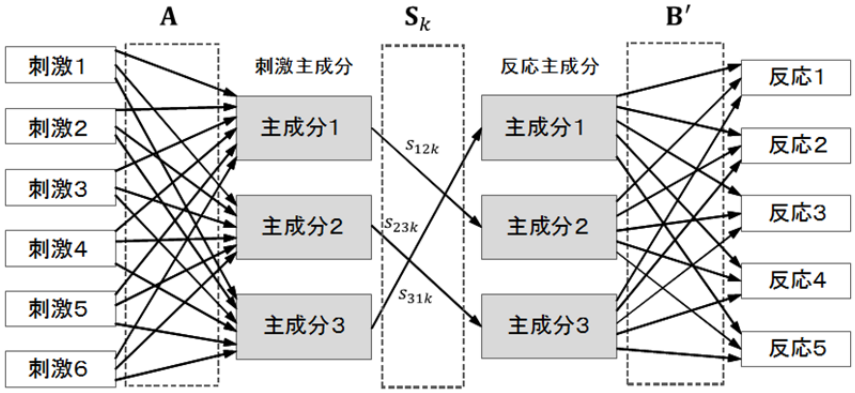
個体差表現を伴う三相主成分分析法の研究開発

山口　奈津実

「相」はデータの構造を表す言葉の 1 つであり，分析の対象となる組の数を表す．本論文で扱う 3 相データは，複数の評定者を対象に Semantic Differential（SD）法を行った場合に得られる（概念）×（形容詞尺度）×（評定者）のような 3 次元配列データであり，第 1 相（以下，「刺激」）と第 2 相（以下，「反応」）から成るデータ行列を，第 3 相(以下，「個体」)ごとに重ね合わせたものとして捉えることができる．このような 3 次元配列のデータを扱う分析法は，解として，個体間で共通する値と共に，共通解への重みを個体ごとに算出する．この重みが個体差として解釈できる．

個体差表現を伴う分析法の 1 つである三相主成分分析法は，3 相データの潜在的な成分を見つけ出す分析法の総称である．個体差表現の役割を持つ核行列にどのような制約を設けるかによって，各手法は特徴付けられる．中でも，Tucker2（Kroonenberg & de Leeuw, 1980）は，核行列に制約を持たない最も包括的なモデルであり，個体差を細かく表現することができる．しかし，計算する際の情報量が多く，個体が多くなるほど解釈も難しくなる．一方，PARAFAC（Parallel Factor Analysis; Harshman, 1970）は，強い制約を持つモデルであり，計算と解釈が比較的容易である．しかし，表現できる個体差に限界があるため，データから得られる情報を十分に要約できていない場合がある．本研究では，これら双方の利点を生かすための中間的モデルを提案する．刺激主成分と反応主成分を一対一で結び付け，PARAFAC の利点である計算と解釈の容易さを保ちつつ，その結び付き方を制限しないことで，個体差をより豊かに表現する（図参照）．なお，図中の$\mathbf{A}$は（刺激）×（刺激主成分），$\mathbf{B}$は（反応）×（反応主成分），$\mathbf{S}_k$は核行列を個体ごとにスライスした（刺激主成分）×（反応主成分）の行列を，それぞれ表す．



図：提案モデルのネットワーク表現例

解の推定には交互最小二乗法を使用する．数値実験の結果，提案手法は，局所解に陥ることは多いものの，十分な試行回数が成されれば，再現性の高い解を得られることが確認された．また，提案手法を用いて実データを解析した結果，Osgood & Luria（1954）の多重人格データでは，似た特徴を持つ人格を見つけ出すことができ，Adachi（2011）の色評定データでは，多個体の中から特殊な評定基準を持つ個体を見つけ出すことができた．提案手法の良さを大いに発揮し得るような構造を持つ実データは，決して多くはない．そのため，色評定データで示したような，提案手法の応用的な適用場面，及びそこでの解の特徴について探求していくことが今後の課題である．（行動データ科学）