利用者向け講座

柔らかい人工知能を目指して（その5）

ニューラルネットワーク

古　橋　　武

I．はじめに

私たちの脳は大変に優れた情報処理能力を有しています。視覚，聴覚，触覚，味覚，嗅覚の5感の刺激に対して反応する刺激反応系なのですが，大変に高度な学習能力，認識・判断能力を持っています。この脳の持つ機能をコンピュータにより実現することができることはとてもすばらしいことに思われます。人工ニューラルネットワークは脳内情報処理の基本単位であるニューロンの工学モデルを用いてネットワークを構成し，脳の機能をコンピュータ上に実現しようとするものです。その基本機能は任意の非線形連続関数の近似能力と学習能力です。人工ニューラルネットワーク(以下，単にニューラルネットワークと呼びます)を文字認識，音声認識，ロボット制御などに応用する研究がここ10年来盛んになされてきていますが，ニューラルネットワークは認識・制御のアルゴリズムを学習により自動獲得できることが最大の利点です。本稿では，まず，人工ニューロン(以下，単にニューロンと呼びます)の基本機能と学習法，単一ニューロンの限界について述べ，次に，ニューラルネットワークの能力と誤差逆伝播(Error Back Propagation，BP)法について解説し，ニューラルネットワークにより人間の制御ノウハウ獲得のシミュレーション例を紹介します。

脳の機能はあまりにも高く，人工ニューラルネットワークは脳の機能の極一部しか実現できていません。今後の研究の展開が楽しみな分野です。

II．ニューロンとは

ニューラルネットワークは生物の脳細胞を工学的に模擬したものです。図1はニューロンの模式図を示します。ニューロンは，細胞体と呼ばれる本体部分，本体から突き出た樹状突起および軸索からなります。ニューロンは情報処理素子であり，樹状突起および細胞体の表面で入力信号を受け，軸索から出力信号を出します。軸索の先は何本にも枝分かれして，その末端がそれぞれ他のニューロンの樹状突起もしくは細胞体の表面と結合しています。この結合部分はシナプスと呼ばれます。一つのニューロンの出力情報は，シナプスを介して他の多くのニューロンへと伝えられます。脳の中では多数のニューロンが複雑に結合してさまざまな情報処理を行っています。

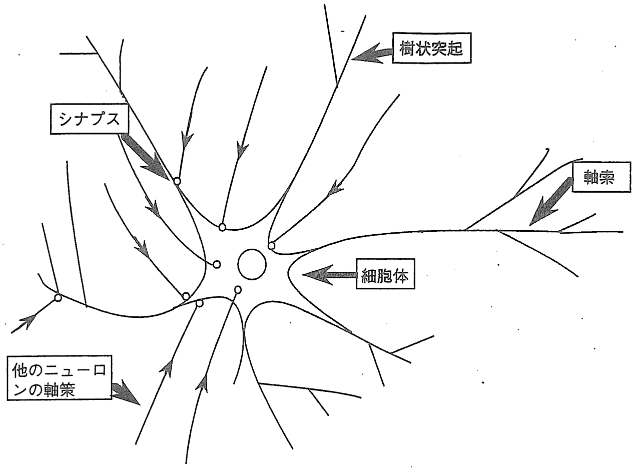


図１　ニューロンの模式図

1．ニューロンのモデル

図２はニューロンの工学モデルを示します。図は４入力１出力のニューロンの例です。細胞体の部分は円で表され，他のニューロンからの刺激 $x_1 \sim x_4$ はシナプス結合 $w_1 \sim w_4$ を介してニューロンに入力されます。出力 y は非線形関数 f により

$$y = f(X) \tag{1}$$

$$X = w_1 x_1 + w_2 x_2 + w_3 x_3 + w_4 x_4 \tag{2}$$

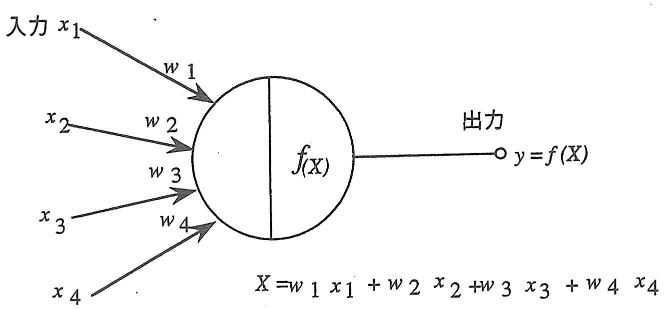


図２　ニューロンの工学モデル

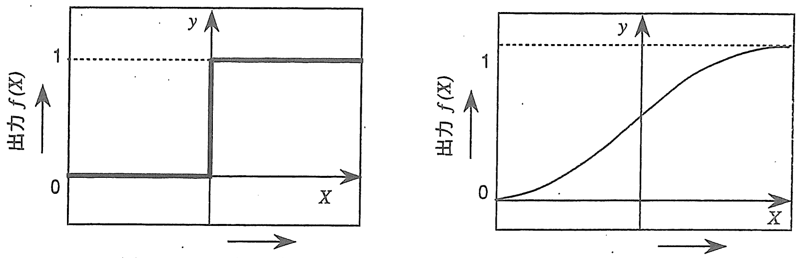
と与えられます。この非線形関数 f としては，しきい値関数やシグモイド関数が用いられます。図３ (a) はしきい値関数を，(b) はシグモイド関数を表します。しきい値関数は

$$f(x) = \begin{cases} 1, X \geqq 0 \\ 0, X < 0 \end{cases} \tag{3}$$

と与えられるステップ関数であり，シグモイド関数は

$$f(X) = 1/(1 + exp(-X)) \tag{4}$$

と表される，しきい値関数の一種です。



(a) しきい値関数

$$f(X)=\begin{cases}1, X\geqq 0\\ 0, X<0\end{cases}$$

(b) シグモイド関数

$f(X)=1/(1+exp(-X))$

図3　ユニットの応答関数（f）の例

このニューロンモデルは入出力間の非線形な写像を行います。多数のニューロンモデルを組み合わせることで，複雑な非線形写像をも実現できそうです。事項では，まず，単一のニューロンについてその写像能力を見ていきます。

