

総合解説

ニューラルネットワークの応用と今後の発展

Application of Neural Networks and Its Prospect

1. 核融合・プラズマ研究への応用に関する全般的解説

竹田 辰興

電気通信大学

(原稿受付：2006年1月25日)

核融合・プラズマ研究へのニューラルネットワークの応用について全般的に解説する。特に、理工学分野の研究開発にニューラルネットワークが利用されるに至った背景に重点を置いて、今後の応用分野の拡大なども考慮に入れて解説する。

Keywords:

neural network, error back propagation, generalization, time series prediction, computerized tomography, residual minimization training

1.1 はじめに

ニューラルネットワークの核融合・プラズマ研究への応用に関しては、2002年に本誌に「プラズマ実験解析・制御へのニューラルネットワークの応用」という解説記事を掲載した[1]。当初、ニューラルネットワークの応用は「従来法と同等の結果が比較的簡単に得られる」というレベルのもが多かったが、現在では「従来法より優れた結果が得られる」というレベルのものも多くなった。今回は、理論的話題と具体的に良好な結果を出している応用例について総合解説を組んだ。全般的解説に関しては包括的理解を助けるために前回の解説と重複した内容も含んでいる。

1.1.1 ニューラルネットワークとは

本総合解説で扱う「ニューラルネットワーク（神経回路網）」は、最初、脳の働きを解明するモデルとして考えられたシステムで「生物学的」ニューラルネットワークそのものではない。正しくは「人工ニューラルネットワーク（Artificial Neural Network）」と呼ぶべきであろう。しかし、「生物学的」ニューラルネットワークと混同する心配もないと考えられるので、以下の記述では「人工」を取り除いて、単に、ニューラルネットワークと呼ぶ。ニューラルネットワークの当初の研究の動機は、このように生物の脳の機能と動作の解明にあったが、実際には、始めから、これ以外にもフィルター理論、多変量解析、統計数学、関数近似、最適化等々理工学の諸分野と深い関係を保って発展を遂げて

author's e-mail: takeda@im.uec.ac.jp

きた。現在でも、ニューラルネットワークの研究はこれらの分野にまたがった研究である。

脳の機能との関係で最初に大きな期待が持たれたニューラルネットワークは1960年に発表されたRosenblattによるパーセプトロン[2]であるが、1969年にMinsky & Papertによってパーセプトロンの機能についての悲観的示唆が与えられた[3]ため研究はやや下火になった。この後もニューラルネットワークの研究者による重要な研究がいろいろとなされているが、1980年代中期にホップフィールドネットワークを使った物理システムの解析[4]と多層ニューラルネットワークのための学習アルゴリズム(Error Back-Propagation Method：誤差逆伝播法)の再発見[5]がなされるに至ってニューラルネットワーク研究に大きな転機が訪れた。この頃から、ニューラルネットワークそのものを対象とした研究だけでなくニューラルネットワークを応用する研究が爆発的に増加する。

ニューラルネットワークとは何であるかということを厳密に定義することは困難である。最も標準的には、複数入力で1出力の単純な計算ユニットであるニューロンが複数個集まって構成されたネットワークで、「学習」と呼ばれるプロセスによりその機能を発揮するものがニューラルネットワークであると考えてよいであろう。上記のホップフィールドネットワークや多層ニューラルネットワークはこのようなネットワークの典型例である。ニューラルネッ

トワークの進化は著しく、毎週のように新しいネットワークや改良版ネットワークが提案されているとも言われている。

ニューラルネットワークの重要な性質の一つが「学習する写像装置」であるが、このような性質を持つ代表的なニューラルネットワークが本総合解説の主要なテーマである多層ニューラルネットワークである。特に、中間層が一層だけでその中間層を構成する J 個のニューロンの活性化関数がシグモイド型の関数であるような3層ニューラルネットワークを考えれば、ニューラルネットワークによる写像関数を具体的に次のように表すことができる。

$$z_k = \sum_{j=1}^J w_{jk} \sigma \left(\sum_{i=1}^I v_{ij} x_i + v_{0j} \right) + w_{0k}, \quad k = 1, 2, \dots, K \quad (1)$$

$$\sigma(x) = \frac{1}{1 + e^{-x}} \quad (2)$$

ここで、 x_i は i 番目の入力データ、 z_k は k 番目の出力層ニューロンからの出力データ、 v_{ij} および w_{jk} は入力層-中間層間、および中間層-出力層間の結合路に割り当てられたウエイトで、 σ は中間層ニューロンの活性化関数(ここではロジスティック関数)である。なお、ここでは出力層ニューロンは活性化関数を持たないものとした。また、 v_{0j} および w_{0k} はバイアスまたはオフセットと呼ばれロジスティック曲線の立ち上がりの位置を定める変数で学習の際はウエイトと同様に扱うことができる。入力層、中間層、出力層のニューロンの数はそれぞれ、 I 、 J 、 K である。学習過程とは v_{ij} および w_{jk} を決めるプロセスである。ニューラルネットワークの出力データ z_k が入力データ x_i とウエイトベクトル \vec{v} 、 \vec{w} の関数 $z_k = f_k(\vec{x}; \vec{v}, \vec{w})$ として解析的に表されていることは特に重要である。

