

投票行動の計量的地域分析の試行と評価¹

舟木基通

辻 光宏

関西大学総合情報学研究科

関西大学総合情報学部

投票行動の地域特性の構造を同定する際に、日本においては、主成分解析やクラスタリングによる選挙区の分類に基づいて、解析されてきた。一方、欧米では、ここ 10 年ほどの間にコレスポネンス解析の社会科学分野への適用が進められ、投票行動への適用により地域特性と政党支持との結びつきを視覚的にも表現できるようになってきた。この報告では、日本での全国都道府県あるいは京都府市町村での投票行動データを対象に、コレスポネンス解析と Ward 法凝集型階層的クラスタリングとの連動により、投票行動の地域構造特性を試行的に明らかにする。

Trial and Its Evaluation for the regional analysis to investigate voting results in Japan
Motomichi Funaki, Mitsuhiro Tsuji
Kansai University

The development of the regional analysis has been performed in recent years to investigate voting results by use of Correspondence Analysis in Europe.

In this paper we are reporting that we applied the Correspondence Analysis for regional analysis of the voting results in Japan and that we evaluated the efficiency of the integrated statistical system with Correspondence Analysis and Clustering.

1. 序に代えて

日本において投票行動の地域構造同定問題、すなわち各選挙区での各政党の支持率に対して、選挙区の地域特性と政党支持率との比較検討は、主に主成分解析を中心に行われてきた。そのアプローチ方法の代表的なものは、各選挙区の地域特性を人口や産業別就業者数をもとにいくつかのグループに分類し、それらのグループごとに各政党支持率との関係を見るものであった。

一方、フランスを中心にして開発が進められた統計解析手法であるコレスポネンス解析は、最近急速に利用が広がり、欧州では投票行動への適用も進んでいる[3]。

本報告では、日本での 1989 年、1992 年、1995 年の参議院比例代表制選挙での各政党得票数を元にして、コレスポネンス解析と Ward 法凝集型階層的クラスタリングとを統合した統計解析[1]により、解析を行なった。これは、政党と選挙区とを同時に低次元で表示するビジュアルイゼーション（視覚化）に秀でたものである。

¹ 本研究は、2000 年度関西大学在外研究 (the Department of Mathematics, University of Essex) に基づき、2002 年度関西大学重点領域研究による成果である

2. データ

対象のデータは、縦列に各都道府県、横方向に政党別得票数を並べた2元表形式のデータである。1989年、1992年、1995年の参議院選挙での各都道府県での比例代表得票数である。政党は、自民党、社会党、公明党、共産党、民社党とした。

日本での従来のアプローチは、得票数を得票率にデータ加工した上で主成分解析などを行ってきたが、ここでは得票数そのものを扱う。

3. コレスポンデンス解析

コレスポンデンス解析は、フランスを中心に発展し、近年に特に注目を集めてきた2元表データを総合的に評価できる手法である。そこでは、縦列と横列のプロフィール（比率パターン）に注目し、それらの情報量をできるかぎり損失することなく、縦列データと横列データを同時に低次元で表現することをめざしている。

主成分解析は変動を定量的に表現する手法であるために、もとのデータの数値の大きさの影響が強いが、コレスポンデンス解析では比率パターンの類似したものを近い場所に配置する傾向が強く、もとのデータの数値の大きさを均す特性がある。このために、各選挙区ごとの得票数の定量的な格差が大きい得票数データには適した手法である。

4. Ward法凝集型階層的クラスタリング

コレスポンデンス解析は、2元表データを総合的に評価することに秀でた統計解析手法であるが、低次元表現する際の情報圧縮での歪みや外れ値の感度の歪みに注意すべきである。すなわち、もとのデータの潜在構造の中で、コレスポンデンス解析結果でじゅうぶんには表現できていない部分を抽出し、正当に評価する必要がある。それを補完するのがWard法凝集型クラスタリングである[4]。

凝集型階層的クラスタリングは、データの潜在構造にあるグループを連続空間ではない尺度で検出することができる。特にWard法はコレスポンデンス解析との整合性がとれたアプローチである。

