2004 年 12 月 22 日

知能型システム論: 多層パーセプトロンの動作

喜多 一

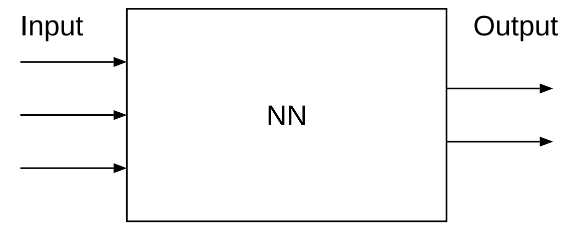

図 1　ニューラルネットワークによる情報処理

1　ニューラルネットワークによる情報処理

動物の脳・神経系を模したネットワークをニューラルネットワーク (Neural Network, 以下 NN と略す) と呼び，これを用いた情報処理である． その構成は以下のようなものである：

- 神経細胞 (neuron) を模した非線形特性を有する処理ユニットを相互に接続し，
- その一部に外部からの入力を与え，一部の処理ユニットの出力をシステムの出力として利用する，
- 処理ユニット間の結合が神経細胞間のシナプス結合に相当し，シナプス伝達効率に相当する結合重みが NN の可調整パラメータとなる．
- このパラメータを例題に基づいて調整 (学習) し，求められる情報処理機能を獲得させる．

NN による情報処理の特性として

- アナログ的な情報処理

- 並列的な情報処理
- 学習による機能獲得

などが挙げられる．

NN はその処理ユニットやネットワークの構成 (アーキテクチャ)，学習アルゴリズムの構成により種々のものが提案されているが，多層パーセプトロン (Multi-Layer Perceptron, MLP) や自己組織化マップ (Self-Ogranizing Map) はその代表例である．

2　多層パーセプトロン

2.1　処理ユニットのモデル

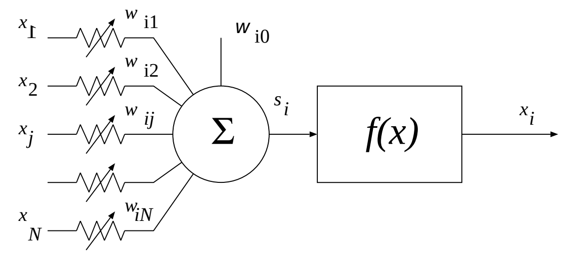

図 2　処理ユニット

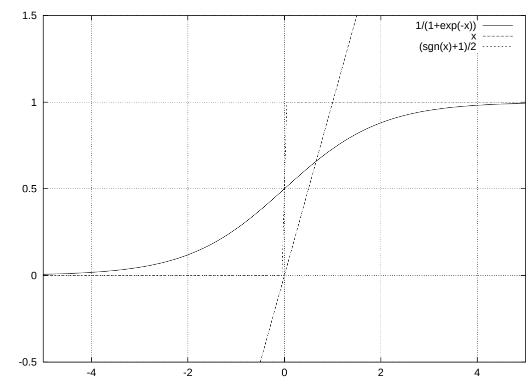

図 3　ユニットの特性

今，N 個の処理ユニットを考え，その出力を x_i とする．多層パーセプトロンで用いる処理ユニットでは以下のように他のユニットの出力の加重和に対して非線形の処理を行う：

$$s_i = \sum_{j=1}^{N} w_{ij} x_j + w_{i0} \tag{1}$$

$$x_i = f(s_i) \tag{2}$$

ここで，s_i は処理ユニット内での中間的出力 (神経細胞の膜電位に相当する)．また，非線形関数 $f(x)$ の特性として，よく用いられるものに

$$\text{シグモイド関数} \qquad f(x) = \frac{1}{1+\exp(-x)} \tag{3}$$

がある．これは

$$\text{ステップ関数} \qquad f(x) = \frac{\mathrm{sgn}(x)+1}{2} = \begin{cases} 0 & \text{for} \quad x < 0 \\ 1 & \text{for} \quad x \geq 0 \end{cases} \tag{4}$$

を微分可能な関数で近似したものである *．このほか，(5) のような線形関数が用いら

* これにより次回紹介する誤差逆伝播学習が可能になる．

れる．

$$f(x) = x \tag{5}$$

なお，(1) 式は恒等的に出力 1 を出す仮想のユニット x_0 を考えれば形式的に定数項を省いた表現が行える：

$$s_i = \sum_{j=0}^{N} w_{ij} x_j \tag{6}$$

2.2　多層パーセプトロンの構成と動作

多層パーセプトロン (multi-layer perceptron, MLP) は 2.1 で述べた処理ユニットを入力側から出力側へと信号の流れが 1 方向になるように結合したネットワークである.

- 入力信号を受取り「そのままユニットの出力」とするユニットを**入力ユニット (input unit)**,
- ネットワークの出力とするユニットを**出力ユニット (output unit)** と呼び,
- 直接は外部と入出力を行わず，入力ユニットと出力ユニットの間に介在する**ユニットを隠れユニット (hidden unit)** と呼ぶ.

