平成28年度前期
電子情報工学科5年生

人工知能

ニューラルネットワーク

中山　謙二

1

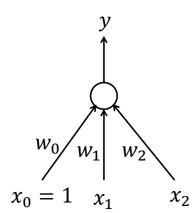
1．パーセプトロン

◆ 特徴

- 1個のニューロンからなるニューラルネットワーク
- 線形分離可能な問題（＊）に適用
  （＊）1個の超平面で分離される問題
  - 2次元データ　→　直線
  - 3次元データ　→　平面
  - 4次元以上のデータ　→　超平面

2

◆ 入出力関係



- 入力ポテンシャル
  $u = w_0x_0 + w_1x_1 + w_2x_2$
  $x_0 = 1$
- 出力信号
  $y = f(u)$
- 非線形関数：$f(\,)$
  (1) ステップ関数
  (2) シグモイド関数
  (3) tanh( )

$$y = f(u) = \begin{cases} 1, u > 0 \\ -1 \; or \; 0 \;\; u < 0 \end{cases} \quad (1)$$

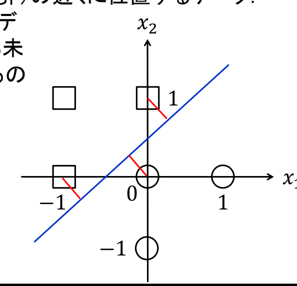
$$y = f(u) = \frac{1}{1 + e^{-u}} \quad (2)$$

$$y = f(u) = \frac{1 - e^{-u}}{1 + e^{-u}} \quad (3)$$

3

◆ 直線（境界）の決め方

- 1本の直線で□と○を分離する．
- 直線は境界データ（＊）からの距離が最大となるように決める．（＊）直線（境界）の近くに位置するデータ．
- 上記の決め方は学習データの回りに分布する未学習データを分類するのに最も有効である．
  （汎化能力）



4

◆ 直線の式について

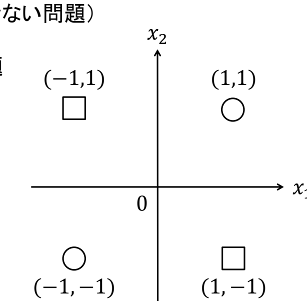
- 入力ポテンシャル
  $u = w_0x_0 + w_1x_1 + w_2x_2, \qquad x_0 = 1$
- 直線（境界線）の式$(u = 0)$
  $$w_0x_0 + w_1x_1 + w_2x_2 = 0$$
  $$x_2 = -\frac{w_1}{w_2}x_1 - \frac{w_0}{w_2}$$
- 直線の式を$x_2 = ax_1 + b$としたときのw_iの計算式
  $w_2 =$ 正の任意の値,　$w_1 = -w_2a, \qquad w_0 = -w_2b$
- $0 < w_2$とすることにより,
  直線（境界線）より上側にあるデータ　→　$0 < u$
  直線（境界線）より下側にあるデータ　→　$0 > u$

5

2．階層形ニューラルネットワーク

パーセプトロンで分離できない問題
（1本の直線で分離できない問題）

線形分離不可能な問題
↓
階層形ニューラル
ネットワーク



6

2層形ニューラルネットワーク

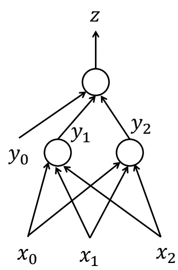
◆ 入出力関係

隠れ層（ニューロン）

$u_1 = w_{10}x_0 + w_{11}x_1 + w_{12}x_2 \cdots (1)$
$u_2 = w_{20}x_0 + w_{21}x_1 + w_{22}x_2 \cdots (2)$
$x_0 = 1$
$y_1 = f(u_1)$
$y_2 = f(u_2)$

出力層（ニューロン）

$v = w_0 y_0 + w_1 y_1 + w_2 y_2 \cdots (3)$
$y_0 = 1$
$z = f(v)$



7

演習問題1

1. 式(1)と(2)の直線($u_1 = u_2 = 0$)と境界データの距離が最大となるようにw_{ji}を決め，直線を図示せよ.
2. 1で求めたw_{ji}を用いたときのy_1, y_2を求めて，プロットせよ（横軸：y_1，縦軸：y_2）. □と○で表示する.
但し，非線形関数は次のステップ関数とする.
$$f(u) = \begin{cases} 1, & u > 0 \\ -1, & u < 0 \end{cases}$$
3. 式(3)の直線($v = 0$)と境界データの距離が最大となるようにw_jを決めよ. また，この直線を図示せよ.
4. 1と3で求めたw_{ji}, w_jを用いたときのzの値を求めよ.