5. 統合するための統計グラフィックス

コレスポンデンス解析によって表示できるプロット図には、選挙区（都道府県）と支持政党とが同時に表現される。地域特性の独自性が強い都道府県は別個に表示されるので支持政党との対比も可能であるが、地域特性が共通している都道府県は多数あるために重なって表示されるために詳細を判断することが困難である。また、凝集型階層的クラスタリングで得られるグループは、都道府県からなる樹状図で表現される。

ここでは、コレスポンデンス解析の結果のプロット図にクラスタリング経過を追加する動的プロットを行い、特に近接した地域特性をもつ都道府県を早期に発見して支持政党との対比を可能にした。

6. 解析結果

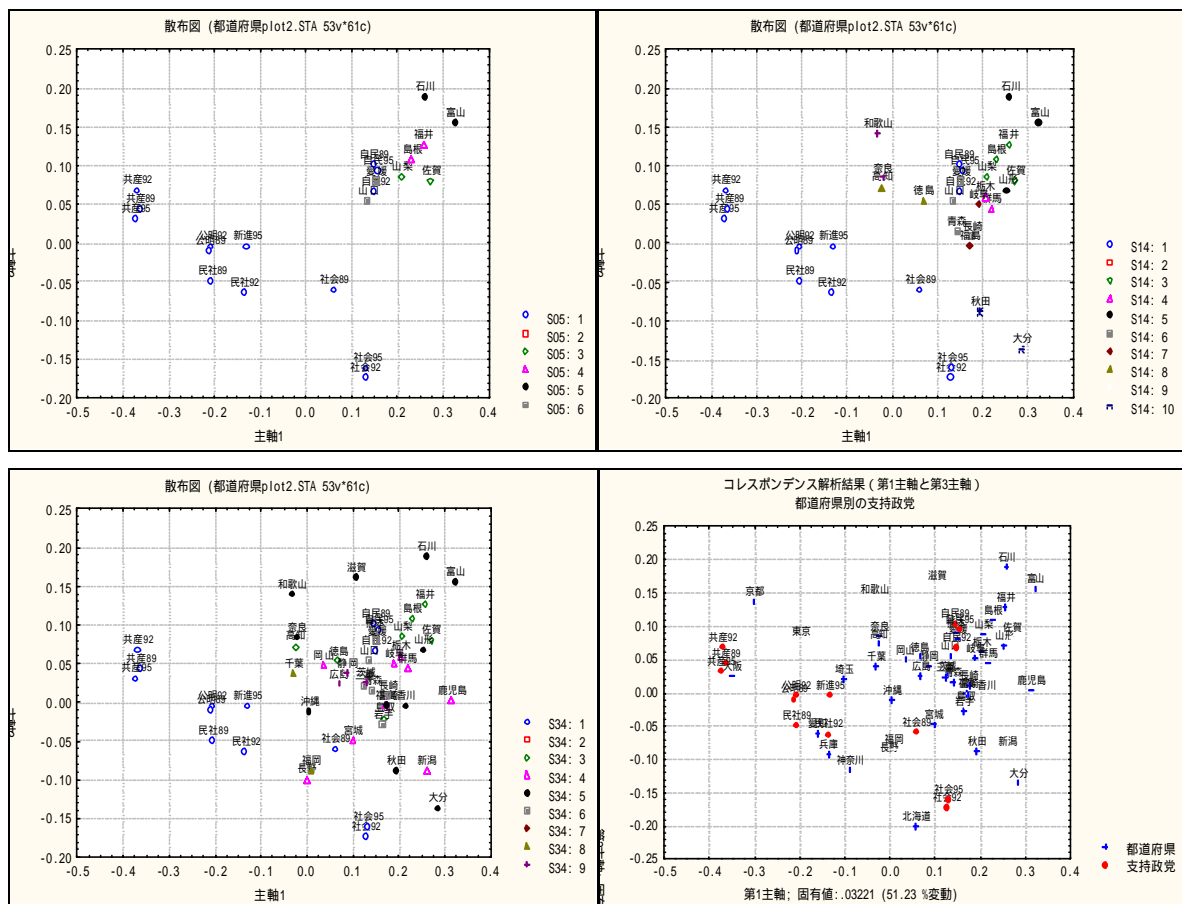
(1) コレスポネンス解析結果

コレスポネンス解析の結果，寄与率は第1主軸が53.93%，第2主軸が27.88%，第3主軸が10.30%であり，3つの主軸で全体の92.12%までが表されている。