実験や観測で得られたデータを処理する過程で、 N 個の測定データ $(x, y) = (x_i, y_i)$ ($i = 1, 2, \dots, N$) について (x, y) の関係 $y = f(x)$ を求めることがしばしば必要になる。この関係は、多項式や三角多項式あるいは他の基底関数系を用いた多項式で近似することが可能である。一方、この関係を3層ニューラルネットワークによって近似することも可能である。前者の場合、近似関数はあらかじめ定まった基底関数の線形結合で表されるのに対して、式(1)から容易にわかるように、ニューラルネットワークの場合には各基底関数は学習の結果定まるシグモイド関数である。このために、前者は「固定基底関数系による関数近似」、後者は「可変基底関数系による関数近似」と呼ばれる。直交関数系等の固定基底関数系による関数近似によって任意の関数が任意の精度で近似できるのと同様、多層ニューラルネットワークによっても任意の関数が任意の精度で近似できることが証明されている[6]。問題は、ニューラルネットワークによる関数近似が固定基底関数によるものに比べて優れているかどうかということである。この問題に理論的に答えることは必ずしも容易ではない。この点については多くの研究がなされているが、本総合解説では第2章の萩原による解説[7]において議論されている。理論的証明は未だ不十分であるとはいえ、経験的・実験的にはニューラルネッ

トワークによる関数近似が固定基底関数系による関数近似よりもよい結果を与えることがしばしば報告されている[8]。特に実際的な問題では、多次元であって測定点が十分とれない場合(「次元の呪い」の克服)や測定データにノイズが含まれておりオーバーフィッティングを避けることが重要な課題に対してはニューラルネットワークの優位性が理論的にも示唆されている[9]。

核融合・プラズマ研究の分野でも1990年頃からニューラルネットワークの応用がなされているが他の分野に比べて特に多いとは思えない。物理科学系の分野に限るならば原子核・素粒子実験や天体物理・天文学の分野では実験データや観測データの解析のためにニューラルネットワークの応用研究が進んでおり、種々の研究会が開かれ雑誌の特集号なども発行されている[10]。これは、扱うデータの性格の違いによると思われるが、ただ単にニューラルネットワークの理解の程度の違いのためかもしれない。ニューラルネットワークを使うことに懐疑的な立場からの議論として次の2点がよく挙げられる。則ち、(1)直交関数系などを使った関数近似で十分なのにニューラルネットワークを使う必要はないのではないか、ということと(2)関数近似した結果はブラックボックス的で「解析」したことにならないのではないか、ということである。前者については、上に述べたように「ニューラルネットワークがより良く関数近似できる」ことが経験的事実として多数報告されており、部分的には理論的にも証明されていることが答えになる。後者については、確かに、ニューラルネットワークによる関数近似からはフーリエ展開で得られる周波数スペクトルのようなわかりやすい解析結果は得られないが、可変基底関数系とはもともとそういうもので、ニューラルネットワークによって「複雑なものを複雑なままに解析している」と答えることができる。現在、着実な研究が精力的に行われており理論的にもかなりの進展が見られているので、今後色々な応用分野でニューラルネットワークは有効に利用されるものと思われる。

1.1.2 ニューラルネットワークの種類と機能

既に述べたようにニューラルネットワークの種類は多岐にわたり、また、ニューラルネットワークと呼ぶべきか否かが微妙なシステムも少なくない。最近では、ニューラルネットワーク関連の優れた教科書も数多く出版されているので、詳細はこれらにゆずる。ニューラルネットワーク全般については[11, 12]等、学習、汎化の問題を中心としたニューラルネットワークの数理的側面に関しては[13, 14]等がある。ここでは、基本的な性質にのみ着目してその種類と機能について記す。まず、最初に、ネットワーク構造の観点から見れば、「多層ニューラルネットワーク」と「相互結合型ニューラルネットワーク」に分類することができる。多層ニューラルネットワークにおいては、ニューロンはいくつかの層の上に配置され、これらの層は入力層、中間層(隠れ層、複数)、出力層の順序で並べられている。ある層のニューロンは隣の層のニューロンすべてと結合しておりこれ以外には結合路を持たない。各結合路にはウエイト値が割り当てられている。データは入力側から出力側に

