論文要旨

システム制御情報学会論文誌, 4 巻, 1 号, pp. 1~10 (1991)

可能性と必然性の概念に基づくニューラルネットの学習アルゴリズム

石渕 久生・田中 英夫

本研究では、2 群判別の問題を取り上げ、ニューラルネットによる可能性分析と必然性分析の方法を提案する。まず、可能性と必然性の概念に基づく判別分析の方法を示す。つぎに、可能性判別分析および必然性判別分析のためのニューラルネットの学習アルゴリズムを提案する。提案手法では、ニューラルネットの結合強度の学習は、重み付き 2 乗誤差の総和が最小になるように行なわれる。このとき、データ点ごとの重みを学習と共に徐々に変化させることにより、可能性あるいは必然性を考慮した判別分析を行なうことができることを示す。さらに、例題を用いて提案手法の説明を行なう。

システム制御情報学会論文誌, 4 巻, 1 号, pp. 11~20 (1991)

ニューラルネットを用いたリモートセンシングデータのパターン認識

吉田 知司・大松 繁・寺西 大

本論文はリモートセンシングデータのパターン認識問題へニューラルネットを用いる手法を提案している。これによって、誤認識の割合を減少させかつ従来の Bayes 法による手法と比較して識別誤差のバラツキの少ないデータ処理結果が得られることを実データに基づいて定量的に検証している。さらに、雲の影の補正を行なって分類する手法に対しても本論文の手法が有効であることを示す。

システム制御情報学会論文誌, 4 巻, 1 号, pp. 21~27 (1991)

階層型神経回路モデルによる時空間パターンの認識

伊藤 崇之

時空間パターンを認識できるニューラルネットワークモデルを提案する。モデルは特徴抽出部と統合部の二つの部分からなっている。前者は、入力信号の立ち上がり、立ち下がりなど、時間に固有の特徴を抽出する。後者は、 P 細胞と T 細胞の 2 種類の細胞からなっている。 P 細胞は特徴の統合を行なう。 T 細胞は、メモリ機能を持つ。これら 2 種類の細胞層を交互に継続的に接続し、位置不変性と変形の許容能力も実現している。コンピュータシミュレーションによって、モデルがパターンの時間伸縮や、位置ずれ、変形を許容できることを示した。

システム制御情報学会論文誌, 4 巻, 1 号, pp. 28~36 (1991)

チェビシェフネットワークによるパターン領域の識別

生天目 章・上田 延寿

コネクショニストモデルを用いて、2 次元領域のパターンを識別する問題について考察する。適当な数の中間ユニットを配置した従来の多層ネットワークによる識別ではパターン領域が複雑になるにつれてその識別能力は低下し、中間ユニットの数や学習例を増大させてもその識別性能は改善されないことを示す。また、一般に非線形の複雑な決定領域におけるパターン識別システムをコネクショニストモデルにより実現するためにはシグモイドのような単調な非線形関数のユニットだけでは不十分でチェビシェフ関数のような非単調関数を持つユニットとの組み合わせが必要であることを示す。従来のモデルに代わるものとしてチェビシェフネットワークという新しい多層ネットワークアーキテクチャを提案する。パターン領域の境界曲線をチェビシェフネットワークに学習させ、その出力値の論理結合により識別する方法および領域の境界曲線近傍における対象領域内外のいくつかのパターンをチェビシェフ関数により非線形変換し、それらを正負の学習例として学習・識別させる二つの方法について提案する。いずれの方法においても新しく提案したチェビシェフネットワークによる識別方法は、一般にパターン領域の複雑性に依存しない優れた識別性能を有することを示す。

システム制御情報学会論文誌, 4 巻, 1 号, pp. 37~47 (1991)

フィードバック誤差学習による閉ループシステムの学習制御

五味 裕章・川人 光男

本論文では神経回路モデルを非線形フィードバックコントローラに利用するときの学習法を提案する。学習後には制御対象の非線形性が補償され、参照モデルで規定される応答に従うようになる。

ここでは、この学習スキームの収束性を確率微分方程式の平均化とリアプノフの方法で証明し、有効性を倒立振り制御のシミュレーションの結果により示す。また、この学習スキームと動物の小脳における姿勢・歩行の学習制御との関連を議論する。

チェビシェフネットワークによるパターン領域の識別*

生天目 章**・上 田 延 寿***

Pattern Classification with Chebychev Networks*

Akira NAMATAME** and Nobuhisa UEDA***

A connectionist classifier with a new network architecture, Chebychev networks, is proposed for classifying continuous patterns in non-convex decision regions. The conventional connectionist classifier with monotonic sigmoid functions as its units is shown to face the problems of scaling and generalization. The new multi-layer networks, Chebychev networks, consist of nonmonotonic, nonlinear Chebychev functions as their input units, and monotonic sigmoid functions as their hidden and output units. The proposed connectionist classifier consists of two modules, a Chebychev networks module and a classification module. The Chebychev networks module transforms a complex pattern space into a high-order simple space. The classification module works on this transformed simple space.

