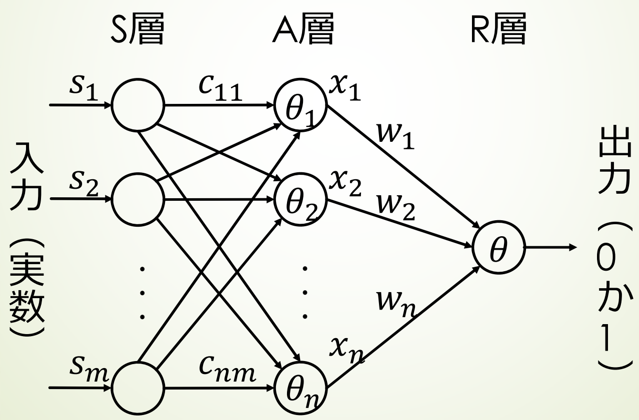
機械学習（3回目）

創域理工学部　情報計算科学科
桂田　浩一

前回の復習（単純パーセプトロン）

- S-A層間は学習する要素がない
  ⇒A-R層間のみを学習



3

前回の復習（学習アルゴリズム（１））

1. 入力パターンベクトル $\boldsymbol{s}_p = (s_{p1}, \cdots, s_{pm})$ と教師信号 t_p $(p = 1, \cdots, P)$ の組を用意する．（P は学習用データ数）

2. 結合荷重 $\boldsymbol{w} = (w_1, \cdots, w_n, w_{n+1})$ の初期値をランダムに小さな値に設定する．さらに学習率 η $(0 < \eta \leq 1)$ を設定する．

4

前回の復習（学習アルゴリズム（２））

3. 学習用データから一つの入力ベクトル $\boldsymbol{s}_p = (s_{p1}, \cdots, s_{pm})$ を選び，$\boldsymbol{s}_p$ に対するA層の各ノードの出力 x_{pj} を次の式で計算する．

$$x_{pj} = \begin{cases} 1 & (\sum_{i=1}^{m} c_{ji} s_{pi} - \theta_j \geq 0) \\ 0 & (\sum_{i=1}^{m} c_{ji} s_{pi} - \theta_j < 0) \end{cases}$$

$\boldsymbol{x}_p = (x_{p1}, \cdots, x_{pn}, 1)$ とする

4. $\boldsymbol{x}_p$ からR層の出力 Out_p を次の式で計算する．

$$Out_p = \begin{cases} 1 & (\boldsymbol{w} \cdot \boldsymbol{x}_p \geq 0) \\ 0 & (\boldsymbol{w} \cdot \boldsymbol{x}_p < 0) \end{cases}$$

5

前回の復習（学習アルゴリズム（3））

5. Out_p と t_p を用いて次の式で $\boldsymbol{w}$ を更新する.

$$\boldsymbol{w} \leftarrow \boldsymbol{w} + \eta(t_p - Out_p)\boldsymbol{x}_p$$

(注)$x_{n+1} = 1$なので$\boldsymbol{x}_p \neq \boldsymbol{0}$

6. 全ての $\boldsymbol{s}_p$ に対して $\boldsymbol{w}$ が変化しなければ終了.
そうでなければ 3.～ 5. を繰り返す.

9/18/2023

6

本日の内容

- 学習可能性について
- 学習できることの証明

9/18/2023

単純パーセプトロンの学習可能性（1）

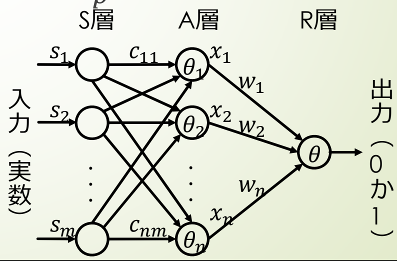
7

- A層の出力 $\boldsymbol{x}_p$ を二つのクラスに分ける.

$$\begin{cases} X^+ \text{: R層から1を出力すべきもの}\ (t_p = 1\text{のもの}) \\ X^- \text{: R層から0を出力すべきもの}\ (t_p = 0\text{のもの}) \end{cases}$$

次に，X^+ に含まれる $\boldsymbol{x}_p$ について $\boldsymbol{w} \cdot \boldsymbol{x}_p > 0$，
X^- に含まれる $\boldsymbol{x}_p$ について $\boldsymbol{w} \cdot \boldsymbol{x}_p < 0$ となる $\boldsymbol{w}$ を考える.

