

機関番号：32619
 研究種目：基盤研究 (C)
 研究期間：2008～2010
 課題番号：20500214
 研究課題名 (和文) 区間演算から多峰性/ファジィ分布演算へ-不確実システムの包括的数値解析-
 研究課題名 (英文) From interval arithmetic to multi-peak/fuzzy distribution arithmetic - comprehensive numerical analysis for uncertain systems -
 研究代表者
 神澤 雄智 (KANZAWA YUCHI)
 芝浦工業大学・工学部・准教授
 研究者番号：00298176

研究成果の概要 (和文) : クラスタリングを不確実モデル化した問題に対する従来法を確率論的側面から見直し、対象データ、もしくは、その特徴空間における状態が正規分布として与えられた条件の下でクラスタリング、主成分分析、もしくは線形クラスタリングを行うために、クラスタリング手法の基となる最適化問題中の目的関数に、データの不確実性を表す変量の2ノルムの2乗をコスト関数として加えた、新たな目的関数を最小化することによって、新たなクラスタリング手法を提案した。

研究成果の概要 (英文) : In this research, conventional clustering methods for uncertain data are reviewed from the probabilistic point of view, new clustering methods are proposed for clustering, PCA, or linear clustering in the data space or in the feature space, which are obtained by solving the optimization problem adding squared Euclidean norm of the value of uncertainty for data to the original objective functions.

交付決定額

(金額単位：円)

	直接経費	間接経費	合計
2008年度	1,300,000	390,000	1,690,000
2009年度	1,000,000	300,000	1,300,000
2010年度	900,000	270,000	1,170,000
年度			
年度			
総計	3,200,000	960,000	4,160,000

研究分野：精度保証付き数値計算とファジィクラスタリング
 科研費の分科・細目：情報学・感性情報学・ソフトコンピューティング
 キーワード：多峰性分布，ファジィ分布

1. 研究開始当初の背景

本研究代表者は、長らく精度保証付き数値計算を用いた非線形解析に関する研究に従事しており、また、クラスタリングに関する研究を始めてきていた。これらの研究課題はいずれも決定論的モデルを対象としている一方で、現実の問題を対象とするために、モデル化の基となる現象に潜む不確実性や与えられたデータに潜む人間の主観を考慮してこれまでの研究成果を新しく展開していきたいと考えるようになった。

区間解析は、丸め誤差を考慮するための道具としてだけでなく、現象の数理モデル化における物理量のバラツキを考慮するためにも有用である。例えば、多段論理回路を構成するトランジスタ特性がバラついた時の遅延時間が、モンテカルロ法よりもずっと少ない計算量で、区間解析によって高精度に計算されているように、区間解析は、バラツキ具合を考慮しなければ(すなわち、一様分布と仮定すれば)従来行われてきたモンテカルロ法にとって代わる手法と云える。本研究代表者

は十数年にわたって区間解析やその周辺問題について研究を続けており、成果の一つは電子情報通信学会で論文賞を受賞している。

クラスタリングは与えられたデータを幾つかのグループに分類する手法の一つであり、パターン認識をはじめとする様々な場面で活用されている。データが物理量の測定から得られたり人間の主観を数値化したものであれば当然、不確実性を考慮する必要がある。本研究代表者は不確実性を表す変数を元の最適化問題の変数として扱い、データがある程度の誤差を含んでいても安定にクラスタを形成する手法を提案したが、データのバラツキ具合によるクラスタリング結果の影響について考慮できていなかった。

これまでになされてきた区間解析を確率解析的視点から捉えると、台が有限区間で一様分布として表される変数を扱ってきたことになる。そこで、区間解析を拡張して、一般的な分布として表される変数を扱えるようにしたいと考え始めた。独立に正規分布で表される確率変数の線形演算ならばわざわざ新しい手法を提案しなくとも母数に関する演算を行えば済む話だが、非線形演算や、入力変量の分布が多峰性を有したり確率分布の枠組を越えたファジィ分布の場合には最早、出力の分布の形状については簡単な場合を除いて理論的解析のしようがない。従来、データの不確実性は正規分布とそれを基とする確率過程で表現されて解析されてきているが、礫岩・礫層の粒度分布や人体の画像から姿勢を推定する際に一部の部位が自己隠蔽されている場合の曖昧性は多峰性分布となることが報告されているように、扱うデータの不確実性が必ずしも正規分布で表されるとは限らなかった。

一方で、ブートストラップ法のようなモンテカルロ法によって、入力変量の分布が複雑な場合でも出力の分布についてある程度の知見が得られるものの、入力変量の膨大な標本に対する演算を行うことになるし、その知見が得られるための標本数には何ら根拠はない。確率微分方程式は、動的な不確実モデルとして近年に目覚しく研究されており、数理生物学などの科学の分野では勿論のこと、金融の分野でもその有用性が認められている。しかし、これまでに提案されているその数値解法はすべて、あくまでも標本点を固定したときに得られる解の標本点であり、系の持つ特性を調べるためにはやはりモンテカルロ法のように、多くの標本点に対する解の標本を得なければならないし、ここでも、その知見が得られるための標本数には何ら根拠はない。