2．ニューロンの能力

ニューロンの基本能力は入力空間の分割にあります。図4 (a) は2入力1出力のニューロンの例です。ニューロン内部の非線形関数は (3) 式のしきい値関数とします。今，

$$w_1 = w_2 = 1 \tag{5}$$

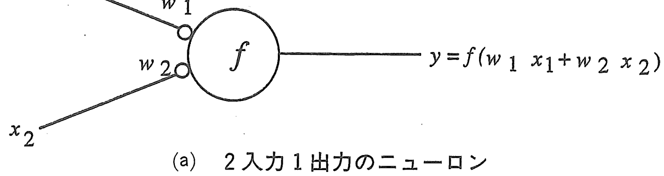
とすると，ニューロンへの入力 X は

$$X = x_1 + x_2 \tag{6}$$

となります。ニューロンへの入力が0のところで出力値は0と1の間をステップ的に変化します。すなわち，

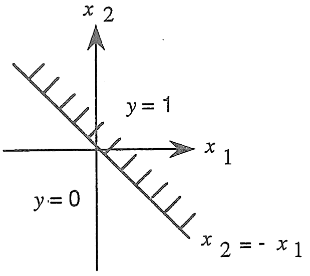
$$X = x_1 + x_2 = 0 \tag{7}$$

を境界線として，$x_2 \geqq -x_1$の領域で出力 $y=1, x_2 < -x_1$の領域で $y=0$ となります。図4 (a) のニューロンは x_1, x_2の入力空間を2分することができます。この様子を図4 (b) に示します。このニューロンは入力空間を原点を通る任意の直線で2分できます。



(a)　2入力1出力のニューロン

図4　ニューロンの能力(I)



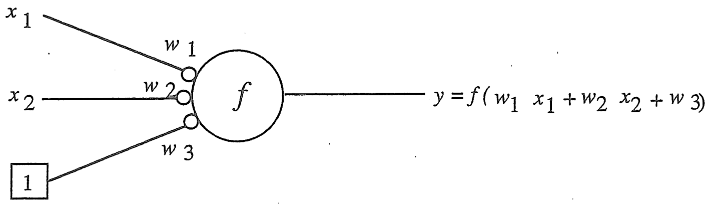
(b)　入力空間の分割（原点を通る直線）

図４　ニューロンの能力(I)

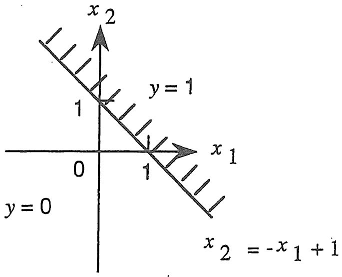
原点を通らない直線で入力空間を２分することは，図５ (a) に示す３入力１出力のニューロンにより実現できます。入力 x_1, x_2 に，新たに１の値を出力するユニットが追加されています。ここで，

$$w_1 = w_2 = 1, w_3 = -1 \tag{8}$$

とすると，図５ (b) のように 原点を通らない直線により入力空間を２分できます。



(a)　３入力１出力のニューロン



(b)　入力空間の分割（任意の直線）

図５　ニューロンの能力(II)

このニューロンにより，例えば，ＡＮＤ，ＯＲの基本演算が実現できます。図６はその様子を示します。図６ (a) は図５と同じ３入力１出力のニューロンです。図６ (b)，(c) はＡＮＤ，ＯＲの各演算の真理値表と x_1, x_2 の入力空間をニューロンが分割する様子をそれぞれ示します。ニューロンの入力 x_1, x_2 は任意の連続値をとり得ますが，ここでは０と１の２値しかとらないこととします。

ＡＮＤ演算は図６ (b) の真理値表に示すように，入力 x_1, x_2が共に１のときのみ出力が１となり，それ以外では０となります。この演算をニューロンに行わせるには，結合荷重 w_1, w_2, w_3 を例えば

$$w_1 = w_2 = 1, w_3 = -1.5 \tag{9}$$

とすればよいでしょう。入力空間は

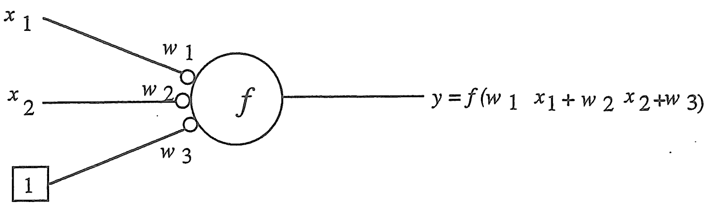
$$x_1 + x_2 = 1.5 \tag{10}$$

の直線により2分されます。図6 (b) の右側にこの様子を示します。この直線より上側では出力は1です。図の●では出力が1であり，○では出力が0であることを示します。$x_1 = x_2 = 1$ のとき $y = 1$ であり，x_1, x_2の少なくとも一つが0のとき $y = 0$ であり，ＡＮＤ演算が確かに行われることが分かります。

ＯＲ演算は入力 x_1, x_2が共に0のときのみ出力が0となります。この演算を実現するにはＡＮＤ演算の分割線を下側へ少し平行移動すればよいことが分かります。例えば

