

複素バックプロパゲーション・ネットワークにおける 重みパラメータと決定表面の構造†

新 田 徹^{††} 赤 穂 昭 太 郎^{†††}
秋 山 泰^{††*} 古 谷 立 美^{††}

近年、階層型ニューラルネットワークの学習方式として提案されたバックプロパゲーション学習アルゴリズム (BP) が注目され、様々な分野に応用されている。複素パターンに対する学習アルゴリズムとして提案された複素 BP は、その学習パラメータがすべて複素数となっており、実数値を用いる通常の BP を複素数に拡張したものであると言える。複素 BP には、従来の BP には見られない 2 次元運動学習能力が備わっていることは既に報告した。本稿では、ネットワークアーキテクチャの観点から、複素 BP の基本特性を解析的、実験的に調べ、従来の BP との差異を明確にしたので報告する。得られた主要な結果は、次のとおりである。(1) 複素 BP ネットワークにおける重み係数は、従来の BP ネットワークにおけるそれとは異なり、2 次元運動に関係した制約を持っており、学習は基本的にその制約を保ちつつ行われる。(2) 複素ニューロンの実部の決定表面と虚部の決定表面とは互いに直交しており、決定領域を均等に 4 つに分割する。3 層の複素 BP ネットワークにおける決定表面は基本的にこの構造を内包しており、中間ニューロンへの総入力十分大きいときに直交する。この意味で、複素 BP は複素パターンに対する自然な学習アルゴリズムであると思われる。これらの特性は意図されたものではなく、複素数への拡張の結果として自ずと現れてきたものであることに注意されたい。

1. はじめに

近年、階層型ニューラルネットワークの学習方式として提案されたバックプロパゲーション学習アルゴリズム⁹⁾ (以下、実 BP と略す) が注目され、様々な分野に応用されている^{11, 3)}。実数値を用いる通常の BP の複素数版である複素バックプロパゲーション学習アルゴリズム⁴⁾ (以下、複素 BP と略す) は、その重み係数としきい値がすべて複素数値となっており、複素パターンに対する学習を想定しているものである。複素 BP には、2 次元運動学習能力 (図形変換能力) が備わっている^{4), 6)} など、実 BP には見られない様々な性質を持っていることが予想される。

本稿では、ネットワークアーキテクチャの観点から、複素 BP の基本特性を解析的、実験的に調べ、実 BP との差異を明確にしたので報告する。以下、複素 BP が適用されるネットワークを複素バックプロパ

ゲーション・ネットワーク (略して、複素 BP ネットワーク) と呼ぶことにする。

得られた主要な結果は、次のとおりである。(1) 複素 BP ネットワークにおける重み係数は、実 BP が適用されるネットワーク (実 BP ネットワーク) におけるそれとは異なり、(線形代数でいうところの) 2 次元運動に関係した制約を持っており、学習は基本的にその制約を保ちつつ行われる。(2) 複素ニューロンの実部の決定表面と虚部の決定表面とは互いに直交しており、決定領域を均等に 4 つに分割する。3 層の複素 BP ネットワークにおける決定表面は基本的にこの構造を内包しており、中間ニューロンへの総入力十分大きいときに直交する。

この意味で、複素 BP は複素パターンに対する自然な学習アルゴリズムであると思われる。これらの特性は意図されたものではなく、複素数への拡張の結果として自ずと現れてきたものであることに注意されたい。

第 2 章において、複素 BP の概略について述べる。第 3 章では、複素 BP ネットワークの解析を行い、決定表面の性質などの基本特性について調べる。それらの計算機実験の結果を第 4 章で示し、最後に第 5 章で結論と今後の課題について述べる。

2. 複素 BP の概要

複素 BP⁴⁾ について簡単に説明する。適用される

† Structure of Weight Parameters and Decision Boundary in Complex Back-propagation Network by TOHRU NITTA (Computational Models Section, Computer Science Division, Electrotechnical Laboratory), SHOTARO AKAHO (Mathematical Informatics Section, Information Science Division, Electrotechnical Laboratory), YUTAKA AKIYAMA and TATSUMI FURUYA (Computational Models Section, Computer Science Division, Electrotechnical Laboratory).

†† 電子技術総合研究所情報アーキテクチャ部計算機研究室

††† 電子技術総合研究所情報科学部情報数理研究室

* 現在 京都大学化学研究所

Institute for Chemical Research, Kyoto University

ネットワークは多層型であり、複素パターンが入力され、複素パターンが出力されることを想定している。ニューロン間の結合の重みおよび各ニューロンが持つしきい値はすべて複素数であり、各ニューロンの出力関数 $f_C: \mathbf{C}^1 \rightarrow \mathbf{C}^1$ は、

$$f_C(z) = f_R(x) + if_R(y), \quad z = x + iy, \quad (1)$$

ただし、 $f_R(u) = 1/(1 + \exp(-u))$ である (\mathbf{C} は複素数全体の集合)。すなわち、ニューロンの出力値の実部および虚部は、それぞれ、ニューロンへの入力の総和 z の実部 x 、虚部 y に関するシグモイド関数値である。また、学習アルゴリズムは基本的に教師パターンとの誤差を最小(極小)ならしめる方向に学習パラメータ(重みおよびしきい値)の修正を行う最急降下法である。学習収束性については、実 BP の理論的基礎である学習識別理論²⁾を複素数に拡張することによって確認されている。学習則は、実 BP のものと形が似通っており、文字どおり実 BP の複素数版であると言える。

