

学位論文題名

情報量の数理的構造および学習問題に
関する情報工学的研究

学位論文内容の要旨

近年における情報処理の高速化、大容量化などには、高い効率を実現するための情報伝送、高い圧縮率を達成するためのデータ圧縮などの技術の進歩が重要な役割を果たしてきた。情報理論において、データが持つ確率構造や、そのデータを生成している情報源のパラメータについての情報を手がかりとして、通信路符号化理論や情報源符号化理論により、それらの技術の理論的限界が示され、それを目標とする研究がなされてきた。この研究によって、データを効率よく表現する符号化からデータの情報構造を見だし、効率のよいデータ圧縮を実現してきた。さらに、これらの研究は、数理統計学、ニューラルネットワーク、計算論的学習理論などに応用され、数多くの成果が得られた。それらは、データの情報構造および確率空間や情報量の数理的構造と、推論、推定、学習、識別などの工学的応用との対応によって得られる成果である。このような背景から、確率空間や情報量の数理的構造に関する基礎的研究とその情報工学的応用との関連性は、より広範囲な工学への応用のためにも重要な基礎的課題であると言える。

Kullback 情報量は、二つの確率分布の差異を表すだけではなく、符号化、データ圧縮、仮説検定などにおいて重要な役割を果たし、また、数理統計でも様々な情報量が用いられている。Csiszár は、様々な情報量を f -divergence と呼ばれる一つのクラスとして表現することで一般化し、無限集合上の確率分布全体での位相に関する性質について研究した。しかし、その情報量は、適用条件が明らかにされない形で表現されているため、情報工学における応用研究が少なく、一般的な情報量を用いることのメリット、有用な情報量のクラスやその性質などは、明らかにされていない。

本研究は、このような問題意識をもとに、情報量の持つ数理的構造に関する研究と情報工学での有効な応用とその意味を考察することを研究目的としている。

本論文では、情報量の持つ構造に関する基礎的考察を行ない、情報量の評価不等式や新たな情報量の提案およびそれらの既存の情報量に対する関係などを明らかにし、さらに、その構造を計算論的学習理論の問題に適用した結果を示す。本論文の構成は、以下のとおりである。

Chapter 1 では、本研究の背景や意義、およびその位置について述べている。ま

ず、エントロピーなどの種々の情報量の定義やその性質について触れ、工学や統計などにおける情報量の重要性を述べる。そして、一般化情報量として定義された f -divergence に関する数理構造や計算論的学習理論への応用といった本研究の意義やその位置および本論文の構成について述べている。

Chapter 2 では、まず、 f -divergence について、その導入の経緯および性質について述べ、そして、 f -divergence に関して新しく得られた結果について紹介する。さらに、それらの結果と、これまでに示された Csiszár や Vajda の結果との比較を行う。この研究によって得られた結果は、Vajda が f -divergence の単調性を用いて証明した不等式と比較して、確率分布に関する制限がないことや、証明法が非常に見通しがよく、簡単であり、さらに、それを用いることで、 f -divergence の最大値を直接導出できるという特徴を持つ。そして、情報量と三角不等式の関連性について、凸関数 f が三角不等式を満たすための必要十分条件を導出し、また、すべてのクラスの f -divergence について成り立つ弱い意味での三角不等式を提案する。

Chapter 3 では、新しい情報量として Hermite-Hadamard divergence を導入し、その性質を、いくつかの定理や補題を用いて示している。次に、その正值性などの基本性質や用いる凸関数に関する不定性を示し、Kullback 情報量および Hellinger distance との関係性を明らかにする。そして、凸関数のクラス上の変換に対して成り立つ性質と Hermite-Hadamard divergence との関連性についても考察する。

Chapter 4 では、計算論的学習理論における統計的な学習モデルに、情報量の評価不等式を適用し、学習の損失関数に f -divergence を用いた場合に拡張したときの学習評価、必要最小事例数とそのオーダーを導出し、情報量の数理構造との関係を論じている。