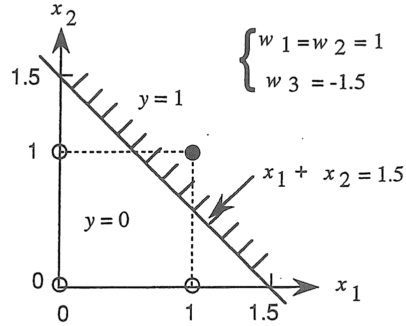
$$w_1 = w_2 = 1, w_3 = -0.5 \tag{11}$$

とすれば，図6（ｃ）に示すように，ＯＲ演算が実現できます。



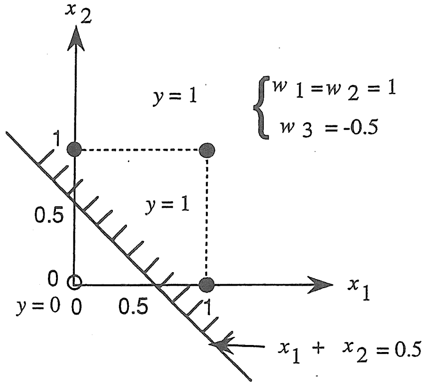
(a)　3入力1出力のニューロン

x_1	x_2	y
0	0	0
0	1	0
1	0	0
1	1	1



(b)　AND 演算

x_1	x_2	y
0	0	0
0	1	1
1	0	1
1	1	1

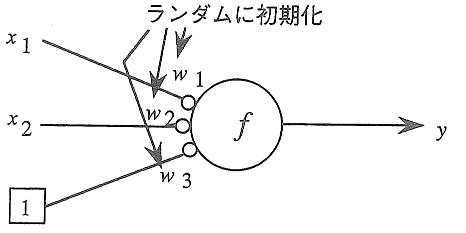


(c)　OR 演算

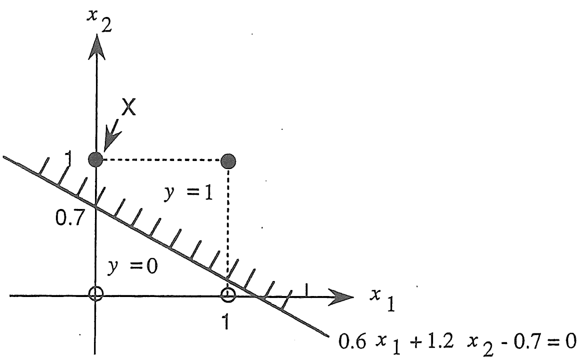
図6　単一ニューロンによる AND，OR，NOT

3．ニューロンの学習

前項の単一ニューロンのＡＮＤ，ＯＲ演算のための結合荷重は事前にこちらから決定して与えました。ここでは，結合荷重を自動的に求める学習法について述べます。ＡＮＤ演算やＯＲ演算程度の簡単な問題ならば，この学習法を知らなくても解けますが，問題が複雑になった場合には学習法は有効な手段となります。



(a)　結合荷重の初期化



(b)　初期の分割線

図７　ニューロンの初期化

ＡＮＤ演算を学習により実現します。まず，図７に示すようにニューロンの結合荷重 w_1, w_2, w_3 を乱数により決定します。例えば，$w_1 = 0.6, w_2 = 1.2, w_3 = -0.7$ と与えられたとします。このニューロンが入力空間を分割する直線を図７(b) に示します。$x_1 = 0, x_2 = 1$ のとき，出力 $y = 1$ であり，この点の出力がＡＮＤ演算とは異なっています。

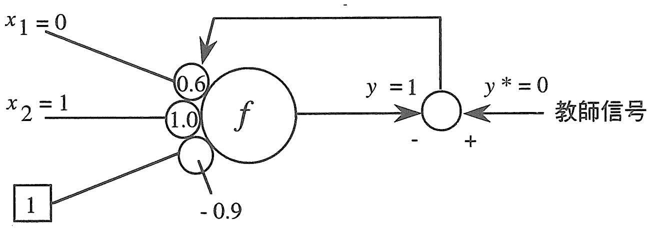
天下り的ではありますが，ニューロンの学習を取り敢えず次式により行うものとします。この式は，ニューロンの結合荷重をニューロンの出力の誤差を利用して更新するものです。

$$w_i = w_i + \eta(y^* - y)x_i \qquad i = 1, 2, 3 \tag{12}$$

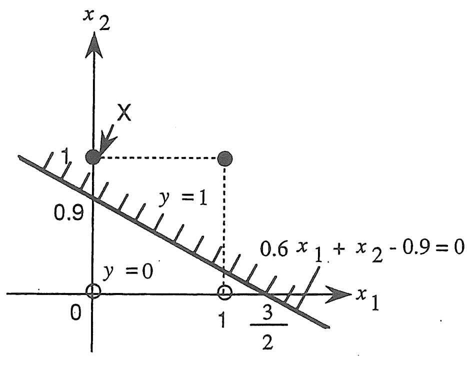
ただし，y^*はニューロンが出力すべき正しい値であり教師信号と呼ばれます。y はニューロンの実際の出力であり，ηは学習率と呼ばれる定数であります。また，$x_3 = 1$ であります。今，学習率 $\eta = 0.2$ とします。$x_1 = 0, x_2 = 1$ のとき，$y^* = 0, y = 1$ と誤差があるので，各結合荷重は

$$
\begin{aligned}
w_1 &= 0.6 + 0.2 \times (0 - 1) \times 0 = 0.6 \\
w_2 &= 1.2 + 0.2 \times (0 - 1) \times 1 = 1.0 \\
w_3 &= -0.7 + 0.2 \times (0 - 1) \times 1 = -0.9
\end{aligned} \tag{13}
$$

と更新されます。ＡＮＤ演算の他の３通りの入力に対しては出力の誤差は０であるので，結合荷重はなんら修正されません。図８は以上の学習後のニューロンの結合荷重と新しい分割線を示します。まだ，$x_1 = 0, x_2 = 1$ の点の誤差は改善されてはいません。以上を，１回の学習と呼びます。



(a)　結合荷重の更新



(b)　修正された分割線

図８　ニューロンの学習（第１回目）

そこで，もう１回の学習を行います。その結果，$w_1 = 0.6, w_2 = 0.8, w_3 = -1.1$ となり，図９のニューロンが得られます。このニューロンはＡＮＤ演算を行えるようになったので，学習を終了とします。