It is shown the problems of scaling and generalization can be overcome with the new connectionist classifier with Chebychev networks, and the classification performance is increased significantly compared with the conventional connectionist classifier.

1. はじめに

知覚情報処理の研究テーマは多岐にわたるが、パターン識別に関する研究は共通の基本的課題である。パターン識別問題は、また信号処理、診断、および意思決定などの工学的問題の共通する課題でもあり、これまでに多くの研究が行われてきた。統計的手法による従来のパターン識別は、パターンの特徴ベクトルと各パターンの属するクラスの代表的サンプルとの適合度を測定し、その適合度が最もよいクラスをそのパターンの帰属するクラスとして識別している。コネクショニストモデルによるパターン識別は、提示されるパターンに応じその識別性能を逐次改善できるといった学習能力に特徴があるだけでなく、各パターンの分布に関する統計的仮定をおく必要がなく、識別能力、ロバスト性にも優れたパターン識別が可能である¹。しかしながら、数多くの現実的パターン識別問題にコネクショニストモデルを適用してい

くためにはネットワークの規模の問題や学習アルゴリズムの効率化の観点からの改善が必要である²。多層ネットワークにバックプロパゲーション学習アルゴリズムを適用することにより、与えられた一般に複雑な入出力関係を実現するための優れた学習結果が得られることが知られている^{3,4}。しかしながら、学習例として与えられなかった入出力関係についてどれだけ正確に学習できるか（汎化能力）についてこれまでのところ十分に解明されていない⁵。汎化能力、すなわち未知のデータに対する識別能力をいかに高めるかは学習研究の共通の基本的課題である^{6,7}。

本研究では、従来のコネクショニストモデルにおいて広く用いられているシグモイド型の単調な非線形型のユニットに非単調・非線形関数であるチェビシェフ関数のユニットを組み合わせた新しいネットワークのアーキテクチャ（これをチェビシェフネットワークと呼ぶ）によるパターン識別法について提案する。無限のパターンを持つパターン空間をそのパターンの特徴に応じていくつかのクラス（カテゴリー）に分類する問題を対象とする。従来の教師付き学習による学習法は、それぞれ複数の領域に属する代表的パターンを例題として識別システムに提示し、その識別能力が逐次改善されるようなプロセスをうまく設計する方法がとられてきた。ここではそのよ

* 原稿受付 1990. 5. 1

** 防衛大学校 情報工学科 National Defense Academy ; 1-10-20, Yokosuka, Kanagawa 239, JAPAN

*** 防衛大学校 理工学研究科 National Defense Academy ; 1-10-20, Yokosuka, Kanagawa 239, JAPAN

Key Words : chebychev neural networks, multi-layer networks, pattern classification, interval representation, generalization.

うな学習戦略だけではなく学習後のネットワークを識別のための一種の評価関数として用いる識別方法についても提案する。第1の方法では、入力された未知パターンの特徴ベクトルを識別するための論理空間に非線形変換するニューラルネットワークモジュールとそのパターンの帰属するカテゴリーを識別するための論理表現されたいくつかのルールを持つ識別モジュールの二つのモジュールで識別システムを構成する。ニューラルネットワークモジュールは、各領域間の境界曲線を多層ネットワーク上に表現するように学習が行なわれる。しかしながら、一般にパターン空間の境界領域は複雑な非線形構造を持つため、チェビシェフネットワークという新しいネットワークアーキテクチャを用いて学習を行なう。パターンの識別は、ニューラルネットワークモジュールの出力値により発火される識別モジュールのルールにより論理的に行なわれる。第2の方法では、2次元座標を持つ各パターンに対しチェビシェフ関数で非線形変換を施し、それを学習例として多層ネットワークにパターン領域を学習させる方法で、学習プロセスは従来のアプローチと同様に逐次その識別性能を改善していく方法である。

