

## 第8章

### 結言

---

本論文は大きく第I部と第II部とから成り、ファジィクラスタリングに関する研究をまとめたものである。第I部では正則化の概念をファジィ化に導入したファジィクラスタリングに関する研究成果をまとめた。第II部では、ファジィクラスタリング手法を用いた、データの部分線形構造発見のための研究を中心にした研究成果をまとめた。それぞれの結果と今後の課題については以下の通りである。

#### 8.1 正則化によるファジィクラスタリング

---

第I部では、fuzzy  $c$ -means 法を crisp  $k$ -means 法の正則化であるという視点に立って fuzzy  $c$ -means 法を考察し、新たに fuzzy  $c$ -means 法とは異なる 2 種類の正則化によるクラスタリング手法を提案した。さらにそれらの手法についてクラスタリング結果を示すとともに、分類関数を導入してそれぞれのクラスタリング手法の性質を比較検討した。

fuzzy  $c$ -means 法は crisp  $k$ -means 法をファジィ化することにより、クラスタリング結果にあいまいさの表現を導入した。しかし、クリスプなクラスタリング結果をファジィ化するとき、あいまいさをどのように与えるかにはさまざまな扱い方が考えられる。あいまいさを取り除いたときには crisp  $k$ -means 法のクリスプ解の近似となるが、クラスタリング結果

にあいまいさを積極的に活用するファジィクラスタリングでは、あいまいさそのものに意味がある。あいまいさの取り扱い方としてこれまで fuzzy c-means 法という 1 種類しか存在しなかったが、本研究ではこれにはもっと多くの方法が存在することを示し、fuzzy c-means 法以外のファジィクラスタリング手法として新たにエントロピー正則化法と 2 次正則化法という 2 つの正則化による手法を提案した。

ファジィクラスタリング手法が複数存在するとなると、どの手法を用いるのがよいのかということが新たな問題となる。本研究ではこれに対して、各クラスタリング手法の異なる性質について新たに分類関数を導入することによって各種法の性質を調べた比較研究を行った。これにより各クラスタリング手法にはそれぞれの分類関数には異なる特徴があることが示された。エントロピー正則化法は、fuzzy c-means 法と遠方の点でのあいまいさの扱いが大きく異なる。fuzzy c-means 法では、無限遠点ではすべてのクラスに等しく分類されるが、エントロピー法では遠方の点の方がはっきりと分類される性質があることが示された。しかし、それぞれのクラスタリング結果から得られる領域の分割はともに等しく、クラスタの中心を含む Voronoi 図の領域とも等しいことなどが命題により示された。2 次正則化法では、遠方での性質はエントロピー法に近いが、分類関数が区分線形となり、よりクリスピーな手法に近い手法であるといえる。

この研究に関する今後の研究課題としては、以下のようなものが考えられる。

1. 正則化によって crisp  $k$ -means 法をファジィ化し、fuzzy c-means 法とは異なる 2 種類のファジィクラスタリング手法を示したが、新たに別のファジィ化によるクラスタリング手法を考える。
2. 本研究でクラスタリング手法の比較のために導入した分類関数を用い、さらに各クラスタリング手法の性質の研究を進める。
3. fuzzy c-means 法の応用として研究されている手法への正則化によるクラスタリング手法の導入を検討する。

この中で 1 番目に関しては、現在のところ考えることは難しいと思われる。もちろん、これ以外にもクラスタリング結果にあいまいさを考えることは可能であると思われるし、積極的に考えるべきである。しかし実際に計算を行うことができ、しかもそのあいまいさに意味がなければクラスタリング手法としては十分ではない。これらの条件を満たす新たな手法を考

えることは難しいのではないかとと思われる。2番目については、さらなる研究が必要であると考えている。複数のファジィクラスタリング手法を使い分けるためには、それぞれの手法の特徴を詳しく調べる必要がある。本研究でも十分に各手法の性質について調べているが、さらに詳しく研究を行うべきである。3番目は、第II部の最後のセクションで線形係数の決定に正則化を用いる手法の提案を行っているが、これ以外にも多くの応用が考えられる。ファジィクラスタリングは、回帰モデルや非線形モデル等にも応用されており、そういった問題について正則化を用いることを検討することが必要である。

