ニューラルネットワーク情報処理

神戸大学自然科学研究科知能科学専攻人工知能講座　村尾　元

平成5年8月10日(金)

1　脳の情報処理

1.1　解析と実現における2つの流れ

マクロな側面 論理学的・認知心理学的な視点から、脳を高次の情報処理機構として解析を行なう立場である。つまり記号を用いた逐次直列方式として脳の思考過程を解析/具現しようというもので、エキスパートシステムの推論エンジンなどはその過程において工学的に成功した代表的な例と言える。

ミクロな側面 生理学的な視点から、脳を分子機械として解析を行なう立場である。つまり、**神経細胞(ニューロン)**や更にはDNAレベルでの情報処理を解析し、それらから脳の情報処理を捕えようというものである。**人工ニューラルネットワーク**(Artificial Neural Network: しばしばANNと略される)はその工学的応用の代表的な例である。

1.2　生体ニューロン

ニューロンは図1に模式的に示すような構造をしている。複数のニューロンは**軸索**と**樹状突起**の**シナプス結合**により結合しており、ニューラルネットワークを構成する。一つのニューロンは複数のニューロンよりシナプス結合を通じて樹状突起に電気的信号を受ける。複数の信号の総和がある閾値を越えた時、そのニューロンは軸索を通じて他のニューロンにパルス信号を送る。

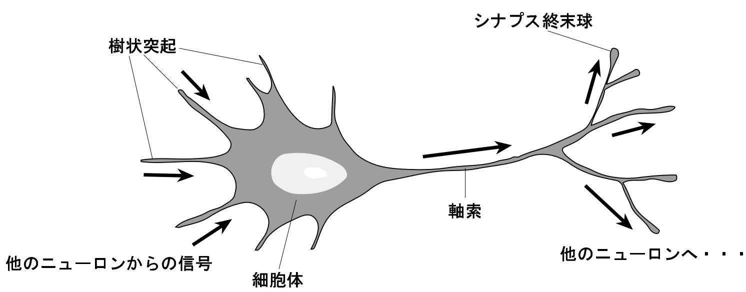


図1 生体ニューロンの構造

1.3　ニューラルネットワークとしての脳

ニューロンは人間の場合には50 種類程あり、その大きさや形状など細かな差異はあるとはいえ、その動作はほぼ同じ原理に従っている。脳は約 10^{10} 個という膨大な数のニューロンからできているが、やはりニューロン単位での構造的な差異は少ない。にも関わらず、ヒトの脳は図 2に示すように幾つかの機能単位に分かれている。つまり、ニューロンが複雑に結合したニューラルネットワークという構造レベルにおいての差異により機能の分化が現れると考えられる。また各機能単位は逐次的に動作し、それぞれの処理をこなしていくため、表面上は直列式の論理型思考を行なっているように見えるが、その実、多数のニューロンが独立に動作する並列処理によって機能している。

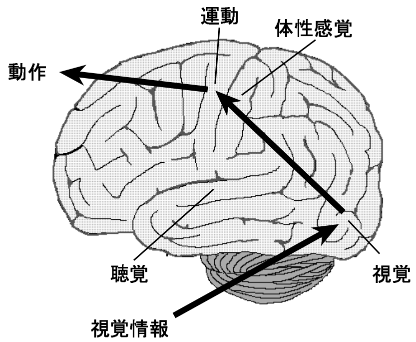


図 2　脳の機能的構造

1.4　学習する情報処理機械

ヒトの学習能力は脳に司られていると言っても過言ではないが、ニューラルネットワークとして脳を見た場合、大きく分けて二つの学習形態があると考えられている。

1. **シナプス結合レベルの学習**　ヒトの短期記憶は、ニューロン間のシナプス結合の強度の変化が関係していると考えられている。ニューラルネットワークにおける学習の一つの形態がこれである。シナプス結合は比較的容易に変化させることが出来、過った学習の修正が容易である。

2. **構造レベルの学習**　生体においては、ニューラルネットワークの構造も外的要因により変化する。これはヒトの長期記憶の要因の一つと考えられている。

2　人工ニューラルネットワーク

2.1　ニューロンの数理モデル

ニューロンを情報処理素子と見れば、図 3に示すような、多入力 1 出力の素子として表すことが出来る。多くのニューロンでは入力信号 $x_1, \ldots, x_n$ 及び出力 y は [0, 1] の間の数値を

取り、シナプス結合強度に相当する結合荷重 $w_1, \ldots, w_n$ 及び閾値 h は適当な実数値を取る。

ニューロンは、他のニューロンより重み付き入力信号 $w_i x_i$ を受け取り、その総和より閾値 h を減じた値を応答関数 f により変形し出力する。応答関数 f をステップ関数とした場合には、重み付き入力信号の総和が閾値を越えた場合に 1 を、それ以外では 0 を出力し、生体ニューロンに類似した動作を行なう。応答関数 f をシグモイド関数とした場合には生体ニューロンのパルス頻度に類似した入出力応答を得られる。