2. 問題の設定

無限に存在するパターンをいくつかの有限個のクラスに分類する問題について考察する。それぞれのパターンは、2次元空間のベクトル (x_1, x_2) によって特徴づけられ、これらの特徴ベクトルはある正規化を施すことにより2次元のパターン空間 $K = \{(x_1, x_2) ; 0 \leq x_1, x_2 \leq 1\}$ を形成するものとする。パターン空間はいくつかの領域に分割されており、同じ領域に属するパターンは同一クラスに分類される。この領域をクラスの決定領域という。すなわち各パターンはその特徴ベクトル (x_1, x_2) がどの決定領域に位置するかにより、そのパターンの属するクラスを識別することができる。パターン空間 K 内の各決定領域はクラスの境界曲線により分離され、それらは重なり合わないものとする。すなわち、それぞれのパターンはいづれか一つのクラスに属するものとする。Fig. 1 (a), (b) に示すパターン空間の例では、各パターンは二つのクラス $\{c_1, c_2\}$ のどちらかに属し、Fig. 1 (a) は $(1/2, 1/2)$ を中心とする半径 $1/3$ および $1/6$ の二つの円が境界曲線として二つのクラスを分離している。Fig. 1 (b) は $x_2 = 0.5 + 0.3 \cdot \sin(2\pi x_1)$ を境界曲線として二つのクラスを分離している。ここで対象にするような一般に複雑な境界をもつ決定領域に属するパターンを学習・識別する問題（より一般的には n 次元パターン空間上における超平面による識別問題）は、パターン認識、制御、診断および意思決定といった広範な分野においてそれぞ

Fig. 1 Shading of each figure denote decision regions for class C_1 and class C_2 . Samples from each class are distributed uniformly over those separated decision regions

れ重要な共通の問題である。

3. 従来の多層ネットワークの規模の複雑性と学習能力の関係

学習の対象領域の複雑度が増すにつれ、多層ネットワーク学習に必要な学習例は多くなる。一方、これらの多くの望ましい入出力関係（学習例）を学習するためには多くの信号処理ユニットおよびそれらの間の結合結線が必要になり、したがってネットワークの規模の複雑性の問題を生じさせる。ネットワークの規模の複雑性は、学習速度の低下や局所解への収束の可能性の増大など、学習アルゴリズムの劣化を引き起こすことが指摘されている⁸。本節では、学習領域の複雑性と多層ネットワークの規模およびネットワーク学習能力との関係について定量的に明らかにする。

3.1 ネットワークの複雑性と学習能力の指標

一般にネットワークの複雑性を表わす指標としてネットワークのユニットの総数やユニット間の結線数などが用いられる⁹。ここでは、中間層のユニット数 (h) をネットワークの複雑性を表わす指標として用いる。また、ネットワーク学習の識別能力の指標として、学習能力と汎化能力の二つを用いる。学習能力とは学習後のネットワークが学習例（望ましい入出力関係）をいかに再現しているかを表わす指標で、汎化能力とは学習例として用いなかった未知の入出力データに関する学習後のネットワークの指標である。

3.2 多層ネットワークの構成と学習法

2次元空間に連続的に分布したパターンを識別するための基本的な多層ネットワークのアーキテクチャは、入力層に $[0, 1]$ の連続値をとる二つの入力ユニットを配置し、パターン空間 K のそれぞれの座標軸 (x_1, x_2) の値を対応させる。また、分類すべきパターンの各クラスに対応した出力ユニットを出力層に配置し、入出力層間にいくつかの中間ユニットを配置するものである^{10,11}。

このようなネットワークによるパターン識別を1次ネットワークによる識別法と呼び、そのネットワーク構成をFig. 2に示す。中間層とそのユニットの数は対象領域の複雑度に応じてアドホックに決定する必要がある。

Fig. 1 (a), (b)に示すパターン空間 K から任意のデータを抽出し、学習用データベースと検証用データベースを構築する。いずれのデータベースに含まれるデータもその特徴を表わすベクトル (x_1, x_2) およびそれが属するクラス名 $\{c_1$ または $c_2\}$ をもつ。学習データベースの学習例を用いた学習後のネットワークに学習データベースの学習例を再度提示する。そしてその特徴ベクトル (x_1, x_2) に対する学習後のネットワークの出力値とその特徴ベクトルをもつパターンが帰属するクラスとを比較し、学習データベクトルの再現率を求める (学習能力)。つぎに学習に用いなかった検証用データベースを用いて同様に汎化能力を評価する。多層ネットワークの学習戦略は、パーセプトロン型学習や記号処理モデルによる帰納的学習と同様にそれぞれのクラスに属するデータの組 (学習例) を繰り返し提示し、ネットワークがこれを正しく分類するようネットワークの構造を修正していく方法である。

Fig. 2 An architecture of a classification system with first-order networks

3.3 ネットワークの規模とその識別能力の検証

Fig. 1 (a) パターン空間からそれぞれ任意の50点を抽出し、これを学習例として学習データベースを構築した。また学習後の帰納性を評価するために学習例に用いなかった任意の1,000点を抽出して、検証用データベースを構築した。バックプロパゲーションを用いて、Fig. 2に示す1次ネットワークアーキテクチャを学習させた。学習回数は10,000回で、学習定数は変動型とした(0.1~0.7)。学習後のネットワークに学習データベースの学習例および検証用データベースのデータを提示し、ネットワークの識別能力を評価した。Fig. 3は