## 8.2 線形ファジィクラスタリング

第II部では、fuzzy c-means法の応用として線形ファジィクラスタリング手法に関連した研究をまとめた。ここでの目的は多次元空間上に分布するデータをユーザに客観的に提示することである。ここでは、多次元データの異なる次元構造による線形ファジィクラスタリング手法、及びそれを用いたモデリングに関する提案を行った。また、第I部で述べた正則化の概念を線形ファジィクラスタリングの次元係数に用いた手法を提案した。

本研究ではデータ解析の見地からクラスタリング手法をとらえ、多次元上のデータに対して異なる次元の線形構造をとらえた線形ファジィクラスタリング手法の提案を行なった。さらに観測データには誤差やはずれ値が含まれていることが多い。この問題についてもDaveらによって提案されているノイズクラスタの適用を試み、簡単な数値例を用いてクラスタリング結果が誤差やはずれ値の影響を低く抑えることができたことを示した。

さらに、ファジィクラスタリングによるモデリングを行うために、楕円型メンバシップ関数で表される新しいタイプのファジィモデルを提案した。大規模システムのファジィモデルによるモデリングに関する問題の1つは、モデルを構築するためにはその対象をよく知る必要があるということである。ここではファジィモデリングにおける課題として、システムの部分的な構造をとらえたサブモデル群をいかに発見するかを考え、ファジィクラスタリング手法を用いた新たなファジィモデリングおよびシミュレーション手法を提案した。従来のファジィモデリングでは、事前に変数間の構造モデルを仮定するが、複雑で大規模なシステ

ムに対しては、構造を探求するためにデータからモデルを構築してみるという側面がある。また多くの場合、数値データは一様に存在せず、変数の選択や、サブモデルの担当領域の決定が困難である。こういった状況でモデリングを実行するためには、対象をよく理解する必要がある。この章で提案する楕円型ファジィモデルは、変数間の関係をデータに忠実に表現しようとするもので、最終モデルというよりは、対象をよりよく理解するための道具である。実際のモデリングの例として、日本の河川に関する適用例を紹介した。

最後に、線形ファジィクラスタリング法の次元係数の決定に、正則化の概念を導入した新たな定式化による手法を提案した。1990年に Dave によって提案された適応型 (adaptive) 手法では、次元係数の決定を各クラスタのファジィ散布行列の固有値を用いて決定する。しかしこれはファジィクラスタリングの2段階アルゴリズムの考え方に合致せず、目的関数の単調現象性が失われていることを指摘し、その具体的な例を示した。そして次元係数の決定に正則化の概念を導入することにより、その問題を解消した。また本手法によるいくつかの数値例によるクラスタリング結果を示した。

この研究に関する今後の研究課題としては、以下のようなものが考えられる。

1. クラスタリングのパラメータ、メンバシップの初期値によりクラスタリング結果が左右されることがあり、多次元データに対してパラメータをいかに設定するか。
2. 次元係数の決定に正則化の概念を導入したが、このときの正則化にエントロピー正則化など他の正則化による手法と比較検討する。
3. 楕円型ファジィモデルは、ユーザーにデータを客観的に提示するモデルであるが、これから if-then 型の高木・菅野型ファジィモデルの構築をサポートするシステムをめざす。

まず1番目の問題は本手法に限らないが、クラスタリングのパラメータ設定の問題は困難な問題である。人が同じデータを見ても人によってその見方が異なるように、パラメータによって調整ができることは必要である。このパラメータをもう少しわかりやすく設定する方法を考える必要がある。次に2番目については、早急に研究を進める予定である。次元係数だけでなく、メンバシップにも色々な正則化をもちいた組み合わせにより、その比較検討を行いたい。最後に3番目は、まだすぐには実現できないが、トータルにデータをユーザーに客観的に提示を行い、そこで詳細な高木・菅野型ファジィモデルの構築をサポートしてゆけるシステムをめざしてゆきたい。