Softmax function.

class가 2개 일때.

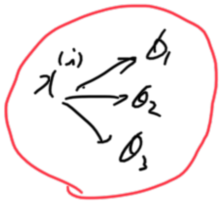
$$\frac{p_j}{1-p_j} \Rightarrow logit(p_j) = \log_e\left(\frac{p_j}{1-p_j}\right) = z = \theta^T x.$$

class가 k개일 때.

j번째 사건이 일어날 확률에 대해서,
X와 θ_j의 linear combination

$$\frac{p_j}{p_k} \Rightarrow logit(p_j) = \log_e\left(\frac{p_j}{p_k}\right) = z_j = X^T\theta_j \longrightarrow$$ θ는 j개만큼 나옴.

가장 높은 것 고름.



$$\log_e\left(\frac{p_j}{p_k}\right) = z_j \Rightarrow \frac{p_j}{p_k} = e^{z_j} \Rightarrow \sum_{j=1}^{k}\frac{p_j}{p_k} = \sum_{j=1}^{k} e^{z_j}$$

$$\Rightarrow \frac{1}{p_k}\underbrace{\sum_{j=1}^{k} p_j}_{1} = \sum_{j=1}^{k} e^{z_j} \Rightarrow \frac{1}{p_k} = \sum_{j=1}^{k} e^{z_j} \Rightarrow p_k = \frac{1}{\sum_{j=1}^{k} e^{z_j}}$$

$z_j = X\theta_j$ 이고, j는 k개(클래스의 개수) 만큼 존재할 수 있다.

이때, $\frac{p_j}{p_k} = e^{z_j}$ 에서 $p_k = \frac{1}{\sum_{j=1}^{k} e^{z_j}}$ 이므로, $\frac{p_j}{\frac{1}{\sum_{j=1}^{k} e^{z_j}}} = e^{z_j} \Rightarrow p_j = \frac{e^{z_j}}{\sum_{j=1}^{k} e^{z_j}}$

$$p_j = \frac{e^{z_j}}{\sum_{j=1}^{k} e^{z_j}} = \frac{e^{z_j}}{\sum_{j=1}^{k} e^{x^T\theta_j}} \quad (\because z_j = X^T\theta_j)$$

p_j는 각각의 사건이 일어날 확률.
ex) 개일 확률, 고양이일 확률, 사자일 확률.

$$\theta = \begin{bmatrix}\theta_1\\ \theta_2\\ \vdots\\ \theta_j\end{bmatrix} = \begin{array}{c} \begin{array}{cccc} x_1 & x_2 & \cdots & \end{array} \\ \begin{bmatrix} w_{10} & w_{11} & \cdots & w_{1i}\\ w_{20} & w_{21} & \cdots & w_{2i}\\ \vdots & \vdots & \ddots & \vdots\\ w_{j0} & w_{j1} & \cdots & w_{ji}\end{bmatrix}\end{array}$$

$$\begin{array}{ccc} \text{개} & \text{고양이} & \text{사자} \end{array}$$
$$\underbrace{\begin{bmatrix} 0.2 & 0.3 & 0.5 \end{bmatrix}}_{\downarrow \; 1}$$

ex) 변수 7개, class가 3개면

$(7+\underbrace{1}_{w_0}) \times 3 = 24$개의 w를 구해야 함.