2. 研究の目的

本研究課題は、変量のバラツキが多峰性を持ったリファジィ分布のときにモンテカル

ロ法より効率的に、包括的に解析できるような解析手法を提案することである。まず、確率変数に対する四則演算や初等関数などの基本演算アルゴリズムを定義する。初等的な確率理論において、確率変数 $x \in X$ に演算 $T: X \rightarrow Y$ を施した結果の確率変数 $y \in Y$ の密度関数 f_Y は、 x の確率密度関数 f_X を用いて表されることを直接的に利用して数値計算することは丸め誤差や打ち切り誤差の影響が大きくタブーとされてきたが、近年の精度保証の技術の発展はこのタブーを打ち破ることができると考えられている。精度保証付き数値計算の思想に基づいた厳密な実装を目標とするが、当面の実験環境を早急に整えるために単純な演算モデルの構築から始めて徐々に理想演算モデルの構築を行っていく。次に、実装した演算ライブラリを用いて各種問題に適用して、精度と計算量の両側面から理論値や従来のモンテカルロ法と比較していく。線形問題は理論解析可能であり、本手法を絶対評価するために実験する。静的非線形問題については、本申請者がこれまでに研究してきた決定論モデル(非線形方程式の求解問題、非線形最適化問題、クラスタリング)を不確実モデル化して得られる問題を対象としてモンテカルロ法と比較していく。動的非線形問題についてはまず、決定論モデルとして常微分方程式に対して本申請者が研究してきた手法を確率解析の立場で焼き直して数値計算アルゴリズムを構築し、Rumelin法などの確率常微分方程式のための数値スキームと比較していく。本研究課題では当座、各確率変数が独立に分布することを仮定するが、より現実に即した問題を扱えるようにするために、各確率変数が互いに相関を有する場合や、動的問題については非ウィーナ過程の場合を想定してさらなる研究発展を目指す。

3. 研究の方法

【平成 20 年度】

平成 20 年度は、確率変数の台を有限区間に限定し、かつ、解析対象を静的問題に限定して本研究課題の有効性を理論・実験の両側面から実証していく。具体的には、以下に挙げる各研究内容について、おおよそ番号順に実施する。

(1). 既に構築してある有限区間台近似分布演算アルゴリズムを、線形演算について適用して、理論値との比較を行う。特に、分布自体の近似と演算の近似それぞれについて、近似精度を高めるにつれて理論値に収束していくことを理論・実験の両側面から実証していく。

(2). (1) と平行して、有限区間台近似分布演算アルゴリズムを、本申請者がこれまで研究してきた決定論モデル(非線形方程式の求解問題、非線形最適化問題、クラスタリング)

について不確実モデル化した問題に適用する。既に小規模・単純な問題については幾つか適用して良好な結果が得られているが、より大規模・複雑な問題について系統だって実験していく。

(3). 有限区間台近似分布演算アルゴリズムを拡張して、有限区間台精度保証アルゴリズムを行う。その本質は、これまで行ってきた非線形常微分方程式の解の数値的存在検証法の研究の中で、関数の多項式近似に関連して、解である関数を上下から精度良く少ない情報量で包み込む技術を応用して、様々な不確実分布を表す確率密度関数を同様に包み込みながら演算を行う技術を構築することである。

(4). 構築した有限区間台精度保証分布演算アルゴリズムを、(1)と同様に線形問題に適用して理論値と比較し、また、(2)と同様に非線形問題に適用して、従来法との比較を行う。経費の使用については、計算用コンピュータが中心になる。その他に、旅費は、研究打合せの他に、成果発表の目的で使用する。また、方法論、理論的研究では、比較的早期に、研究成果を公にすることが可能である。よって、本年度中に、数回の学会発表を行うこととする。また、研究への協力の半分については資料整理やデータ入力などの比較的一般的な業務に充てるが、残り半分については、システム作業、実験結果処理など、ある程度専門的知識をもつ実験補助者に充てる。実験補助者は主に当研究室に所属する大学院生から雇用する。謝金は、成果の英文化のための指導に充てる。

4. 研究成果

平成 20 年度は、研究代表者が前年度までに研究してきたクラスタリングを不確実モデル化した問題を確率論的側面から見直し、新たなクラスタリング手法を提案した。具体的には、従来、クラスタリング対象データが特徴空間上の点として与えられているという仮定を緩和して、データがある領域内を自由に動き回り得るという条件の下でクラスタリングを行う手法が提案されていたが、この手法を確率論的側面から見直すと、対象データがある有限台上一様分布として与えられた条件の下でクラスタリングを行うことと解釈することができた。この解釈を基に、対象データが正規分布として与えられた条件の下でクラスタリングを行うために、クラスタリング手法の基となる最適化問題中の目的関数に、データの不確実性を表す変量の 2 ノルムの 2 乗をコスト関数として加えた、新たな目的関数を最小化することによって、新たなクラスタリング手法を提案した。また、この手法にカーネル関数を導入し、分類境界が非ボロノイ形状となる場合にも適用でき

るように拡張した。

平成 21 年度は、研究代表者が前年度までに研究してきたクラスタリングを不確実モデル化した問題をデータ解析手法の一つである主成分分析ならびにそのファジィクラスタリング版であるファジィ線形クラスタリングに応用した。具体的には、従来、主成分分析ならびにファジィ線形クラスタリング対象データが特徴空間上の点として与えられているという仮定を緩和して、データがある領域内を自由に動き回り得るという条件の下で主成分分析もしくはファジィ線形クラスタリングを行う手法を提案した。この手法を確率論的側面から見直すと、対象データがある有限台上一様分布として与えられた条件の下でクラスタリングを行うことと解釈することができることを前年までの研究成果から予想でき、さらにこの解釈を基に、対象データが正規分布として与えられた条件の下で主成分分析もしくはファジィ線形クラスタリングを行うために、各手法の基となる最適化問題中の目的関数に、データの不確実性を表す変量の 2 ノルムの 2 乗をコスト関数として加えた、新たな目的関数を最小化することによって、新たなクラスタリング手法が提案できる展望を得ることができた。

平成 22 年度はまず、研究代表者が前年度までに研究してきたクラスタリングを不確実モデル化した問題を、陽写像を用いたカーネルファジィクラスタリングに応用した。具体的には、従来、特徴空間上の挙動は不明なままに特徴空間における内積値だけが与えられている条件下でクラスタリングしてきたために特徴空間上の不確実性をも直接には扱うことができなかつたのに対して、陽写像を用いることによって従来法の特性を損なうことなく、特徴空間上の不確実性を定量的に扱いつつクラスタリングを行う手法を提案した。本年度はさらに、従来ではデータの不確実性をデータがある領域内を自由に動き回り得るという条件下で考えていたのに対して、不確実性を表す変量の 2 乗損失をクラスタリングの目的関数に付加することによる、新たなクラスタリング手法を提案した。この手法を確率論的側面から見直すと、対象データがある正規分布として与えられた条件の下でクラスタリングを行うことと解釈することができることが分かりつつある。この解釈を基に、対象データが正規分布とは異なる他の確率分布として与えられた条件の下でファジィクラスタリングを行うために、各手法の目的関数にデータの不確実性を表す何かしらのコスト関数を加えることによって、新たな手法が提案できる展望を得ることができた。

5. 主な発表論文等

(研究代表者、研究分担者及び連携研究者に

は下線)

〔雑誌論文〕(計 6 件)

- ① Yuchi Kanzawa, Yasunori Endo, Sadaaki Miyamoto, Fuzzy c-Means Algorithms for Data with Tolerance Using Kernel Functions, IEICE Trans. Fundamentals, Vol. E91-A, No. 9, 2008, pp.2520-2534
- ② Yuchi Kanzawa, Yasunori Endo, Sadaaki Miyamoto, Fuzzy Classification Function of Standard Fuzzy c-Means Algorithms for Data with Tolerance Using Kernel Function, Springer Lecture Notes in Artificial Intelligence, LNAI5285, 2008, pp.122-133
- ③ Yuchi Kanzawa, Yasunori Endo, Sadaaki Miyamoto, Fuzzy Classification Function of Fuzzy c-Means Algorithms for Data with Tolerance, 知能と情報, Vol.21, No.4, 2009, pp.438-451
- ④ Yasunori Endo, Yasushi Hasegawa, Yukihiro Hamasuna, Yuchi Kanzawa, Fuzzy c-Means Clustering for Uncertain Data Using Quadratic Penalty-Vector Regularization, Journal of Advanced Computational Intelligence and Intelligent Informatics, Vol.15, No.1, 2011, pp.76-82
- ⑤ Yuchi Kanzawa, Yasunori Endo, Sadaaki Miyamoto, Semi-Supervised Fuzzy c-Means Algorithm by Revising Dissimilarity Between Data, Journal of Advanced Computational Intelligence and Intelligent Informatics, Vol.15, No.1, 2011, pp.95-101
- ⑥ Yuchi Kanzawa, Yasunori Endo, Sadaaki Miyamoto, Indefinite Kernel Fuzzy c-Means Clustering Algorithms, Lecture Notes in Artificial Intelligence, LNAI6408, pp.126-138