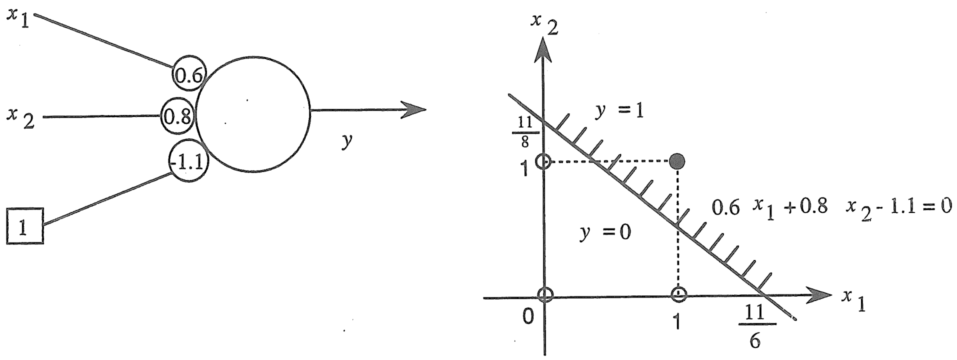


図９　ニューロンの学習（第２回目）

４．学習法の定式化

前項では学習法を (12) 式により与えたところ，ＡＮＤ演算の例題では学習がうまくいきました。ここで，(12) 式で何故よかったのか，学習法の定式化を行い，種明かしをします。次式の誤差 E を考えます。

$$E = \frac{1}{2}(y^* - y)^2 \tag{14}$$

誤差 E は，図 10 に示すように，ニューロンの出力値 y が教師信号 y^* と一致したときに 0 となり，その他では 0 より大きな値となります。

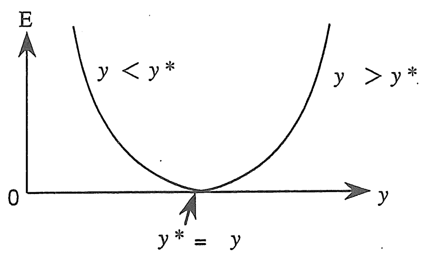


図 10　誤差Ｅ

学習法の基本的な考え方は，誤差 E を減少させる方向に結合荷重を更新するというものです。そこで，結合荷重を増減のどちらに変化させたら誤差 E を減少させられるかを求めます。図 11 は 3 入力 1 出力のニューロンを再度示します。w_1の修正について考えます。w_1の変化が E にどれだけの変化を与えるかは，w_1の変化→ニューロンの入力($u = w_1x_1 + w_2x_2 + w_3$)の変化→ニューロンの出力($y = f(u)$)の変化→ E の変化をみればわかります。式で表現すれば

$$\begin{aligned}\frac{\partial E}{\partial w_1} &= \frac{\partial E}{\partial y}\cdot\frac{\partial y}{\partial u}\cdot\frac{\partial u}{\partial w_1}\\ &= (y^* - y)\cdot\frac{\partial f}{\partial u}\cdot x_1\end{aligned} \tag{15}$$

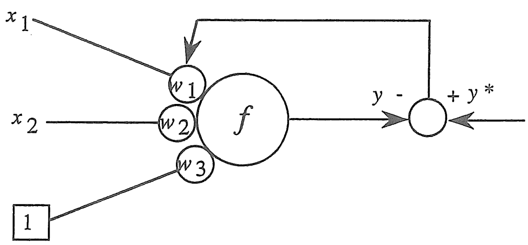


図11　学習法

と表されます。左辺は w_1 の変化に対する E の変化量を表し，右辺はその変化量を，（E の変化量）／（y の変化量），（y の変化量）／（u の変化量），（u の変化量）／（w_1 の変化量）に分解しています。右辺は前述の矢印を反対にたどっています。

したがって，結合荷重の更新量 ΔE は，誤差 E を減少させる方向にとればよいので

$$\Delta w_1 \propto -\frac{\partial E}{\partial w_1} \tag{16}$$

比例係数を η として，

$$w_1 = w_1 + \eta(y^* - y) \cdot \frac{\partial f}{\partial u} \cdot x_1 \tag{17}$$

と求められます。$\frac{\partial f}{\partial u} = 1$ とすれば前項で天下り的に与えた学習法の式と一致します。前項の学習法は，実は，この結果を先取りしたものでありました。

ここで，$\frac{\partial f}{\partial u}$ の項について考えます。前項で用いたニューロンの内部関数 f はステップ関数であり，$\frac{\partial f}{\partial u} = 1$ ではありません。(12) 式の学習法では必ずしも誤差 E を減少させることを保証してはいませんでした。誤差 E の減少を保証するには，f は微分可能である必要があります。そこで，内部関数 f として (4) 式のシグモイド関数を用いることとします。このとき，

$$\begin{aligned}\frac{\partial f}{\partial u} &= \frac{\partial}{\partial u} \cdot \frac{1}{1 + exp(-u)} \\ &= f_{(u)}(1 - f_{(u)}) \\ &= y(1 - y)\end{aligned} \tag{18}$$

となるので

$$\begin{aligned}w_1 &= w_1 + \eta(y^* - y)y(1 - y)x_1 \\ &= w_1 + \eta\delta x_1\end{aligned} \tag{19}$$

$$\delta = (y^* - y)y(1 - y) \tag{20}$$

と求められます。この結合荷重の更新則によれば，誤差 E を減少できることが保証されます。

学習法の定式化ができたところで，学習率についてコメントをしておきます。(19) 式の学習法は結合荷重の更新の方向を与えてくれるが，更新量までは決められません。更新量は学習率によって決定することになりますが，学習率が大き過ぎると1回の学習による結合荷重の更新量が大きくな

り過ぎてしまい，結合荷重は行きつ戻りつと振動的になります。場合によっては，結合荷重が大きく変わり過ぎて，誤差は一向に減らないことになってしまいます。逆に学習率が小さ過ぎると，1回の学習における結合荷重の更新量が少なく，学習はなかなか進まないことになります。学習率には適度な値が存在します。

