

クラスター分析を用いたラフ集合解析

1 はじめに

ラフ集合 [1] を用いて選好の決定ルールを併合することで、より確実に消費者から選好されるように選好適合度を上げようとする試みが行われている。しかしこの手法では効率的に精度の高い特徴を抽出することは難しいという問題や、項目数が増えると解析が急に難しくなるという欠点がある。本論文では、クラスター分析の手法を用いてある程度消費者を特定の物事に関する考え方や価値観の似た集団に分類し、その集団からそれぞれ決定ルールを抽出する方法を提案する。この方法によって従来より詳しく、より精度の高い特徴の抽出を目的としている。

2 ラフ集合による決定ルール

ルールを導くために目的関数となる決定属性を設ける。その属性値を明示したものを決定表と呼び、この表が示すものを決定ルールと呼ぶ。しかし決定ルールの条件部では情報が多すぎるため、必要な最小の決定ルールの条件部で説明できることが求められる。その絞り込まれた決定ルールを極小決定ルールと呼ぶ。本論文では選好に関する極小決定ルールを選好ルールと呼ぶ。

3 多人数ルール条件部併合システム

森ら [2] は多人数に適用できる手法を以下に与えた。被験者を (S_1, S_2, \dots, S_k) としたとき

手順 1 各被験者の選好ルール条件部を、決定クラスの中でその条件部の占める割合 $C.I.$ の高いものから順に 100 個ずつ抽出。

手順 2 S_1, S_2 の選好ルール条件部を総当りで併合し、選好併合ルール条件部を得る。

手順 3 得られた選好併合ルール条件部の中から、条件部の大きさ (ルール条件部を構成するカテゴリー数) の小さいものから順に 100 個抽出。

手順 4 抽出した 100 個の選好併合ルール条件部と S_3 の選好ルール条件部を総当りで併合する。

手順 5 手順 3~4 を被験者 S_k まで繰り返す。

手順 6 最終的に得られた選好併合ルール条件部の $S.C.I.$ を求める。

以上のアルゴリズムからなるシステムを多人数ルール条件部併合システムと呼ぶ。

4 提案手法

提案する手法を以下に示す。

手順 1 検証する事柄について、被験者の考え方や価値観が分析できるような質問事項を作成。ただし用いる質問は YES または NO で答えられる質問にしておく。

手順 2 手順 1 で作成した質問事項を用いて被験者にアンケートを行う。

手順 3 手順 2 で得られるデータを階層型クラスター分析の n 次元の点とする。被験者が YES と答えた質問には 1 を、NO と答えた質問には 0 を、対応する変数に入力する。ここで階層型クラスター分析で用いられる点の次元数は質問事項数とする。

手順 4 手順 3 で入力した点を用いて階層型クラスター分析を行う。これによって、被験者を、検証する事柄に対する考え方や価値観の似たクラスターに分類する。

手順 5 再び被験者に対して実際に検証する事柄のサンプルをいくつか見せ、それぞれに対する選好を尋ねる。

手順 6 得られたデータより被験者の選好、非選好ルール条件部を取り出す。

手順 7 手順 4 で得られたクラスター毎に、被験者の選好併合ルール条件部を計算する。

5 実験

階層型クラスター分析におけるクラスター間の類似度 (距離) 5 種類を用いて提案手法の実験を行う。実験によって得られた結果の例を以下に示す。

表 1: ウォード法: グループ 5

多人数併合ルール条件部	$S.C.I.$
A2B3C8E3H1	0.750000
A2B3C8E3L1	0.750000
A2C8E3G2H1	0.750000
B5C8D4	0.500000
A3B5G1N2	0.500000

6 結論

本論文では、階層型クラスター分析を用いて被験者を価値観の似たいくつかのクラスターに分類し、そのクラスター毎に多人数ルール条件部併合システムを用いて被験者の特徴抽出を行う方法を提案した。

参考文献

- [1] Z.Pawlak, "Rough sets," Internat. J. Inform. Comput.Sci., Vol. 11, No. 5, pp. 341-356, 1982.
- [2] 森典彦, 田中英夫, 井上勝雄, "ラフ集合と感性データからの知識獲得と推論," 海文堂, 2004.