



Title	Contributions to the Exploratory Analysis of Compositional Data
Author(s)	Maria, Cristina Martin
Citation	大阪大学, 1997, 博士論文
Version Type	
URL	https://hdl.handle.net/11094/40676
rights	
Note	著者からインターネット公開の許諾が得られていないため、論文の要旨のみを公開しています。全文のご利用をご希望の場合は、 〈a href="https://www.library.osaka-u.ac.jp/thesis/#closed"〉 大阪大学の博士論文について 〈/a〉 をご参照ください。

The University of Osaka Institutional Knowledge Archive : OUKA

<https://ir.library.osaka-u.ac.jp/>

The University of Osaka

氏 名	マリア クリスティナ マルティン Maria Cristina Martin
博士の専攻分野の名称	博 士 (理 学)
学 位 記 番 号	第 1 3 3 9 2 号
学 位 授 与 年 月 日	平 成 9 年 9 月 12 日
学 位 授 与 の 要 件	学位規則第4条第1項該当 基礎工学研究科 数理系専攻
学 位 論 文 名	Contributions to the Exploratory Analysis of Compositional Data (構成比データの探索的データ解析)
論 文 審 査 委 員	(主査) 教 授 白旗 慎吾 (副査) 教 授 稲垣 宣生 教 授 後藤 昌司 助教授 磯貝 恭史

論 文 内 容 の 要 旨

健康食品中の栄養素の含有率とか岩石中の鉱物の含有率などを示すのに用いられる構成比データは各成分の和が1になるという制約を持つため、通常の無制約の状況で適用される多変量解析手法をただちに利用するわけにはいかない。構成比データの位置の推定をどう行い、分散構造をどう捉えるかが問題になる。

我々は最初に第2章でこの分野において本質的な貢献を行った Aitchison 博士の接近法の簡単な紹介を行い、構成比データ固有の取り扱いの困難さの所在と我々の立場を明確にする。

ついで第3章では、構成比データの位置の推定の問題を取り扱い、多変量モードの推定法を提案する。さらに多変量モードの性能及びそのほかの各種の推定量(幾何平均などの平均値、中央値、M推定量)の性能の比較検討を行う。

第4章以降では、構成比データの分散構造を捉えるのに、構成比データ間の距離を用いた接近法を提案する。第4章では、よく知られた3つのデータ集合を用いて、7つの距離について、6種類のクラスター分析法を適用して、解析結果の比較からどの距離とどのクラスター法が構成比データに適しているかを明らかにする。

第5章では構成比データのバイプロット表現の問題を、3種類の距離を用いて取り扱う。ユークリッド距離を用いた場合、バイプロット表現はユークリッド空間内に得られ、その座標は直線座標であるが、考察している構成比データ間の距離はユークリッド距離ではない。ここでは構成比データに対して、非ユークリッド距離を用いてユークリッド空間に座標表現を得る問題を取り扱い、構成比データに対する非線形バイプロット法を提案する。

最後の第6章で我々の仕事の総括を行い、これからの研究の発展性について述べる。

論 文 審 査 の 結 果 の 要 旨

鉱石中の鉱物の含有率などを表現するのに用いられる構成比は各成分の和が1という制約のある多変量データであるため、その解析には通常の無制約な場合に適用される多変量解析手法を直接には用いることができない。構成比デ

ータの位置母数の推定をどう行い、かつデータ間の距離構造をどう捉えるかが問題となる。構成比データ解析では Aitchison の方法論が大きな貢献を為しているが、彼の方法論は相関構造に特化しており、位置母数や距離構造の研究は不十分である。本論文では、第 1 章では全体の概観、第 2 章では Aitchison の方法論の紹介、構成比データ解析固有の問題の困難さおよび著者の立場について述べている。研究の成果は第 3 章以降で述べられている。

第 3 章では構成比データの位置母数の推定の問題を扱っている。多変量モードの推定法を提案し、さらにその性能および幾何平均等の各種の平均、中央値、M 推定量との性能比較を行っている。データには色々の型があるが、その多くで提案された推定量が最良であることをシミュレーションにより示している。第 4 章では、構成比データのクラスター構造を捉えるために距離を考察している。良く知られた重要な 3 つのデータに対し、7 つの距離、6 種類のクラスター分析法を適用し、解析結果の比較からどの距離とクラスター分析法の組み合わせが構成比データ解析に有用かを明らかにしている。第 5 章では構成比データのクラスター分析法におけるグラフ表現法として知られているバイプロット表現で、特に非ユークリッド距離を用いる非線形バイプロット法を提案している。最後の第 6 章では研究の総括及びこれからの発展性について述べている。

以上のように本論文は構成比データ解析における推定論およびクラスター分析に寄与するものであり、博士(理学)の学位論文として価値あるものと認める。