土木学会第54回年次学術講演会（平成11年9月）

IV-289

リッジ・パラメータの同定手法に関する一考察

東京大学大学院　正会員　堤　盛人
東京大学大学院　学生会員　井出 裕史
東京大学大学院　正会員　清水 英範

1. はじめに

回帰モデルのパラメータ推定において説明変数が高い相関をもつ多重共線性が存在している状況は、解の安定性が損なわれるという意味で非適切である。非実験科学の分野ではデータの受動性が顕著であり、多重共線性に悩まされることが非常に多い。リッジ回帰は多重共線性に対する代表的な適切化手法の一つであるが、その適用においては、折衷の役割を担うリッジ・パラメータ（折衷パラメータ）をどのように同定するかが重要な問題となる。本研究では、リッジ・パラメータの同定についての既存研究を概観し、実際のデータを用いた実証比較を行う。

2. リッジ・パラメータの同定に関連する既存研究

以下yを被説明変数、x_j（$j=1,2,\cdots,m$）を説明変数、β_j（$j=0,1,2,\cdots,m$）をパラメータとする通常の回帰モデルを扱う。

$$y = X\beta + \varepsilon \tag{1}$$

$$y=(y_1,\cdots,y_i,\cdots,y_n)^t,\quad \beta=(\beta_0,\beta_1,\cdots,\beta_m)^t,\quad \varepsilon=(\varepsilon_1,\cdots,\varepsilon_i,\cdots,\varepsilon_n)^t,\quad X=\begin{pmatrix}1 & x_{11} & \cdots & x_{m1}\\ \vdots & \vdots & \ddots & \vdots\\ 1 & x_{1n} & \cdots & x_{mn}\end{pmatrix}$$

nはデータの数（$>m+1$）であり、tはベクトル及び行列の転置を表す。また、誤差項εは平均0、分散均一の通常の仮定を満たすものとする。Hoerl and Kennardによって提唱されたリッジ回帰では、リッジ・パラメータと呼ばれる適当なスカラーcと単位行列Iを用いて、パラメータの推定量が次式で与えられる。

$$\beta_R(c)=\left(X^tX+cI\right)^{-1}X^ty \tag{2}$$

リッジ回帰については批判も多いが、数理科学一般においてTikhonovの適切化として広く知られている方法に他ならず、予測値の安定化を図るうえでは自然な方法である。

リッジ回帰の適用に際しては、リッジ・パラメータcをどのように同定するかが問題となる（パラメータについて何らかの事前情報を仮定すると、ベイズの定理を用いてリッジ・パラメータの推定が可能な場合もあるが、一般にはこの種の事前情報を持っていることは稀である）。

Hoerl and Kennard (1970) は、リッジ・トレースという、リッジ・パラメータcとパラメータの値β_iをプロットした図を用い、パラメータβ_iの値が安定するようなcを採用することを提唱した。しかし、この基準は極めて曖昧で分析者の恣意性が介入する余地が大きく、しかもそれによって求めたパラメータの推定値がどのような統計的性質を持つかは不明である。

Allen（1974）による折衷パラメータの選択方法は、交差確認法の利用によって求められた予測値の誤差の期待値であるPRESS (*Prediction Sum of Squares*) を最小にするものであり、リッジ回帰ではPRESSは次式で与えられる。

$$PRESS(c)\equiv\frac{1}{n}\sum_{i=1}^{n}\left([X\beta^{(i)}(c)]_i-y_i\right)^2 \tag{3}$$

（$[\cdot]_i$はベクトルのi番目の要素を表す演算子とする。）

AllenのPRESSは説明変数の単位系（座標系）に依存し、次式で表される行列$A(c)$が対角行列に近づくと解が不安定になるという問題が生じる［Mallows（1973）］。

$$A(c)\equiv X\left(X^tX+cI\right)^{-1}X^t \tag{4}$$

そこでMallows（1973）は、PRESSを誤差分散で基準化したJ_pの期待値として定義される次式のC_pを最小にするリッジ・パラメータを採用することを提唱した。

$$C_p(c)\approx\frac{1}{\sigma^2}\sum_{i=1}^{n}\left([X\beta(c)]_i-y_i\right)^2+tr\{A(c)\}^2-tr\{I-A(c)\}^2 \tag{5}$$

一方、Wahba（1977）・Golub et al.（1979）はAllenのPRESSを拡張し、座標依存性の問題点を克服したGCV (*Generalized Cross-Validation*)を提案した。リッジ回帰においては、GCVは次式で与えられる。

$$GCV(c)\equiv\frac{1}{n}\left\|(\mathrm{I}-A(c)y)\right\|^2\Big/\left[\frac{1}{n}tr(\mathrm{I}-A(c))\right]^2 \tag{6}$$

以上の詳細については、参考文献も含めて堤・清水・井出(1998)を参照されたい。

GCVは折衷パラメータの変化に対して鈍感な場合があるとの指摘がある［Varah（1983），細田・北川（1993）］。そこでHansen（1992）は、GCVに比べて頑健な折衷パラメータの同定法として*l*-curveと呼ばれるものを用いた手法を提案した。*l*-curveとは、折衷パラメータcを$[0,\infty)$の