以下では、実 BP ネットワークにおけるニューロンを実ニューロン、複素 BP ネットワークにおけるニューロンを複素ニューロンと呼ぶことにする。

3. 複素 BP ネットワークにおける決定表面の性質

本章では、主として複素 BP ネットワークにおける決定表面の性質について調べ、実 BP ネットワークのものとの比較を行う。

3.1 実ニューロンの重みパラメータ

まず、実ニューロンの入力の結合重みの基本的な構造について確認しておく。重み $w_k \in \mathbf{R}^1$ ($1 \leq k \leq n$) (\mathbf{R} は実数全体をあらわす) およびしきい値 $\theta \in \mathbf{R}^1$ を持つ n 入力実ニューロンを考える。出力関数 $f_R: \mathbf{R}^1 \rightarrow \mathbf{R}^1$ は $f_R(u) = 1/(1 + \exp(-u))$ とする。このとき、入力 $x_k \in \mathbf{R}^1$ ($1 \leq k \leq n$) が与えられると、実ニューロンは $f_R\left(\sum_{k=1}^n w_k x_k + \theta\right)$ を出力する。つまり、実ニューロンは実数直線(1次元)上の n 個の点 x_1, \dots, x_n を原点との距離に関して、それぞれ、 w_1, \dots, w_n 倍し、それらを1次元実ベクトルとみなして加えた結果得られるベクトルの示す点をさらに θ だけ移動させたと解釈することができる(図1)。その後、非線形変換 f_R を施したものを出力値としている。このように、実ニューロンは基本的に1次元上の点列の移動を司っており、重みパラメータ w_1, \dots, w_n は互いに独立であって、何らの影響も与え合わない。ここで、実ニューロ

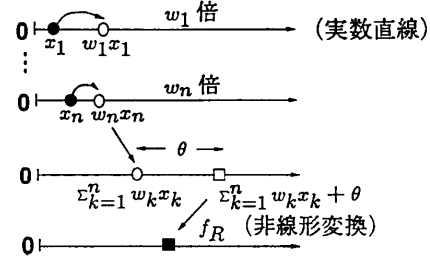


図1 実ニューロンにおける処理のイメージ
Fig. 1 An image of the processing in a neuron which is used in standard BP method.

ンの出力値は非線形変換 f_R が施されることによって得られるため、実ニューロンの出力値に関して上記の幾何学的解釈が厳密に成り立つのは非線形変換 f_R が線形変換とみなせるとき、すなわち、実ニューロンへの入力の総和の絶対値 $\left| \sum_{k=1}^n w_k x_k + \theta \right|$ が小さな値であるときに限ることに注意しなければならない。

3.2 複素ニューロンの重みパラメータ

次に、複素ニューロンの入力の結合重みの基本的な構造について調べる。重み $w_k = w_k^r + iw_k^i \in \mathbf{C}^1$ ($1 \leq k \leq n$) およびしきい値 $\theta = \theta^r + i\theta^i \in \mathbf{C}^1$ を持つ n 入力複素ニューロンを考える。入力 $x_k + iy_k \in \mathbf{C}^1$ ($1 \leq k \leq n$) が与えられると、この複素ニューロンの出力 $X + iY$ は、

$$\begin{aligned} X + iY &= f_C\left(\sum_{k=1}^n (w_k^r + iw_k^i)(x_k + iy_k) + (\theta^r + i\theta^i)\right) \\ &= f_R\left(\sum_{k=1}^n (w_k^r x_k - w_k^i y_k) + \theta^r\right) \\ &\quad + if_R\left(\sum_{k=1}^n (w_k^i x_k + w_k^r y_k) + \theta^i\right) \end{aligned} \quad (2)$$

となる。ゆえに、 n 入力複素ニューロンは図2に示す2つの $2n$ 入力実ニューロンと等価である。複素ニューロンの出力の実部 X に相当する実ニューロンを“実部”ニューロン、虚部 Y に相当する実ニューロンを“虚部”ニューロンと呼ぶことにする。ここで、

$$\begin{aligned} \begin{bmatrix} X \\ Y \end{bmatrix} &= F \left[\begin{array}{cc|cc} w_1^r & -w_1^i & \dots & w_n^r & -w_n^i \\ w_1^i & w_1^r & \dots & w_n^i & w_n^r \end{array} \begin{bmatrix} x_1 \\ y_1 \\ \vdots \\ x_n \\ y_n \end{bmatrix} + \begin{bmatrix} \theta^r \\ \theta^i \end{bmatrix} \right] \\ &= F \left(|w_1| \begin{bmatrix} \cos \alpha_1 & -\sin \alpha_1 \\ \sin \alpha_1 & \cos \alpha_1 \end{bmatrix} \begin{bmatrix} x_1 \\ y_1 \end{bmatrix} + \dots \right. \\ &\quad \left. + |w_n| \begin{bmatrix} \cos \alpha_n & -\sin \alpha_n \\ \sin \alpha_n & \cos \alpha_n \end{bmatrix} \begin{bmatrix} x_n \\ y_n \end{bmatrix} + \begin{bmatrix} \theta^r \\ \theta^i \end{bmatrix} \right), \end{aligned} \quad (3)$$

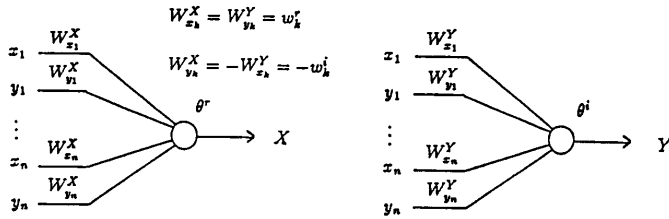


図 2 複素ニューロンと等価な実ニューロン
Fig. 2 Two neurons to be equivalent to a neuron which is used in complex BP method.