向かって順次計算されていく。各ニューロンに入力されるデータは結合路のウェイトを使って重み付き和が計算され活性化関数によって非線形変換された後出力される。相互結合型ニューラルネットワークは、ニューロン相互の結合に層構造を持たないようなネットワークである。1つのニューロンが他のすべてのニューロンと結合しているものを完全結合ニューラルネットワークというが、ホップフィールドニューラルネットワークはこのネットワークの一例である。多層ニューラルネットワークにフィードバック結合路のついたものもこの範疇に入る。もっとも、多層ニューラルネットワークを構成した後に構造最適化を行い、一部の結合路や一部のニューロンを削除するという場合もあるので、この分類の定義は必ずしも厳密ではない。

次に重要な分類は、学習を行う際に、既知の正しい出力データ（教師データ）を使うか（教師あり学習）、そのようなデータを使わないか（教師なし学習）という区別である。具体的問題を分析してみるとわかるが、実は、「教師あり学習」と「教師なし学習」の違いが判然としない場合も存在する。単純にその特徴をまとめれば、「教師あり学習」では、学習の結果、答を教えてない入力データに対しては内挿によって答を出し、「教師なし学習」では入力データとして与えた情報だけから入力パターン（一組の入力データ）をいくつかのクラスターに分類する目的で使うことが多い。

他にも、いろいろな観点からニューラルネットワークの分類を行うことが可能であるが省略する。

ニューラルネットワークの最も基本的な性質は、上に述べたように、学習によってデータ $((\vec{x}, \vec{y}); \vec{x}: \text{入力データ}, \vec{y}: \text{出力データ})$ から関数関係を構築できることである。この時、入力データ、出力データともに、連続値でも離散値でもよい。入出力変数やウェイトが複素数値をとる複素ニューラルネットワーク [15] も最近注目を集めている。写像が複素関数で表されるような問題については効率的に扱うことができる。このネットワークでは活性化関数の選び方に注意が必要である。ニューラルネットワークの優れた機能を示す時によく引用される例に Sejnowsky らによる NETtalk がある [16, 17]。これは、アルファベット 7 文字分の入力チャンネルと 1 音素分の出力チャンネルを持つネットワークである（空白と句読点を含むアルファベットは 29 bit のいずれかが 1 で他は 0 になるようにして表してある）。ニューラルネットワークは $7 \times 29 = 203$ ニューロンからなる入力層、80 ニューロンからなる中間層、26 ニューロンからなる出力層によって構成されている。その機能は、7 文字分の中心にあるアルファベットに対応する音素を出力することである。色々な条件の実験が行われたが、例えば、約 1,000 語の学習で 90% を超える正答率が得られており、ルールベースの方法よりも良い結果が得られている。この例題は 2 つの重要な意味を持っている。まず、第一に、このニューラルネットワークは、膨大な数の英単語の中のアルファベットをどのように発音するかの統計的サマリーをしているということである。英語のような自然言語においてはアルファベットと音素は一対一には対応していないし、アルファベットの位置を考慮してもそれらの間

には厳密な関数関係があるわけではない。その対応関係は、ある厳密な関数関係にノイズが入ったものと考えられることができる。もう一つの重要な点は、入力チャンネル数の増加に伴う「次元の呪い」がないことである。このニューラルネットワークは 7 次元入力データ空間内に離散的に既知出力を持つ点（学習データ）が与えられているときの内挿問題と考えられ、通常の回帰問題では次元の増大につれて必要な学習データの数は指数関数的に増大するが可変基底関数を使っているニューラルネットワークの場合にはこのようなことは起こらない。

1.1.3 ニューラルネットワークの学習と汎化

ニューラルネットワークの学習は、通常、適切に選んだ目的関数の非線形最適化問題に帰着される。また、後に述べる残差最小化学習ニューラルネットワーク等を除いて、多層ニューラルネットワークの目的関数はネットワーク出力と教師データの差の 2 乗の和にとるのが普通である。最適化のアルゴリズムはいろいろあるが、どのような場合にも手軽に使えるのが誤差逆伝播法である。これは、勾配法そのもので 1 次収束をするアルゴリズムなので多数の反復回数を必要とするがロバストな方法である。問題や定式化によっては数十万回の反復計算を必要とする場合もある。入力パターンの提示とウェイト更新のタイミングの違いによってバッチ法とオンライン法の違いがある。学習のために用意されたすべての入力パターンについて出力値を求めてから目的関数を計算してウェイトを更新するのがバッチ法であるのに対して、1 パターンを入力・計算する度にウェイト更新をするのがオンライン法である。オンライン法では各更新時に必ずしもシステムの最適解に近づく方向にウェイトが更新されるわけではないが局所最小解に陥り難く収束が比較的早い。しかし、一般に、オンライン法と高次アルゴリズムを組み合わせることは困難である。