Fig. 3 Network classification performance for the problem in Fig. 1 (a) as a function of the number of hidden units

中間ユニット数とネットワークの識別能力との関係を示したものである。Fig. 3から明らかなようにネットワーク規模の増加に伴い学習能力 (学習例に対する識別能力) は向上するが汎化能力については向上はない。また、この傾向はパターン空間の決定領域が複雑になるにつれより顕著になる。

パターン空間の決定領域の複雑性が増すに伴いネットワーク規模を増大する必要があることは指摘されているが、対象領域やネットワークの規模の複雑性がネットワークの学習能力にどのように影響を及ぼすかについては現在のところ十分に解明されていない。学習例として与えられた入出力関係の正しい再現だけでなく、その背景にある入出力関係の背景にある規則性を見だし、未知パターンに対しても満足な識別結果が得られること、すなわち汎化能力の向上がコネクショニストモデルをより大規模で現実的問題に適用するためには不可欠である。しかしながら、Fig. 2に示すような従来から用いられている1次のネットワークアーキテクチャではその識別能力に限界があることがわかった。

3.4 学習戦略の工夫

多層ネットワーク学習の学習能力の向上を図るために数多くの方法が提案されている^{12,13}。最も単純な方法は学習例を増やすことである。すなわち、学習の対象領域が複雑になるほど、満足のいく学習結果を得るために多くの学習例を必要とするのは当然といえよう。3.4のシミュレーションでは学習例の数を50に固定しているが、ネットワークの規模を一定 (中間ユニット数: 7個) とし学習例の数を増加した場合、Fig. 1 (a) のパターン空間に対する1次ネットワークの識別能力をFig. 4に示す。この結果から明らかなように学習例の増大により、1次ネットワークの汎化能力は学習能力に近づく傾向にあるが、学習能力は学習例の増大により低下する。ネッ

Fig. 4 Network classification performances for the problem in Fig. 1 (a) as a function of the number of training examples

トワークの規模を一定として学習例を増大させると学習例に対する入出力関係を表現するための自由度が制約されるために十分な学習が行なわれなかった（中間ユニットが3～5個の間では局所解に収束し、十分な学習を行うことができないために識別性能が低い）。したがって、学習能力および汎化能力を含めたネットワークの識別能力はネットワーク規模を一定とした場合、学習例の増大によって改善することは困難であることがわかった。識別能力改善のための他の方法として学習例の適切な選択法がある。学習例の選び方によって学習プロセスが著しく影響されることが、特に記号処理モデルによる学習研究で報告されている^{14,15}。コネクショニストモデルによる学習研究においてはニアミスという概念を取り入れた境界領域からの学習例の選択という方法が提案されている^{16,17}。この学習戦略は一般に数多くのパターンが存在するパターン空間から学習例としてデータを抽出する際、それらをパターン空間の中から任意に抽出するのではなく、それぞれのパターン空間を分類する各決定領域の境界領域近傍から抽出するものである。このような戦略に基づく学習により高い汎化能力の実現が可能であることが報告されている。しかしこれらの研究は、いずれもパターン空間が線形分離可能問題に限定されており、一般に中間層が必要であるような線形分離不可能な問題に対するこのような学習例の抽出法の有効性は未確認である。

Fig. 1 のようなクラス c_1 と c_2 の境界曲線が一般に複雑な非線形となるような領域を対象とし、各クラスの境界の近傍領域から学習例を選択した場合（境界パターンによる学習法）および任意に学習例を抽出した場合（任意のパターンによる学習法）について学習後のネットワークの汎化能力により比較した。それぞれの学習戦略により Fig. 1 (b) から 50 個の学習例を選びそれを 1 次のネットワーク（中間ユニット数：7）に学習させた。学

Table 1 Network classification performances with two strategies of choosing training examples in Fig.1 (b)

Strategy	Random strategy	Boarder strategy
Classification rate (%)	82	70

習回数は 10,000 回である。学習後のネットワークに学習例にない Fig. 1 (b) から任意に抽出した 1,000 個のパターンを提示し、ネットワークの汎化能力を検証した結果を Table 1 に示す。この結果から明らかなように任意のパターンを抽出する学習戦略の方が汎化能力に優れていることがわかる。このように学習例としてその問題領域に応じて意図的に選択する方法は一般に学習すべきパターンが数多く存在し、そのパターン空間が線形分離可能な領域となるような問題については適当でないといえる。