ただし、 $F\left(\begin{bmatrix} x \\ y \end{bmatrix}\right) = \begin{bmatrix} f_R(x) \\ f_R(y) \end{bmatrix}$, $\alpha_k = \arctan(w_k^i/w_k^r)$ ($1 \leq k \leq n$) である。

式(3)において、 $|w_k|$ は 2次元平面における点 (x_k, y_k) の原点との距離に関する縮小・拡大、 $\begin{bmatrix} \cos \alpha_k \\ \sin \alpha_k \end{bmatrix}$ は原点を中心とした正の方向 (反時計回り) への α_k 度の回転、 $[\theta^r \ \theta^i]$ (θ^r はベクトル x の転置をあらわす) は平行移動を意味する。このように、 n 入力複素ニューロンは、入力された 2次元情報 (複素数) それぞれに対して “運動” という線形変換を施している (図 3)。すなわち、式(3)は 2次元平面上の “運動” を基本的に含んでいると言える。ただし、複素ニューロンの出力値 $[X \ Y]$ は、非線形変換 F が施されることによって得られるため、 $[X \ Y]$ に関して上記の幾何学的解釈が厳密に成り立つのは、非線形変換 F が線形変換とみなせるとき、すなわち、複素ニューロンへの入力の総和 (複素数) の実部および虚部のそれぞれの絶対値が共に小さいときに限ることに注意しなければならない。

前節において見たように、実ニューロンは基本的に実数直線 (1次元) 上の点列の移動を司っており、実ニューロン間の重みパラメータは互いに独立であって、何らの影響も与え合わない。複素ニューロンは先に見たように基本的に 2次元平面上の “運動” を司

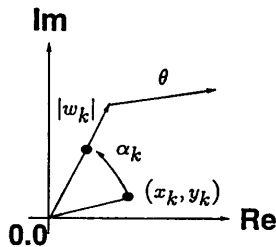


図 3 複素信号に対する 2次元運動のイメージ
Fig. 3 An image of the two dimensional motion for complex signals.

ており、学習においてはその運動の調整を行うものと解釈することができる。それに伴って、複素ニューロンには、

- (“実部”ニューロンへの入力値の実部 x_k に対する重み)
- = (“虚部”ニューロンへの入力値の虚部 y_k に対する重み),
- (“実部”ニューロンへの入力値の虚部 y_k に対する重み)
- = - (“虚部”ニューロンへの入力値の実部 x_k に対する重み)

という制約が重みに課されており (図 2 参照)、学習はその制約を保ちつつ行われる。見方を変えると “実部”ニューロンと “虚部”ニューロンはそれらの重みパラメータを通じて互いに影響を与え合っていると言える。このように、基本的に、数直線 (1次元) 上の移動を司っていた実 BP は、複素数に拡張されることにより、平面 (2次元) 上の運動に関した構造を有するものへと変質していることがわかった。次節で述べるように、重みパラメータのこの構造は決定表面の直交性として現れる。

3.3 複素 BP ネットワークにおける決定表面の直交性

決定表面 (decision boundary) とは、学習識別系 (たとえば、実 BP) がパターンを切り分ける境界線のことであり、一般に超曲面から成るものである。実ニューロンから構成されるニューラルネットワークの決定表面については、Lippmann により、中間層の個数が 0~2 個の 3 種類に分けて実験的に調べられている⁹⁾。そこで、本節では複素 BP ネットワークの決定表面について解析的に調べる。

3.3.1 単一複素ニューロンの場合

まず、単一の複素ニューロン (中間層 0 個) の決定表面について調べる。

重みを $w = [w_1 \dots w_n] = w^r + iw^i$, $w^r = [w_1^r \dots w_n^r]$, $w^i = [w_1^i \dots w_n^i]$, しきい値を $\theta = \theta^r + i\theta^i$ とする。出力関数 $f_1: C^1 \rightarrow C^1$ は、簡単のために、

$$f_1(z) = 1(x) + i1(y), \quad z = x + iy, \quad (4)$$

ただし、 $1(u) = \begin{cases} 1 & (\text{if } u \geq 0) \\ 0 & (\text{if } u < 0) \end{cases}$ とする。このとき、複素パターン $z = [z_1 \dots z_n] = x + iy$, $x = [x_1 \dots x_n]$, $y = [y_1 \dots y_n]$ が入力されると複素ニューロンは、

$$1\left([\begin{matrix} w^r & -w^i \end{matrix}] \begin{bmatrix} x \\ y \end{bmatrix} + \theta^r\right)$$

$$+i1\left([\text{'w}^i \text{'w}^r]\begin{bmatrix} x \\ y \end{bmatrix} + \theta^i\right) \quad (5)$$