各主軸について，地域特性を表わす指標データとの相関を取り，意味付けを試みた。第1主軸と正の相関が大きいものは，持ち家率(0.72)，世帯人口(0.60)であり，負の相関が大きいものは，人口集中地区人口(-0.72)，人口(-0.71)，小売業商店数合計(-0.71)，第2次産業就業者数(-0.70)，第3次産業就業者数(-0.69)であり，都会度に関係している。第2主軸に正の相関がやや大きいものは，失業率(0.44)，世帯密度(0.29)であり，負の相関としてやや大きいものは，就業率(-0.38)，工業製造品年間出荷額(-0.34)であり，製造業集中度に関係している。第3主軸と正の相関がやや大きいものは，持ち家率(0.32)，世帯人口(0.28)であり，負の相関がやや大きいものは，第1次産業就業者数(-0.72)，農林・漁業事業所数(-0.71)，総面積(-0.43)であり，農村に関係している。

(2) クラスタリングとの統合結果

第2主軸は，民社党と愛知県との密接な関連を表現しているものの，他の傾向を説明することが難しいので，第1主軸と第3主軸に注目して解析を進めた。

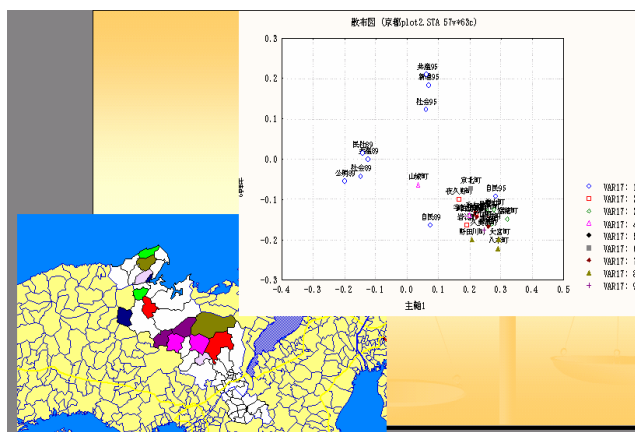


最初に，自民党の周辺にさまざまな都道府県が小さなグループを生成する。第2に，社会党周辺にも小さなグループが発生する。次に，自民党と社会党との共通部分として都市圏以外のグループが発生後，特定の都市圏が野党の周辺に分散している。

7. 結びに代えて

投票行動の計量的地域特性を調べる上で、人口集中度や産業別就業者などに基づく地域特性を加味して、地域特性と支持政党との関連性を探索することが可能になった。

本報告では、都道府県レベルの選挙区を対象としたが、政治学で今後注目されるのは各都道府県内の市区町村レベルでの同様の分析である。京都府の市区町村レベルでの分析も試みたので、研究会当日に報告する予定である。



最近の投票行動の分析では、無党派層の検出が非常に大きな課題である。本研究の今後の課題ととらえ、最近の参議院選挙にも取り組んでいきたいと考えている。

8. 参考文献

- [1] 辻 光宏(2002). 対応解析とクラスタリングとの統合解析システムによる政治分析の試み. 人工知能学会研究会資料 SIG-KBS-A201, 141-144.
- [2] 舟木基通 (2002). 社会科学への対応解析の応用. 関西大学総合情報学研究科修士論文
- [3] F. Fehlen (1998), The cloud of candidates: exploring the political field, in *J. Visualization of Categorical Data*, ed. by Blasius and M. Greenacre, 159-170, Academic Press.
- [4] L. Lebart (1994), Complementary use of correspondence analysis and cluster analysis. *Correspondence Analysis in the Social Sciences*, ed. by M. Greenacre and J. Blasius, 162-178, Academic Press.