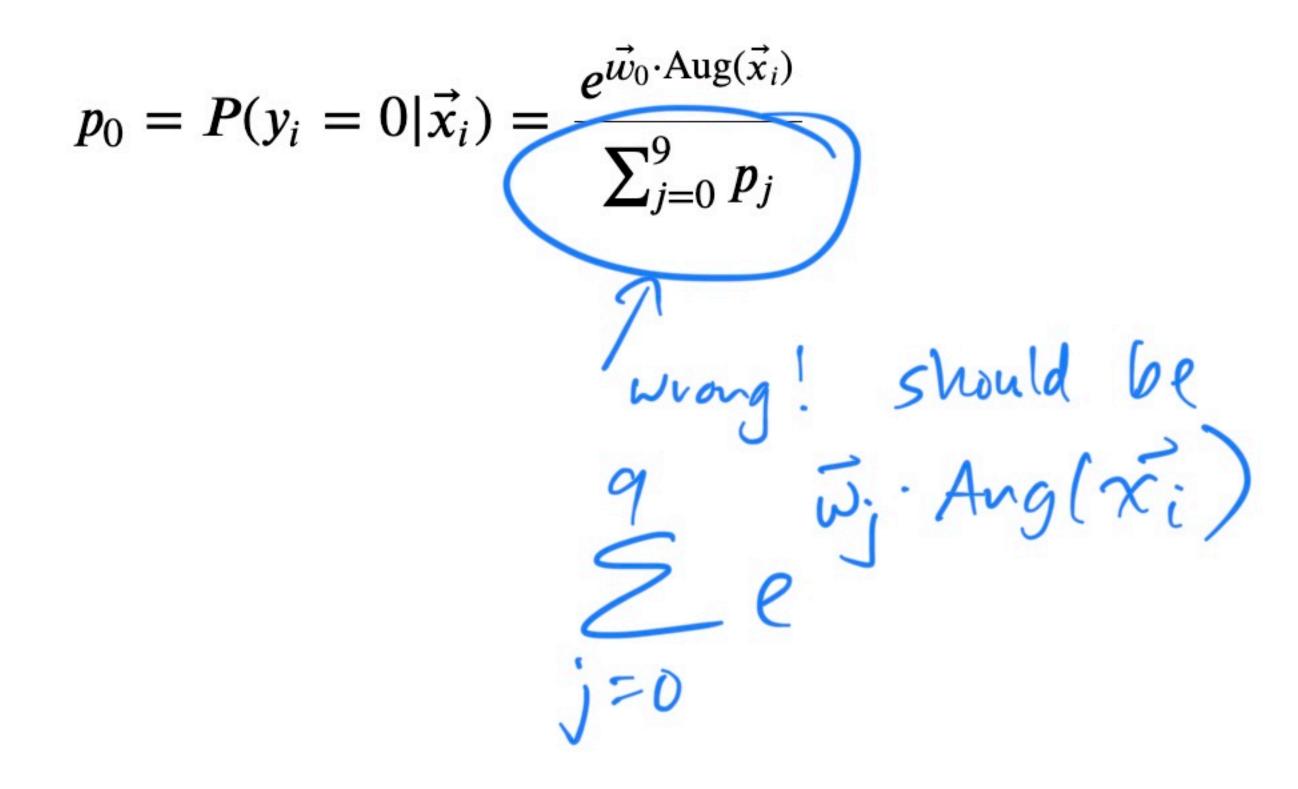






Let  $p_j$  represent the modelled probability of class j, given a feature vector. Note that  $j \in \{0, 1, \dots, 9\}$ .

Then, for instance:











The softmax function is a generalization of the logistic function to multiple dimensions.

Suppose  $\vec{z} \in \mathbb{R}^d$ . Then, the softmax of  $\vec{z}$  is defined element-wise as follows:

