

学 位 論 文 題 名

ファジィクラスタリング手法の開発と

その応用に関する研究

学位論文内容の要旨

データのもつ情報を構造モデルや探索的方法により要約し、データの変動要因を明らかにすることはデータ解析法の様々な手法に共通した目的である。近年情報量の増大に伴いデータ解析法の果たす役割がますます重要となっている。クラスター分析はデータを類似性に基づき分類することによって内在する構造を明らかにしようとするものであり、データ解析において最も基本的な分析手法の一つである。

一方 L.A. Zadeh によってファジィ理論が提唱されて以来、制御理論のみならず多くの分野でその応用可能性が論じられている。データ解析の分野においても“あいまいさ”を考慮した構造モデルや“あいまいさ”を含むデータに対する分析法の研究が数多くなされファジィデータ解析なる分野も形成されつつある。クラスター分析においても従来排反的なクラスターの構成が中心であったのに対して分類対象の集合上のファジィ部分集合としてのファジィクラスタの構成法が注目されている。またデータの観測技術の向上や計算機の処理能力の増大によって、従来の 2-way データ(個体×変量)のみならず 3-way データ(個体×変量×時点)さらに一般には multi-way データを解析するための理論や手法の開発が要求されている。このような背景の下で、本論文はファジィクラスタリングの新しい手法、すなわち多基準ファジィクラスタリング手法および加法的ファジィクラスタリングモデルを提案し、それらの有効性、妥当性を論じたものである。

本論文は以下の 5 章から構成されている。

第 1 章では、クラスター分析について、その背景と分析の目的あるいは対象となるデータおよびその性質を論じた。さらに従来のクラスタリング手法について概説し、本研究の目的について述べた。

第 2 章では、ファジィ理論について解説し、ファジィ集合論とクラスタリングの関連性および従来のファジィクラスタリングに関する研究について述べ本研究の立場を明確にした。

第 3 章では、多基準ファジィクラスタリング手法の提案を行ない、その応用として 3-way データのクラスタリング問題について論じた。このクラスタリング問題は各時点におけるクラスタリング基準をまとめて多基準とみなし、各時点に共通なファジィクラスタを構成する多基準最適化問題として定式化できることを示した。一般に多基準最適化問題における解はパレート最適解として得られる。多基準ファジィクラスタリング手法においては求められるファジィクラスタがパレート最適解となる。3-way データを個体と変量とからなる一つの 2-way データにまとめて分析を行なうこともできるが、この場合時点の差異を無視し、暗に同質性を仮定してクラスタリングを行なうこととなる。また各時点ごとに個別にクラスタリングを行なった場合得られた結果の時点間の関連性を論ずることは困難である。ここでは

各時点に共通なクラスターを生成する, すなわち各時点での個体間の類似性のみならず各時点を通じての観測値の変化パターンの類似性をも加味した分類手法を提案した. また従来の多基準ハードクラスタリング問題では避けることのできない組合せ最適化問題をファジィ集合の概念を用いた多基準ファジィクラスタリング問題において回避できることは本手法の利点の一つと考えられる. さらに, 多基準最適化問題におけるいくつかの定理を応用し 3-way データのクラスタリング問題における新たな定理, すなわちパレート最適解がファジィクラスターとなるための十分条件を与える定理を証明した. ここで提案した手法の有効性を応用的側面から立証するために子供の体格の成長を測定したデータを用いて個体を分類する数値実験を行なった. 通常, この種のデータは個体間の類似性のみを考慮して分類されているが, これを 3-way データとしてクラスタリングすることにより個体間の体格の類似性に基づく分類のみならず成長の経時的変化における類似性を反映した分類となっていることを示した.

第4章では, ファジィ集合の概念を導入した新しいファジィクラスタリングモデルを提案した. 加法的ファジィクラスタリングモデルにおいて, クラスターとはある性質を共有する個体の集合として定義され, 個体が共通の性質を有することは個体間の類似性を増加させるものとし, いくつかの性質を共有する場合にはそれらが類似性に対して加法的に寄与するものと仮定されている. 類似性データの尺度に応じて三つのモデル, 単純加法的ファジィクラスタリングモデル, 重複を考慮した加法的ファジィクラスタリングモデル, 類似度の順序のみを考慮した加法的ファジィクラスタリングモデルを提案した. 単純加法的ファジィクラスタリングモデルは類似性が間隔尺度で測定されている場合のモデルである. またファジィクラスターは元来重複の概念を含んでいるが, 個体の各クラスターへの帰属度の総和が1であるという条件を緩和して, 重複の仕方をより柔軟に考慮した, つまりどのクラスターにも属す個体やどのクラスターにも属さない個体をも許す(重複を考慮した)加法的ファジィクラスタリングモデルを構築し, 類似性の構造がより明確に表現可能であることを示した. さらに類似度が間隔尺度レベルで与えられるならば通常最小二乗法的にモデルを当てはめることができるが, 順序尺度で与えられる場合には単純な最小二乗法ではモデルを同定できない. そこで類似度の順序情報のみが意味をもつ場合単調回帰原理に基づく類似度の順序のみを考慮した加法的ファジィクラスタリングモデルを提案しその有効性を検証した. 上記のモデルに加えてクラスター間の類似性という概念をファジィクラスタリングモデルに導入することによって類似性の構造をより詳細に表現し, 非対称な類似性に対する加法的ファジィクラスタリングモデルを構築し, その有効性を示した.

また観測データにある構造を仮定してモデルを当てはめる場合, 観測データが仮定された構造を真に持っているときそれが手法によって再現できるかというモデルの再現性あるいは観測データの真の構造からのずれに対して手法がどの程度ロバストであるかということが重要である. これに対して本論文で提案したファジィクラスタリングモデルに関する再現性やロバスト性がシミュレーション実験によって確認された. さらにクラスタリングモデルにファジィ集合の概念を導入することによって離散的な最適化問題を連続的な最適化問題として扱うことができモデルの推定が容易となることを示した. 本手法の応用例として親族を表す言葉の主観的判断に基づく類似関係を分析することによりその類似構造を明らかにした. さらに非対称な類似性に対するモデルの有効性および妥当性を検討するために子供のソシオメトリックデータを用いファジィクラスターの観点から好嫌関係の構造を明らかにした.

第5章では, 結論としてクラスタリング手法におけるファジィ理論の有効性について論じ, 本論文において開発された多基準ファジィクラスタリングおよび加法的ファジィクラスタリングモデルを通じて得られたファジィクラスタリングに関する新知見をまとめ, さらにそれらの理論的発展性について述べ今後の研究の指針を示した.

学位論文審査の要旨

主 査 教 授 佐 藤 義 治

副 査 教 授 新 保 勝

副 査 教 授 伊 達 惇

学 位 論 文 題 名

ファジィクラスタリング手法の開発と

その応用に関する研究

データのもつ潜在構造を陽に表現し、内在する特性を明らかにすることはデータ解析における共通の目的である。クラスター分析は類似性に基づきデータを分類しその潜在的特性を明らかにすることを目的としたデータ解析の基本的手法である。一方近年、工学の分野においては人間の思考あるいは言語における曖昧さを積極的に認めそれを数理的解析に反映させる手段としてファジィ理論が広く用いられている。クラスター分析においても従来の排反的クラスター構成手法に比して、分類基準に含まれる曖昧さを考慮し、境界の曖昧なクラスターを構成するファジィクラスタリング手法が現実に即した解析手法として注目されている。

本論文は、多基準ファジィクラスタリング手法を確立し、従来未開拓の分野である 3-way データに対するファジィクラスタリング手法およびファジィクラスタリングモデルについて新たな手法を提案し、それらの有効性、妥当性を論じたものである。

第 1 章では、クラスター分析の目的について述べ、対象となるデータの性質とそこで用いられる類似度との関連性について要約している。さらにデータ解析におけるクラスター分析の位置付けを行ない従来のクラスタリング手法との関連において本研究の意義と目的を明らかにしている。

第 2 章では、ファジィ理論を概説し、この理論のクラスタリング手法への適用を意味論的あるいは数理的に解説することによりファジィクラスタリング手法の諸定義を行なっている。また従来のファジィクラスタリング手法に関する研究について論じ本研究の立場とその必要性を明確にしている。

第 3 章では、3-way データに対するクラスタリング手法は一般的に多基準ファジィクラスタリング手法と捉えることができ、さらに 3-way データにおいて各時点に共通した個体のクラスターを得る問題は多基準最適化問題として定式化できることを示している。一般に多基準最適化問題における解はパレート最適解として得られることが知られているが多基準ファジィクラスタリング手法に関してパレート最適解を与える新たな定理を証明することにより個体間の類似性のみならず観測値の経時的变化の類似性をも加味した分類手法を提案している。この手法ではファジィ集合

のメンバーシップ関数の連続性を用いることにより、従来困難とされていた組合せ最適化問題を回避できることを明らかにし、さらに実測データを用いた数値実験結果により多基準ファジィクラスタリング手法の有効性を応用的側面から立証している。

第4章においてはいくつかの加法的ファジィクラスタリングモデルを提案しクラスタリングモデルの立場から分類手法を論じている。とくに類似性データの尺度を考慮することにより三つのモデル、単純加法的ファジィクラスタリングモデル、重複を考慮した加法的ファジィクラスタリングモデル、類似度の順序のみを考慮した加法的ファジィクラスタリングモデルを提案し、これらのモデルの特性について比較、検討している。この結果、各クラスターへの帰属の総和が1であるという条件を緩和し帰属の重複を考慮した加法的ファジィクラスタリングモデルにおいては、データの類似構造がより少数のクラスターで明確に表現可能であることを明らかにしている。さらにクラスター間の類似性という概念をファジィクラスタリングモデルに導入することによってこのモデルの拡張を計り、非対称な類似性にも適用可能な加法的ファジィクラスタリングモデルを構築し、その有効性を示している。またモデルが有すべき基本的性質である再現性やロバスト性については、シミュレーション実験によりそれらが確認されている。さらに実測データによる数値実験によりデータのもつ類似構造が明らかとなることを示し、このモデルの実用性を保証している。

第5章では、本研究において新たに提案されたファジィクラスタリングの諸手法を通じて得られた新見解はファジィクラスタリング手法の有効性を示唆するものであると結論し、その発展性を述べることにより今後の研究課題を明示している。

これを要するに、著者は、多基準ファジィクラスタリングおよび加法的ファジィクラスタリングモデルを開発しその性質を解明することによりクラスタリング手法におけるファジィ理論の有効性に関する有益な新知見を得ており情報工学の進歩に貢献するところ大なるものがある。

よって著者は、北海道大学博士（工学）の学位を授与される資格あるものと認める。