III．ニューラルネットワークとは

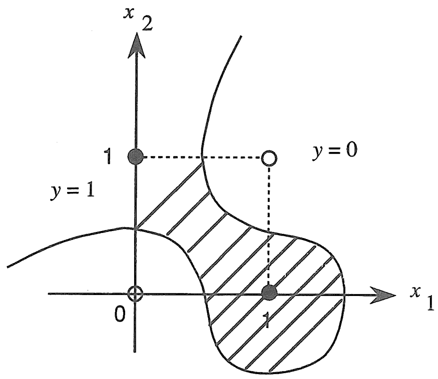
前節では，単一のニューロンを取り上げ，その能力と学習法について説明しました。単一のニューロンを複数用いてネットワークを構成すれば，機能の向上が期待できます。本節では，単一ニューロンの能力の限界とニューラルネットワークの能力について述べ，ニューラルネットワークの有力な学習法であるバックプロパゲーション法(Back Propagation 法：BP 法)を紹介します。

1．単一ニューロンの限界

単一ニューロンは入力空間を直線により2分する能力を持つことが前節で明らかとなりました。しかし，単一ニューロンは曲線もしくは複数の直線で入力空間を分割することはできません。このことを端的に示すために Exclusive OR(ＸＯＲ)問題を考えましょう。この問題の真理値表は図 12(a) に示すように，x_1 と x_2の値が一致したときに出力 $y=0$ であり，両者が異なるとき出力は1です。このような入出力関係を実現するには，例えば，図 12(b) に示すような曲線により入力空間を分割しなければなりません。このような非線形な分割問題は，単一のニューロンでは解くことができません。

x_1	x_2	y
0	0	0
0	1	1
1	0	1
1	1	0

(a) 真性値表



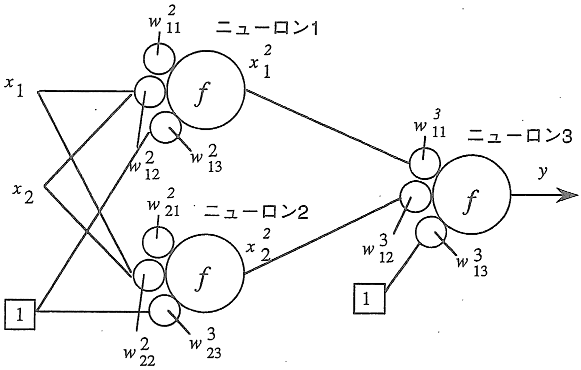
(b) 入力空間の分割

図 12　Exclusive OR 問題

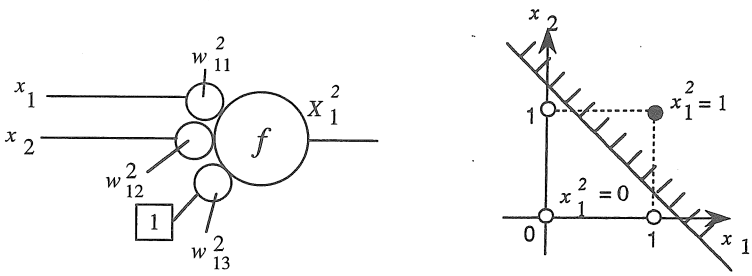
2．ニューラルネットワークの能力

ニューロンを複数用いてネットワークを構成します。図 13(a) はニューロンを3個用いてニューラルネットワークを構成した例です。このネットワークはＸＯＲ問題を解くことができます。ネットワーク内のニューロンに番号をつけ，それぞれニューロン1，2，3とします。ニューロン1，2

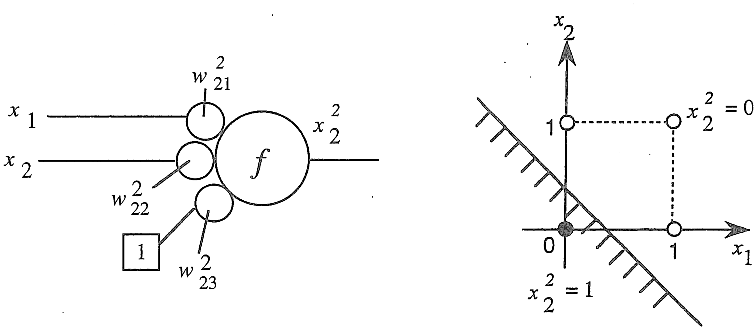
の各記号には，これらニューロンがネットワークの中間にあることを示すために，右肩に添え字2を付しています。ニューロン3は出力ニューロンであり，右肩の添え字は3としています。ニューロン1を図13(b)に示すようにAND回路とし，ニューロン2，3をNAND回路(AND回路＋NOT回路)とする(同図(c)(d))ことで，ニューラルネットワークの入力空間を同図(e)のように2本の直線で3つの領域に分割することができます。