なお，ユニット群を「層」として考えて，それぞれ入力層，出力層，隠れ層と呼ぶ．多層パーセプトロンでは隠れユニットの特性としてシグモイド特性を用いることにより，複雑な入出力関係が表現できる.

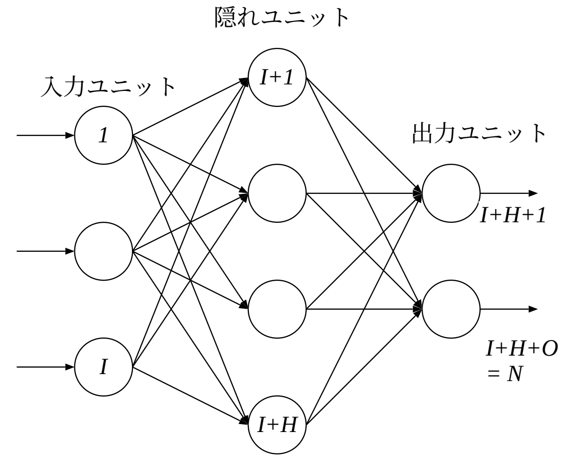

図 4　多層パーセプトロンの構成

すべてのユニットの特性が線形であるならば，多層化しても線形変換の繰り返しは線形変換にすぎないことから，実現できる特性は限られている.

多層パーセプトロンの構造は入力側から出力側に順にユニット番号を与えるとすれば，ユニット番号の大きいユニットからそれ以下のユニットへの結合が無いこと，すなわち

$$w_{ij} = 0 \quad \text{for} \quad j \geq i \tag{7}$$

と表現できる．今，

- 入力ユニットの添字を $1, \cdots, I$,
- 隠れユニットの添字を $I+1, \cdots, I+H$,
- 出力ユニットの添字を $I+H+1, \cdots, I+H+O(=N)$ とすると，

とすると，多層パーセプトロンの動作は次のように計算できる

1. 入力の値を入力ユニットの出力 $x_i, i = 1, \cdots, I$ に設定する．
2. 隠れユニット，出力ユニットの値をユニット番号の小さいものから順次，(1) ∼ (4) 式に従って計算する．
3. 出力ユニットの出力 $x_i, i = I+H+1, \cdots, N$ をネットワークの出力とする．

w_{ij} を行列と考えれば，これは行列 $W = \{w_{ij}\}$ が下三角行列となることである．隣接層間のみに結合を考える場合はさらに，ブロック構造になる．

自分より番号の大きいユニットの出力は $w_{ij} = 0 \quad \text{for} \quad j \geq i$ より影響しない．このために，各ユニットの出力が計算可能となる．

2.3 多層パーセプトロンの表現能力

多層パーセプトロンについては

- (少なくとも) 隠れユニットを 1 層持つ 3 層構造とし
- 隠れユニットの非線形特性としてシグモイド関数を採用し,
- 出力ユニットには出力制限のない線形関数を採用し,
- 十分な数の隠れユニットを用いれば,

任意の連続関数を任意の精度で近似できることが知られている.

ただし，この性質は「存在定理」であり，与えられた関数を近似する結合重みを「学習により獲得できること」は「保証していない」ことに注意を要する.

これと関連して，入出力を 2 値 $\{0, 1\}$ としたとき，少なくとも 3 層構造を持ち，隠れユニット，出力ユニットの特性をステップ関数とし，十分な数の隠れユニットを用いれば任意の論理関数が実現できることも知られている.

演習

1. 図 4 のような構成の NN に表 1 のような入力を与えた場合について，隠れユニット，出力ユニットについて，それぞれ s_i, x_i を求めて表を完成させよ．ただし，
   - ユニットの円内の数字がユニット番号である．
   - ユニットの非線形特性は隠れユニット，出力ユニットともにステップ関数とする．
   - ユニット間の接続の横に書かれた数字が結合重みの値である．定数項はユニットへの線の横に書かれたものである．
2. この NN はどのような論理関数を実現したものとなっているか．
3. 処理ユニットの特性を線形 $f(x) = x$ とした場合，この論理関数は実現可能か?

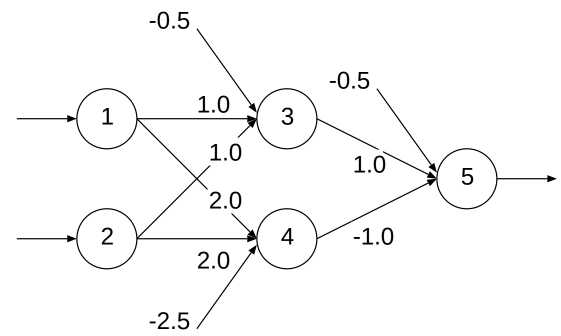

図 5　演習課題ネットワーク

表 1　提示パターン

パターン	x_1	x_2	s_3	x_3	s_4	x_4	s_5	x_5
1	0	0						
2	0	1						
3	1	0						
4	1	1						