非対称データの多次元尺度法に関する研究

辻井　岳

社会現象や心理現象を扱う様々な学問分野において，二者関係を示すデータが数多く存在し，特に二者間の類似性，混同率，連関性，心理的距離などは距離データに置き換えることができる．多次元尺度法（Multi Dimensional Scaling：MDS）は，対象間の非類似性データを距離に変換して，潜在する少数の次元や背後にある認知構造を明らかにしたい時に利用する多変量データ解析手法の一つである．特に，非対称データを解析する手法を狭義に非対称 MDS と呼ぶ．また多次元尺度法を変形した解析手法として，行と列の対象が異なるデータ行列に適用される多次元展開法が存在する．非対称 MDS はモデルへの制約が厳しく，汎用性は低い代わりに，事前に解釈のしやすい前提を定めることができる．逆に，多次元展開法を用いて非対称データを解析した場合，データの非対称性による制約がなく，モデルとしての汎用性は高い分，行と列の対象が同じデータを解析する際には，有意義な解釈が得られにくい．

本研究では行と列に同一の対象が並ぶ非対称データ行列を解析することを目的に，こうした先行研究の問題点を改善するモデルとして，自己類似制約付き多次元展開法を提案した．

以下は制約なしの多次元展開法のモデルである．

$$min\sum_{i,j}^{n}\left(q_{ij}-\left\|\mathbf{x}_i - \mathbf{y}_j\right\|\right)^2$$

で表される．ここでq_{ij}は n 個の行の対象（i=1, …, n）と n 個の列の対象（j=1, …, n)間の非類似度，$\mathbf{x}_i$，$\mathbf{y}_j$ は対象 i と対象 j を空間布置する際のそれぞれの位置ベクトルである．多次元展開法に制約を与えた提案手法の目的関数は以下の通りである．

$$f(\mathbf{x}_i, \mathbf{y}_j \mid q_{ij}, \alpha)=\sum_{i,j}^{n}\left(q_{ij}-\left\|\mathbf{x}_i - \mathbf{y}_j\right\|\right)^2+\alpha\sum_{i}^{n}\left\|\mathbf{x}_i - \mathbf{y}_i\right\|^2$$

αは任意の正の実数であり，提案手法では多次元展開法のモデルに，同一対象間の距離の二乗和の値がより小さくなるよう制約を与えた．これによりモデルの汎用性を落とさずに，より直観的な解釈がしやすい空間布置を行うことを目指した．

提案手法の有用性を検証するため，世代間職業移動データ，ブランド交換データ，格付け推移率データを用いた実データ解析を行った．その結果，対象間の推移の方向を解釈することが可能になり，直観的に解釈しやすい空間布置が可能になることが示され，提案手法の有用性が示された．今後の課題として，制約となるαの値を決定する基準や，非類似性データから距離データへのデータ変換手法の提案，そして新たな制約の検討などが考えられる．（行動統計科学）