8

階層形ニューラルネットワークの学習

◆ 線形モデルで考える

$$y(n) = w_1(n)x_1(n) + w_2(n)x_2(n)$$
$$e(n) = d(n) - y(n) \qquad d: \text{目標値}$$
$$J(w) = e^2(n) = (d(n) - y(n))^2$$

誤差関数$J(w)$は$w_1(n), w_2(n)$の2次形式（放物曲面）

最急降下法

$$w_i(n+1) = w_i(n) - \frac{\mu}{2}\cdot\frac{\partial J(w)}{\partial w_i(n)}$$

（板書で説明）

9

2層形ニューラルネットワークの学習アルゴリズム

◆ 隠れ層→ w_j →出力層

$$w_j(n+1) = w_j(n) - \frac{\mu}{2}\cdot\frac{\partial J(w)}{\partial w_j(n)}$$

$J(w) = e^2$
$e = d - z \qquad d$：目標値
$z = f(v)$
$v = w_0 y_0 + w_1 y_1 + w_2 y_2, \qquad y_0 = 1$
$$\frac{\partial J(w)}{\partial w_j} = \frac{dJ(w)}{de}\cdot\frac{de}{dz}\cdot\frac{dz}{dv}\cdot\frac{\partial v}{\partial w_j}$$
$$= 2e\cdot(-1)\cdot f'(v)\cdot y_j$$
$$w_j(n+1) = w_j(n) + \mu\cdot e(n)\cdot f'(v(n))\cdot y_j(n)$$

10

◆ 入力層→ w_{ji} →隠れ層

$$w_{ji}(n+1) = w_{ji}(n) - \frac{\mu}{2}\cdot\frac{\partial J(w)}{\partial w_{ji}(n)}$$

$J(w) = e^2$
$e = d - z$
$z = f(v)$
$v = w_0 y_0 + w_1 y_1 + w_2 y_2, \qquad y_0 = 1$
$y_j = f(u_j)$
$u_j = w_{j0}x_0 + w_{j1}x_1 + w_{j2}x_2$
$$\frac{\partial J(w)}{\partial w_{ji}} = \frac{dJ(w)}{de}\cdot\frac{de}{dz}\cdot\frac{dz}{dv}\cdot\frac{\partial v}{\partial y_j}\cdot\frac{\partial y_j}{\partial u_j}\cdot\frac{\partial u_j}{\partial w_{ji}}$$
$$= 2e\cdot(-1)\cdot f'(v)\cdot w_j\cdot f'(u_j)\cdot x_i$$
$$w_{ji}(n+1) = w_{ji}(n) + \mu\cdot e(n)\cdot f'(v(n))\cdot w_j(n)\cdot f'(u_j(n))\cdot x_i(n)$$

11

◆ 非線形関数の微分

➢ シグモイド関数

$$f'(v) = \left(\frac{1}{1+e^{-v}}\right)' = \frac{e^{-v}}{(1+e^{-v})^2}$$
$$= \frac{1}{1+e^{-v}}\left(1 - \frac{1}{1+e^{-v}}\right) = z(1-z)$$

➢ tanh( )

$$f'(v) = \left(\frac{1-e^{-v}}{1+e^{-v}}\right)' = \frac{e^{-v}(1+e^{-v}) + (1-e^{-v})e^{-v}}{(1+e^{-v})^2}$$
$$= \frac{2e^{-v}}{(1+e^{-v})^2} = \frac{1}{2}\left(1 + \frac{1-e^{-v}}{1+e^{-v}}\right)\left(1 - \frac{1-e^{-v}}{1+e^{-v}}\right)$$
$$= \frac{1}{2}(1+z)(1-z)$$

特徴：ニューロンの出力（FF計算：後述）で計算できる.