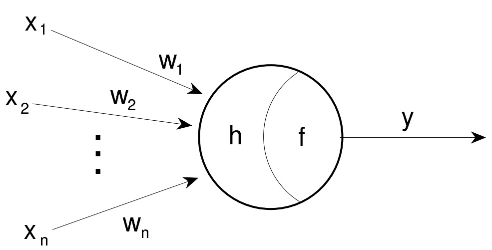


図 3　ニューロンの数理モデル

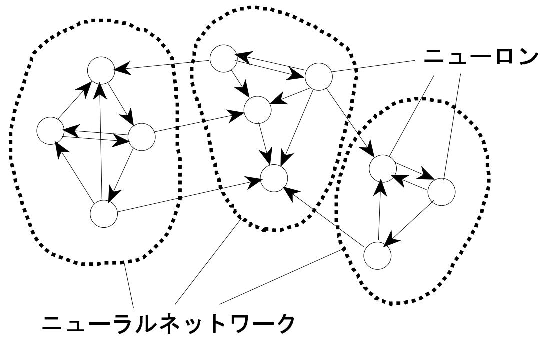


図 4　人工ニューラルネットワークモデル

2.2　ニューラルネットワークモデル

脳内には約 10^{10} 個のニューロンが存在しているが、全てのニューロンが相互に結合している訳ではない。図 2 に示される機能単位に対応したモジュール構造になっていたり、下位のニューラルネットワークから上位のニューラルネットワークに信号が伝播していく階層構造になっていたりする。このような脳を手本とするならば、人工ニューラルネットワークの一般的なモデルとしては、図 4 に示されるようなニューラルネットワークが妥当であると考えられる。実際に提案されているニューラルネットワークモデルでは、このような複雑なモデルはまだ少なく、比較的シンプルで単一の処理を司る形態の物が多い。人工ニューラルネットワークに対する一般的な問題として以下のような物が考えられる。

結合形態 ニューロン同士をどのように結合するかは基本的な問題である。考えられる要因としてはモジュール構造・階層構造・フィードバック結合などがある。

同期の問題 脳においてはニューロンは非同期に動作しているが、人工ニューラルネットワーク特に計算機上でシミュレートされた物では同期をどのように取るかがネットワークの特性にも影響を与える。多くの人工ニューラルネットワークでは、全てのニューロンが同期して動作するものとされている。

ニューロンの特性 脳内には幾つかの特性の異なるニューロンが存在する。しかし、人工ニューラルネットワークにおいては、問題の簡単化の為にも、単一のニューロンを用いてネットワークを構築することが多い。

2.3　ニューラルネットワークの学習

現在の人工ニューラルネットワークにおいては、ニューラルネットワークの構造つまりニューロン間の結合の有無を学習することは困難であり、実際提案されているものも非常に少ない。ほとんど全ての典型的ニューラルネットワークでは構造は学習せず、ニューロン間の結合荷重のみの学習を行なう。結合荷重の学習方法には大きく分けて次のような 2 種類の物がある。

教師あり学習 ニューラルネットワークの出力と、その目標 (**教師信号**と呼ばれる) を比較し、その結果からニューロン間の結合荷重を更新する。代表的な物にバックプロパゲーション学習法がある。

教師無し学習 ニューラルネットワークの出力を自己評価し、その結果からニューロン間の結合荷重を更新する。代表的な物に自己組織化マップがある。

3　多層フィードフォワード型ニューラルネットワーク

3.1　ニューラルネットワークの構造

本章では図 5 に示すような多層フィードフォワード型ニューラルネットワークについてその学習方法及び応用例について述べる。この型のニューラルネットワークは、同じ性質を持ったニューロンにより 1 つの層を形成し、各層を 1 方向に結合したものである。一般的には層内のニューロン同士の結合や逆方向の結合 (フィードバック結合) は考えない。ニューラルネットワークは入力層から入力信号を受け取り、それを隠れ層に伝播し、出力層より出力信号を出力する。

学習方法としては Hebb 則を用いた教師無し学習やバックプロパゲーション法による教師あり学習がある。次節では、この形式のニューラルネットワークの学習法として一般的な、バックプロパゲーション法について簡単に説明する。

