

R 言語に基づく行列因子分析法のソフトウェアの構築

世ノ一 はるか

因子分析は、因子(潜在変数)を説明変数、観測変数を従属変数とした因果モデルに基づく分析である。複数の観測変数の値の大小は、それらに共通する原因といえる共通因子(潜在変数)の値の大小によって説明され、共通因子によっては説明されずに残る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。こうしたモデルに基づき、共通因子と観測変数の関係を表す因子負荷量と、各変数に独自の変動の大きさを表す独自分散が、因子分析で推定されるパラメータとなる。

新たな提案手法である行列因子分析(MDFA)では、モデルパートと残差が独自に定まるので、各観測変数についてモデルの適合度を評価できる。そのため性格や能力を測るテストなど、個体の個人属性を同定したい場合に適しており、人間科学において今後役立つと考えられる。そうした MDFA を、フリー且つオープンソフトウェアの統計解析向けのプログラミング言語である R 言語を用いて実装し、人間科学分野での使いやすさを向上することが本研究の目的である。

実装したのち、実際のデータをプログラムに適用して解析を行った。一つ目のデータは学生 351 人に対して行われた 24 変数知能テストデータ(<https://astro.temple.edu/~alan/MMST/datasets.html>) である。(Holzinger & wineford, 1939)。二つ目のデータは 190 名を対象とした 25 項目の性格検査のデータである Big Five(ビッグファイブ; Goldberg, 1990, 1992)を用いた。因子数はスクリーテストを用いて決定した。また MDFAと比較するために、通常の因子分析に VARIMAX 回転を加えたものも示した。これらの結果を比較すると、いずれのデータに関しても、通常の因子分析と MDFA で因子負荷量や独自分散がほぼ一致していた。そのうえ、MDFA の最大の特徴であるモデルパートと残差の出力も行った。両データとも、残差はほとんどの値が極めて小さくなっていたが、Big Five については残差の大きい部分が見られた。これは、性格検査において本心ではない作為的な回答をした可能性があることを示している。

本研究により、R を用いて MDFA によって人間科学のデータを分析することが可能となった。本研究のプログラムを用いることで、通常の因子分析の結果とほとんど一致する因子負荷量や独自分散が出力できたうえ、MDFA の最大の特徴であるモデルパートと残差の出力にも成功した。今後は、アルゴリズムの加速化など、より利便性を高めるための改善が望まれる。(行動統計科学)