

学位論文題名

確率的ニューラルネットワークの
情報量の解析および学習効率化の研究

学位論文内容の要旨

確率的性質を持つ情報工学の研究対象においては、多次元の確率変数に関する分布を考える場合、複数の確率変数によって定義される情報量を用い、確率変数どうしの依存関係、パラメータ推定の限界、および符号化の限界等が定量的かつ定性的に理解されている。このような研究対象をモデル化する際に、確率分布の広いクラスを含む代表的な分布族である指数型分布族は、パラメトリックな統計モデルとして広く用いられている。二値指数型分布族は、二つ以上の確率変数の間の交互作用を考慮に入れたものとして、最も単純な統計モデルであり、高次の交互作用を考慮した一般化された確率的な動作をするニューラルネットワークの平衡分布の集合、あるいは各情報源が二状態の多元的な情報源における確率分布の集合と一致する。

一般に、情報量それ自体の性質に関する研究は広く行なわれている。しかし、このモデルではパラメータと情報量の間密接な関係が存在しているにもかかわらず、高次の交互作用の影響を反映した情報量と高次のパラメータとの依存関係を詳しく解析した研究は見受けられず、Kullback 情報量を用いた確率的ニューラルネットワークの学習の研究においても、計算機シミュレーション等が多く、理論解析はまだ十分に行なわれていない。

本研究は、このような問題意識をもとに、確率的ニューラルネットワークが実現する確率分布である二値指数型分布族における統計量・情報量とパラメータとの間の依存関係を解析し、パラメータとこれらの情報量との依存関係における、互いに双対的な関係にある単調増加と単調減少という性質、および変化率の半順序関係等の特徴的な新しい事実を明らかにすること、さらにパラメータ推定のための勾配法の最良化によって学習を効率化することについて考察することを研究目的とする。論文の構成は、以下のようになっている。

第1章では、確率的な性質を有する情報工学的対象とその背景について述べ、本研究の位置付けと目的について述べる。

第2章では、第3章以降での議論の準備として、代表的な確率分布の族である指数型分布族に関する一般論について述べ、確率分布と統計的モデル、指数型分布族・混合型分布族と確率分布の空間における幾何学構造の関係について述べる。さらに、統計的推定において重要である Fisher 情報量、十分統計量等について述べる。

第3章では、確率的に動作するネットワークであるボルツマンマシンの自然な拡

張を考えることにより、ボルツマンマシンが実現する確率分布を高次の交互作用を考慮した形で一般化して取り扱う。この確率分布の全体は、状態集合が $\mathcal{X} = \{0,1\}^n$ である自然パラメータと双対な関数が二値指数型分布族に一致することから、これに関して定義される統計量（期待値、Fisher 情報行列の要素、正規化された Fisher 情報行列の要素）に対して、自然パラメータ（ニューラルネットワークにおける閾値パラメータ、結合パラメータ）との関係を考察する。特に、相関係数が所定の包含関係をみたすような自然パラメータの変化に対して単調に変化するという著しい特徴を明らかにする。同時に、単調増加と単調減少という性質が、互いに双対的な関係にあることを示す。

第4章では、第3章で考察した二値指数型分布族のうち、最も一般化された次数 n の二値指数型分布族（次数 n の Higher-Order ネットワークが実現する平衡確率分布の全体、または状態集合上の正值確率分布の全体に一致）に対して定義される期待値パラメータに対する Fisher 情報行列、および正規化された Fisher 情報行列の要素と、期待値パラメータとの関係を調べる。特に、自然パラメータに対する相関係数の単調性と対照的な集合条件で、期待値パラメータに対する偏相関係数に単調性が現れるという著しい特徴を明らかにする。同時に、単調増加と単調減少という性質が、やはり互いに双対的な関係にあることを示す。

第5章では、統計量と同様に確率分布に関して定義される量である情報量（Shannon の結合エントロピ、条件付きエントロピ、相互情報量、条件付き相互情報量、および結合エントロピで正規化された相互情報量）が、自然パラメータにどのように依存するかについて調べる。特に、第3章で明らかにされた、自然パラメータと統計量の関係・性質と類似の性質が情報量に対しても成立し、特定の条件付きエントロピ、条件付き相互情報量が所定の包含関係をみたすような自然パラメータの変化に対して単調に増加ないし減少するという事実を明らかにする。またそれ自体に単調性がない結合エントロピ・相互情報量に関して、相関係数のアナログとして定義された結合エントロピで正規化された相互情報量が、やはり所定の包含関係をみたすような自然パラメータの変化に対して単調に増加ないし減少するという事実を明らかにする。