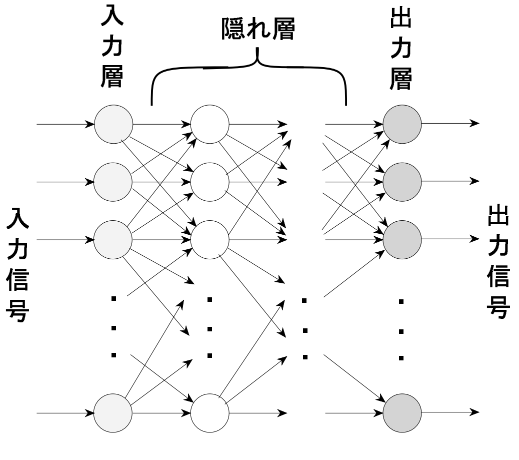


図 5　多層フィードフォワード型ニューラルネットワーク

3.2　バックプロパゲーション法

バックプロパゲーション法(Back Propagation)は、Rumelhart らにより 1986 年に提案された多層フィードフォワード型ニューラルネットワークに対する教師あり学習法である。この方法を用いることにより、単純な動作と教示の繰り返しによって非線形の入出力変換をニューラルネットワークに学習させせることができる。ここでは図 6に示すような、3 層のニューラルネットワークを例にバックプロパゲーション法について説明する。

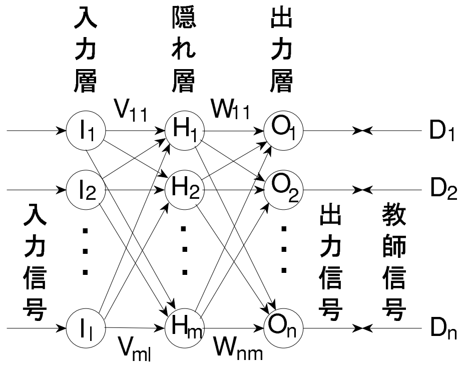


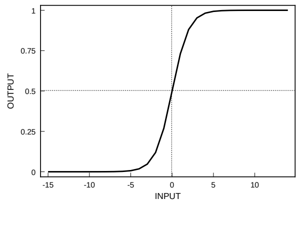
図 6　3 層フィードフォワード型ニューラルネットワーク

各層のニューロンは次のようなダイナミクスで動作する。

$$\begin{cases} H_j(t) = f(\sum_i^l V_{ji}(t)I_i(t) \quad \theta_j(t)) \\ O_k(t) = f(\sum_j^m W_{kj}(t)H_j(t) \quad \phi_k(t)) \end{cases}$$

$I(t)$, $H(t)$, $O(t)$ はそれぞれ時刻 t における入力ニューロン、隠れニューロン、出力ニューロンの出力で、$V(t)$, $W(t)$ はそれぞれ各層間の結合荷重、$\theta(t)$, $\phi(t)$ は閾値である。入力ニューロンは入力層への入力信号を隠れ層の各ニューロンへ分配するだけのファンアウトニューロンである。ニューロンの入出力応答関数 $f()$ は

$$f(x) = \frac{1}{1 + \exp(\quad Tx)}$$



で表される**シグモイド関数**である。Tはシグモイド関数の傾きを決定する定数で、生体ニューロンの場合の温度に対するアナロジーとなっている。

さて、出力ニューロンに対しては、出力の目標として教師信号 $D(t)$ が与えられる。ある入力 p に対するニューラルネットワークの出力の誤差 $E_p(t)$ を

$$E_p(t) = \frac{1}{2}\sum_{k}^{n}(D_k(t) \quad O_k(t))^2$$

とする。最急降下法を用いると、出力誤差 $E_p(t)$ を最小にする為の、時刻 t における結合荷重 $V(t), W(t)$ の更新量は以下のように表される。

$$\begin{cases} \Delta_p V_{ji}(t) = \quad \alpha \dfrac{\delta E_p(t)}{\delta V_{ji}(t)} \\ \Delta_p W_{kj}(t) = \quad \beta \dfrac{\delta E_p(t)}{\delta W_{kj}(t)} \end{cases}$$

ここでα, βは $0 < \alpha, \beta \leq 1$ となる定数である。これを計算すると

$$\begin{cases} \Delta_p V_{ji}(t) = \quad \dfrac{\beta T}{\alpha}(1 \quad H_j(t))I_i(t)\displaystyle\sum_k \Delta W_{jk} W_{jk} \\ \Delta_p W_{kj}(t) = \quad \beta T(D_k(t) \quad O_k(t))O_k(t)(1 \quad O_k(t))H_j(t) \end{cases}$$

となる。これを用いて次の時刻の結合荷重を求める。

$$\begin{cases} V_{ji}(t+1) = V_{ji}(t) + \Delta V_{ji}(t) \\ W_{kj}(t+1) = W_{kj}(t) + \Delta W_{kj}(t) \end{cases}$$

