

回帰分析における整合性評価とデータ分割

愛知学院大学経営学部 田中浩光

1.はじめに

回帰分析では、回帰係数の最小二乗推定値の解釈上の問題が指摘されている。説明変数値間の相関係数値が高い場合に惹起する、多重共線性問題(あるいは準共線性問題)はその1つである。この対策の1つとして、回帰標本上での整合性を確保することを第1とし、その上で回帰性を得る推定回帰式の構成が考えられる(田中(2011))。また、回帰分析の結果、とくに回帰係数の推定値において、相関係数値などの基礎情報との突き合わせを通して整合性が確保できる方法についての考察がみられる(田中(2013))。本報告では、回帰データ(行列)に着目し、列方向に対応する説明変数、行方向である観測個体での分割の視点にもとづき、整合性方式と評価について整理する。本報告では、説明変数の個数が2個、3個とする場合の重回帰分析を取り上げる。

2.整合性を診る知識

- ・ 相関情報と回帰分析結果(回帰係数の推定値)の乖離：準共線性の問題(多重共線性)
- ・ 奇異な、回帰分析結果(回帰平方和に関する、周辺と同時の関係)：suppressionの生起
- ・ データ分割に伴う観測変数集合間(観測個体集合間)における回帰分析結果の妥当性

3.回帰データ(行列)の分割

- ・ 説明変数(列方向)：説明変数集合間における回帰分析結果(回帰係数値、回帰平方和)
- ・ 観測個体(行方向)：観測個体集合間における回帰分析結果(回帰係数値、回帰平方和)

参考文献

- ・ Suich, R. & Derringer, G.C. (1977). Is the regression equation adequate? one-criteria. *Technometrics*, 19, 213-216.
- ・ Hamilton, D. (1987). Some times $R^2 > r_{y1}^2 + r_{y2}^2$, correlated variables are not always redundant. *The American Statistician*, 41, 2, 129-132.
- ・ Shieh, G. (2001). The inequality between the coefficients of determination and the sum of squared simple correlation coefficients. *The American Statistician*, 55, 2, 122-124.
- ・ 田中浩光(2003). 回帰分析における Suppression と準共線性, 2003年度統計関連学会連合大会講演報告集.
- ・ 田中浩光(2007). 回帰データの分割, 2007年度統計関連学会連合大会講演報告集
- ・ 田中浩光(2011). 適切な推定回帰式と整合性問題について, 2011年度統計関連学会連合大会講演報告集・田中浩光(2013). 回帰分析における整合性問題に関する一考察, 2013年度日本経営数学会研究大会報告集