この $\boldsymbol{w}$ は全ての $\boldsymbol{x}_p$ に対して正解を出力する



単純パーセプトロンの学習可能性（2）

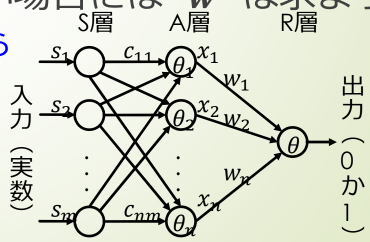
8

- $\boldsymbol{w} \cdot \boldsymbol{x}_p = w_1 x_{p1} + \cdots + w_n x_{pn} + w_{n+1} x_{p(n+1)}$
であり，$\boldsymbol{w} \cdot \boldsymbol{x}_p = 0$ は $\boldsymbol{w}$ を法線ベクトルとする $n+1$ 次元の超平面を表す.

このような $\boldsymbol{w}$ は平面 $\boldsymbol{w} \cdot \boldsymbol{x}_p = 0$ で必ず $\boldsymbol{w} \cdot \boldsymbol{x}_p > 0$ (X^+) と $\boldsymbol{w} \cdot \boldsymbol{x}_p < 0$ (X^-) の領域を分けることから，X^+ と X^- の領域が一つの平面で分けられない場合には $\boldsymbol{w}$ は求まらない.

X^+ と X^- が上のように平面で分けられることを「線形分離可能」という

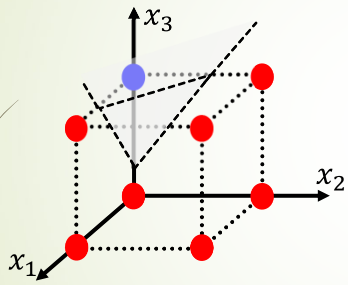
「線形分離可能」でなければ学習できない



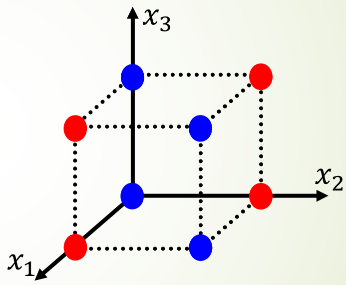
9

単純パーセプトロンの学習可能性（3）

線形分離可能な例



線形分離不可能な例



$X^+ = x_1 \oplus x_2$

X^+

X^-

9/18/2023

$w_1 x_{p1} + \cdots + w_n x_{pn} + w_{n+1} x_{p(n+1)} = 0 \quad \Rightarrow \quad w_1 x_{p1} + \cdots + w_n x_{pn} = \theta$

10

パーセプトロンの収束定理（1）

（定理）X^+ と X^- が線形分離可能であれば

$\boldsymbol{w}_{t+1} = \boldsymbol{w}_t + \eta(t_p - Out_p)\boldsymbol{x}_p$ （tは更新回数）

に従って結合荷重を更新すれば有限回で X^+ と X^- を正しく分離する．

（証明） X^+ と X^- を線形分離する $\boldsymbol{w}^*$ （すなわち正解を与える $\boldsymbol{w}^*$）と学習途中の $\boldsymbol{w}_t$ について，$\boldsymbol{w}^*$ と $\boldsymbol{w}_t$ のなす角を θ とするとき G を次式で定義する． （ただし一般性を失うことなく$\|\boldsymbol{w}^*\| = 1$と正規化されているとする）

$$G = \frac{\boldsymbol{w}^* \cdot \boldsymbol{w}_t}{\|\boldsymbol{w}_t\|} = \cos(\theta) \leq 1$$

9/18/2023

11

パーセプトロンの収束定理（2）

$G = \frac{\boldsymbol{w}^* \cdot \boldsymbol{w}_t}{\|\boldsymbol{w}_t\|} = \cos(\theta) \leq 1$

簡単化のため初期の荷重 $\boldsymbol{w}_0 = \boldsymbol{0}$ と仮定する.