(a) ニューラルネットワーク

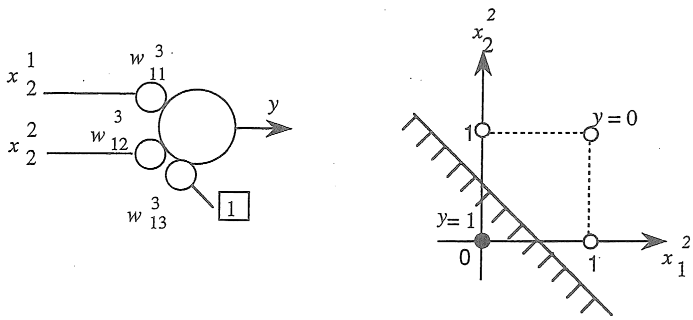


(b) ニューロン1

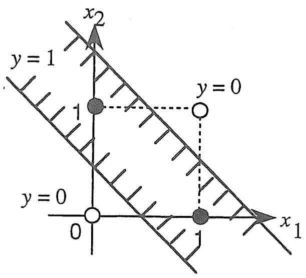


(c) ニューロン2

図13 ニューラルネットワークによるXOR問題の解法



(d)　ニューロン3



(e)　ニューラルネットワークの入出力関係

図13　ニューラルネットワークによるXOR問題の解法

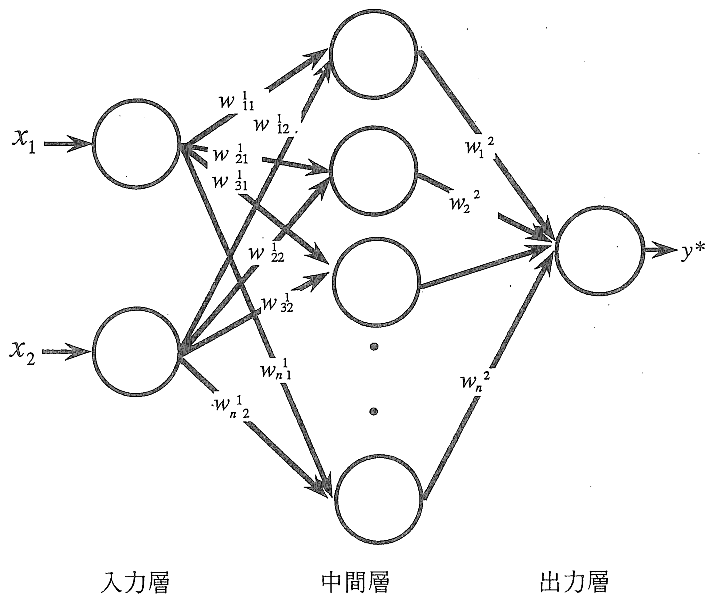


図14　3層ニューラルネットワーク

ニューロンを複数組み合わせることで，単一ニューロンではできなかった入力空間の非線形な分割を実現できることを示しました。より多くのニューロンを組み合わせれば，より複雑な写像を実

現できます。ニューラルネットワークはニューロンの組み合わせ次第でさまざまな構成が可能です。本稿では，最も基本的な構成の一つである3層ニューラルネットワークについて解説をします。図14は3層ニューラルネットワークの構成例です。2入力1出力で，入力層，中間層，出力層の3層からなるニューラルネットワークです。各ニューロンの結合荷重は○印を省略して，層間の結線の上に記してあります。各ニューロンの内部関数にシグモイド関数を用いることで，このニューラルネットワークの非線形写像能力はとても高いものとなります。中間層に任意の数のニューロンを用いることを許せば，図14のニューラルネットワークは，入力 x_1, x_2 と出力 y との間の任意の連続関数を学習により獲得できます [1]。この証明は本稿の範囲を超えるので，興味のある読者は文献 [1] を読まれることを勧めます。この保証が得られることで，ニューラルネットワークにより非線形写像問題を解くことは，少なくとも解の無い問題を解こうとしているわけではないという安心感を与えてくれます。

3．バックプロパゲーション法

3層のニューラルネットワークが任意の連続関数の写像が可能であることは保証されましたが，その写像をどのようにして実現すればよいのかが示されたわけではありません。ネットワーク内の結合荷重の決定法が重要となります。前項では，ＸＯＲ問題の解法のために各結合荷重を発見的に与えていました。問題とそれを解くためのニューラルネットワークが複雑となれば，各結合荷重を発見的に決めることはほとんど不可能となります。

ニューラルネットワークの結合荷重は学習により決定します。学習法の基本的な考え方は，II. 4項で説明した単一ニューロンの学習法と同じです。すなわち，出力誤差を減少させる方向に各結合荷重を少しずつ変化させます。図13(a)のニューラルネットワークを例として取り上げます。図15は図13(a)のネットワークの一部を拡大して示します。w^2_{11}の更新の方向は（14）式の誤差 E を減少させる方向です。w^2_{11}の変化がネットワークの各部の値に与える影響は，w^2_{11}の変化→ニューロン1の入力 ($u^2 = w^2_{11}x_1 + w^2_{12}x_2 + w^2_{13}$) の変化→ニューロン1の出力 x^2_1の変化→ニューロン3の入力 ($u^3 = w^3_{11}x^2_1 + w^3_{12}x^2_2 + w^3_{13}$) の変化→ニューロン3の出力 y の変化→誤差 E の変化となります。式では

$$
\begin{aligned}
\frac{\partial E}{\partial w^2_{11}} &= \frac{\partial E}{\partial y} \cdot \frac{\partial y}{\partial u^3} \cdot \frac{\partial u^3}{\partial x^2_1} \cdot \frac{\partial x^2_1}{\partial u^2} \cdot \frac{\partial u^2}{\partial w^2_{11}} \\
&= -\delta^3 w^3_{11} x^2_1 (1 - x^2_1) x_1
\end{aligned}
\tag{21}
$$

と求められます。したがって，結合荷重 w^2_{11}は

$$
\begin{aligned}
w^2_{11} &= w^2_{11} + \eta \delta^3 w^3_{11} x^2_1 (1 - x^2_1) x_1 \\
&= w^2_{11} + \eta \delta^2 x_1
\end{aligned}
\tag{22}
$$

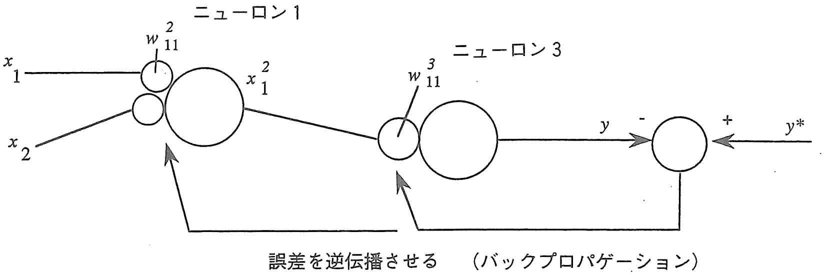


図15　バックプロバケーション法

により更新します。η は学習率です。また，δ^2, δ^3は誤差項です。δ^3はニューロン3の結合荷重の更新則((19)，(20)式参照)における値です。ニューロン1の結合荷重の更新にはニューロン3の結合荷重の誤差項 δ^3が利用されます。

