

chapter 5 性能解析

2021/5/3

パーセプトロンの学習定理

(復習)

- 入力ベクトル $x \in \mathbb{R}^m$ およびパラメータ $w \in \mathbb{R}^m$ に対するパーセプトロンの出力

$$\text{sign}(w^T x) \quad (1)$$

- パラメータの初期値: $w^{(0)} = 0$
- 学習データ $x \in \mathbb{R}^m, y \in \{-1, 1\}$ を受け取り、
 $\text{sign}(w^T x) \neq y$ ならば、次の更新則に従いパラメータを更新する

$$w^{(t+1)} = w^{(t)} + yx \quad (2)$$

パーセプトロンの学習定理

- 性能解析に関連する定義

定義 (学習データのマージン)

学習データ $\{x^{(t)}, y^{(t)}\}_{t=1, \dots, N}$ について、 $y^{(t)} u^T x^{(t)} \geq \gamma$ を満たし $\|u\| = 1$ であるベクトル u が存在するとき、学習データはマージン γ で線形分離可能であるという。

定義 (学習データの半径)

学習データ $\{x^{(t)}, y^{(t)}\}_{t=1, \dots, N}$ の半径を $R = \max_t \|x^{(t)}\|$ とする。

パーセプトロンの学習定理

定理 (パーセプトロンの学習定理)

学習データについて、半径が R でありマージン γ で分離可能であるならば、この学習データに対するパーセプトロンの更新回数は高々 $(R/\gamma)^2$ 回である。

パーセプトロンの学習定理

Proof.

k回目の更新を考える。

$$\begin{aligned}w^{(k+1)}_u &= w^{(k)}_u + y^{(k)}_u T x^{(k)} \\ &\geq w^{(k)}_u + \gamma\end{aligned}$$

ここで、 $w^{(1)} = 0$ より、
 $k = 1$ のとき

$$w^{(2)}_u \geq \gamma$$

$k = 2$ のとき

$$\begin{aligned}w^{(3)}_u &\geq w^{(1)}_u + y^{(1)}_u T x^{(1)} + \gamma \\ &\geq 2\gamma\end{aligned}$$

パーセプトロンの学習定理

Proof.

と再帰的に導くと、

$$w^{(k+1)}_u \geq k\gamma \quad (3)$$

また、 $y^{(k)} w^{(k)T} x^{(k)} < 0, y^2 \|x^{(k)}\|^2 \leq R^2$ より、

$$\begin{aligned} \|w^{(k+1)}\|^2 &= \|w^{(k)} + y^{(k)} x^{(k)}\|^2 \\ &\leq \|w^{(k)}\|^2 + R^2 \end{aligned}$$

ここで、先ほどと同様再帰的に導くと、

$$\|w^{(k+1)}\|^2 \leq kR^2 \quad (4)$$



パーセプトロンの学習定理

Proof.

(3), (4)およびコーシー・シュワルツの定理より、

$$\begin{aligned} kR^2 &\geq \|\mathbf{w}^{(k+1)}\|^2 \|\mathbf{u}\|^2 \\ &\geq \|\mathbf{w}^{(k+1)} \mathbf{u}\|^2 \\ &\geq k^2 \gamma^2 \end{aligned} \tag{5}$$

(5)より、

$$k \leq (R/\gamma)^2 \tag{6}$$



パーセプトロンの学習

- パーセプトロンの更新回数の上限は学習データの半径とマージンの大きさにのみ依存する
- よって、高次元データであってもスパースであれば更新回数を抑えることができる

線形分離可能でない場合のパーセプトロンの学習定理

ref. Freund and Schapire (1999)

定義 (学習データのペナルティ)

$\|u\| = 1$ であるベクトル u と $\gamma > 0$ が
与えられたとき、学習データのペナルティを次のように定義する。

$$d^{(t)} = \max\{0, \gamma - y^{(t)} u^T x^{(t)}\}$$

定義 (学習データのペナルティのノルム)

ペナルティノルムを次のように定義する。

$$D = \sqrt{\sum_{i=1}^n d^{(t)2}} \quad (7)$$

線形分離可能でない場合のパーセプトロンの学習定理

定理 (線形分離可能でない場合のパーセプトロンの学習定理)

学習データの半径を R 、 $\gamma > 0$ のときのペナルティノルムを D とする。このとき、パーセプトロンの更新回数は高々 $(R + D/\gamma)^2$ 回である。

線形分離可能でない場合のパーセプトロンの学習

Proof.

ここでは $D > 0$ の場合を考える。