残差最小化学習 [18] は、これまで述べてきた方法と同じネットワーク構成を用いた方法であるが、教師データの代わりにネットワークの入出力関係を決定する支配方程式を与えて学習を行う点が違っている。この方法を使う際には多層ニューラルネットワークの入出力関係が（通常は数値的な関係であらわされているが）、実は解析的な式で与えられているということが重要である。このおかげで、出力変数を入力変数で何回でも微分することが可能である。たとえば、3 層ニューラルネットワークによる写像 (式 (1)) については次のように出力変数 z_k の入力変数 x_ℓ による 1 階微分が求められる。

$$\frac{\partial z_k}{\partial x_\ell} = \sum_{j=1}^J w_{jk} v_\ell \sigma \left(\sum_{i=1}^I v_{ij} x_i + v_{0j} \right) \quad (3)$$

通常、ニューラルネットワークの出力データはノイズを含む測定値を近似するものであるから数値的に微分を行うことはほとんど不可能である。しかし、上記のような解析的表現を使うならば数値微分を行うことによる誤差の拡大を避けられるのでいろいろな応用が考えられる。残差最小化学習ニューラルネットワークでは、学習に際して微分方程式の残差の 2 乗を目的関数に付け加えて学習を行う（もと

もとの目的関数は微分方程式の初期条件や境界条件のみを教師データとする)。この目的関数を最小化するように学習したニューラルネットワークは使った微分方程式の解となる。もちろん、これだけでは、ただ単に微分方程式を解く手段[19, 20]にすぎないので他の効率的な微分方程式解法には太刀打ちできない。そこで、目的関数に含まれる「初期条件や境界条件を教師データとする項」を「測定データあるいは観測データを教師データとする項」とすれば、この方法はデータ同化問題やパラメータ評価問題を解くための有効な方法となる。また、この方法における目的関数を微分方程式に限らず一般の方程式に拡張することもできる。たとえば、積分では、微分と違って離散化による誤差の拡大はないので数値積分を使った残差最小化学習を行うことによってトモグラフィ型の問題を解くことが可能になる。

ところで、ニューラルネットワークに学習させる目的は、学習データについてではなく未学習データについて「正しい」答を出すことである。この能力が、ニューラルネットワークの汎化 (generalization) 能力であって、学習能力を高めるためにニューラルネットワークを複雑にしたり (ニューロン数を増やす) 反復計算を増やしすぎたりすれば、汎化能力が低下して未学習データを適切に処理できなくなる。汎化能力を高めて prediction risk (未学習データに対する総合的誤差) を減少させるための様々な実用的方法が提案されており、理論的研究も盛んに行われているが、ここでは省略する。

1.2 多層ニューラルネットワークの核融合・プラズマ研究への応用

すでに述べたように、ニューラルネットワークにはいろいろな種類があり、それぞれに特徴的な機能を持っている。しかし、今まで核融合・プラズマの研究分野で応用されたネットワークは限られている。これらについて解説する。

1.2.1 関数当てはめ問題

ニューラルネットワークの応用の多くを占めるのは、その「内挿・関数近似能力」あるいは「分類能力」を陽に使ったものであろう。実験科学においては、システムを支配する変数 \bar{X} のある値に対して、そのシステムの出力として変数 \bar{Y} の値が測定され、この測定値から \bar{X} の値を解析する必要がしばしば起こる。 \bar{X} から \bar{Y} を求めるプロセスは微分方程式などによって数理モデルとして定式化されている場合も少なくない。 \bar{Y} から \bar{X} を求める問題は「逆問題」とよばれ、通常、悪条件問題である。これは、 \bar{Y} から \bar{X} を求める時に拘束条件が少なく \bar{X} が一意的に求まらないというのが大きな理由である。そこで、 \bar{X} から \bar{Y} の対応関係 (\bar{X}, \bar{Y}) を表す具体例を n 個 $\{(x_i, y_i), i = 1, 2, \dots, n\}$ 用意して、この対応関係を \bar{Y} から \bar{X} の対応関係であると読み替える。このようにして、 \bar{Y} を多層ニューラルネットワークの入力とし、 \bar{X} を出力とすれば、この逆問題が解けることになる。対応関係 $\{(y_i \rightarrow x_i), i = 1, 2, \dots, n\}$ は有限個の点についてしか定義されていないが、これをニューラルネット

ワークによって学習すれば、その汎化能力によって未学習の点についても答が得られる。これは、「内挿」あるいは「関数当てはめ」である。このプロセスは他の関数当てはめ法を使っても原理的には可能であるがニューラルネットワークによって一層効率的に実行が可能となると推測される。学習データを用意する際に、既知の支配方程式を解いてデータの組を用意することも、何らかの方法で既に解析されている実験データや観測データの組を用いることもできる。この方法は多層ニューラルネットワークの応用方法として最も単純なもので色々な分野で数多くの応用例を見つけることができる。