もし，3層ニューラルネットワークのようにニューロン1の入力側にも別のニューロンがあれば，このニューロンの持つ結合荷重は誤差項δ^2, δ^3 を利用して更新されます。すなわち，結合荷重の更新に際して，誤差 E はこれらの誤差項を介してニューラルネットワークの出力側から入力側へと伝播されていきます。このゆえに，以上の学習法は，誤差逆伝播法と呼ばれます。

Ⅳ．ニューラルネットワークによる制御

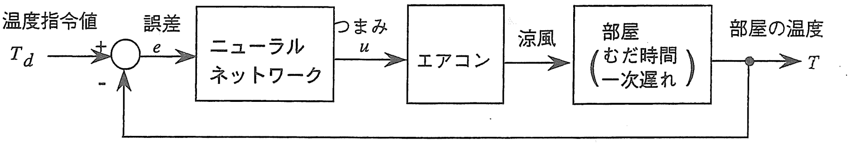
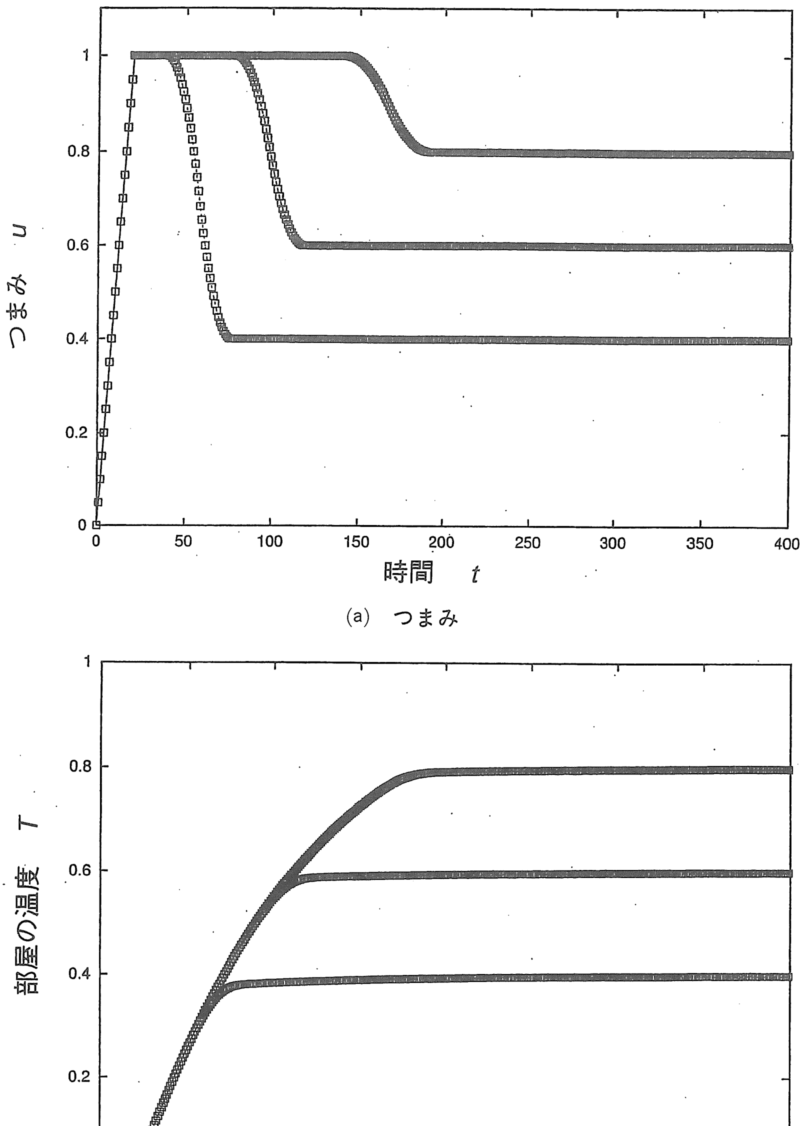


図16　部屋の温度のニューラルネットワークによる制御

3層ニューラルネットワークにより部屋の暖房制御を行います。部屋はむだ時間を含む一次遅れ系です。図16はニューラルネットワークを組み込んだ制御系を示します。ニューラルネットワークの学習は，ベテランが実際に部屋の温度制御を行ったときのつまみと温度に関するデータを用います。図17はベテランの制御結果です。指令値 $T_d = 0.4, 0.6, 0.8$ に対してそれぞれ速やかに指令値に追従する応答を得ています。この制御結果はコンピュータシミュレーションにより得られたものであり，計算刻みを1として，むだ時間 $t_d = 100$，時定数 $t_s = 100$ としています。ニューラルネットワークは，図18に示すように，このデータから誤差 e と誤差の変化分 Δe を取り出して入力とし，つまみの変化分Δu を教師信号としてＢＰ法により学習を行います。$e, \Delta e, \Delta u$ の符号をすべて正負反転させたデータについても学習を行い，負の指令値変化に対しても同様の制御が行えるようにします。ニューラルネットワークの入力層，中間層，出力層の各ニューロン数は，それぞれ，2，15，1とします。