3.3 バックプロパゲーション法の応用

バックプロパゲーション法を用いて学習を行なう多層フィードフォワード型ニューラルネットワークは、今日最も利用されているニューラルネットワークであり、その応用例も多い。特に手書き文字認識を始めとするパターン認識、画像処理、制御工学への応用例は多い。ここでは幾つかの例を挙げる。

1 **気象衛星画像を用いた降水量推定**

気象衛星ひまわりの赤外画像データを入力、その直下の地域における降水量を出力とする 3 層フィードフォワードニューラルネットワーク (図 7) を構成し、バックプロパゲーション法により学習を行なう。

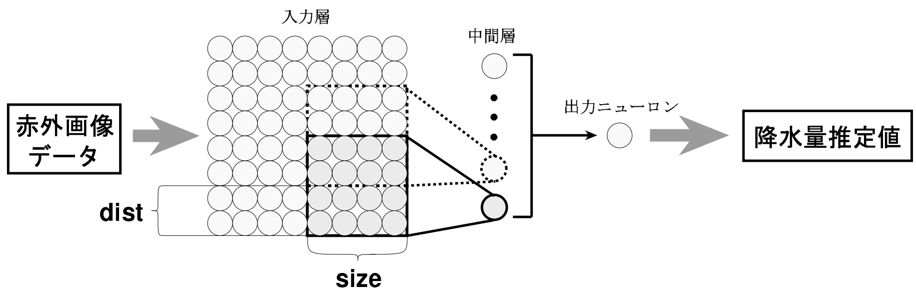


図 7 降水量推定に用いるニューラルネットワーク

学習後のニューラルネットワークによるテスト結果を図 8 に示す。実線が実際の降水量で点線がニューラルネットワークによる出力である。

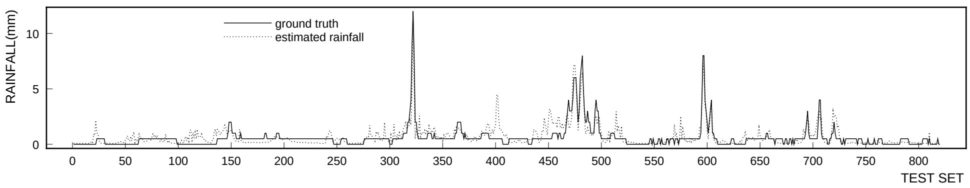


図 8　気象衛星画像による降水量推定結果

2　マニュピレータ手先の位置制御

図 9 に示すような 3 層フィードフォワード型ニューラルネットワークにより、システム(ここではマニュピレータ) の順システム(↔逆システム) を学習する。学習のブロック図は図 10 に示す。

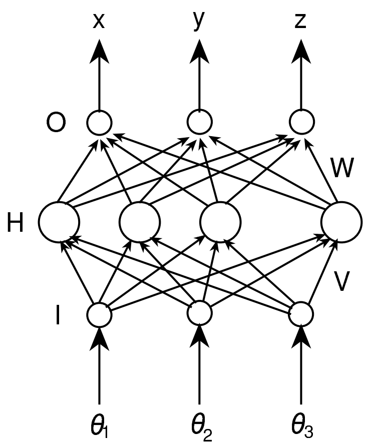


図 9　マニュピレータの順システムを学習するネットワーク

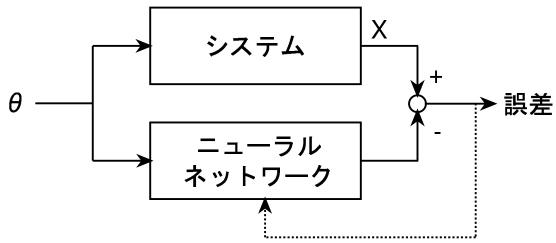


図 10　順システムの学習

十分に学習した後にニューラルネットワークより、以下のような式を用いてヤコビアンを求め、システムを制御する。

$$\frac{\delta y}{\delta \theta_1} = f_y'(x_y) \sum_i W_{yi} f_i'(x_i) V_{i1}$$

ここで、$f()$ はニューロンの応答関数、x はニューロンへの入力の重み付き和、W_{yi} は隠れ層と出力層間の結合荷重、V_{i1} は入力層と隠れ層間の結合荷重である。

4 ホップフィールド型ニューラルネットワーク

4.1 ニューラルネットワークの構造

J J Hopfield が1982 年に提案したホップフィールド型ニューラルネットワークは、図 11 に示すようなニューラルネットワークであり、以下のような構造上の特徴を持つ。

- 対称的な相互結合
  全てのニューロンは自分を除く他の全てのニューロンと結合しており、その結合荷重は対称 ($w_{ij} = w_{ji}$) となっている。
- ニューロンの非同期動作
  全てのニューロンは以下のダイナミクスに従って、独立に並列的に動作している。

$$
\begin{aligned}
v_i(t) &= f(x_i(t)) \\
x_i(t) &= \sum_{i \neq j} w_{ij} v_j(t) + \theta_i
\end{aligned}
$$

