様々な形式で回答された調査データの統計解析手法の紹介

総合情報部 社会情報学科 黒田正博・森裕一

Keywords: 最適尺度化、数量化、名義データ、順序データ、交互最小二乗法

1. 研究目的

社会調査や消費者行動等に関する調査において，様々な観点から質問項目は設定されている．そのため，回答して得られるデータの形式も多様である．統計学では，性別や職種等の項目に関する回答データを名義データ，満足度や成績評価等に関しては順序データ，そして年齢や試験の得点等を数値データと呼び，これらを区別する．データ解析において，統計解析手法は目的とデータの形式に合わせて選択される．このため，データ形式が混在している場合，従来の統計解析手法を直接適用することは困難である．

そこで，ある基準によってデータを変換することで，すべてのデータを同じ形式に統一し，統計解析をおこなうことを考える．そのための方法の1つに，名義データと順序データを最適尺度化し数値データに変換する方法がある．これを数量化といい，回帰分析による予測，主成分分析のデータの次元縮約と特徴抽出，あるいはクラスター分析によるグループ分けといった多様な統計解析手法によるデータ解析が可能になる．

また，名義データや順序データが文字列であった場合でも数値得点に変換されるため，これらを量的に評価することが可能になる．

2. 最適尺度化による名義データと順序データの数量化と統計解析手法との連携

名義データおよび順序データの尺度最適化による数量化の基準として，最小二乗基準が用いられる．名義データは何の制約も課されずに数量化をおこなうのに対して，順序データについては，回答項目に順序がついているため順序制約を考慮した数量化をおこなう．

上記の3つの形式から構成される調査の回答データ $\boldsymbol{X}=(\boldsymbol{X}_{名義}, \boldsymbol{X}_{順序}, \boldsymbol{X}_{数値})$ が得られているとする．このとき，$\boldsymbol{X}_{名義}$と$\boldsymbol{X}_{順序}$を数量化したものを$\boldsymbol{W}_{名義}=\mathbf{G}_1\boldsymbol{a}_1$，$\boldsymbol{W}_{順序}=\mathbf{G}_2\boldsymbol{a}_2$と書く．ここで，$\mathbf{G}_1$と$\mathbf{G}_2$は回答パターンを示す行列であり，$\boldsymbol{a}_1$と$\boldsymbol{a}_2$は$\boldsymbol{X}_{名義}$と$\boldsymbol{X}_{順序}$を$\boldsymbol{W}_{名義}$と$\boldsymbol{W}_{順序}$に数値変換するための値（パラメータ）である．数量化された回答データを$\boldsymbol{W}=(\boldsymbol{W}_{名義}, \boldsymbol{W}_{順序}, \boldsymbol{X}_{数値})$で表すことにすると，

$$f(\boldsymbol{X},\boldsymbol{W}) = \| \boldsymbol{X}-\boldsymbol{W} \|^2$$

を最小化する$(\boldsymbol{a}_1, \boldsymbol{a}_2)$を見つける問題に帰着される．ここで，$\boldsymbol{W}_{名義}$あるいは$\boldsymbol{W}_{順序}$に文字列が含まれる場合は，これらを数値にコード化したものを用いる．

実際の計算では，最適尺度化による数量化データ$\boldsymbol{W}$の計算と，それをベースにおこなう統計解析手法のパラメータ計算を繰り返しおこなう．これらの反復計算には次に示す交互最小二乗法が用いられる：

- **ステップ1**：最適尺度化による$(\boldsymbol{a}_1, \boldsymbol{a}_2)$の推定と数量化データ$\boldsymbol{W}=(\boldsymbol{W}_{名義}, \boldsymbol{W}_{順序}, \boldsymbol{X}_{数値})$の計算
- **ステップ2**：$\boldsymbol{W}$に対する統計解析手法のパラメータの計算

上記の計算ステップを予め設けた収束条件に達するまで繰り返す．このとき，**ステップ2**の統計解析手法は，解析目的に合わせて選択することができる．

従来の統計解析手法で得られる解析結果に加え，最適尺度化による名義データおよび順序データを数量化することにより，回答者の意識等に関して新たな知見を得ることが期待できると考えている．

連絡先 TEL: 086-256-9741, FAX: 086-256-9741, E-mail: kuroda@soci.ous.ac.jp