核融合研究の分野で多層ニューラルネットワークを関数当てはめ問題に応用した例は、Lister らによってなされたトカマクプラズマの磁気プローブデータからプラズマの MHD 諸量を求めた論文[21]が最初であろう。トラス型核融合プラズマ閉じ込め装置のプラズマの理想 MHD 平衡の計算コードはほぼ確立しているのでプラズマ平衡を規定するパラメータを与えることでプラズマ周辺に置かれた電磁気信号測定プローブの出力は計算可能である。これらの電磁気信号から得たデータをニューラルネットワークの入力データとしてプラズマの MHD 平衡に関するデータを求めることができる。Lister らは円形断面プラズマのモデル平衡についてのテスト計算をしたほか、Doublet III を対象としたモデル平衡について計算を行い有効性を確かめた。その後、同様の方法を他の核融合装置のプラズマに適用した場合の研究が色々なされている[22-24]。この方法は、解析手法としておむね確立した感があるが、制御への適用を念頭に信頼性向上を目指した改良の研究も行われている[25, 26]。

この方法のもう一つの応用例は、関数の未知パラメータ推定に関するもので、具体例としては、荷電交換中性粒子エネルギースペクトル[27]や不純物分光分析スペクトル[28]の分布のパラメータ推定がある。この場合関与する関数は、ガウス関数で、一つの単色スペクトルピークについて、位置(波長, エネルギー), 振幅, 半値幅の3つのパラメータを推定すれば関数が決定される。実際に測定されるデータは複数の単色スペクトルピークの重なりであったり、分解能の低い測定器による少数データからなるスペクトルであったりするので、これらのスペクトルを整形して理想的なガウス関数に戻すのがこの解析の目的である。同様に関数のパラメータ値を推定する問題に、最初 Lagin らによって発表された静電プローブの電流電圧特性からプラズマの電子密度と電子温度を推定する問題がある[29, 30]。この場合の電流電圧特性は、電子密度と電子温度をパラメータとして含む非線形な関係式である。

プラズマの輸送過程や閉じこめの実験解析に関連して、多変数のデータベースや閉じこめ比例則の作成のためにニューラルネットワークを利用した例も報告されているがこれらも関数当てはめ問題の範疇に入る応用であろう[31-33]。もっとも、他の応用例ももとをたどればニューラルネットワークの関数近似能力を利用したものとも考えられるので、この分類はかなり便宜的なものである。

1.2.2 時系列予測問題

ニューラルネットワークを用いたもう一つの重要な応用は予測問題への適用である。典型的なのは時系列予測であるが、必ずしも時系列予測の形をとらない場合もある。ニューラルネットワークで将来を予測する場合、時系列データが如何にランダムに見えようとも「それらは決定論的プロセスの結果であり現在までのデータの中にシステムの状態を指定するのに必要な情報はすべて含まれている」ことが重要である。単純な例を挙げるならば、あるシステムの状態の時間発展が n 次元状態空間の中の軌道として表される時に軌道上の1点の状態 $\vec{r}(t)$ から Δt だけ将来の点の状態 $\vec{r}(t+\Delta t)$ を予測するには、 t_i ($i=1, 2, \dots, n$) について $\vec{r}(t_i)$ を入力し出力データを計算し、これを「真の値」 $\vec{r}(t_i+\Delta t)$ と比較して学習すればよさそうである。一般の測定では、多次元状態量であっても、時系列データとしては1変数についてしか得られない場合が多い。しかし、このような場合でも、1変数の時系列データの解析を行いシステムの状態空間軌道のアトラクターのフラクタル次元 d を求めて、 $n \geq 2d+1$ となるように n をとれば n 次元空間に軌道を埋め込むことができる[34]。そこで、 $(x(t_{m-n+1}), x(t_{m-n+2}), \dots, x(t_m))$ を入力し、 $x(t_{m+1})$ を予測する方法は初期から良く行われてきた。学習は、いろいろな t_m に対してこれら2つのデータの関係をニューラルネットワークに記憶させることである。一方、この問題に残差最小化学習ニューラルネットワークを用いることもできる。まず、システムの状態空間の次元 n が計算できればそれぞれの座標に対応する変数 y_j ($j=1, 2, \dots, n$) に対して次のような発展方程式を仮定する。

$$\frac{dy_j}{dt} = f_j(y) \quad (j=1, 2, \dots, n) \quad (4)$$

測定された時系列データから n 個の時間遅れ座標に対する時系列データを構成して、このデータと上記の発展方程式から目的関数を構成して残差最小化学習を行う。この場合、解析システムは、 y_j に対応する n 個と $f_j(y)$ に対応する n 個のニューラルネットワークの合計 $2n$ 個から構成される。