ここで、$v_i(t)$ は時刻 t におけるニューロン i の出力、$w_{ij}(= w_{ji})$ はニューロン i と j の間の結合荷重、θ_i は閾値である。入出力応答関数 $f()$ は一般的にはステップ関数を利用する。シミュレーションの際には、各ニューロンの動作を実際に並列計算機上に並列にインプリメントすると良い。逐次計算を行なう場合には、乱数でニューロンを 1 つ選び更新するという動作を繰り返せば良い。

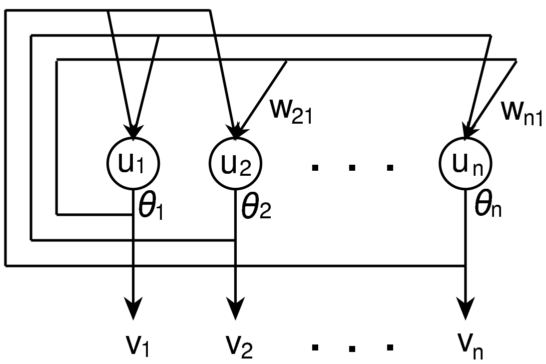


図 11 ホップフィールド型ニューラルネットワーク

4.2 エネルギー最小値探索

ホップフィールド型ニューラルネットワークは全く学習を行なわないニューラルネットワークであり、初期に設定した結合荷重は変化することがない。ホップフィールド型ニュー

ラルネットワークは、ニューロンの出力と結合荷重によって

$$E(t) = \quad \frac{1}{2}\sum_{i \neq j} w_{ij} v_i(t) v_j(t) \quad \sum_i \theta_i v_i(t)$$

ように定義されるエネルギーの最小値を探索するニューラルネットワークである。ほとんどの組み合わせ最適化問題は、多項式の最小化問題に変換でき、さらに多項式よりエネルギーを導き出すことにより、ホップフィールド型ニューラルネットワークで解を導き出すことができる。例えば、巡回セールスマン問題[1]などが代表的な例として取り扱われている。

4.3 応用例:文字の連想記憶

ホップフィールド型ニューラルネットワークを連想記憶として用いる。つまり、幾つかのパターンを結合荷重に埋め込んでおき、入力パターンに最も近い記憶パターンを出力させようというものである。

今、記憶させたい M 個のパターンの m 番目のパターンの、i 番目のニューロンに対応する値を v_i^m とし、このときのニューラルネットワークの結合荷重を

$$w_{ij} = w_{ji} = \sum_{m=1}^{M} (2v_i^m \quad 1)(2v_j^m \quad 1)$$

とする。ある制約条件の元では、これで定義された結合荷重により、記憶された M 個のパターンがエネルギー関数の最小値の場合にあたることが証明できる。これによりパターンがホップフィールド型ニューラルネットワークに記憶される。

さて、ネットワークが 16 個の 2 値のニューロンからなっており、それらが 4×4 の正方形に並べられているとする。ここに 3 つのパターンを記憶させ、それを適当な入力より連想させた場合の結果を図 12に示す。

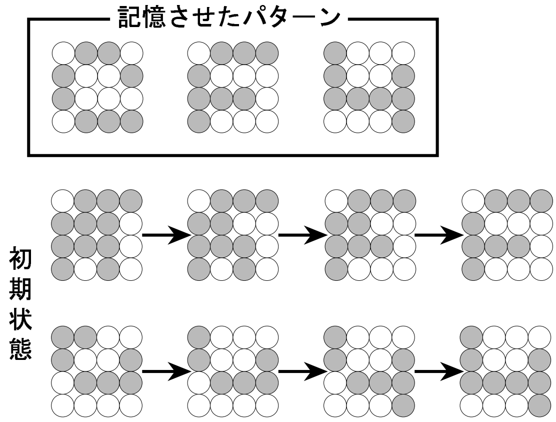


図 12　パターンの連想記憶

[1] 幾つかの都市に対して、それらをそれぞれ 1 度だけ訪れて、その行程が最短になるような都市の訪問順を求める問題

5 自己組織化マップ (Self-Organizing Map)

5.1 自己組織化マップの構造

自己組織化マップは Kohonen により 1982 年に提案された、ニューロン間の**側抑制**結合と Hebb 則による学習を統合した、教師無し学習を行なうニューラルネットワークで、図 13 に示すように 2 次元格子状に並んだニューロンにより構成されている。各ニューロンは内部に入力ベクトルと同次元の参照ベクトルを持っている。入力ベクトルが自己組織化マップに与えられると、全ての参照ベクトルが入力ベクトルと比較され、最も近い参照ベクトルを有するただ一つのニューロンが発火 (1 を出力) する。