を出力する。ゆえに、複素ニューロンの出力値の実部は $[\text{'w}^i \text{'w}^r]$ を法線ベクトルとする超平面 (決定表面) $[\text{'w}^i \text{'w}^r]\begin{bmatrix} x \\ y \end{bmatrix} + \theta^i = 0$ を境にして定まる。

つまり、超平面 ≥ 0 ならば 1, そうでなければ 0 となる。同様に、複素ニューロンの出力値の虚部は $[\text{'w}^i \text{'w}^r]$ を法線ベクトルとする超平面 (決定表面) $[\text{'w}^i \text{'w}^r]\begin{bmatrix} x \\ y \end{bmatrix} + \theta^i = 0$ を境にして定まる。ここで、これら 2 つの決定表面の法線ベクトルの内積をとると $[\text{'w}^i \text{'w}^r]\begin{bmatrix} \text{'w}^i \\ \text{'w}^r \end{bmatrix} = 0$ となり、2 つの決定表面は直交しているということがわかる。

次に、式(1)で示される出力関数を持つ場合について調べる。複素パターン $z = [z_1 \dots z_n] = x + iy$, $x = [x_1 \dots x_n]$, $y = [y_1 \dots y_n]$ が入力されると複素ニューロンは、

$$X + iY = f_R\left([\text{'w}^i \text{'w}^r]\begin{bmatrix} x \\ y \end{bmatrix} + \theta^i\right) + if_R\left([\text{'w}^i \text{'w}^r]\begin{bmatrix} x \\ y \end{bmatrix} + \theta^i\right) \quad (6)$$

を出力する。ここで、常数 $C^R, C^I \in (0, 1)$ を任意に選び、

$$X(x, y) = f_R\left([\text{'w}^i \text{'w}^r]\begin{bmatrix} x \\ y \end{bmatrix} + \theta^i\right) = C^R, \quad (7)$$

$$Y(x, y) = f_R\left([\text{'w}^i \text{'w}^r]\begin{bmatrix} x \\ y \end{bmatrix} + \theta^i\right) = C^I \quad (8)$$

とおく。式(7)は n 入力複素ニューロンの出力値の実部に関する決定表面である。すなわち、入力信号 $(x, y) \in \mathbb{R}^{2n}$ は、超曲面 (式(7)) を境界線にして、2 つの決定領域 (decision region) $\{(x, y) \in \mathbb{R}^{2n} | X(x, y) \geq C^R\}$, $\{(x, y) \in \mathbb{R}^{2n} | X(x, y) < C^R\}$ に分類される。同様に、式(8)で示される超曲面は n 入力複素ニューロンの虚部に関する決定表面である。そこで、式(7)および式(8)のそれぞれの法線ベクトル $H^R(x, y)$, $H^I(x, y)$ を求めると、

$$H^R(x, y) = \left(\frac{\partial X}{\partial x_1} \dots \frac{\partial X}{\partial x_n} \frac{\partial X}{\partial y_1} \dots \frac{\partial X}{\partial y_n} \right) = f'_R\left([\text{'w}^i \text{'w}^r]\begin{bmatrix} x \\ y \end{bmatrix} + \theta^i\right) \cdot [\text{'w}^i \text{'w}^r], \quad (9)$$

$$H^I(x, y) = \left(\frac{\partial Y}{\partial x_1} \dots \frac{\partial Y}{\partial x_n} \frac{\partial Y}{\partial y_1} \dots \frac{\partial Y}{\partial y_n} \right)$$

$$= f'_R\left([\text{'w}^i \text{'w}^r]\begin{bmatrix} x \\ y \end{bmatrix} + \theta^i\right) \cdot [\text{'w}^i \text{'w}^r] \quad (10)$$

となり、それらの内積は 0 となるから、実部に関する決定表面と虚部に関する決定表面は互いに直交していることがわかる。

実ニューロンは入力された実パターンを 2 つの (0 および 1 という値で代表される) クラスに分類する。一方、複素ニューロンは入力された複素パターンを 4 つの (0, 1, i , $1+i$ という値で代表される) クラスに分類する。先の議論により、複素ニューロンの決定表面は 2 つの直交した超曲面 (超平面) から構成されており、領域を 4 つに均等に分割しようとする。つまり、複素ニューロンは複素パターンに対する自然な決定表面を有していると言える。

3.3.2 3層ネットワークの場合

次に、3層ネットワーク (中間層 1 個) の決定表面について調べる。入力ニューロン L 個、中間ニューロン M 個、出力ニューロン N 個の 3層ネットワークを考える。 $w_{ji} = w_{ji}^r + iw_{ji}^i$ を入力ニューロン i と中間ニューロン j との間の結合の重み、 $v_{kj} = v_{kj}^r + iv_{kj}^i$ を中間ニューロン j と出力ニューロン k との間の結合の重み、 $\theta_j = \theta_j^r + i\theta_j^i$ を中間ニューロン j が持つしきい値、 $\gamma_k = \gamma_k^r + i\gamma_k^i$ を出力ニューロン k が持つしきい値とする。出力関数は、式(1)で示されるものとする。このとき、複素パターン $z = [z_1 \dots z_L] = x + iy$, $x = [x_1 \dots x_L]$, $y = [y_1 \dots y_L]$ が入力されると、中間ニューロン j への入力の総和 U_j は、

$$U_j = U_j^r + iU_j^i = \left[\sum_{i=1}^L (w_{ji}^r x_i - w_{ji}^i y_i) + \theta_j^r \right] + i \left[\sum_{i=1}^L (w_{ji}^i x_i + w_{ji}^r y_i) + \theta_j^i \right] \quad (11)$$