核融合・プラズマ研究に関しては、時間遅れ座標系[34]の考え方に基づくディスラプション予測がHernandezら[35]によって最初に試みられ、また、状態を表す変数の値が必要なだけ測定されているとする考え方に基づく方法はWroblewskiら[36]およびSenguptaら[37]によって行われた。また、その後、各地のトカマクについてニューラルネットワークを用いたディスラプション予測に関する研究が多数行われた。現在では、ニューラルネットワークをディスラプション予測に使う方法の信頼性が高まり完成の域に近づいている。この問題に関しては本総合解説の第4章の芳野の解説[38]に詳しい。

残差最小化学習ニューラルネットワークを用いて予測を行う方法については、今までのところ、核融合・プラズマ研究の分野では応用されていない。時系列データを発生するモデル方程式としてローレンツ方程式を採用して時間遅

れ座標系の方法で長期予測を行うときの性能を調べた。許容振幅誤差を10-20%の範囲にとると、通常の時間遅れ座標系に基づく方法に比べて2倍程度の長期にわたって予測が可能であるとの結果が得られており、実際の実験データへの適用が期待される。

1.2.3 トモグラフィ型の問題

トモグラフィは物体の周辺に置かれた測定器によって物体内部の物理量の2次元分布あるいは3次元分布を非破壊的に推定する技術である。ここでは、測定される量は物体内部を横切る経路に沿って当該物理量に何らかの形状因子の掛かった量を線積分した値であるとする。核融合・プラズマ研究にかかわる問題として考えるとき、たとえばマイクロ波等の経路に沿う位相遅延を測定して求めた電子密度分布の線積分値から電子密度を再構成する問題、プラズマ内部の高温原子が発生する軟X線の線積分値を外部に置いた測定器で観測し内部の温度分布を再構成する問題等、応用の可能性は枚挙にいとまない。

核融合・プラズマ研究に関してトモグラフィにニューラルネットワークを適用した例としては、DemeterがMT-1Mトカマクの軟X線を32チャンネルのMCPカメラで観測してベレット入射による不純物分布の解析を行ったものが報告されている[39]。この解析は、5つのパラメータ(振幅、2次元位置、2次元幅)を持つガウス型の不純物分布を仮定し、これらのパラメータの値をいろいろ変えて32チャンネルの出力を計算したデータを使って学習用データの組を作りネットワークに学習させて逆変換を行う。これは1.2.1で述べた関数当てはめ問題の一つであって、通常のとモグラフィの方法ではない。

ニューラルネットワークを用いた本格的なトモグラフィは、残差最小化学習ニューラルネットワークによって可能になる[18]。この方法では、解くべき物理量分布を1つのニューラルネットワーク出力で表す。このネットワークに位置ベクトルを入力するとその場所の物理量分布が出力されるので、測定経路を指定して数値積分のための分点を与えればネットワーク出力から物理量の線積分値が数値的に計算できる。この線積分値が測定値にあうように学習を行えば、ニューラルネットワークが物理量の分布を表す関数になる[40]。いくつかの付加的アイデアを加えることによって、少ない積分経路や偏った積分経路の場合にも適用可能なトモグラフィ解析が可能になる。測定経路の方向が一方のみの場合その方向の分布情報は測定からは得られない、しかし、等高線の位置・形状がMHD平衡量等比較的少ないパラメータで表現できてこれらの量が式やデータベースの形で利用できる時には、マイクロ波干渉計によるプラズマ密度測定のように測定経路が1方向だけの場合にもこの方法は適用可能である[41]。本小特集の第3章には、実際のGPSデータを使った電離圏プラズマ密度分布の3次元トモグラフィを中心にした残差最小化学習ニューラルネットワークによるトモグラフィについての馬の解説[42]がある。

1.3 まとめと討論

ニューラルネットワークを理工学分野の研究に応用するときには有用と思われる若干の知識について記述した。また、現在までに核融合・プラズマ研究の分野でなされたニューラルネットワークの応用について概観した。これらの中で特に重要だと思われる課題、すなわち、理論的バックグラウンド、トモグラフィ型の問題への応用、ディスプレイ予測への応用については、それぞれ、本総合解説の第2章[7]、第3章[42]、第4章[38]に解説してある。

