

平均2乗誤差指標によるリッジパラメータに関する 影響力評価

竹内 秀一

Several influence measures have been proposed to assess the influence of observations on ridge regression. Influence measures based on the case deletion procedure that are used in the case of linear regression, such as Cook's distance or the likelihood distance, have been extended to the case of ridge regression. Most of these influence measures are derived from benchmarks based on statistical or mathematical distance criteria. In the case of a typical linear regression, these measures are the primary methods for conducting such assessments. However, in the case of ridge regression, we must consider not only assessments of the influence of observations but also assessments of the influence of the ridge parameter, a feature unique to the case of ridge regression. Therefore, in the present paper, we consider an ancillary influence measure for the latter assessment.

The present study is a detailed investigation of the properties of an ancillary influence measure derived from benchmarks based on the mean square error criterion. The mean square error criterion is a standard statistical benchmark used to assess the regression coefficients in ridge regression, and it is also frequently used to identify an optimal ridge parameter. To investigate the properties of this ancillary influence measure, we derive a derivative function by computing partial derivatives with respect to the ridge parameter. As a result of this investigation, we discovered that the derivative function may be expressed as the sum of the original function and functions associated with it. In addition, we plot the behavior of the derivative function for each observation in order to obtain a clear understanding of variations with respect to the ridge parameter. Our results demonstrate that assessments of the influence of the ridge parameter are a useful, albeit ancillary and supplemental, complement to the case deletion procedure for assessing the influence of observations.

1 はじめに

リッジ回帰 (ridge regression) における観測値の影響力評価において、いくつかの診断統計量 (influence measure) が提案されている。線形回帰 (linear regression) の場合と同様に、観測値除去法 (case deletion procedure) に基づく診断統計量として、Cook の距離や尤度距離などがリッジ回帰の場合に拡張されている。これらの診断統計量の多くは統計的あるいは数学的な距離規準に基づく指標から導入されている。通常の線形回帰の場合、観測値の影響力を評価するためには、観測値除去法に基づくこうした診断統計量を中心にデータ解析が行われる。しかしながら、リッジ回帰の場合は、観測値の影響力評価に加えて、リッジ回帰固有の問題であるリッジパラメータに関する影響力評価についても考慮する必要がある。そこで、リッジパラメータに関する影響力を評価するための補助診断統計量について検討する。

本研究では、Takeuchi [4] によって提案された補助診断統計量の性質について詳しく調べる。この補助診断統計量は、各種の推定量を評価するうえでよく利用される平均2乗誤差 (mean square error, 以下「MSE」) に基づく指標から導入された。MSE 規準自体はリッジ回帰における回帰係数を評価するための典型的な統計指標であり、最適なリッジパラメータを特定するためにもよく利用される。

この補助診断統計量の性質を調べるために、リッジパラメータについての偏微分により導関数を導く。この結果、元の関数やそれに関連した関数の和として、導関数を表現できることがわかった。また、リッジパラメータに関する変化を明確に捉えるために、導関数の挙動を示すグラフを観測値ごとに描く。これにより、観測値除去法による観測値の影響力評価に加えて、補助的あるいは補完的にはなるが、リッジパラメータに関する影響力評価に役立つことも示された。

本論文の構成は以下のとおりである。第2節では線形回帰およびリッジ回帰における各種の基本的な統計量を与える。また、リッジ回帰における補助診断統計量も定義する。第3節において、補助診断統計量の性質を調べるために、リッジパラメータについての偏微分により導関数を導き、その挙動をグラフ化する。第4節は全体のまとめと今後の課題である。