4. 非線形識別曲線を明示的に用いたパターン識別

Fig. 2 に示す従来の 1 次のネットワークアーキテクチャでは、学習例以外のデータに対しては識別能力が低いこと、およびこのことは対象とするパターン領域が複雑になると顕著になることが明らかになった。識別能力向上のために学習例やネットワークの規模の増大もあり効果がなかった。対象とする問題領域の複雑性に対する識別能力の改善はコネクショニストモデルをより現実的な大規模問題に適用していく上で解決しなければならぬ大きな問題である。コネクショニストモデルにおける規模と汎化能力の問題は学習アルゴリズムであるバックプロパゲーションよりもむしろ Fig. 2 に示すような従来のネットワークアーキテクチャと学習戦略に問題があると思われる。このような観点から本節ではパターン識別のための新たなネットワークアーキテクチャとネットワーク学習戦略によるこの問題の解決法について提案する。

4.1 複雑なパターン空間の論理空間への変換

Fig. 1 の例に示すようなパターン空間は、各クラスの決定領域の境界曲線を識別関数として用いることにより論理的に分類することができる。一般に、境界曲線に対し一組の関数の系 $\{f_k(x_1)\}$ を用いることにより、パターン空間上の複雑な領域は識別のための簡単な論理空間に変換することができる。このような変換を境界曲線による論理変換と呼ぶことにする。Fig. 1 (a), (b) のパターン空間の論理空間への変換例を Fig. 5 (a), (b) に示す。一般に複雑な非線形構造をもつ境界曲線を精度よ

Fig. 5 Transformation of a complex decision spaces into logical decision spaces

く学習できれば、これを用いて論理変換された空間上での識別は論理的に精度よく、また極めて容易に行なうことができる。

4.2 チェビシェフネットワークによる非線形境界曲線の学習

一般に複雑な境界曲線はいくつかの非線形関数の系により構成される。したがって、多層ネットワークにより非線形関数を精度よく学習できればそれぞれの非線形関数を学習したいくつかのネットワークを合成することにより境界曲線を表現することができる。一般に、任意の連続関数は中間層を1層持つ3層ネットワークで近似できることが証明されている¹⁷。しかしながらそのためのネットワークをどのように構築するかについての一般的な方法論はない。著者は、非線形関数を学習するための新しいネットワークアーキテクチャとしてチェビシェフネットワークを提案した¹⁸。チェビシェフネットワークとは、パターン空間上の入力変数に関するいくつかのチェビシェフ関数に対応する入力ユニットを配置し、さらに1層の中間層をもつ3層ネットワークを構成する。このようなネットワークアーキテクチャを用いることによって、入出力関係の背景にある複雑な非線形構造について精度よく学習が可能なることを示した。チェビシェフ関数は、次式で表わされる。

$$T_n(x) = [\cos\{n \arccos(2x-1)\} + 1]/2$$

$$n = 0, 1, 2, \dots, k$$

チェビシェフ関数ネットワークのアーキテクチャを Fig. 6 に示す。Fig. 6 において第1層と第2層間の結合係数

Fig. 6 An architecture of Chebyshev networks

は、チェビシェフ変換のための固定結線で、結合係数を同定するための学習は第2層と第4層の間で行なわれる。

4.3 新しい学習戦略に基づく学習法

新しい学習戦略では、まず学習例としてそれぞれのクラスを分離している境界曲線上のデータを抽出し、これを用いてチェビシェフネットワークのアーキテクチャをもつ多層ネットワークに学習させる。境界曲線を多層ネットワークを用いて表現することにより、一般に複雑なパターン空間を各パターンの帰属するクラスの判定が容易となる論理空間に変換し、論理変換された論理空間上で各パターンを論理的に識別する。このような戦略に基づいて構築されたパターン識別システムの性能は、多層ネットワークによる境界曲線の学習精度のみに依存しており、極めて精度の高いパターン識別が可能となる。本節で提案するパターン識別システム構成を Fig. 7 に示すが、これはパターン空間の各決定領域間の境界曲線を学習するためのニューラルネットワークモジュールと

Fig. 7 An architecture of a classification system with boundary curves

パターンの帰属性について判定するためのいくつかの明示的ルールをもつ識別モジュールとから構成される。一般にパターン空間には複数の境界曲線が存在するのでニューラルネットワークモジュールはそれぞれの境界曲線を学習したネットワークの複合モジュールとして構成される。また識別モジュール内の各識別ルールは、ニューラルネットワークモジュールの出力値により発火される。

5. チェビシェフネットワークによる逐次学習識別法

学習研究における基本的な学習戦略は複数のクラスにおける学習例を逐次システムに提示し、その学習例を正しく分類できるようになるまで識別システムの学習プロセスを変更する逐次修正法である。しかしながら、前節で提案した識別法はこのような逐次改善能力を有していない。また、パターン決定領域が多次元のため、境界曲線の位置情報が容易に、求められない場合には適用が困難である。

