パターン認識

学籍番号：　　　　　　　　　　　　　　　　氏名：

2 クラス識別ではシグモイド関数を使って事後確率を回帰する。シグモイド関数は次のとおり：

P(y=1|x) + P(y=0|x) = 1

$$\sigma(a) = \frac{1}{1+\exp(-a)} \tag{1}$$

ロジスティック回帰を用いる 2 クラス識別器は、式 (1) に $a = \boldsymbol{w}^\top \boldsymbol{x}$ を代入した次式で事後確率を表現する。

xがgivenのとき
yは0 or 1

$$p(y=1|\boldsymbol{x}) = \frac{1}{1+\exp(-\boldsymbol{w}^\top \boldsymbol{x})}$$

学習で推定するパラメータは $\boldsymbol{w}$ である。もう一方のクラスに属する確率は次式で表される：

$$p(y=0|\boldsymbol{x}) = 1 - p(y=1|\boldsymbol{x}) = \frac{\exp(-\boldsymbol{w}^\top \boldsymbol{x})}{1+\exp(-\boldsymbol{w}^\top \boldsymbol{x})}$$

多クラス識別のときにはソフトマックス関数により事後確率を表現する。K クラス識別問題で、k 番目のクラスに属する事後確率を $p(y=k|\boldsymbol{x})$ で表すことにする。ソフトマックス関数は次式のように表すことができる：

$$\mathrm{SM}_k(\boldsymbol{x}) = \frac{\exp(a_k)}{\sum_{i=1}^{K} \exp(a_i)} \tag{2}$$

(非線形)

式 (2) に $a_k = \boldsymbol{w}_k^\top \boldsymbol{x}$ を代入した次式で事後確率を表現する。

パラメータ
データ

クラスk用

wはクラスごとに用意

$$\boxed{p(y=k|\boldsymbol{x}) = \frac{\exp(\boldsymbol{w}_k^\top \boldsymbol{x})}{\sum_{i=1}^{K} \exp(\boldsymbol{w}_i^\top \boldsymbol{x})}}$$

学習で推定するパラメータは $\boldsymbol{w}_1, \boldsymbol{w}_2, \ldots, \boldsymbol{w}_K$ である。

例えば $K=2$ のときのソフトマックス関数は次のとおりとなる。

$$p(y=1|\boldsymbol{x}) = \mathrm{SM}_1(\boldsymbol{x}) = \frac{\exp(\boldsymbol{w}_1^\top \boldsymbol{x})}{\exp(\boldsymbol{w}_1^\top \boldsymbol{x}) + \exp(\boldsymbol{w}_2^\top \boldsymbol{x})}$$

$$p(y=2|\boldsymbol{x}) = \mathrm{SM}_2(\boldsymbol{x}) = \frac{\exp(\boldsymbol{w}_2^\top \boldsymbol{x})}{\exp(\boldsymbol{w}_1^\top \boldsymbol{x}) + \exp(\boldsymbol{w}_2^\top \boldsymbol{x})}$$

$\mathrm{SM}_1(\boldsymbol{x})$ の分子分母を $\exp(\boldsymbol{w}_1^\top \boldsymbol{x})$ で割ると次式のように、シグモイド関数と同じ形になる。

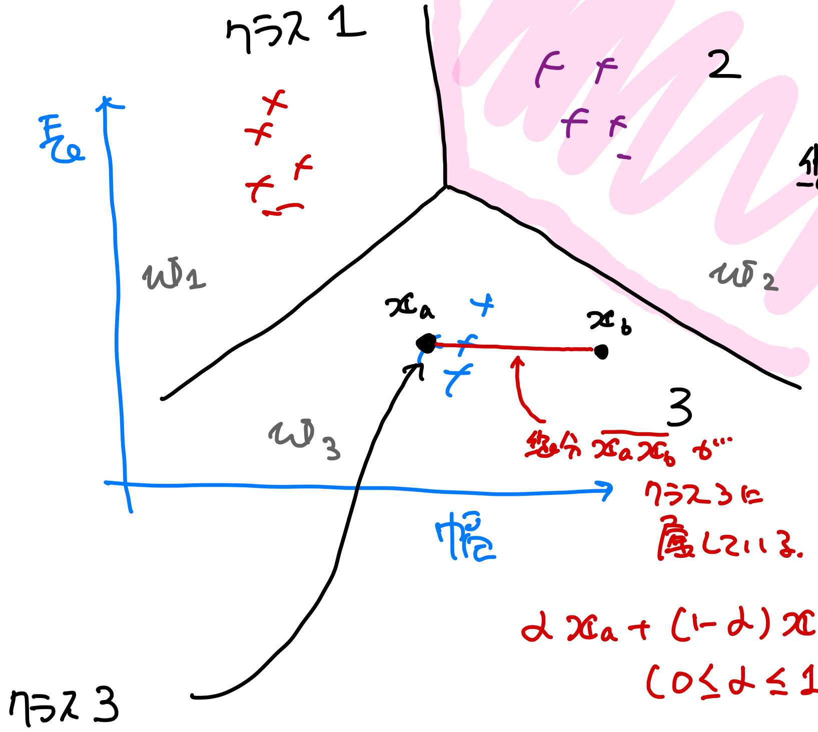
$$p(y=1|\boldsymbol{x}) = \frac{1}{1+\exp(-\boldsymbol{w}^\top \boldsymbol{x})}$$

ただし、$\boldsymbol{w} = \boldsymbol{w}_1 - \boldsymbol{w}_2$ である。同様に次式が成立する

$$p(y=2|\boldsymbol{x}) = \frac{\exp(-\boldsymbol{w}^\top \boldsymbol{x})}{1+\exp(-\boldsymbol{w}^\top \boldsymbol{x})}$$

上記の観点で、ソフトマックス関数はシグモイド関数の自然な拡張である。

問い 前項のソフトマックス関数を使って多クラス識別器を構築した。すなわち、事後確率が最大のクラスを出力する器械を作った。このとき、データ $\boldsymbol{x}$ の空間に占める各クラスの領域が凸になることを証明せよ。凸であるとは、二つのデータ $\boldsymbol{x}_a$ と $\boldsymbol{x}_b$ がいずれもクラス k に属するとき、その 2 つを内挿するデータ $\alpha\boldsymbol{x}_a + (1-\alpha)\boldsymbol{x}_b$ $(0 \leq \alpha \leq 1)$ も必ずクラス k に属することと同値である。



線形識別がつくる
各クラスの領域は凸.

$\alpha x_a + (1-\alpha) x_b$ がクラス3.
$(0 \leq \alpha \leq 1)$

$$P(y=3 \mid x) > P(y=1 \mid x)$$
$$P(y=2 \mid x)$$

$$\frac{\exp(w_3^T x)}{\sum_{k=1}^{3} \exp(w_k^T x)} > \frac{\exp(w_i^T x)}{\sum_{k=1}^{3} \exp(w_k^T x)} \qquad i \neq 3$$

線形.

$$w_3^T x > w_i^T x$$

$$\alpha \times w_3^T x_a > w_i^T x_a \quad (i \neq 3)$$
$$(1-\alpha) \times w_3^T x_b > w_i^T x_b$$

$$w_3^T(\alpha x_a + (1-\alpha) x_b) > w_i^T(\alpha x_a + (1-\alpha) x_b)$$

が $0 \leq \alpha \leq 1$ で必ず成立する //