2 定義

本節では、まず線形回帰およびリッジ回帰における各種の基本的な統計量を与える。つぎに、リッジ回帰における補助診断統計量を示す。

2.1 線形回帰およびリッジ回帰

ここでは、線形回帰モデルとして、

$$y = X\beta + \varepsilon$$

を考える。このとき、 y は $n \times 1$ の目的変数ベクトル、 X は $n \times p$ のフルランクの説明変数行列、 β は $p \times 1$ の回帰係数ベクトル、そして ε は $n \times 1$ の誤差ベクトルであり、正規分布 $N(0, \sigma^2 I_n)$ に従うものとする。ただし、 I_n は n 次の単位行列を表す。また、 β の最小2乗推定量は $\hat{\beta} = (X'X)^{-1}X'y$ として得られ、誤差分散 σ^2 の不偏推定量は $\hat{\sigma}^2 = e'e/(n-p)$ となる。ただし、「 $'$ 」は行列あるいはベクトルの転置を表し、 e は残差ベクトルであり、 $e = y - X\hat{\beta} = (I_n - H)y$ である。このとき、 H は説明変数行列 X から構成されるハット行列 (hat matrix) $H = X(X'X)^{-1}X'$ であり、その第 i 対角成分 h_{ii} がてこ比である。このてこ比については、 $1/n \leq h_{ii} < 1$ とする。

線形回帰の1つの代替的方法としてリッジ回帰がある (Groß [2] や Hoerl and Kennard [3] などを参照)。リッジ回帰における回帰係数ベクトル β の推定量 (以下「リッジ推定量」) を、リッジパラメータ $k (\geq 0)$ を導入することにより、

$$\hat{\beta}^R = (X'X + kI_p)^{-1}X'y$$

と定義する。すると、最小2乗推定量の場合と同じく、残差ベクトル e^R は、 $e^R = y - X\hat{\beta}^R = (I_n - H^R)y$ となる。ただし、リッジ回帰におけるハット行列 H^R は $H^R = X(X'X + kI_p)^{-1}X'$ であり、その第 i 対角成分 h_{ii}^R がリッジ回帰におけるてこ比である。このとき、 $0 < h_{ii}^R < 1$ である。

また、第 i 番目の観測値を除去したときの回帰係数ベクトル β の最小2乗推定量は

$$\hat{\beta}_{(i)} = (X'_{(i)}X_{(i)})^{-1}X'_{(i)}y_{(i)}$$

と定義される。ただし、添字の (\cdot) は n 個の観測値の中から除去される観測値の番号を表す。通常の最小2乗推定量の場合と同様に、第 i 番目の観測値を除去したときのリッジ推定量は

$$\hat{\beta}_{(i)}^R = (X'_{(i)}X_{(i)} + kI_p)^{-1}X'_{(i)}y_{(i)}$$

となる。

特に、 $k=0$ とすれば、 $\hat{\beta}^R = \hat{\beta}$ 、 $\hat{\beta}_{(i)}^R = \hat{\beta}_{(i)}$ 、あるいは $H^R = H$ など、リッジ回帰において定義される統計量が通常の線形回帰の統計量を包含していることがわかる。

2.2 補助診断統計量

通常の線形回帰における観測値に対する影響力評価では、Cookの距離や尤度距離 (Cook and Weisberg [1] などを参照) などの診断統計量が利用される。けれども、リッジ回帰においては、こうした観測値除去法に基づく診断統計量だけでは、観測値の影響力評価において不十分な場合があるので、これを補完する意味で補助診断統計量がいくつか提案されている。ここでは、リッジパラメータに対する影響力を評価するための補助診断統計量として、

平均2乗誤差指標によるリッジパラメータに関する影響力評価

MSE 指標 (MSE index) を取り上げる.

リッジ推定量に基づく補助診断統計量 MI_i を以下のように定義する (Takeuchi [4] を参照).

$$MI_i \equiv \text{MSE}(\hat{\beta}^R) - \text{MSE}(\hat{\beta}_{(i)}^R) \quad (2.1)$$

ただし, $\hat{\beta}^R$ および $\hat{\beta}_{(i)}^R$ の MSE は, 回帰係数ベクトルの真の値 β との差に基づく平方距離の期待値として算出され, それぞれ

$$\text{MSE}(\hat{\beta}^R) = E\{(\hat{\beta}^R - \beta)'(\hat{\beta}^R - \beta)\} \quad (2.2)$$

および

$$\text{MSE}(\hat{\beta}_{(i)}^R) = E\{(\hat{\beta}_{(i)}^R - \beta)'(\hat{\beta}_{(i)}^R - \beta)\} \quad (2.3)$$

である. 特に, $k=0$ のとき, 通常の線形回帰における最小2乗推定量の MSE と同じになり, (2.2) 式および (2.3) 式が $\text{MSE}(\hat{\beta})$ および $\text{MSE}(\hat{\beta}_{(i)})$ とそれぞれ一致する.

