主成分分析と多次元展開法の比較に関する研究

土方　啓一郎

主成分分析（PCA）と多次元展開法（MDU)を比較すると，両者はともに，元データの行と列に並ぶ異なる二つの対象の空間布置を求めるための探索志向の分析法である（足立, 2006)．一方でPCAは内積モデル，MDU は距離モデルという相違点も存在するが，関数としての次数から内積は距離の特殊ケースであるといえる．本研究の目的はこのように多くの共通点とモデルの関連性をもちつつもこれまで比較されてこなかった PCA と MDU の比較を行うことである．しかし，MDUは退化解により合理的な解が得られにくく，かつその回避策の多くはモデルや目的関数が複雑であったり元データに干渉していたりするため，他の多変量解析法との比較が難しかった．そこで本研究ではPCAでしばしば用いられる制約条件をMDUのパラメータに課し，それにより退化解を回避しつつPCAと比較可能なモデルを想定した．

実際の分析では，元データとして(対象)×(変数)の 2 相 2 元親近性データ行列に対して，PCA と MDU に基づいた二つの分析法を PCA アプローチと MDU アプローチとして適用し，得られた解の比較を行った．このとき，行に並ぶ対象の布置を表す座標行列が列中心化され，かつその共分散行列が単位行列に一致するという制約条件を両アプローチに課した．特に後者の制約によりMDUアプローチでの退化解を防ぐことができると考えた．MDU アプローチの最適化アルゴリズム上では，制約を課したパラメータの解を，制約を満たすようPCA アプローチの解から流用した．

実データ解析の結果を比較すると，制約を課していない列に並ぶ変数の布置に関して，Adachi（2003）と同様に PCA アプローチの解は MDU アプローチに比べて原点付近に密集し，変数間の差が分かりにくいという結果が得られた．これは両アプローチのモデルに設定した切片の解法の違いによるものだと考えられる．また，適合度比較を行ったところ MDU アプローチの方がより良い適合度を得られる可能性があることが示唆された．このことは内積という関数の線形性と距離という関数の非線形性によるものと考えられる．さらに，PCAアプローチとMDUアプローチの内積モデルと距離モデルというモデルの比較から，MDU アプローチの方が得られた布置の読み取りが容易で，かつ直観的に行えることを述べた．本研究ではMDUアプローチの制約を課したパラメータの解を，PCAアプローチの解から流用した．そのため，MDUアプローチで制約を課したパラメータの解法を開発し，この制約が本当に退化解の回避に役立つのか検証することが今後の課題となる．（行動統計科学）