となる。よって、その出力 H_j は、

$$H_j = H_j^r + iH_j^i = f_R(U_j^r) + if_R(U_j^i) \quad (12)$$

である。さらに、出力ニューロン k への総入力 S_k は、

$$S_k = S_k^r + iS_k^i = \left[\sum_{j=1}^M (v_{kj}^r H_j^r - v_{kj}^i H_j^i) + \gamma_k^r \right] + i \left[\sum_{j=1}^M (v_{kj}^i H_j^r + v_{kj}^r H_j^i) + \gamma_k^i \right] \quad (13)$$

である。ゆえに、その出力 O_k は、

$$O_k = O_k^r + iO_k^i = f_R(S_k^r) + if_R(S_k^i) \quad (14)$$

となる。ここで、常数 $C^R, C^I \in (0, 1)$ を任意に選び、

$$O_k^r(x, y) = C^R, \quad (15)$$

$$O_i(\mathbf{x}, \mathbf{y}) = C_i \quad (16)$$

とおく。式(15)および式(16)は、それぞれ、3層の複素 BP ネットワークにおける出力ニューロン k の実部および虚部に関する決定表面である。これら2つの超曲面の法線ベクトル $H^R(\mathbf{x}, \mathbf{y})$, $H^I(\mathbf{x}, \mathbf{y})$ は、それぞれ、

$$H^R(\mathbf{x}, \mathbf{y}) = \left(\frac{\partial O_k}{\partial x_1} \dots \frac{\partial O_k}{\partial x_L} \frac{\partial O_k}{\partial y_1} \dots \frac{\partial O_k}{\partial y_L} \right), \quad (17)$$

$$H^I(\mathbf{x}, \mathbf{y}) = \left(\frac{\partial O_k}{\partial x_1} \dots \frac{\partial O_k}{\partial x_L} \frac{\partial O_k}{\partial y_1} \dots \frac{\partial O_k}{\partial y_L} \right) \quad (18)$$

であり、これらの内積は、

$$\begin{aligned} H^R(\mathbf{x}, \mathbf{y}) \cdot H^I(\mathbf{x}, \mathbf{y}) &= \frac{\partial O_k}{\partial x_1} \frac{\partial O_k}{\partial x_1} + \dots + \frac{\partial O_k}{\partial x_L} \frac{\partial O_k}{\partial x_L} \\ &\quad + \frac{\partial O_k}{\partial y_1} \frac{\partial O_k}{\partial y_1} + \dots + \frac{\partial O_k}{\partial y_L} \frac{\partial O_k}{\partial y_L} \end{aligned} \quad (19)$$

で与えられる。ここで、任意の $1 \leq i \leq L$ に対して、

$$\begin{aligned} &\frac{\partial O_k}{\partial x_i} \frac{\partial O_k}{\partial x_i} + \frac{\partial O_k}{\partial y_i} \frac{\partial O_k}{\partial y_i} \\ &= \frac{\partial f_R(S_k)}{\partial S_k} \frac{\partial f_R(S_k)}{\partial S_k} \\ &\quad \cdot \left[\sum_{j=1}^M \left(v_{k,j}^R w_{j,i}^R \frac{\partial f_R(U_j^R)}{\partial U_j^R} - v_{k,j}^I w_{j,i}^I \frac{\partial f_R(U_j^I)}{\partial U_j^I} \right) \right] \\ &\quad \cdot \left[\sum_{j=1}^M \left(v_{k,j}^I w_{j,i}^I \frac{\partial f_R(U_j^I)}{\partial U_j^I} + v_{k,j}^R w_{j,i}^R \frac{\partial f_R(U_j^R)}{\partial U_j^R} \right) \right] \\ &\quad - \frac{\partial f_R(S_k)}{\partial S_k} \frac{\partial f_R(S_k)}{\partial S_k} \\ &\quad \cdot \left[\sum_{j=1}^M \left(v_{k,j}^I w_{j,i}^I \frac{\partial f_R(U_j^I)}{\partial U_j^I} - v_{k,j}^R w_{j,i}^R \frac{\partial f_R(U_j^R)}{\partial U_j^R} \right) \right] \\ &\quad \cdot \left[\sum_{j=1}^M \left(v_{k,j}^R w_{j,i}^R \frac{\partial f_R(U_j^R)}{\partial U_j^R} + v_{k,j}^I w_{j,i}^I \frac{\partial f_R(U_j^I)}{\partial U_j^I} \right) \right]. \end{aligned} \quad (20)$$

よって、法線ベクトルの内積は必ずしも0になるわけではなく、実部と虚部のそれぞれの決定表面は互いに単純に直交するとは言えない。しかしながら、式(20)を注意深くみると、任意の $1 \leq j \leq M$ に対して、

$$\frac{\partial f_R(U_j^R)}{\partial U_j^R} = \frac{\partial f_R(U_j^I)}{\partial U_j^I} \quad (21)$$

