

General ML Question

2. Write the equation of soft-max operation.

Express the corresponding

(log-likelihood) loss incurred when the true class label index is y .

soft-max is sigmoid on n dimensions

$$\begin{aligned} &\searrow \text{Sigmoid}(x) = \frac{1}{1 + e^{-x}} \\ &\swarrow \text{Softmax}(x) = \frac{e^{x_i}}{\sum_{i=1}^n e^{x_i}} \end{aligned}$$

log loss : $D_{KL}(q \parallel p) = q \log q - \underbrace{q \log p}$

$\begin{matrix} \text{true} & \text{prediction} \\ \downarrow & \swarrow \end{matrix}$

$\min - q \log p \rightarrow \underline{\max q \log p}$

$$= \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix} \log \begin{pmatrix} e^{x_1} / \sum_{i=1}^n e^{x_i} \\ \vdots \\ e^{x_n} / \sum_{i=1}^n e^{x_i} \end{pmatrix}$$