

回帰診断と整合性問題

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、推定回帰式が標本の抽出・説明変数の選定など観測過程の影響を受ける。多重共線性問題(あるいは準共線性問題)など、回帰データの観測過程に纏わる解釈上の問題が指摘されている。通常、適切な推定回帰式として、所謂、回帰診断の諸方法を通して得ることになる。しかし、これらの診断方法は、主として推測に焦点をおくものであり、必ずしも実地の解釈に沿うものでない。一方、観測特性間の相関情報に基づいて、解釈上の整合性に重きをおいた診断方法が提示されている(田中(2012))。方法の眼目は、準共線性の現象として特定された現象が生起する領域に着目するものである。本報告では、準共線性現象の生起領域(田中(2012))に基づいた整合性指標を提示する。整合性の程度として、相関係数と回帰係数の符号の関係を通して、4段階で階層的に分類する。提示する整合性指標に対し、回帰標本の崩れとなる観測過程の諸作業、とくに説明変数の選定に関係する準共線性の影響について数値的に吟味・考察する。

2.整合性指標

本報告では、2 個の説明変数からなる回帰分析に限定する。観測変数間の相関係数をそれぞれ r_{y1} 、 r_{y2} 、 r_{12} とする。説明変数を単一にするときの回帰係数の推定値を b_1 、 b_2 とし、説明変数を 2 個とするときの回帰係数の推定値を $b_{1.(1,2)}$ 、 $b_{2.(1,2)}$ とする。

(1) 解釈上の整合性(準共線性の現象が生起する領域の排反)

相関情報と回帰係数の対応関係に着目し、最小二乗推定値間の符号の変化に注意する。

$$\begin{aligned} \cdot r_{12} > 0 &\rightarrow b_1 b_2 > 0, b_{1.(1,2)} b_{2.(1,2)} > 0 \\ \cdot r_{12} < 0 &\rightarrow b_1 b_2 < 0, b_{1.(1,2)} b_{2.(1,2)} < 0 \end{aligned}$$

(2) 整合性指標

相関係数と回帰係数の符号の関係を通して、準共線性現象の生起領域を階層的に分割する。

参考文献

- ・ 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.
- ・ 田中浩光(2003).最小二乗推定回帰係数の幾何学的解釈—抑制と準共線性—,愛知学院大学論叢経営学研究,12(4),87-97.
- ・ 田中浩光(2009).回帰分析における Shieh の H 関数について,愛知学院大学論叢経営学研究,19(1),19-33.
- ・ 田中浩光(2011).適切な推定回帰式と整合性問題, 2011 年度統計関連学会連合大会講演報告集.
- ・ 田中浩光(2012).線形回帰分析における準共線性と現象の特定,日本経営数学会誌,33(1/2),57-63..