すなわち、

$$\begin{aligned} f_k^R \left[\sum_{i=1}^L (w_{j,i}^R x_i - w_{j,i}^I y_i) + \theta_j^R \right] \\ = f_k^I \left[\sum_{i=1}^L (w_{j,i}^I x_i + w_{j,i}^R y_i) + \theta_j^I \right] \end{aligned} \quad (22)$$

が成り立つならば、内積=0となることがわかる。一般に、 $f_k^R(u_1)$ と $f_k^I(u_2)$ は $|u_1|$ および $|u_2|$ がある

大きな値以上であれば、ほぼ同じ値をとるものとみなせる。ゆえに、任意の $1 \leq j \leq M$ に対して、十分大きな正数 K_1, K_2 が存在して、

$$\begin{aligned} |U_j^R| &= \left| \sum_{i=1}^L (w_{j,i}^R x_i - w_{j,i}^I y_i) + \theta_j^R \right| \\ &= \left| \sum_{i=1}^L |w_{j,i}| |z_i| \cos(\alpha_{j,i} + \beta_i) + \theta_j^R \right| \\ &> K_1, \end{aligned} \quad (23)$$

$$\begin{aligned} |U_j^I| &= \left| \sum_{i=1}^L (w_{j,i}^I x_i + w_{j,i}^R y_i) + \theta_j^I \right| \\ &= \left| \sum_{i=1}^L |w_{j,i}| |z_i| \sin(\alpha_{j,i} + \beta_i) + \theta_j^I \right| \\ &> K_2, \end{aligned} \quad (24)$$

ただし、 $\tan(\alpha_{j,i}) = w_{j,i}^I / w_{j,i}^R$, $\tan(\beta_i) = y_i / x_i$ が成り立つとき、2つの決定表面は互いに直交する。つまり、任意の $1 \leq j \leq M$ に対して、中間ニューロン j への総入力(複素数)の実部および虚部の絶対値が共に十分大きいとき、2つの決定表面は互いに直交する。あるいは、さらに、 $|z| = \sqrt{|z_1|^2 + \dots + |z_L|^2}$ が十分大きい値をとる (\mathbf{x}, \mathbf{y}) において、2つの決定表面は互いに直交する可能性が高いと言える。

このことを定理および系の形でまとめておく。

定理

すべての中間ニューロンへの総入力(複素数)の実部および虚部の絶対値が共に十分大きいとき、実部と虚部のそれぞれの決定表面は互いに直交する。

系

$|z|$ が十分大きい値をとる (\mathbf{x}, \mathbf{y}) において、実部と虚部のそれぞれの決定表面は互いに直交する可能性が高い。

4. シミュレーション

本章では、3層の複素 BP ネットワークにおける決定表面の特性について計算機実験を通じて調べ、実 BP ネットワークのものとの比較を行う。出力関数は、シグモイド関数をベースにした複素関数(式(1))である。

実験には、表1に示す学習パターンを使い、学習パラメータの初期値は -0.3 から $+0.3$ の間の乱数により設定する。学習は誤差

$$\sqrt{\sum_{k=1}^N |T_k - O_k|^2} \quad (25)$$

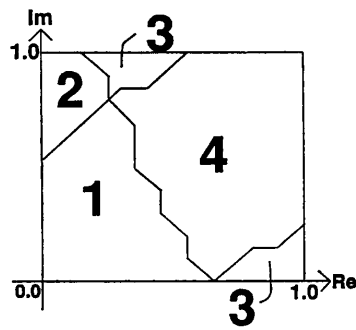
が 0.10 となったときに収束(終了)したものとみなし、収束の規範とする(ただし、 $O_k, T_k \in \mathbb{C}$ は、それ

表 1 学習パターン
Table 1 Learning patterns.

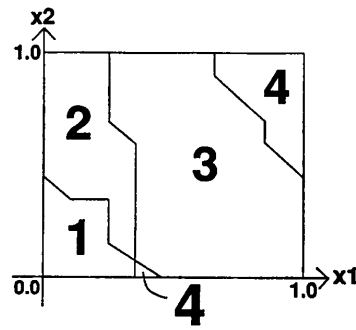
パターン番号	入力パターン	出力パターン
1	0	0
2	i	1
3	1	$1+i$
4	$1+i$	i

ぞれ、出力ニューロン k の出力値およびその教師信号であり、 N は出力ニューロンの個数をあらわす。また、表 1 に示す 4 種類のパターンをひと通り提示することで 1 回の学習と数えることとする。

ネットワークは、入力 1 出力 1、中間ニューロンは 3 から 6 までの 4 種類、学習率は 0.5 とした。学習終



(a) 複素 BP
(a) Complex BP.



(b) 実 BP
(b) Standard BP.

図 4 決定表面 (3 中間ニューロン)

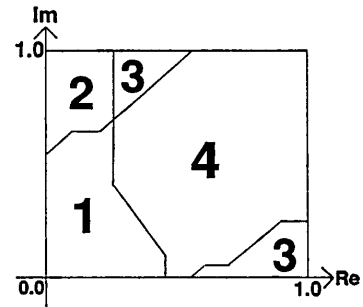
図 4 ~ 8 における数字の意味

- 1: 実部 0.0~0.5, 虚部 0.0~0.5
- 2: 実部 0.5~1.0, 虚部 0.0~0.5
- 3: 実部 0.5~1.0, 虚部 0.5~1.0
- 4: 実部 0.0~0.5, 虚部 0.5~1.0

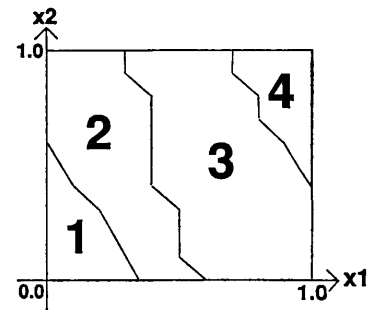
Fig. 4 Decision boundary (three intermediate neurons).