(a)　つまみ

(b)　部屋の温度

図 17　ベテランの制御結果

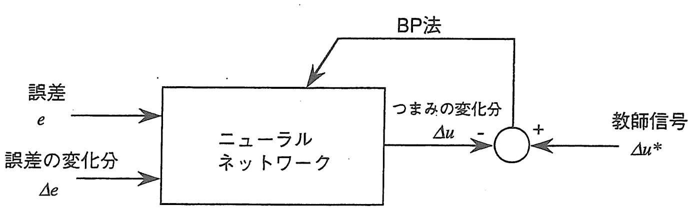
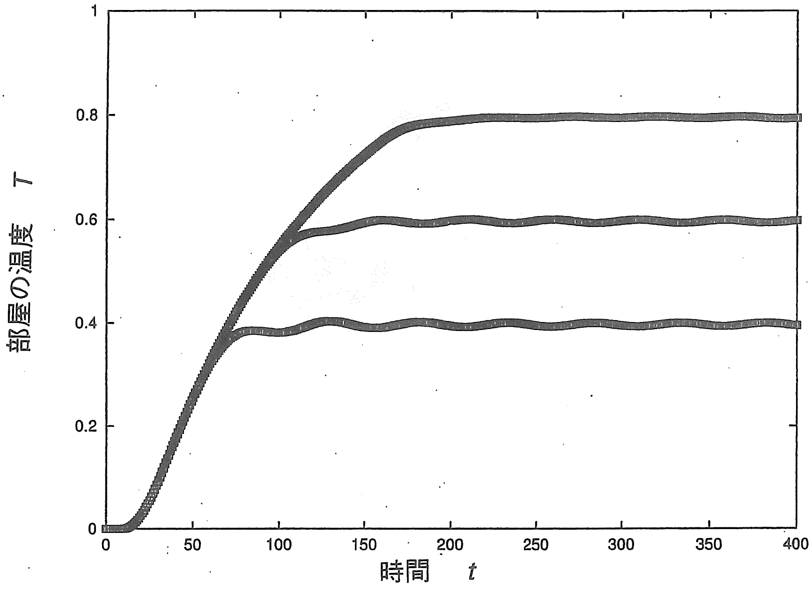
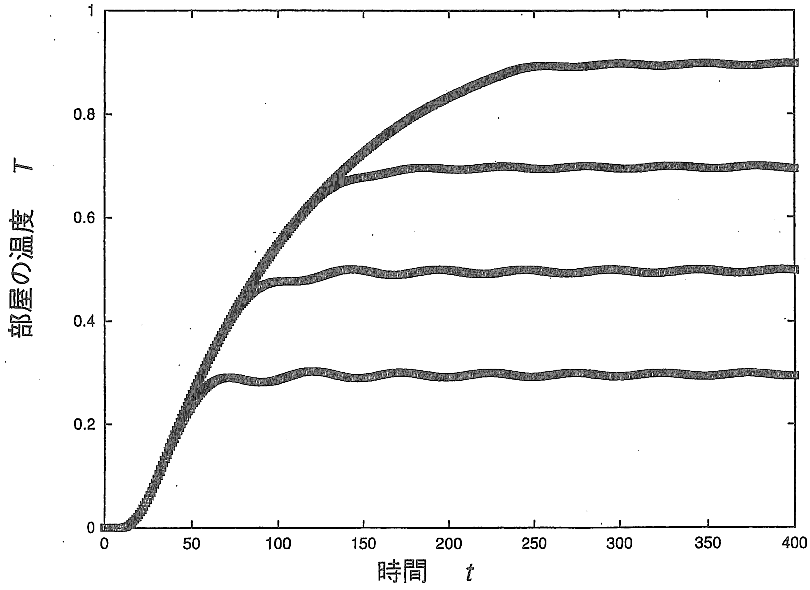


図 18　ニューラルネットワークの学習

学習後のニューラルネットワークを図16の制御系に組み込み制御を行います。エアコンにはΔuを積算した値uが与えられます。制御結果を図19に示します。同図(a)は学習した指令値に対する応答です。ニューラルネットワークはベテランと同様の制御を行うことができました。同図(b)は未学習の指令値に対する応答です。ニューラルネットワークは未学習の指令値に対しても内挿，外挿により良好な制御を行うことができています。



(a)　学習した指令値に対する応答



(b)　未学習の指令値に対する応答

図19　学習後のニューラルネットワークによる制御結果

ニューラルネットワークが学習により獲得した制御ルールは，その入出力の写像を見ることである程度理解できます。図20は学習後のニューラルネットワークの入出力曲面です。この曲面からは，誤差 e が大きく ($e\ is\ PB$)，誤差の変化分Δe が零 ($\Delta e\ is\ ZO$) のときはつまみを大きく回し ($\Delta u\ is\ PB$)，誤差 e が零で ($e\ is\ ZO$)，誤差の変化分Δe が減少中 ($\Delta e\ is\ NB$) のときはつまみを戻す ($\Delta u\ is\ NB$) 等のルールを読み取ることができます。

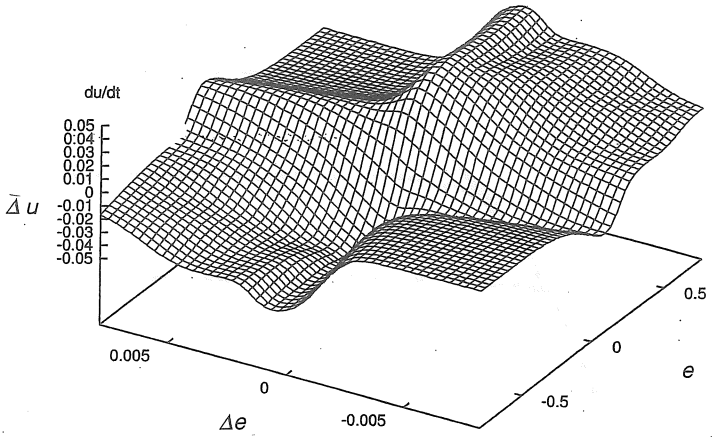


図20　学習後のニューラルネットワークの入出力曲面

V。まとめ

ニューラルネットワークは学習により任意の非線形連結関数を近似する能力を持っています。本稿では，ニューロンとニューラルネットワークの基本的な機能について解説しました。ニューラルネットワークに関する文献は膨大であり，インターネット上でキーワード検索を行えば，大変に豊富な情報が収集できます。本稿が，読者の皆さんの関心を呼ぶことができれば幸いです。

参考文献

(1) K. Funahashi : On the Approximate Realization of Continuous Mapping by Neural Networks, Neural Network, Vol. 1, pp. 183-192 (1989)

（ふるはし　たけし：名古屋大学工学研究科電子情報学専攻）