これらの問題を解決するために従来のネットワークアーキテクチャにチェビシェフ多項式のユニットを追加した Fig. 8 にネットワークアーキテクチャを持つ識別システムを提案する。前節と異なるのはパターンの境界線ではなくパターン領域のシルエットを識別に利用する点である。シルエットを学習させるには学習例の作成法に工夫が必要である。まず、それぞれのクラスを分離している境界曲線の近傍からほぼ対称位置にある二個の点を抽出し、この座標値と各クラスに応じた出力値を合わせて二個一組の学習例とする。この手法により一組の学習例の間にパターン決定面を固定できる。したがって、チェビシェフネットワークのアーキテクチャをもつ多層ネットワークに上記の方法で抽出した学習例を用いて学

習させることによりシルエットを表現することができる。Fig. 1 (b) のようにチェビシェフネットワークモジュール1個でシルエットを表現できる場合はそのクラスに対応した出力ユニット1個を識別モジュールに配置することによりパターン識別システムを構成できるが、Fig. 1 (a) のようなシルエットを表現するのにチェビシェフネットワークモジュールを2個以上使う場合は各チェビシェフネットワークモジュールの出力に基づいて各クラスを識別できるように識別モジュールをさらに学習させる必要がある。Table 2 は Fig. 1 (a) の問題においてチェビシェフネットワークモジュールの各出力 (z_1, z_2) に応じてパターンのクラス (c_1, c_2) を正しく識別するように識別モジュールを学習させるための学習例である。Table 2 の学習例により識別モジュールを学習させ、それにチェビシェフネットワークモジュールを融合することにより第2の方法によるパターン識別システムを構成できる。このような識別システムは前節と同じような高い識別能力が期待でき、また学習例による逐次改善能力も備えている。

Table 2 Training instances for the classification module

Outputs of Chebyshev Networks Modules		Outputs of Classification Module
z_1	z_2	$C_1 (1)$ or $C_2 (0)$
0	0	0
1	0	1
1	1	0

6. シミュレーション

6.1 非線形識別曲線を明示的に用いたパターン識別法

4. および 5. で提案したように非単調なチェビシェフ関数と単調なシグモイド関数を組み合わせた新しいコネクションモデルによるパターン識別法の有効性について Fig. 1 の例を用いて検証する。

パターン認識の境界曲線を識別関数に用いる方法では、Fig. 1 (a) および (b) の境界線上から任意にそれぞれ 50 および 20 個の学習例を抽出し、5 次までのチェビシェフ多項式を入力ユニットとして持つチェビシェフネットワーク (中間ユニット数: 4 個) に学習させ (学習回数: 2,000 回)、学習後のネットワークの出力値をプロットしたのが Fig. 9 で、境界曲線を精度よく学習していることがわかる。識別モジュールの各識別ルールは、前節で示した変換によって得られる論理空間上で記述される。

Fig. 8 An architecture of classification system with the incremental learning property

Fig. 9 Learning achieved with Chebyshev networks trained with the boundary curves in Fig. 1 (a), (b)

たとえば Fig. 1 (a) の例は 8 個の識別ルールで構成される。このようにして構築された Fig. 7 に示すパターン識別システムに、ある任意のパターンの特徴ベクトル $(x_1, x_2) \in K$ が提示されるとニューラルネットワークモジュールは、入力変数 x_1 境界曲線上の値 x_2^* (ネットワーク出力値) を識別モジュールに出力する。識別モジュールはネットワーク出力値 x_2^* と特徴ベクトルの要素 x_2 を比較してその入力パターンの帰属するクラスを識別する。3. での学習能力の検証と同じように Fig. 1 (a), (b) のパターン空間から任意に 1,000 個のパターンを抽出し、汎化能力を検証した。この結果を Table 3 に示す。これより従来の 1 次ネットワークを用いた学習戦略に比べ識別能力が画期的に向上することがわかった。

Table 3 Classification performances of a Chebyshev method for 1,000 arbitrarily chosen test patterns

Problem Region	Classification Rate (%)	
	Boundary Line	Silhouette
Fig. 1 a	98.4	98.1
Fig. 1 b	99.4	97.1

6.2 チェビシェフネットワークによる逐次学習識別法

上記の識別法は一般に複雑なパターン領域の境界線をチェビシェフネットワークに学習させ、それを非線形の識別関数として明示的に用いてパターン識別を行なう方法である。第 2 のアプローチでは、従来のコネクションモデルと同じく逐次その性能を改善するような学習戦略を用いる。