Meaning of numerals in Figs. 4-8.

- 1: Real part 0.0~0.5, Imaginary part 0.0~0.5
- 2: Real part 0.5~1.0, Imaginary part 0.0~0.5
- 3: Real part 0.5~1.0, Imaginary part 0.5~1.0
- 4: Real part 0.0~0.5, Imaginary part 0.5~1.0



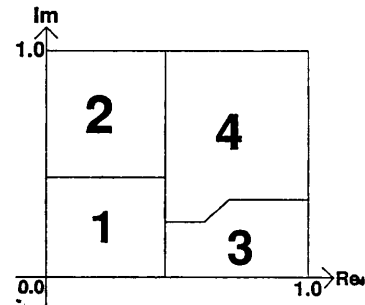
(a) 複素 BP
(a) Complex BP.



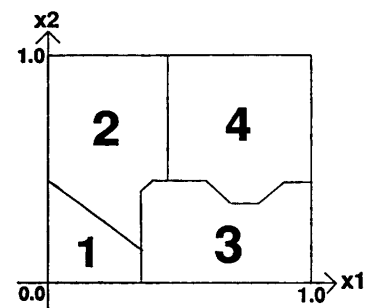
(b) 実 BP
(b) Standard BP.

図 5 決定表面 (4 中間ニューロン)

Fig. 5 Decision boundary (four intermediate neurons).



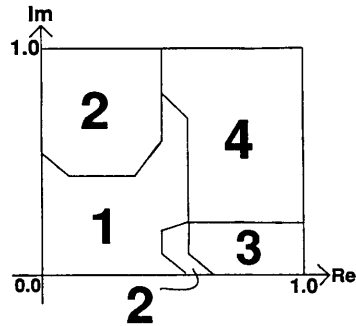
(a) 複素 BP
(a) Complex BP.



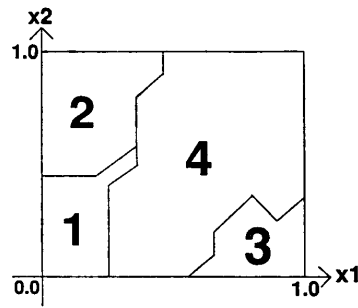
(b) 実 BP
(b) Standard BP.

図 6 決定表面 (5 中間ニューロン)

Fig. 6 Decision boundary (five intermediate neurons).



(a) 複素 BP
(a) Complex BP.

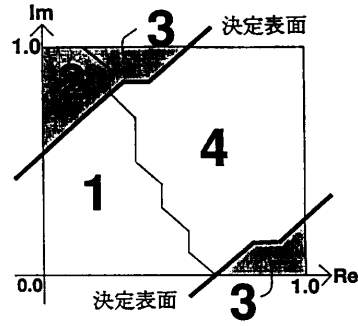


(b) 実 BP
(b) Standard BP.

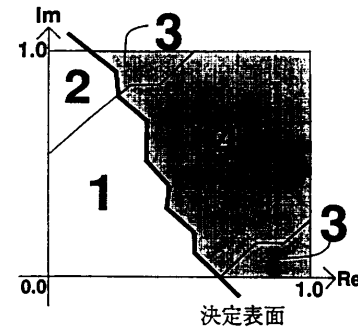
図7 決定表面 (6 中間ニューロン)
Fig. 7 Decision boundary (six intermediate neurons).

了後、テストパターンとして、実部、虚部、共に、0.0 から 1.0 まで 0.1 刻みの値 (合計 $11 \times 11 = 121$ パターン) を入力した。その結果を図 4 (a)~図 7 (a) に示す。また、比較のために行った実 BP の結果が図 4 (b)~図 7 (b) である。複素 BP と異なるのは入力 2 出力 2 というネットワークを使い、入出力データの実部を入出力ニューロン 1、虚部を入出力ニューロン 2 に対応させた点である。図 4~図 7 において、領域 1 はテストパターンに対するニューラルネットワークの出力の実部がオフ (0.0~0.5) かつ虚部がオフであったことを示す。同様に、領域 2 は実部オン (0.5~1.0)、虚部オフ；領域 3 は実部オン、虚部オン；領域 4 は実部オフ、虚部オンであったことを示す。

図 4~図 7 において、複素 BP ネットワークの決定表面は実 BP ネットワークのものとは比べて見るとある一定の構造を持っているように見受けられる。特に、図 4 (a) は実部の決定表面 (領域 2 および領域 3 が成す領域とその他の領域との境界線) と虚部の決定表面 (領域 3 および領域 4 が成す領域とその他の領域との境界線) とが直交する傾向が見られ (図 8 参照), 3.3



(a) 出力の実部がオンである領域 (領域 2+領域 3)
(a) Region in which the real part of the output value of the neural network is "ON" (0.5~1.0) (region 2+region 3).