ここまでのニューラルネットワークはすべて多層型であるが、他のタイプのニューラルネットワークの核融合・プラズマ研究への応用可能性等についても簡単に触れておく。相互結合型ニューラルネットワークについては、粒子シミュレーションにおける高精度「粒子装荷（ローディング）」に用いた例がある[43]。粒子シミュレーションを行うとき初期条件として粒子の速度分布をなるべく対称性がよくなるように装荷しなければならない。この操作は乱数発生プログラムを用いて行うが、これだけでは必ずしも十分な対称性が得られない。そこで、ニューロン1個が粒子1個に対応するような相互結合ニューラルネットワークを考え、目的関数として粒子速度に関する最小二乗和の形のものを用いれば、ホップフィールドネットワークのウエイト更新アルゴリズムを使うことで目的関数の値が減少する。この応用は、粒子速度分布の高次モーメントの非対称性が0の状態が目的関数最小状態になるようにするアルゴリズムとしてホップフィールドネットワークを使った最適化法の一つの表現である。必ずしもネットワークをイメージしなくともよい。

ニューラルネットワークの研究の初期に、シグマ・パイ($\Sigma\Pi$)ニューロンというニューロンが提案されたことがある[44]。これは、ニューロンへの入力データ x_i に対して、通常行われているように重み付き和 $y = \sum w_i x_i$ のかわりに入力データ x_i の高次の項を含んだ重み付き和 $y = \sum w_i x_i + \sum_j \sum_k w_{jk} x_j x_k$ を考える（原理的には3次以上の項を含めてもよい）。シグマ・パイニューロンは生物学的ニューラルネットワークの機能との関連で考え出されたものであるが、応用の観点からはこのようなニューロンは誤差逆伝播法の反復を早く収束させる効果があることが確認された。

コホネンの自己組織化マップ[45]は、ここまで記述してきたニューラルネットワークとは異なるもので、多くの変数で表されるパターンを比較的低次元（普通、2次元）に写像する次元圧縮法と考えられる。たとえば、関数当てはめ法のところで述べた平衡解析のように磁気プローブデータの組（プローブの数だけ変数）で表されたパターンをプラズマ中心位置を表す2次元面に写像することができる。これはプラズマの代わりに線電流を置いたモデル実験で確認できた。ただし、この写像はトポロジカルな写像であるので2点間の距離についての情報は含まない。実際に使うにはこの点について何らかの工夫が必要である。

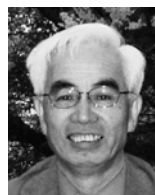
これらの他にも核融合・プラズマ研究の分野でのニューラルネットワークの応用は色々あるものと考えられる。今

後の研究開発に期待したい。

参考文献

- [1] T. Takeda and A. Iyomasa, J. Plasma Fusion Res. **78**, 842 (2002) [in Japanese].
- [2] F. Rosenblatt, Psychological Rev. **65**, 386 (1958).
- [3] M. Minsky and S. Papert, *Perceptron - An Introduction to Computational Geometry* (The MIT Press, Cambridge, 1969).
- [4] J.J. Hopfield, Proc. Natl. Sci. USA, **79** 2254 (1982).
- [5] D.E. Rummelhart, G.E. Hinton and R.J. Williams, Nature (London) **323**, 533 (1986).
- [6] K. Funahashi, Neural Netw. **2**, 183 (1989).
- [7] K. Hagiwara, J. Plasma Fusion Res. **82**, 282 (2006) [in Japanese].
- [8] S. Lawrence, C.L. Giles and A.C. Tsoi, *Proc. Fourteenth Nat. Conf. on Artificial Intelligence, AAAI-97*, AAAI Press, Menlo Park, Calif. (1997) p.540.
- [9] A.R. Barron, Mach. Learn. **14**, 115 (1994).
- [10] Ed. By R. Tagliaferri, G. Longo and B. D'Algenio, A. Incoronato, Neural Network Analysis of Complex Scientific Data: Astronomy and Geosciences, Neural Netw. **16**, 2003 Special Issue Nos.3-4 (2003).
- [11] R. Rojas, *Neural Network: A Systematic Introduction* (Springer, Berlin, 1996).
- [12] S. Haykin, *Neural Networks: A Comprehensive Foundation* (Prentice Hall International Inc., New Jersey, 1999).
- [13] 中野良平:ニューラル情報処理の基礎数理(数理工学社, 東京, 2005).
- [14] 渡辺澄夫:萩原克幸, 赤穂昭太郎, 本村陽一, 福水健次, 岡田真人, 青柳美輝, 学習システムの理論と実現(森北出版, 東京, 2005).
- [15] 廣瀬明:複素ニューラルネットワーク(サイエンス社, 東京, 2005).
- [16] T.J. Sejnowski and C.R. Rosenberg, Complex Syst. **1**, 145 (1987).
- [17] T.J. Sejnowski and C.R. Rosenberg, NETtalk: a parallel network that learns to read aloud, The Johns Hopkins University Electrical Engineering and Computer Science Technical Report, JHU/EECS-86/01.
- [18] T. Takeda, A. Liaqat and M. Fukuhara, *Proc. Internat. Workshop on Modern Science and Technology 2002* (Sept. 19-20, Univ. Electro-Communications, Tokyo, 2002) p.145.
- [19] B.Ph. van Milligen, V. Tribaldos and J.A. Jimenez, Phys. Rev. Lett. **75**, 3594 (1995).
- [20] I.E. Lagaris, A. Likas and D.I. Fotiadis, IEEE Trans. Neural Netw. **9**, 987 (1998).
- [21] J.B. Lister and H. Schnurrenberger, Nucl. Fusion **31**, 1291 (1991).
- [22] L. Lagin, R. Bell, S. Davis, T. Eck, S. Jardin, C. Kessel, J. McEnerney, M. Okabayashi, J. Popyack and N. Sauthoff, *Proc. 17th Sympo. on Fusion Technology*, Rome (1993) Vol.2 p.1057.
- [23] E. Coccorese, C. Morabito and R. Martone, Nucl. Fusion **34**, 1349 (1994).
- [24] R. Yoshino, J. Koga and T. Takeda, Fusion Technol. **30**, 237 (1996).
- [25] Y-M. Jeon, Y-S. Na, M-R. Kim, and Y-S. Hwang, Rev. Sci. Instrum. **72**, 513 (2001).