G の分子について，$\boldsymbol{w}_t$ が更新されるとき，

$$\boldsymbol{w}^* \cdot \boldsymbol{w}_{t+1} = \boldsymbol{w}^* \cdot (\boldsymbol{w}_t + \eta(t_p - Out_p)\boldsymbol{x}_p)$$
$$= \boldsymbol{w}^* \cdot \boldsymbol{w}_t + \eta(t_p - Out_p)(\boldsymbol{w}^* \cdot \boldsymbol{x}_p)$$

ここで $\boldsymbol{x}_p \in X^+$ なら $\boldsymbol{w}^* \cdot \boldsymbol{x}_p > 0$ であり，このとき $t_p = 1$，
$Out_p = 0$ となるので $\eta(t_p - Out_p)(\boldsymbol{w}^* \cdot \boldsymbol{x}_p) > 0$ となる.
（$\boldsymbol{x}_p \in X^-$ のときも同様に $\eta(t_p - Out_p)(\boldsymbol{w}^* \cdot \boldsymbol{x}_p) > 0$）

9/18/2023

12

パーセプトロンの収束定理（3）

η，$t_p - Out_p$，$\boldsymbol{w}^* \cdot \boldsymbol{x}_p$ はそれぞれ定数であることから，
$\boldsymbol{w}^* \cdot \boldsymbol{w}_{t+1} \geq \boldsymbol{w}^* \cdot \boldsymbol{w}_t + \delta$ となる δ $(\delta > 0)$ が存在する.

$\boldsymbol{w}_0 = \boldsymbol{0}$ と仮定すると，
$\boldsymbol{w}^* \cdot \boldsymbol{w}_N \geq \boldsymbol{w}^* \cdot \boldsymbol{w}_0 + N\delta = N\delta$

9/18/2023

13

パーセプトロンの収束定理（4）

$G = \frac{\boldsymbol{w}^* \cdot \boldsymbol{w}_t}{\|\boldsymbol{w}_t\|} = \cos(\theta) \leq 1$

簡単化のため初期の荷重 $\boldsymbol{w}_0 = \boldsymbol{0}$ と仮定する.

G の分母の二乗について，$\boldsymbol{w}_t$ が更新されるとき，

$$
\begin{aligned}
\|\boldsymbol{w}_{t+1}\|^2 &= \boldsymbol{w}_{t+1} \cdot \boldsymbol{w}_{t+1} \\
&= \|\boldsymbol{w}_t\|^2 \\
&\quad +2\eta(t_p - Out_p)(\boldsymbol{w}_t \cdot \boldsymbol{x}_p) \\
&\quad +\eta^2(t_p - Out_p)^2\|\boldsymbol{x}_p\|^2
\end{aligned}
$$

9/18/2023

14

パーセプトロンの収束定理（5）

$\|\boldsymbol{w}_{t+1}\|^2 = \|\boldsymbol{w}_t\|^2 + 2\eta(t_p - Out_p)(\boldsymbol{w}_t \cdot \boldsymbol{x}_p) + \eta^2(t_p - Out_p)^2\|\boldsymbol{x}_p\|^2$

ここで $\boldsymbol{x}_p \in X^-$ とすると，$\boldsymbol{w}_t$ が更新されている（出力が誤っている）ことから $t_p = 0$，$Out_p = 1$ であり $\boldsymbol{w}_t \cdot \boldsymbol{x}_p \geq 0$ となる.

よって $2\eta(t_p - Out_p)(\boldsymbol{w}_t \cdot \boldsymbol{x}_p) \leq 0$ （$\boldsymbol{x}_p \in X^+$ も同様）
ここで $\|\boldsymbol{x}_p\|^2$ の最大値を M とすると，

$$2\eta(t_p - Out_p)(\boldsymbol{w}_t \cdot \boldsymbol{x}_p) + \eta^2(t_p - Out_p)^2\|\boldsymbol{x}_p\|^2 \leq M$$

9/18/2023

15

パーセプトロンの収束定理（6）

$\|\boldsymbol{w}_{t+1}\|^2$
$= \|\boldsymbol{w}_t\|^2 + 2\eta(t_p - Out_p)(\boldsymbol{w}_t \cdot \boldsymbol{x}_p) + \eta^2(t_p - Out_p)^2\|\boldsymbol{x}_p\|^2$

以上より $\|\boldsymbol{w}_{t+1}\|^2 \leq \|\boldsymbol{w}_t\|^2 + M$
$\|\boldsymbol{w}_N\|^2 \leq \|\boldsymbol{w}_0\|^2 + NM = NM$

よって $G \geq \frac{N\delta}{\sqrt{NM}}$
$G \leq 1$ より $1 \geq \frac{N\delta}{\sqrt{NM}}$
$N \leq \frac{M}{\delta^2}$ 　　M，δ は定数より，N は有限である

9/18/2023

16

出題予定の演習課題

- パーセプトロンの問題点について

9/18/2023