キーワード：適切化，リッジ回帰，リッジ・パラメータ，地価分析
連絡先　：　〒113-8656　文京区本郷7-3-1　東京大学大学院工学系研究科　社会基盤工学専攻
TEL：03-5841-6128　FAX：03-5689-7290　e-mail：tsutsumi@planner.t.u-tokyo.ac.jp

範囲で動かすことにより得られる点（‖e‖, ‖β‖）（‖e‖：残差ノルム、‖β‖：推定パラメータのノルム）の軌跡のことであり、Hansen（1992）は l-curve が corner を持つところで折衷パラメータを定めることを提案した。corner の定義は曖昧であったが、細田・北川（1992）は l-curve の曲率 κ が最大になるところと定義した。リッジ回帰においては l-curve の曲率 κ は次式で与えられる。

$$\kappa(c) = \frac{1}{\left(\|e(c)\|^2 + c^2\|\beta(c)\|^2\right)^{\frac{3}{2}}}\left|\frac{\|e(c)\|^2\|\beta(c)\|^2\left(L_1(c)+3cL_2(c)\right)}{\{L_3(c)\}^2} - c\left(c\|e(c)\|^2 + \|\beta(c)\|^2\right)\right| \quad (7)$$

$$L_1(c) \equiv \sum_{j=1}^{m}\frac{\lambda_j(\lambda_j-2c)}{(\lambda_j+c)^4}(s_j\cdot y)^2,\quad L_2(c) \equiv \sum_{j=1}^{m}\frac{\lambda_j}{(\lambda_j+c)^4}(s_j\cdot y)^2,$$

$$L_3(c) \equiv \sum_{j=1}^{m}\frac{\lambda_j}{(\lambda_j+c)^3}(s_j\cdot y)^2 \quad (8)\sim(10)$$

3. 地価データを用いた実証的比較分析

本研究では、堤・清水・井出(1998)と同じく東京都足立区内の平成8年公示地価データ（住宅地）を用いた。紙面の都合で、詳細は参考文献に譲る。

2.に示した手法を適用した結果を図に示す。リッジ・トレースでは c の同定が困難であるが、Allen の PRESS では c =0.12 が与えられた。一方、Mallows の C_p と GCV は、蓑谷（1992）が指摘するようにほぼ同じ値 c = 0.082 を与えたが、l-curve 法ではかなり大きい c = 0.20 が与えられ、細田・北川（1993）の結論とは対照的な結果を得た。すなわち、GCV が折衷パラメータの変化に対して鈍感であるとその有効性への疑問が指摘されるのに対し、むしろ GCV に代わる方法として提案されている l - curve 法の方が過大なリッジ・パラメータを与える可能性を示唆している。紙面の都合上、より詳細な結果や議論については、講演時に示すこととする。

【参考文献】

Hansen, P. C.（1992）Analysis of discrete ill-posed problems by means of the l-curve, *SIAM Review*, Vol.34, No.4, pp.561-580.

細田陽介・北川高嗣（1993）不適切問題に対する MAICE-DP 法による最適正則化法について，『日本応用数理学会論文誌』， Vol.3, No.2, pp.47-58.

Kitagawa, T.（1987）A deterministic approach to optimal regularization -The finite dimensional case-, *Japan Journal of Applied Mathematics*, Vol.4, pp371-391.

蓑谷千凰彦（1992）計量経済学の新しい展開，多賀出版.

堤盛人・清水英範・井出裕史（1998）多重共線性に対する適切化手法とその実証的比較研究 ー地価の回帰分析を例としてー，『応用力学論文集』, Vol. 1, pp.137-145.

Varah, J. M.（1983）Pitfalls in the numerical solution of linear ill-posed problems, *SIAM Journal on Scientific and Statistical Computing*, Vol.4, No.2, pp.164-176.

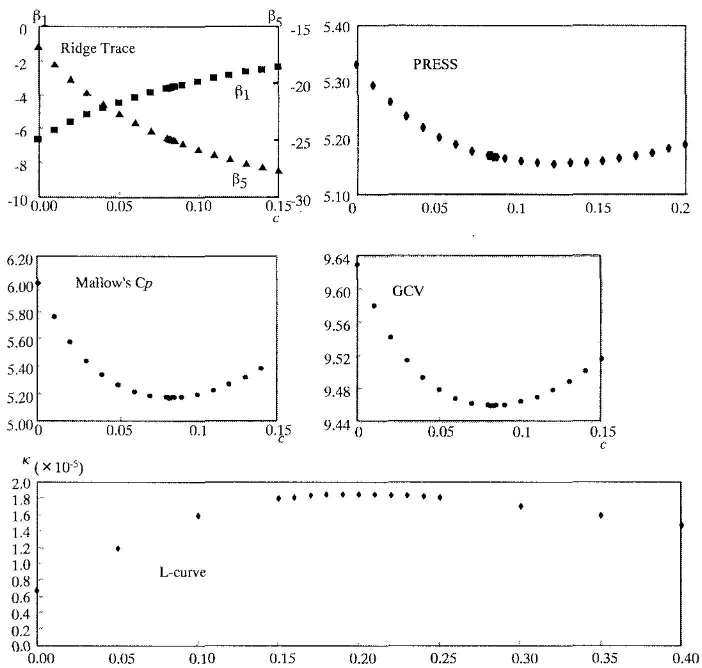

図　各手法における指標の推移（横軸はいずれもリッジ・パラメータ）