第6章では、指数型分布族における自然パラメータ推定のためのパラメータ修正規則（ボルツマンマシンにおける結合・閾値パラメータ決定のための学習に一致）において、Kullback 情報量を目的関数とする、その勾配ベクトルを用いた一次の勾配法、および Kullback 情報量の計量行列である Fisher 情報行列を用いた二次の勾配法を、Kullback 情報量変化の下限を最大にするという意味での最良化を行なう。これによって勾配法のステップ幅を決定し、Kullback 情報量変化の限界（下限と上限）を評価する。このステップ幅には、勾配法を開始する以前に既知である十分統計量を用い、従来求められていた Bianchi の評価よりも改善されることを定量的に示す。さらに、一次の勾配法と二次の勾配法の上限と下限を比較・考察する。

第7章では、まとめとして、第3章から第6章までの結果に関する考察を行ない、今後の課題・研究の方向性について述べる。

学位論文審査の要旨

主 査 教 授 伊 達 惇
副 査 教 授 宮 腰 政 明
副 査 教 授 新 保 勝
副 査 教 授 佐 藤 義 治

学 位 論 文 題 名

確率的ニューラルネットワークの 情報量の解析および学習効率化の研究

確率的性質を持つ情報工学の研究対象においては、多次元の確率変数に関する分布を考える場合、確率変数間の依存関係、パラメータ推定の限界、および符号化の限界等が、複数の確率変数によって定義される情報量を用いて定量的かつ定性的に捉えられる。確率分布の広いクラスを含む代表的な分布族である指数型分布族は、このような研究対象をモデル化する際に、パラメトリックな統計モデルとして広く用いられている。

情報量それ自体の性質に関する研究は広く行なわれているが、情報量と統計モデルのパラメータ間の依存関係を詳しく解析した研究は見受けられず、確率的ニューラルネットワークの学習の研究においても、学習精度、収束の速度等は論じられているが、学習初期段階からの効率に関する理論解析はまだ十分に行なわれていない。

本論文では、このような問題意識をもとに、対象を確率的ニューラルネットワークに限定して、それが実現する確率分布である二値指数型分布族における統計量・情報量とパラメータとの間の依存関係を解析し、特徴的な新しい事実を明らかにする。さらにパラメータ推定のための勾配法による学習効率の限界を改善し、それを最良化する方法について考察する。

第1章では、確率的な性質を有する情報工学的対象とその背景について述べ、本研究の位置付けと目的について述べている。

第2章では、第3章以降での議論の準備として、代表的な確率分布の族である指数型分布族に関する一般論について述べている。

第3章では、確率的に動作するネットワークであるボルツマンマシンの自然な拡張を考えることにより、ボルツマンマシンが実現する確率分布を高次の交互作用を

考慮した形で一般化し、これに関する統計量と自然パラメータとの関係を考察している。特に、相関係数が所定の包含関係をみたすような自然パラメータの変化に対して単調に変化するという著しい特徴を明らかにしている。同時に、対象とする統計量の単調増加と単調減少の現れる条件が、互いに双対的な関係にあることを示している。

第4章では、第3章で考察した二値指数型分布族のうち、一般化された次数の二値指数型分布族に対して定義される期待値パラメータに対する Fisher 情報行列、および正規化された Fisher 情報行列の要素と、期待値パラメータとの関係を調べている。特に、自然パラメータに対する相関係数の単調性と対照的な集合条件で、期待値パラメータに対する偏相関係数に単調性が現れるという著しい特徴を明らかにしている。同時に、単調増加と単調減少という性質が、やはり互いに双対的な関係にあることを示している。

第5章では、統計量と同様に確率分布に関して定義される重要な量である情報量（Shannon の結合エントロピ、条件付きエントロピ、相互情報量、条件付き相互情報量、および結合エントロピで正規化された相互情報量）が、自然パラメータにどのように依存するかについて調べている。特に、第3章で明らかにされた、自然パラメータと統計量の関係・性質と類似の性質が情報量に対しても成立し、特定の条件付きエントロピ、条件付き相互情報量が所定の包含関係をみたすような自然パラメータの変化に対して単調に増加ないし減少するという事実を明らかにしている。

第6章では、指数型分布族における自然パラメータ推定のためのパラメータ修正規則（ボルツマンマシンにおける結合・閾値パラメータ決定のための学習に一致）において、Kullback 情報量を目的関数とする、その勾配ベクトルを用いた一次の勾配法による学習効率、および Kullback 情報量の計量行列である Fisher 情報行列を用いた二次の勾配法による学習効率を、情報量変化の下限を最大にするという意味で最良化している。これによって勾配法のステップ幅を改善し、Kullback 情報量変化の限界（下限と上限）を評価し、従来求められていたものよりも改善された評価を得ることができる。また、この改善される量の限界を求め、最大値を求めている。

第7章では、まとめとして、第3章から第6章までの結果に関する考察を行ない、今後の課題・研究の方向性について述べている。

以上のように、本論文において、確率的ニューラルネットワークの解析と学習効率化を行なったことによって、工学への応用の上で有益な知見を得ているもので、情報工学の発展に寄与するところが大きい。よって、著者は、博士（工学）の学位を授与される資格あるものと認める。