Subject: IFH - Q3

Date: Dec 5th 2021

Attendants

soft - argmax

$$\mathrm{sigmoid}(x) = \frac{1}{1+e^{-x}}$$

$$x \mapsto \mathrm{sigmoid}(x)$$

$$(\infty, \infty) \mapsto [0, 1]$$

これの多次元拡張が

$$\boldsymbol{x} = (x_1, \cdots x_n)^T$$

$$\mathrm{softmax}(\boldsymbol{x}) = \underbrace{\left(\frac{e^{x_i}}{\sum_{i=1} e^{x_i}}, \cdots \frac{e^{x_n}}{\sum_{i=1} e^{x_n}} \right)}_{\text{確率分布のベクトルとして解釈できる}}$$

$T = \frac{1}{\beta}$ (逆温度)

また softmaxの引数に $\mathrm{softmax}(\beta x)$ に対応

$$= \left(\frac{1}{K}, \frac{1}{K}, \frac{1}{K}, \cdots \frac{1}{K} \right) \text{ if } \beta = 0$$

$$= (0, 1, 0 \cdots 0) \text{ if } \beta \to \infty$$

Subject　　　　　　　　　　　　　　　　Date

Attendants

transform as log-loss

$$\hat{y} = \mathrm{softmax}(x) = \left(\frac{e^{x_1}}{\sum_{i=1} e^{x_i}}, \cdots \frac{e^{x_n}}{\sum_{i=1} e^{x_i}} \right)$$

ここで y と $\hat{y}$ の KL 距離を考える
真　予測

$$L = KL(y \| \hat{y}) = y \log \frac{\hat{y}}{y} \quad > 0$$

$$= \underbrace{y \log \hat{y}}_{} - \underbrace{y \log y}_{変えられない}$$

これを最小化したい

$y \log \hat{y}$ は cross-entropy と呼ばれる

$$H(y, \hat{y}) = -y \log \hat{y}$$

$$L = H(y, \hat{y}) = -y \log \hat{y}$$

Subject KL Divergence 復習　　Date

Attendants

$\underbrace{p(x)}_{\text{正解}}$ を 前提　$\underbrace{q(x)}_{\text{推定}}$ の 予測しにくさ

$$D_{KL}(p\|q) = \int_{-\infty}^{\infty} \underbrace{p(x)}_{\text{正解ラベル}} \log \frac{q(x)}{p(x)} dx$$

$$= \underbrace{p(x) \log q(x)}_{\text{これを最小化}} - p(x) \log p(x)$$

クロスエントロピー　$H(p, q) = \underbrace{-p(x) \log q(x)}$

これが最大のとき

$p(x) = q(x)$