- [26] Y.-S. Na, Y.-M. Jeon, S.-H. Hong and Y.-S. Hwang, *Rev. Sci. Instrum.* **72**, 1400 (2001).
- [27] D.R. Baker, R.J. Groebner and K.H. Burrell, *Plasma Phys. Control. Fusion* **36**, 109 (1994).
- [28] C.M. Bishop and C.M. Roach, *Rev. Sci. Instrum.* **63**, 4772 (1992).
- [29] L. Lugin, D. Manos and M. Diesso, *Bulletine of the American Physical Society, Division of Plasma Physics Thirty-Third Annual Meeting, (1991) Vol.36(9), 7T13.*
- [30] O. Barana and G. Manduchi, *Neural Comput. Applic.* **10**, 351 (2002).
- [31] L. Allen and C.M. Bishop, *Plasma Phys. Control. Fusion* **34**, 1291 (1992).
- [32] A. Wakasa, S. Murakami, H. Yamada, M. Yokoyama, H. Maassberg, C.D. Beidler, K. Watanabe, S. Oikawa *et al.*, *J. Plasma Fusion Res. SERIES* **6**, 203 (2004).
- [33] G. Tresset, C.D. Challis, X. Garbet, X. Litaudon, M. Mantinen, D. Mazon and D. Moreau, *29th EPS (Montreux, Switzerland, 17-21 Jun., 2002) p.2.039 (2002).*
- [34] F. Takens, Detecting strange attractors in turbulence. In D.A. Rand and B.S. Young, ed., *Dynamical Systems of Turbulence*, Vol. **898** of *Lecture Notes in Mathematics* (Springer Verlag, Berlin, 1981) p.366.
- [35] J.V. Hernandez, A. Vannucci, T. Tajima, Z. Lin, W. Horton and S.C. McCool, *Nucl. Fusion* **36**, 1009 (1996).
- [36] D. Wroblewski, G.L. Jahns and J.A. Leuer, *Nucl. Fusion* **37**, 725 (1997).
- [37] A. Sengupta and P. Ranjan, *Nucl. Fusion* **41**, 487 (2001).
- [38] R. Yoshino, *J. Plasma Fusion Res.* **82**, 294 (2006) [*in Japanese*].
- [39] G. Demeter, *Rev. Sci. Instrum.* **68**, 1438 (1997).
- [40] X.F. Ma, M. Fukuhara and T. Takeda, *Nucl. Instrum. Methods in Phys. Res. A* **449**, 366 (2000).
- [41] X.F. Ma, M. Fukuhara and T. Takeda, *Nucl. Instrum. Methods in Phys. Res. A* **492**, 178 (2002).
- [42] X.F. Ma, T. Takeda, *J. Plasma Fusion Res.* **82**, 287 (2006) [*in Japanese*].
- [43] J.L.V. Lewandowski, *Phys. Lett. A* **313**, 291 (2003).
- [44] D.E. Rumelhart, J.L. McClelland and PDP Research Group, *Parallel Distributed Processing* (The MIT Press, Cambridge, Massachusetts, 1986).
- [45] T. Kohonen, *Self-Organizing Maps* (Springer-Verlag, Berlin, 2001).



たけだ たつおき
竹田辰興

1968年東京大学大学院理学系研究科物理学専攻博士課程修了。理学博士。日本原子力研究所にて核融合研究に従事，1994年～2006年電気通信大学教授，その間，日本原子力研究所計算科学技術推進センター長，九州大学教授を兼任。専門は，核融合研究および計算科学，最近はニューラルネットワークの応用研究に集中している。趣味は写真撮影と篆刻。