(2.1) 式の MSE の定義式において, (2.2) 式および (2.3) 式における β や MSE の分散項における σ^2 のままでは具体的な数値が算出できないので, β の推定量として $\hat{\beta}$ を代入し, また σ^2 の推定量として $\hat{\sigma}^2$ を代入することによって, MSE の推定量を求めることになる.

この MSE の推定量から, リッジ回帰における観測値の影響力を評価するための補助診断統計量として, 以下のような MSE 指標である MI_i を導くことができる.

$$MI_i = \frac{\hat{\sigma}^2}{k} s_i \left[\left(1 + \frac{s_i}{s_i^*} \right) \left(\frac{1}{s_i^*} - u_i^2 \right) - 2 \left(\frac{1}{s_i^{**}} - u_i u_i^* \right) \right] \quad (2.4)$$

ただし, u_i および u_i^* はそれぞれ

$$u = \frac{1}{\hat{\sigma}} [\text{diag}(\mathbf{I}_n - \mathbf{H})]^{-\frac{1}{2}} (e^R - e)$$

および

$$u^* = \frac{1}{\hat{\sigma}} [\text{diag}(\mathbf{I}_n - \mathbf{H})]^{-\frac{1}{2}} (e^{R^2} - e^R)$$

の第 i 成分であり, また, s_i , s_i^* それに s_i^{**} はそれぞれ

$$\begin{aligned} \mathbf{S} &= \text{diag}(\mathbf{I}_n - \mathbf{H}) [\text{diag}(\mathbf{I}_n - \mathbf{H}^R)]^{-1}, \\ \mathbf{S}^* &= \text{diag}(\mathbf{I}_n - \mathbf{H}) [\text{diag}\{\mathbf{I}_n - (\mathbf{H}^R)^2\}]^{-1}, \end{aligned}$$

それに

$$\mathbf{S}^{**} = \text{diag}(\mathbf{I}_n - \mathbf{H}) [\text{diag}\{\mathbf{I}_n - (\mathbf{H}^R)^3\}]^{-1},$$

の第 i 対角成分であり, \mathbf{S} の第 i 対角成分が $s_i = (1 - h_{ii}) / (1 - h_{ii}^R)$, \mathbf{S}^* の第 i 対角成分が $s_i^* = (1 - h_{ii}) / (1 - h_{ii}^{R^2})$, それに \mathbf{S}^{**} の第 i 対角成分が $s_i^{**} = (1 - h_{ii}) / (1 - h_{ii}^{R^3})$ である. ここで, $e^{R^2} = (\mathbf{I}_n - \mathbf{H}^{R^2})y$ であり, $h_{ii}^{R^2}$ は $\mathbf{H}^{R^2} = (\mathbf{H}^R)^2$ の第 i 対角成分であり, それに $h_{ii}^{R^3}$ は $\mathbf{H}^{R^3} = (\mathbf{H}^R)^3$ の第 i 対角成分である. 加えて, $\text{diag}(\mathbf{A})$ は正方形行列 \mathbf{A} の対角成分のみを取り出し, 非対角成分をすべて 0 にした行列を表す.