境界曲線の近傍からほぼ対象の位置にある 2 個の点を抽出し、それらを一組の正と負の学習例とする学習例の集合を構成するために次のような方法を用いた。すなわ

ち、Fig. 1 (a) それぞれの境界線が存在する x_1 座標の範囲の中から任意に 50 点抽出し、それぞれの x_1 座標値に対応する境界線上の x_2 座標値に ± 0.1 の値を追加した値を持つ (x_1, x_2) 座標値がシルエットの内側に位置するか、または外側に位置するかによって正 (出力値 1) または負 (出力値 0) の学習例として構成する。ここで x_2 座標値に追加した値の選び方は任意ですべての座標値が同じ値によって決定できるように選択されればよい。

このようにして構成した 50 組の正負の学習例の x_1 および x_2 の座標値を 5 次までのチェビシェフ多項式で非線形変換し、それを多層ネットワークにバックプロパゲーションを用いて学習させた (Fig. 1 (a) の例では、入力ユニット数: 10 個, 中間ユニット数: 4 個, 出力ユニット数: 1 個, Fig. 1 (b) の例では、入力ユニット数: 10 個, 中間ユニット数: 2 個, 出力ユニット数: 1 個)。学習後のネットワークにおいて x_1 と x_2 の値を $[0, 1]$ の範囲内でそれぞれ 1/100 間隔で入力し、識別システムが獲得したパターン領域を表示したのが Fig. 10 であり、チェビシェフネットワークにより各領域のシルエットが精度よく獲得されていることがわかる。

Fig. 10 Learning achieved with Chebyshev networks trained with silhouettes of Fig. 1 (a), (b)

前節と同様に学習後のネットワークに Fig. 1 (a), (b) のパターン空間から任意に 1,000 個のパターンを抽出し、そのパターンの帰属性について判定した結果を Table 3 に示す。この結果、前節と同様に従来の 1 次ネットワークを用いた学習戦略に比べ汎化能力が画期的に向上することがわかる。しかし、逐次学習・識別法が非線形識別曲線を明示的に用いる方法に比べて識別率が数 % 低いこともわかる。Fig. 11 に示す識別を誤った検証データを見てみると (c_1, c_2) の境界付近にあいまい性を持った領域が存在し、ここでは正確な識別ができず識別精度が低下することがわかる。これは一組の学習例間の距離によるもので、この間に非線形なパターン決定面の傾きが存在し、ネットワークはクラスを (1, 0) の離散

Fig. 11 Discrimination error data

値で出力するのではなく、シルエットへの距離的な近さを連続値で表わすために識別を誤った。

6.3 チェビシェフネットワークの内部表現能力の解析

3. で記述したようにパターン空間の次元数に応じたユニットを入力層に配置し (n 次元パターン空間では n 個の入力ユニット)、シグモイド型の単調な非線形関数を出力関数に持つユニットを中間および出力層に配置した従来のネットワークアーキテクチャを1次の多層ネットワークという。1次ネットワークは、パターンの入出力関係を学習例として提示するだけで、パターンを識別するための知識 (識別関数の構造) が、多層ネットワークの内部表現として自動的に獲得できるという特徴がある。しかしながら、本論文で対象としたような識別すべきパターン空間が複雑な場合には (すなわち、獲得すべき識別関数の構造が複雑な非線形の場合) 1次ネットワークでは識別のための適切な内部表現が獲得できないことを示した。

無限に多くの識別対象パターンを含むパターン空間の中から、比較的少ない有限個のパターンを抽出し、それをチェビシェフ関数を用いて非単調で非線形な変換を施し、それを学習例として多層ネットワークに学習させることにより、複雑なパターン空間を識別するための内部表現として適切に獲得されることを Fig. 1 (b) の例を用いて示す。Fig. 1 (b) を学習したネットワークの各中間ユニットの出力を u_1, u_2 とし、それぞれを座標軸とする (u_1, u_2) 平面を表わしたのが Fig. 12 である。これにより Fig. 1 (b) の (c_1, c_2) 領域は各々上と下にクラスタされ、線形識別関数 ($u_2 = 0.154 : u_1 = -0.399$) によって (c_1, c_2) に分けられる。チェビシェフネットワークの利点はチェビシェフ多項式による変換により、一般に複雑なパターンは Fig. 12 のように識別が容易な高次のパターン空間に非線形変換され、簡素化された領域上において最終的なパターン識別を行なう点にある。また、チェ

Fig. 12 The outputs (u_1, u_2) of the hidden units of Chebyshev networks. Complex pattern space in Fig. 1 (b) nonlinearly transformed into the simple space at the hidden layer. (C_1 shows the region that output of the output layer is greater than 0.9. C_2 shows the region that output of the output layer is less than 0.1.)

