The University of Osaka
Institutional Knowledge Archive

Title	クラスタリングを伴う因子分析法の研究開発
Author(s)	宇野, 光平
Citation	大阪大学, 2019, 博士論文
Version Type	VoR
URL	https://doi.org/10.18910/72457
rights	
Note	

The University of Osaka Institutional Knowledge Archive : OUKA

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

The University of Osaka

様式３

論　文　内　容　の　要　旨

氏　名　（　宇野　光平　）	
論文題名	クラスタリングを伴う因子分析法の研究開発

論文内容の要旨

因子分析法は心理学をはじめとする多くの分野で用いられる統計的手法であり，応用範囲も幅広い。ところが、因子分析には不定性と呼ばれる重大な問題がある。

不定性は二種類存在し、一つは回転の不定性と呼ばれるものである。回転の不定性の影響によって、因子負荷量は何らかの回転を行わなければ幾らでも解を変形させることができてしまう。一方で、回転の不定性を積極的に活用することで因子負荷量を望んだ形に変化させる手法は心理統計学において主要な問題でもある。今日までに多くの回転基準が提案されており、それらを利用することで分析者は変数と因子の関係性を解釈しやすくなっている。

もう一つの不定性は因子得点の不定性と呼ばれる。回転の不定性は変数と因子の関係性を解釈する際に問題が生じるのに対して、因子得点の不定性は個人と因子との関係に関与する。回転の不定性が多くの研究者によって研究されてきたのに比べ、因子得点の不定性は早くから指摘されていたにもかかわらず取り上げられてこなかった。しかしながら、因子得点によって個人を解釈しようとするためには因子得点の不定性に対処しなくてはならない。

そこで、本論文ではクラスタリングによる因子得点の不定性の解消、さらにいえば不定性を積極的に利用した解釈しやすい因子得点を得るための手法を提案する。手法およびアルゴリズムを提案し、シミュレーションと実データを通して提案手法の有効性を確認した。

さらに、クラスタリングを用いることで因子分析の別の問題を解決することも確認した。因子分析には大きくわけて三種類のモデルが存在し、因子得点を確率変数として扱うかパラメータとして扱うかによって分類できる。その中の母数モデル因子分析は、因子得点の不定性が存在しない稀有な因子分析モデルであるが、尤度が発散してしまうために最尤推定が不可能という問題があった。そこで、パラメータである因子得点にクラスター制約を課すことで、母数モデル因子分析においても最尤推定が可能なようにリパラメトライズした。さらに次元縮約とクラスタリングを同人に達成する提案手法の特徴から、既存の類似手法に比べ分類精度が優位であることが実データ解析を通して判明した。

また、近年統計学において研究が活発なスパース推定が因子分析にもたらす恩恵などを考察した。具体的には解釈の向上、そして回転の不定性の解消によって、従来統一的に扱われることはなかった確認的因子分析と探索的因子分析の良いところをあわせ持つスパース因子分析の特徴などを考察した。

様式7

論文審査の結果の要旨及び担当者

氏　名　（宇野　光平）		
論文審査担当者	（職）	氏　名
	主 査　教授	足立 浩平
	副 査　教授	八十島 安伸
	副 査　准教授	青野 正二
	副 査　教授	狩野　裕（基礎工学研究科）

論文審査の結果の要旨

心理統計学に生まれ育ち、多変量解析の諸方法の中でも重要な位置を占める因子分析は、複数個体から得られた複数の観測変数のデータを分析対象とするが、複数観測変数の変動が少数の共通因子によって説明され、説明されずに残った各変数に独自の変動は、独自因子によって説明されると仮定される。こうした因子に関する各個体の得点(因子得点)の扱い方によって、因子分析のモデルは、変量モデル・母数モデル・行列モデルの3種に分類できる。変量モデルは、共通・独自因子の得点を潜在変数とみなすもので、因子得点をパラメータとして最適推定することはできない。母数モデルは、独自因子を変数とみなすが、共通因子の得点をパラメータとして扱い、その最適値の推定は可能である。しかし、最尤推定ができないという問題を持つ。行列モデルは、共通因子・独自因子の両者をパラメータとみなし、最適な因子得点を数式で表現できるが、解は一意ではない。本論文の主要な貢献は、上記のモデルの中でも、母数モデルと行列モデルに焦点をあて、因子得点がクラスター化される、つまり、各個体が少数の群に類別されるという仮定を組み込むことによって、解釈を促す因子得点を同定する方法を研究開発したことにある。これらを[1]、 [2]として、以下に、研究開発の成果を要約するとともに、[3]として、本論文で論及された因子分析における変数のクラスタリングについて記す。

[1] 母数モデルは、最尤推定ができないという問題を持つが、個体が少数のクラスターに類別され、同じクラスターの個体は同値の共通因子得点を持つという制約を導入すると、最尤推定が可能になることを見出し、クラスター制約つき母数モデルの最尤因子分析法を開発した。これは、最尤推定の実現と、因子得点クラスター化による解釈の促進という2つの効用を、母数モデルに獲得させる方法と言える。真のクラスターが既知である複数種の実データへの適用例では、クラスターの正分類率において、クラスタリングを伴う主成分分析と称せられる関連手法(Reduced *k*-means法・Factorial *k*-means法)を十分に凌ぎ、提案手法の有用性を示した。さらに、主成分分析が有しない独自因子をモデルに持つ因子分析では、個体のクラスター化に寄与しない変数の分散を、独自因子が吸収するという因子分析の効用を示し、主成分分析と因子分析の本質的相違の一旦を考究している点は理論的に意義深い。

[2] 行列モデルは、最適な因子得点の推定値を数式で表現できるが、その解は一意でない。一見短所に見えるこの性質をいわば逆利用して、複数組ある解の中から、共通因子得点が良くクラスター化される解を選出するアルゴリズムを開発した。その挙動の良好性をシミュレーションによって確認し、実データへの適用によって有用性を例証した。このアプローチは、因子負荷行列の解が一意でない事を逆利用した因子回転と同じ着想を、因子得点に関する研究開発に導入した初めての試みであり、今後、本研究を出発点として、種々の因子得点同定法が開発されていくと考えられる点で、本研究はパイオニア的な試みに位置づけられ、高く評価される。

[3] 以上が、個体のクラスタリングであったのに対して、変数のクラスタリングにも論及したもので、その実現に適するスパース因子分析の諸方法を的確に論考し、優れた展望研究と評価できる。

行動科学をはじめ、いかなる分野においても、多数の変数を少数の共通因子に縮約して、かつ、縮約できない独自因子成分を特定し、それらと同時に、個体をクラスタリングする方法論は、ニーズの高いものであり、そのニーズに合った多変量解析法を開発し、さらに開発手法の数理的性質を考究した本論文は、博士（人間科学）の学位授与に十分に値するものと判定された。