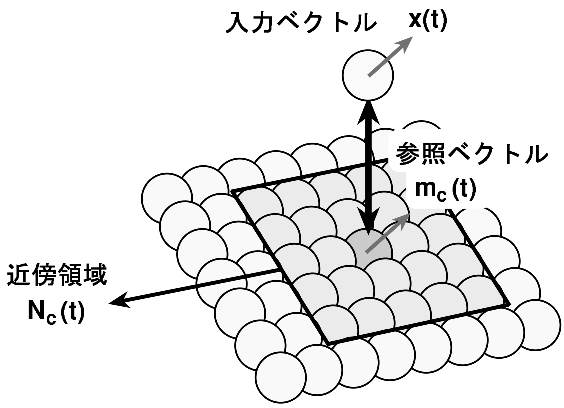


図 13 自己組織化マップ

5.2 自己組織化マップの学習

自己組織化マップの学習は、入力ベクトルに最も近い内部ベクトルを入力ベクトルにさらに近づけることで行われる。すなわち、$\boldsymbol{x}(t)$ を時刻 t における入力ベクトル、$\boldsymbol{m}_i(t)$ $(i = 1, 2, \ldots, m)$ を番目のニューロンの内部ベクトル (m はニューロンの数) とすると、

$$\|\boldsymbol{x}(t) - \boldsymbol{m}_c(t)\| \equiv \min_i \|\boldsymbol{x}(t) - \boldsymbol{m}_i(t)\|$$

となるニューロン c を選び、その内部ベクトル$\boldsymbol{m}_c$を以下の式に従って更新する。このとき更に次式により、選ばれたニューロン c だけでなく、その 2 次元平面上における近傍のニューロン群も同様に更新するため、入力ベクトル空間で近いベクトル同士が 2 次元平面上でも近くなるような写像が実現される。

$$\boldsymbol{m}_i(t+1) = \begin{cases} \boldsymbol{m}_i(t) + \alpha(t)[\boldsymbol{x}(t) - \boldsymbol{m}_i(t)], & \text{for } i \in N_c(t) \\ \boldsymbol{m}_i(t), & \text{for } i \notin N_c(t) \end{cases}$$

ここで$\|\quad\|$は入力ベクトルの次元での距離を示し、$\alpha(t)$ は $0 < \alpha(t) < 1$ を満たす単調減少関数で表される学習係数を表す。$N_c(t)$ はニューロン c を中心とする 2 次元平面上での近傍領域を表し、この範囲も学習と共に減少させる。

この結果、m 個のニューロンは入力ベクトル空間の確率密度に従って分布する。すなわち、入力ベクトルの密なところにはより多くのニューロンが対応し、粗な部分に対応するニューロンは少なくなる。

5.3 自己組織化マップの応用例

自己組織化マップの動作を知るために、2 次元の入力ベクトルデータに対して学習した場合を示す。入力ベクトルは 3 個の正規分布からなり、それぞれのデータ数は 5000,3000,2000 となっている。シミュレーションに用いた自己組織化マップは 10×10 のニューロンからなり、学習回数は各データ毎に 100 回となっている。

図 14で、点は入力データを表している。太い線で結ばれた格子は、各交点がニューロンの参照ベクトルを表し距離が 1 となるニューロン同士がリンクで結ばれている。これはニューロンの近傍関係を示すために、良く用いられる方法である。

学習の結果構成されたマップを見るとデータ数の多いところ、つまり確率密度の高いところにニューロンの参照ベクトルがより多く配置されていることがわかる。また、近傍関係が保存され入力ベクトル空間で近いものは参照ベクトル空間でも近い位置に写像されていることがわかる。