12

◆ 階層形ニューラルネットワークの学習の流れ

＜ステップ1＞
入力層→隠れ層→出力層（フィードフォワード：FF）におけるニューロンの入力ポテンシャルと出力信号の計算

$$u_1(n) = w_{10}(n)x_0 + w_{11}(n)x_1(n) + w_{12}(n)x_2(n)$$
$$u_2(n) = w_{20}(n)x_0 + w_{21}(n)x_1(n) + w_{22}(n)x_2(n)$$
$$x_0 = 1$$
$$y_1(n) = f(u_1(n))$$
$$y_2(n) = f(u_2(n))$$
$$v(n) = w_0(n)y_0 + w_1(n)y_1(n) + w_2(n)y_2(n)$$
$$y_0 = 1$$
$$z(n) = f\big(v(n)\big)$$
$e(n) = d(n) - z(n)$　　　$d(n) =$ 目標値

13

＜ステップ2＞
出力層→隠れ層→入力層（フィードバック：FB）の順に結合重み（係数）の学習

① $w_j(n)$（隠れ層→出力層）の学習
$$w_j(n+1) = w_j(n) + \mu \cdot e(n) \cdot f'\big(v(n)\big) \cdot y_j(n)$$
上式において，$e(n), y_j(n)$はFF計算で求まっている．
$f'(v(n))$もFF計算で求まった$z(n)$により計算できる．

② $w_{ji}(n)$（入力層→隠れ層）の学習
$$w_{ji}(n+1) = w_{ji}(n) + \mu \cdot e(n) \cdot f'\big(v(n)\big) \cdot w_j(n) \cdot f'\big(u_j(n)\big) \cdot x_i(n)$$
上式においても，必要な値は与えられているか，またはFF計算で求まっている．$e(n) \cdot f'\big(v(n)\big) \cdot w_j(n)$は等価的な誤差→誤差逆伝搬法（Error Back Propagation: BP法）

14

3　リカレント形ニューラルネットワーク

ホップフィールドネットワーク

◆ 特徴

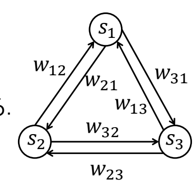
- 結合重みが対象である　　$w_{ji} = w_{ij}$
- 自己ループがない　　　$w_{ii} = 0$
- 状態遷移：1個のニューロンがランダムに選択されて状態を更新する．
- あるエネルギー関数が存在し，ニューロン1個の状態変化において，このエネルギー関数は変わらないか減少する（増加しない）．

15

◆ ニューロンの状態遷移

1個のニューロン(j番目)がランダムに選択され，次式によりその状態(s_j)を更新する．

$$u_j(n) = w_{j1}s_1(n) + w_{j2}s_2(n) + w_{j3}s_3(n), w_{jj} = 0$$
$$s_j(n+1) = \begin{cases} 1, & u_j(n) > 0 \\ s_j(n), & u_j(n) = 0 \\ -1, & u_j(n) < 0 \end{cases}$$



初期状態$s_j(0)$は問題に応じて決める．
（例）
ランダムなパターン
記憶パターン＋雑音

16

◆ 重み係数の設計法

ニューロン数$= N$，ネットワークに記憶するパターンを$\boldsymbol{p}_1 \sim \boldsymbol{p}_M$（記憶パターン数$= M$）とする．$\boldsymbol{p}_k$の要素は$\pm1$である．

$$\boldsymbol{W} = \frac{1}{N}\sum_{k=1}^{M} \boldsymbol{p}_k \cdot \boldsymbol{p}_k^T - \frac{M}{N}\boldsymbol{I}$$

$\boldsymbol{W}$は$N \times N$行列で，j行i列の要素はw_{ji}である．
$\boldsymbol{W}$は対称行列（$w_{ji} = w_{ij}$）であり，かつ，対角要素は零となる（$w_{jj} = 0$）．