〔学会発表〕(計 15 件)

- ① Yuchi Kanzawa, Yasunori Endo, Sadaaki Miyamoto, Fuzzy C-Means for Data with Tolerance Introducing Penalty Term in Feature Space, 2008 International Symposium on Nonlinear Theory and Its Applications, 2008年9月7日, ブダペスト(ハンガリー)
- ② Yuchi Kanzawa, Yasunori Endo, Sadaaki Miyamoto, Entropy Regularized Fuzzy c-Means for Data with Tolerance introducing Penalty Term in Feature Space, Joint 4th International Conference on Soft Computing and Intelligent Systems and 9th International Symposium on advanced

Intelligent Systems, 2008年9月17日, 名古屋(日本)

- ③ 神澤雄智、遠藤靖典、宮本定明、Kernelized Fuzzy c-Means for Data with Tolerance based on Opposite Criteria, 第24回ファジィシステムシンポジウム, 2008年9月3日, 大阪(日本)
- ④ Yuchi Kanzawa, Yasunori Endo, Sadaaki Miyamoto, Fuzzy C-Means for Data with Tolerance by Introducing Penalty Term, 2008 IEEE Conference on Soft Computing in Industrial Applications, 2008年6月27日, 室蘭(日本)
- ⑤ Yuchi Kanzawa, Yasunori Endo, Sadaaki Miyamoto, Fuzzy Classification Function of Fuzzy c-Means Algorithms for Data with Tolerance, 2008 World Congress on Computational Intelligence, 2008年6月1日, 香港(中国)
- ⑥ Yuchi Kanzawa, Yasunori Endo, Sadaaki Miyamoto, Entropy Regularized Fuzzy c-Lines for Data with Tolerance, 2011 IEEE International Conference on Fuzzy Systems, 2009年8月21日, Jeju(Korea)
- ⑦ Yuchi Kanzawa, Yasunori Endo, Sadaaki Miyamoto, Fuzzy c-Lines for Data with Tolerance, 2009 International Fuzzy System Association World Congress / 2009 European Society for Fuzzy Logic and Technology, 2009年7月23日, Lisbon (Portugal)
- ⑧ Yuchi Kanzawa, Yasunori Endo, Sadaaki Miyamoto, A Semi-Supervised Entropy Regularized Fuzzy c-Means, 2009 International Symposium on Nonlinear Theory and Its Applications, 2009年10月21日, 北海道(日本)
- ⑨ 神澤雄智、遠藤靖典、宮本定明、許容付きデータのための1次元線形ファジィクラスタリング、第25回ファジィシステムシンポジウム、2009年7月14日、筑波(日本)
- ⑩ 新宮弘敏、神澤雄智、非線形方程式の解曲線追跡における区間演算とアフィン演算の比較、応用数理学会 2009年度年会、2009年9月30日、大阪(日本)
- ⑪ 新宮弘敏、神澤雄智、非線形方程式の解曲線追跡における区間演算とアフィン演算の比較、電子情報通信学会回路とシステム・非線形問題研究会、2009年9月24日、広島(日本)
- ⑫ Yuchi Kanzawa, Yasunori Endo, Sadaaki Miyamoto, Two Semi-supervised Entropy Regularized Fuzzy c-Means, Joint 5th International Conference on Soft Computing and Intelligent Systems and

9th International Symposium on advanced Intelligent Systems, 2010年12月9日, 岡山(日本)

- ⑬ Yuchi Kanzawa, Yasunori Endo, Sadaaki Miyamoto, Geodesic-Dissimilarity-based Hard and Fuzzy c-Means, Joint 5th International Conference on Soft Computing and Intelligent Systems and 9th International Symposium on advanced Intelligent Systems, 2010年12月9日, 岡山(日本)
- ⑭ Yuchi Kanzawa, Yasunori Endo, Sadaaki Miyamoto, Hard and Fuzzy c-Means Clustering Algorithms with Geodesic Dissimilarity, 2010 International Symposium on Nonlinear Theory and Its Application, 2010年9月8日, クラコフ (ポーランド)
- ⑮ Yuchi Kanzawa, Yasunori Endo, Sadaaki Miyamoto, On Kernel Fuzzy c-Means for Data with Tolerance using Explicit Mapping for Kernel Data Analysis, 2010 IEEE World Congress on Computational Intelligence, 2010年7月21日, バルセロナ(スペイン)

[図書] (計0件)

[産業財産権]

○出願状況 (計0件)

○取得状況 (計0件)

6. 研究組織

(1) 研究代表者

神澤 雄智 (KANZAWA YUCHI)
芝浦工業大学・工学部・准教授
研究者番号: 00298176

(2) 研究分担者

なし ()

(3) 連携研究者

なし ()