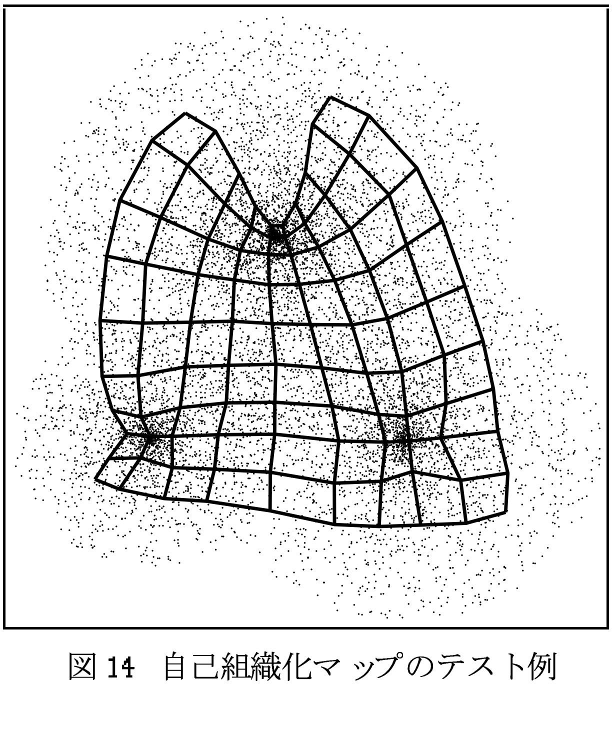


図 14 自己組織化マップのテスト例

6　リカレントニューラルネットワーク

ニューロン間にフィードバック結合を含む場合、そのニューラルネットワークをリカレントニューラルネットワークと呼ぶ。先に述べたホップフィールド型ニューラルネットワークもリカレントニューラルネットワークの特殊な形態である。本章では更に幾つかの代表的なリカレントニューラルネットワークとその学習法を紹介する。

6.1　Elman 型ニューラルネットワーク

図 15 に示すように、3 層フィードフォワード型ニューラルネットワークに文脈層という層を付加し、隠れ層から文脈層に対してフィードバック結合を持つニューラルネットワークを Elman 型ニューラルネットワークと呼ぶ。

文脈層は 1 時刻前の隠れ層の状態を保持し、それにより時系列的な入力に対してのパターン認識能力に優れている。学習は順方向の結合荷重についてのみ行なわれるので、バックプロパゲーション法がそのまま用いられる。

同じような形式のリカレントニューラルネットワークに Jordan 型と呼ばれるものがある。これは中間層ではなく、出力層のコピーを文脈層に保持するニューラルネットワークで、学習法は Elman 型と同じくバックプロパゲーション法を用いる。

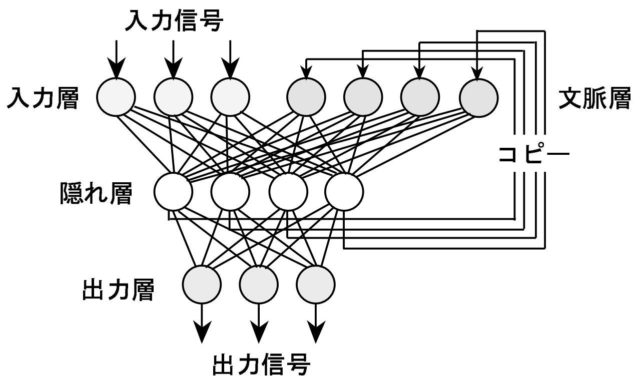


図 15　Elman 型ニューラルネットワーク

6.2　フルリカレントニューラルネットワーク

6.2.1　ネットワークのダイナミクス

全てのニューロンが、それ自身を含む全てのニューロンと相互に結合しているニューラルネットワークをフルリカレントニューラルネットワークと呼ぶ。これはニューラルネットワークの一般形であり、今迄述べてきたニューラルネットワークもフルリカレントニューラルネットワークの特殊な形と言える。

フルリカレントニューラルネットワークにおいてはニューロンは以下のダイナミクスに従う。

$$\begin{cases} x_i(t+1) = \displaystyle\sum_{j \in \mathrm{all\ neurons}} w_{ij}(\tau) y_j(t) + I_i(t) \\ y_i(t) = f_i(x_i(t)) \end{cases}$$

ここで $x_i(t)$ は時刻 t における i 番目のニューロンへの入力の総和、$w_{ij}(\tau)$ は時刻τにおける j 番目のニューロンからの結合荷重、$y_j(t)$ は j 番目のニューロンの出力、$I_i(t)$ はニューロンへの入力信号を表す。$f_i()$ はニューロンの入出力応答関数を表し、通常はシグモイド関数が用いられる。

6.2.2 Real Time Recurrent Learning 学習法

ネットワークの学習法としては Back Propagation Through Time(BPTT) と Real Time Recurrent Learning(RTRL) と呼ばれる方法がある。ここでは RTRL 法について説明する。

ネットワークに対するエラー関数を以下のように定義する。

$$\begin{cases} e_i(t) = \dfrac{1}{2}(y_i(t) \quad d_i(t))^2 & i \in \text{出力ニューロン} \\ E(t) = \displaystyle\sum_i e_i(t) \\ E = \displaystyle\int_0^T E(t)dt \end{cases}$$