ビシェフネットワークを用いることにより識別が容易となるような高次空間への非線形変換のメカニズムは一般にパターン領域に存在せず、多層ネットワークの学習アルゴリズムによって決定されることがわかった。

7. 結論

本研究では、非凸な決定領域を持つ複雑なパターンをコネクショニストモデルにより識別する問題を考察した。パターン空間がより複雑になるとネットワークの規模が増大するだけでなく、学習例として与えなかった未知パターンに対する識別能力が著しく低下することを示した。このようなコネクショニストモデルの規模と汎化能力の問題は、従来のネットワークアーキテクチャとネットワーク学習の学習戦略に起因していることを指摘し、これに代わる新たなアーキテクチャとしてチェビシェフネットワークとその学習戦略を提案した。

チェビシェフネットワークは、従来のシグモイド型の単調で非線形の出力関数を持つユニットとチェビシェフ関数といった非単調で非線形の出力関数を持つユニットを融合したものである。識別対象パターンの座標値をチェビシェフ関数を用いて非単調で非線形の変換を行ない、変換後のパターンを学習例として多層ネットワークに学習させることにより、複雑な非線形識別関数の構造について十分な学習ができることを示した。

チェビシェフネットワークを用いることにより、複雑なパターンを識別するための知識がその内部表現として適切に獲得されることにより、従来のモデルと比較してパターンの識別能力は著しく向上し、パターン空間の複雑性にあまり依存しないことを示した。本研究では無限個のパターンをもつパターン空間を対象としたが、線形分離不可能な有限個のパターンに対しても適用すること

ができる。また、パターン空間が n 次元の場合は境界曲線は一般に n 次の超平面となるが、この場合はそれぞれの次元にチェビシェフ変換を施すことにより本手法が適用可能となる。

参 考 文 献

- 1) P. Lippman, : An Introduction to Computing with Neural Nets ; IEEE ASSP Magazine, April, pp. 4~22 (1987)
- 2) M. Minsky & S. Papert : Perceptron, Expanded Version, MIT Press (1988)
- 3) D. E. Rumelhart, G. E. Hinton & R. J. Williams : Learning Representation by Back-propagation Errors ; Nature, Vol. 323, pp. 533~536 (1986)
- 4) P. J. Werbos, : Generalization of Backpropagation with Application to a Recurrent Gas Market Model ; Neural Networks, Vol. 1, pp. 339~356 (1988)
- 5) S. Ahmad, & G. Tesauro, : Scaling and Generalization in Neural Networks ; Proc. of 1988 Connectionist Models, pp. 3~10, Morgan Kaufmann (1988)
- 6) J. H. Holland, K. J. Holyork, R. E. Nisbett & P. Thagard : Induction, MIT Press (1986)
- 7) R. Michalski, G. J. Carbonell, & T. M. Mitchell : Machine Learning, Vol. 1 & 2, Morgan Kaufmann (1983)
- 8) S. Fahlman, : Faster-learning Variations on Back-propagation ; Proc. of 1988 Connectionist Models, pp. 38~51, Morgan Kaufmann (1988)
- 9) D. E. Rumelhart, Learning and Generalization : The Role of Minimal Networks, 1988 ATR International Workshop (1988)
- 10) W. Huang, & R. Lippman : Comparison between Neural Nets and Conventional Classifier ; Proc. of IEEE First Int. Conf. on Neural Networks, Vol. 4, pp. 485~493 (1987)
- 11) S. Singhal, & L. Wu : Training Multilayer Perceptrons with the Extended Kalman Algorithm ; in Advances in Neural Information Processing Systems 1, pp. 133~140, Morgan Kaufmann (1989)
- 12) T. Grossman, R. Mair & E. Dömany : Learning by Coice of Internal Representations ; in Advances in Neural Information Processing Systems 1, pp. 73~80, Morgan Kaufmann (1989)
- 13) D. Plastis & N. Neifeld : The Emergence of Generalization in Networks with Constrained Representation ; IEEE Int. Conf. on Neural Networks, pp. 371~382 (1988)
- 14) P. H. Winston, : Artificial Intelligence, Addison-Welsey, 2nd Edition (1984)
- 15) J. Amsterdam : Extending the Valiant Learning Model ; Proc. of the 5-th Int. Conf. on Machine Learning, pp. 381~394 (1988)
- 16) K. A. Huyser & M. A. Horowitz : Generalization in Connectionist Networks that Realize Boolean Functions ; Proc. of the 1988 Connectionist Models, pp. 191~200, Morgan Kaufmann (1988)
- 17) G. Cybenko : Continuous Valued Neural Networks with Two-layers are Sufficient, Tech. Report, Tuft Univ. (1988)
- 18) 生天目, 木俣 : ニューラルネットワークの情報表現と非線形関数の同定 ; 情報処理学会論文誌, 30 巻, 10 号, pp. 1274~1280 (1989)