Chapter 5 では、本研究に関するまとめと今後の課題が述べられている。さらに、本研究によって得られる補題や定理の導出するための数学的手段として用いた主要な数学的諸性質を、補遺の形でまとめている。

学位論文審査の要旨

主 査 教 授 伊 達 惇
副 査 教 授 新 保 勝
副 査 教 授 大 内 東
副 査 教 授 佐 藤 義 治

学 位 論 文 題 名

情報量の数理的構造および学習問題に 関する情報工学的研究

情報処理の高速化、大容量化において、高度な情報伝送技術およびデータ圧縮技術が重要な役割を果たしている。データが持つ確率構造の特性を手がかりとした理論的限界が、通信路符号化理論や情報源符号化理論によって示されており、それを目標とする研究が推進されている。さらに、それらの研究とニューラルネットワーク、パターン識別、計算論的学習理論、統計などの分野の研究と共に進められ、数多くの成果が得られている。すなわち、これらの成果は、データの情報構造および確率空間や情報量の数理的構造に関する基礎的研究と、推論、推定、学習、識別などの工学的応用研究との相互関係を通して得られるものであり、両者の関連性に関する考察は、重要な課題と言える。

二つの確率分布の差異を表す情報量としては、Kullback 情報量が、工学的に最も広く適用されているが、他にも様々な情報量が提案されている。Csiszár は、これらの情報量を統一的に扱うために、 f -divergence と呼ばれる情報量を導入し、それを活用して、無限集合上の確率分布全体での位相に関する性質を整理した。しかし、その情報量は、明らかな形で表現されていないため、その後の応用研究において活用されることが無かった。

本論文においては、このような問題意識をもとに、情報量の持つ数理的特性の解明と情報工学における活用を目的として、情報量の持つ構造に関する基礎的考察を行ない、情報量の評価不等式や新たな情報量の提案、およびその既存の情報量に対する関係などを明らかにし、さらに、情報量の数理的特性に基づいて、計算論的学習理論の問題に対する考察を行なっている。

Chapter 1 では、研究の背景について述べ、 f -divergence に関する数理構造や計算

論的学習理論への応用に関する研究の意義やその位置および論文の構成について述べている。そして、エントロピー、Kullback 情報量、相互情報量などの基本的な情報量に関する性質や、符号化や推定におけるそれらの情報量の重要性を述べている。

Chapter 2 では、まず、 f -divergence に関して、その導入の経緯および正値性や Kullback 情報量などの既存の情報量との関係について述べている。そして、その情報量を不変にする性質を示し、この性質を用いて、その情報量の評価に関する新しい結果を導いている。さらに、Csiszár や Vajda の結果との比較を行うことによって、結果の有効性を検証している。この結果は、確率分布に関する制約が無く、導出法が明解であり、それを用いることによって、 f -divergence の最大値を直接導出できるという利点を持っている。また、情報量と三角不等式の関連性についての考察から、 f -divergence が三角不等式を満たすための必要十分条件の導出を行ない、さらに、 f -divergence について成り立つ弱い意味での三角不等式を導出している。

Chapter 3 では、Hermite-Hadamard 情報量という新しい情報量を導入し、その正値性などの基本性質を明らかにし、その情報量を不変にする性質を示している。そして、既存の情報量との関係を明確にするために、Kullback 情報量および Hellinger 距離との関係を示す評価不等式を導出している。

Chapter 4 では、計算論的学習理論における統計的な学習の枠組に、先に主要結果として示した情報量の評価不等式を適用することによって、情報量の数理解造と学習問題との関連性について理論的考察が行なっている。学習の損失関数として f -divergence を用いた場合における学習評価の関係式の導出や必要最小事例数とそのオーダーの算出を行なっている。これらのことから、情報量の数理解造と学習との関連性を明らかにしている。

Chapter 5 では、結びとして、本研究のまとめと今後の研究課題を述べている。さらに、本研究の主要結果として得られている補題や定理の導出するための数学的手段として用いた主要な数学的諸性質を、補遺の形でまとめている。

以上のように、本論文において、情報量の数理的構造の解明と学習問題への適用を行なったことによって、一般化情報量の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。