ここで $d_i(t)$ は教師信号である。このとき、バックプロパゲーション法と同様に最急降下法を用いて結合荷重の変更量を

$$\Delta w_k l(\tau) = \quad \lambda \frac{\partial E}{\partial w_{kl}(\tau)}$$

と計算すると、時間τごとの結合荷重の更新量は

$$\begin{cases} P_{kl}^i(t+1) = f_i'(x_i(t+1))(\delta_{ik} y_l(t) + \displaystyle\sum_j w_{ij} P_{kl}^j(t)) \\ \Delta w_{kl}(\tau) = \quad \lambda \displaystyle\sum_{i \in \mathrm{for\ output\ neuron}} \int_0^T (y_i(t) \quad d_i(t)) P_{kl}^i(t) dt \\ w_{kl}(\tau+1) = w_{kl}(\tau) + \Delta w_{kl}(\tau) \end{cases}$$

と計算される。

6.2.3 リカレントニューラルネットワークの応用例

RTRL 法によるリカレントニューラルネットワークによって、sin 関数の学習を行なった。ニューロン数は 6 で、そのうちの一つに教師信号を与え出力ニューロンとした。図 16 に学習結果を示す。ニューラルネットワークに入力を与えなくとも、自律的に振動を出力するようになっていることが分る。

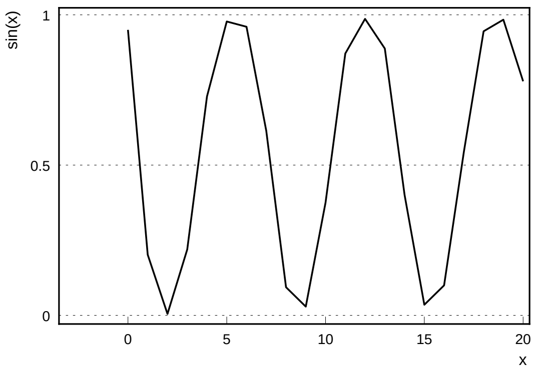


図 16　リカレントニューラルネットワークによる三角関数の学習結果

7　ニューラルネットワーク関連書籍

入門と実習ニューロコンピュータ 中野 馨監修, 飯沼 一元編, ニューロンネットグループ+桐谷 滋著/技術評論社 — バックプロパゲーション法、ホップフィールドネット、ボルツマンマシンについて述べてあり、C 言語のプログラムが掲載されている。非常に簡潔に書かれた入門書であり、数学的根拠にはあまり触れられていない。

ニューロコンピューティング R ヘクト・ニールセン著, 袋谷 賢吉訳/トッパン — 代表的なニューラルネットワークについては網羅してある。ヘクト・ニールセンが教科書として使っていたようだが、一般性を持たせるためにかえって難しくしてある部分もあり若干読みにくい。

神経細胞が行なう情報処理とそのメカニズム 松本 元・大津 展之共編/倍風館 — 生体ニューロン及び脳の既知のことに詳しい。

ニューロコンピューティングの周辺 松本 元・大津 展之共編/倍風館 — ニューラルネットワーク情報処理に関連のあるパターン認識における従来法や並列計算機理論について説明している。

神経回路網の数理 甘利 俊一著/産業図書 — 確率統計場の理論などから数学的にニューラルネットワークをモデル化、解析している。

認知科学選書 22 神経回路モデルとコネクショニズム 甘利 俊一/東京大学出版会 — 甘利色豊かな一冊。代表的なニューラルネットワークについてはほとんど網羅している。

ニューラルコンピュータ～脳と神経に学ぶ～ 合原 一幸著/東京電気大学出版会 — 合原氏の研究から出た一冊で、ヤリイカの生体ニューロンの解析から代表的なニューラルネットワーク、カオスニューロンの話について述べられている。

ニューラルネットワークと脳理論 M A アービブ著, 金子 隆芳訳/サイエンス社 — 脳の話からニューラルネットワークの一般的なモデルについて解析を行なっている。一般的なニューラルネットワークの本と異なりかなり啓発的な本である。

ニューラルネットワーク情報処理 麻生 秀樹著/産業図書 — 代表的なニューラルネットワークについての解説と、記号処理とニューラルネットワークについて述べた書籍。

ニューラルネットワークの物理モデル タマシ・ゲスチ著, 秋葉 巴也訳/吉岡書店 — ホップフィールド型ネットワークの物理学的解説と自己組織化マップについて述べている。

ニューロコンピューティングの基礎理論 日本工業技術振興会ニューロコンピュータ研究部会編 — 代表的なニューラルネットワーク及びその応用について解説している。ニューラルネットワークの基礎学習に適する。

ニューロ・ファジィ・カオス 合原 一幸編著/オーム社 — ニューラルネットワークの中でも比較的新しい技術について簡単に解説している。

パターン認識と学習のアルゴリズム 上坂 吉則・尾関 和彦著/文一総合出版 — パターン認識について総集してある書籍であり、その一部としてニューラルネットワークを用いた場合を解説している。