(b) 出力の虚部がオンである領域 (領域 3+領域 4)
(b) Region in which the imaginary part of the output value of the neural network is "ON" (0.5~1.0) (region 3+region 4).

図 8 決定表面の直交性
Fig. 8 Orthogonality of decision boundary.

節において示した複素 BP ネットワークの決定表面の直交性の性質が反映されているものと思われる。

5. おわりに

複素 BP の基本特性を、ネットワークアーキテクチャの観点から解析的、実験的に調べ、実 BP との差異を明らかにした。

複素 BP には、複素数への拡張の結果として、次のような性質が生じている。複素 BP ネットワークにおける重みパラメータは互いに独立ではなく 2 次元運動を基づいた制約を有しており、学習は基本的にその制約のもとで進められる。複素ニューロンの決定表面は 2 つの直交した超曲面 (超平面) から構成されており、決定領域を均等に 4 つに分割しようとする。これは、0, 1, i , $1+i$ という 4 つのクラスを扱う複素パターンに適した性質である。3 層の複素 BP ネットワークにおける決定表面は基本的にこの性質を内包しており、中間ニューロンへの総入力が十分大きいとき

に直交する。

この意味で複素 BP は複素パターンに対する自然な学習アルゴリズムであり、複素パターンを扱う分野において有効に利用され得るものと期待される。

今後の課題としては、重みパラメータに課せられている 2 次元運動に基づいた制約と 2 次元運動学習能力 (図形変換能力)^{4),6)}との関係の明確化、画像処理および音声認識などへの本格的な応用、学習アルゴリズムの高速化、複素関数の近似能力の解析などが考えられる。また、複素 BP をさらに高次元化した場合に、どのような特性が現れるかについても興味のあるところである。

謝辞 本研究の機会を与えてくださった情報アーキテクチャ部 弓場部長、情報科学部 田村部長、討論に参加いただいた計算機構研究室諸氏、また、有益なコメントをくださった査読者の方々に感謝いたします。特に、3.3 節で述べた結果のうち、シグモイド関数から構成される出力関数を持つ複素 BP ネットワークにおける決定表面に関する結果は、査読者の方のコメントを元にして得られたものです。末筆ながら感謝の意を表します。

参 考 文 献

- 1) 麻生: ニューラルネットワーク情報処理, p. 198, 産業図書, 東京 (1988).
- 2) 甘利: 学習識別の理論, 電子通信学会誌, Vol. 50, No. 7, pp. 1272-1279 (1967).
- 3) 佐竹, 古谷, 下原: ニューラルネットワークによる米の品質判定, 農業施設学会研究発表会講演要旨, pp. 3-4 (1990).
- 4) 新田, 古谷: 複素バックプロパゲーション学習, 情報処理学会論文誌, Vol. 32, No. 10, pp. 1319-1329 (1991).
- 5) 新田, 赤穂, 秋山, 古谷: 複素 BP の基本的性質, 第 43 回情報処理学会全国大会論文集 (分冊 2), 6 G-8, pp. 223-224 (1991).
- 6) 新田, 古谷: 複素 BP の汎化能力, 第 43 回情報処理学会全国大会論文集 (分冊 2), 6 G-9, pp. 225-226 (1991).
- 7) 新田, 赤穂, 秋山, 古谷: 複素 BP における決定表面の構造, 信学技報, NC 91-62, pp. 105-109 (1991).

8) Lippmann, R. P.: An Introduction to Computing with Neural Nets, *IEEE Acoustic, Speech and Signal Processing Magazine*, pp. 4-22 (Apr. 1987).

9) Rumelhart, D. E. et al.: *Parallel Distributed Processing*, Vol. 1, p. 547, MIT press (1986).

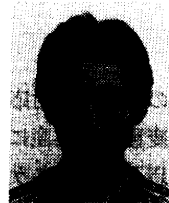
(平成 4 年 1 月 16 日受付)

(平成 4 年 9 月 10 日採録)



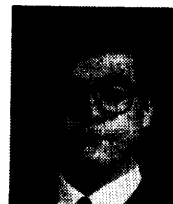
新田 徹 (正会員)

1960 年生. 1983 年筑波大学第 1 学群自然科学類卒業. 1985 年同大学院修士課程修了. 同年日本電気(株)入社. 1990 年同社退社. 同年電子技術総合研究所入所. ニューラルネットワーク, 資源配分理論の研究に従事. 日本 OR 学会会員.



赤穂昭太郎

1965 年生. 1988 年東京大学工学部計数工学科卒業. 1990 年同大学院修士課程修了. 同年, 電子技術総合研究所に入所. 神経回路網などの数理的研究に従事. 神経回路学会, 電子情報通信学会各会員.



秋山 泰 (正会員)

昭和 36 年生. 平成 2 年慶応義塾大学大学院博士課程修了. 工学博士. 同年電子技術総合研究所入所. 平成 4 年より京都大学化学研究所生体分子情報研究部門助教授. IEEE, 電子情報通信学会, 神経回路学会, 生物物理学会, 分子生物学会等各会員.



古谷 立美 (正会員)

昭和 22 年生. 昭和 48 年成蹊大学大学院修士課程修了. 同年電子技術総合研究所入所. 現在同所計算機構研究室長. 工学博士. IEEE 会員.