17

（例題）
$N = 3, M = 2$

$$\boldsymbol{p}_1 = \begin{bmatrix} 1 \\ -1 \\ 1 \end{bmatrix}, \quad \boldsymbol{p}_2 = \begin{bmatrix} -1 \\ 1 \\ -1 \end{bmatrix}$$

$$\boldsymbol{W} = \frac{1}{3}\begin{bmatrix} 1 \\ -1 \\ 1 \end{bmatrix}[1,-1,1] + \frac{1}{3}\begin{bmatrix} -1 \\ 1 \\ -1 \end{bmatrix}[-1,1,-1] - \frac{2}{3}\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$
$$= \frac{1}{3}\begin{bmatrix} 2,-2,2 \\ -2,2,-2 \\ 2,-2,2 \end{bmatrix} - \frac{2}{3}\begin{bmatrix} 1,0,0 \\ 0,1,0 \\ 0,0,1 \end{bmatrix} = \frac{1}{3}\begin{bmatrix} 0,-2,2 \\ -2,0,-2 \\ 2,-2,0 \end{bmatrix}$$

以上より，
$$w_{12} = w_{21} = -\frac{2}{3}, w_{13} = w_{31} = \frac{2}{3}$$
$$w_{23} = w_{32} = -\frac{2}{3}, w_{jj} = 0$$

18

状態遷移の例

n	u_1	u_2	u_3	s_1	s_2	s_3	
0				1	1	1	初期状態
1	0	(−)	0	1	−1	1	
2	+	−	+	1	−1	1	安定状態（記憶パターンを連想）

n	u_1	u_2	u_3	s_1	s_2	s_3	
0				1	1	−1	初期状態
1	(−)	0	0	−1	1	−1	
2	−	+	−	−1	1	−1	安定状態（記憶パターンを連想）

$$\left\{\begin{matrix}[1,1,1]\\ [-1,-1,1]\\ [1,-1,-1]\end{matrix}\right\}\rightarrow [1,-1,1] \qquad \left\{\begin{matrix}[-1,-1,-1]\\ [-1,1,1]\\ [1,1,-1]\end{matrix}\right\}\rightarrow [-1,1,-1]$$

19

エネルギー関数

$$E(n) = -\frac{1}{2}\sum_{i=1}^{N}\sum_{j=1}^{N} w_{ji}s_i(n)s_j(n), \qquad w_{ii} = 0$$

1個のニューロンの状態変化に対して，$E(n)$は「変化しない」か「減少する」のいずれかであり，増加することはない．
その為，ニューロンの状態遷移を繰り返すことにより$E(n)$は極小値に落ち着く．

$E(n)$の極小値に記憶パターンを対応させる（＊）ことにより連想記憶システムを構築できる．
（＊）前述の重み係数の設計法が一つの方法

20

エネルギー関数の例題（3ニューロンの場合）

$w_{12} = w_{21}, w_{13} = w_{31}, w_{23} = w_{32}$であるから

$$E(n) = -[w_{12}s_1(n)s_2(n) + w_{13}s_1(n)s_3(n) + w_{23}s_2(n)s_3(n)]$$

$s_1(n)$が変化したとする．

$$s_1(n+1) = s_1(n) + \Delta s_1(n+1)$$

$$\begin{aligned}E(n+1) &= -[w_{12}s_1(n)s_2(n) + w_{13}s_1(n)s_3(n) \\ &\quad + w_{23}s_2(n)s_3(n)] \\ &\quad - \Delta s_1(n+1)[w_{12}s_2(n) + w_{13}s_3(n)] \\ &= E(n) + \Delta E(N+1)\end{aligned}$$

$u_1(n) = w_{12}s_2(n) + w_{13}s_3(n)$であるから

$$\Delta E(n+1) = -\Delta s_1(n+1)u_1(n)$$

21

エネルギー関数$E(n)$の変化

$$u_1(n) > 0 \rightarrow \Delta s_1(n+1) \geq 0 \rightarrow \Delta E(n+1) \leq 0$$
$$u_1(n) = 0 \rightarrow \Delta s_1(n+1) = 0 \rightarrow \Delta E(n+1) = 0$$
$$u_1(n) < 0 \rightarrow \Delta s_1(n+1) \leq 0 \rightarrow \Delta E(n+1) \leq 0$$

このように，1個のニューロンの状態変化により生じるエネルギー関数の変化$\Delta E(n+1)$は零または負であり，増加することはない．

22

演習問題2

3ニューロンから成るホップフィールドネットワークにおいて，次の2つのパターンを記憶する場合を考える．

$N = 3, M = 2$

$$\boldsymbol{p}_1 = \begin{bmatrix}1\\1\\-1\end{bmatrix}, \qquad \boldsymbol{p}_2 = \begin{bmatrix}-1\\-1\\1\end{bmatrix}$$

1. 結合重みw_{ji}を求めよ．
2. 全てのパターンを初期状態として，状態遷移の過程とその結果（安定状態）を求めよ．
3. 状態遷移の様子を直方体上に示せ．

23

4. $[-1,1,-1]$と$[1,1,-1]$におけるエネルギーEを求めて，比較せよ．同様に$[1,-1,1]$と$[-1,-1,1]$についてもエネルギーを求めて比較せよ．
5. 2と3において，全てのニューロンの状態を同時に変化させたときの結果を示せ．
（参考）ホップフィールドネットワークとは違う方法であり，ある状態に落ち着くことは保証されていない．

24

演習問題3

入力データ$x_1(n), x_2(n)$に対する出力を$y(n)$とし，目標値を$d(n)$とする．勾配法（最急降下法）により$w_1(n), w_2(n)$を更新する式を求めよ．
式は$w_i(n), x_i(n), y(n), d(n)$を用いて表せ．

$$y(n) = w_1(n)x_1(n) + w_2(n)x_2(n)$$
$$e(n) = d(n) - y(n) \cdots \text{誤差}$$
$$J(w) = e^2(n) \cdots \text{誤差関数}$$
$$w_i(n+1) = w_i(n) - \frac{\mu}{2} \cdot \frac{\partial J(w)}{\partial w_i(n)} \cdots (1)$$
$$\frac{\partial J(w)}{\partial w_j(n)} = \frac{dJ(w)}{de(n)} \cdot \frac{de(n)}{dy(n)} \cdot \frac{dy(n)}{dw_i(n)}$$

25

演習問題4

パーセプトロンにおいて，入力データ$x_1(n), x_2(n)$に対する出力を$y(n)$とし，目標値を$d(n)$とする．

$$u(n) = w_1(n)x_1(n) + w_2(n)x_2(n)$$
$$y(n) = f\big(u(n)\big) = \begin{cases} 1, u(n) \geq 0 \\ -1, u(n) < 0 \end{cases}$$
$$e(n) = d(n) - y(n)$$

$w_1(n), w_2(n)$を次式で更新することにより，誤差$e(n)$が減少することを示せ．

$$w_i(n+1) = w_i(n) + \mu e(n) x_i(n), 0 < \mu \cdots (*)$$

（場合分け）
(1) $e(n) > 0, x_i(n) > 0$　(3) $e(n) < 0, x_i(n) > 0$
(2) $e(n) > 0, x_i(n) < 0$　(4) $e(n) < 0, x_i(n) < 0$

26

演習問題5

ホップフィールドネットワークによる連想記憶に関して以下の問に答えよ．但し，ニューロン数を4個とする．Excelプログラムを用いて計算すること．

1. ４個の記憶パターン$p_1 \sim p_4$を決定せよ（次頁に例題）．
2. 記憶パターンが安定度を求めよ．
   安定度＝安定である記憶パターン数／４
3. １６通りの初期パターンに対する収束パターンを求めよ．
4. 収束パターンが記憶パターンである場合は「連想成功」，そうでない場合は「不成功」とし，次の成功率を求めよ．
   成功率＝（成功した初期パターン数）／１６
5. 記憶パターンを変えて２～４を繰り返せ（３通り）．
6. 上記の結果から，記憶パターンの組合せと安定度，成功率の関係について考察せよ（ハミング距離が関係）．

29