3 補助診断統計量の性質

リッジ回帰における補助診断統計量の性質を調べるために、リッジパラメータ k について偏微分することにより導関数を導出する。(2.4) 式で与えられる MI_i の導関数 DM_i は、 $k > 0$ の場合について

$$DM_i \equiv \frac{\partial}{\partial k} MI_i = -\frac{2}{k} \left[\left(1 + \frac{s_i}{s_i^*} \right) MI_i - MI_i^* - RMI_i \right] \quad (3.1)$$

となる(導出過程については付録 A を参照)。ただし、

$$MI_i^* = \frac{\hat{\sigma}^2}{k} s_i \left[\left(1 + \frac{s_i}{s_i^*} \right) \left(\frac{1}{s_i^{**}} - u_i^2 \right) - 2 \left(\frac{1}{s_i^{***}} - u_i u_i^{**} \right) \right] \quad (3.2)$$

および

$$RMI_i = \frac{\hat{\sigma}^2}{k} s_i \left[\frac{1}{s_i^{**}} (1 - s_i u_i^2) - \frac{1}{s_i^{***}} \{ 1 - s_i^{***} (u_i^*)^2 \} \right] \quad (3.3)$$

である。このとき、 u_i^{**} は

$$u_i^{**} = \frac{1}{\hat{\sigma}} [\text{diag}(I_n - H)]^{-\frac{1}{2}} (e^{R3} - e^{R2})$$

の第 i 成分であり、 $e^{R3} = (I_n - H^{R3})y$ である。また、 s_i^{***} は

$$S^{***} = \text{diag}(I_n - H) [\text{diag}(I_n - H^{R4})]^{-1}$$

の第 i 対角成分であり、 $s_i^{***} = (1 - h_{ii}) / (1 - h_{ii}^{R4})$ である。ここで、 h_{ii}^{R4} は $H^{R4} = (H^R)^4$ の第 i 対角成分である。

この結果、(3.1) 式からわかるように、導関数 DM_i の構成要素は、元の関数 MI_i を含み、(3.2) 式のような MI_i^* という (2.4) 式で与えられる MI_i に酷似した部分と、(3.3) 式の残余項 RMI_i であることがわかった。さらに、具体的な性質を調べるために、数値的な検証を行う。

通常の診断統計量である Cook の距離や尤度距離の場合にも適用した数値例として、Takeuchi [4] や竹内 [6] [7] で取り上げている人工データ (Artificial Data) について、補助診断統計量の性質を調べる。このデータは観測値の数が $n=20$ で、説明変数の数は定数項を含めると $p=4+1=5$ である。このデータに対して、 $0 < k \leq 1$ の領域について、 MI_i および DM_i の挙動をグラフ化した結果が、それぞれ図 3.1 および図 3.2 である。また、参考までに、(3.1) 式の DM_i の構成要素である (3.2) 式の MI_i^* および (3.3) 式の RMI_i の挙動についても図 3.3 および図 3.4 にそれぞれグラフ化して示した。なお、20 個の全観測値をグラフ化すると見にくくなるので、ここでは代表的なデータ No.7, No.10, No.14 それに No.17 の 4 つに絞って提示している(参考として全観測値の挙動を示すグラフを付録 C に掲げた)。

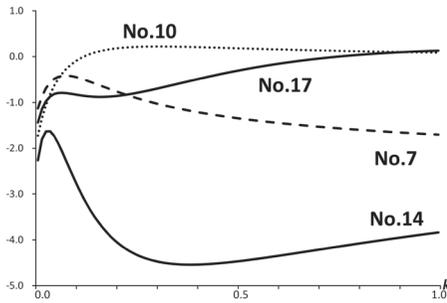


図 3.1 MI_i の挙動

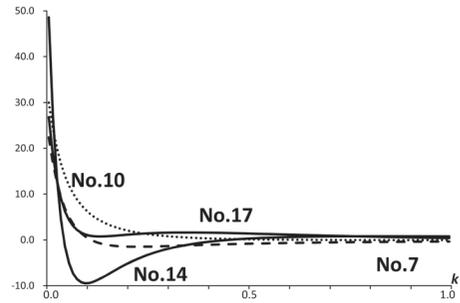


図 3.2 DM_i の挙動

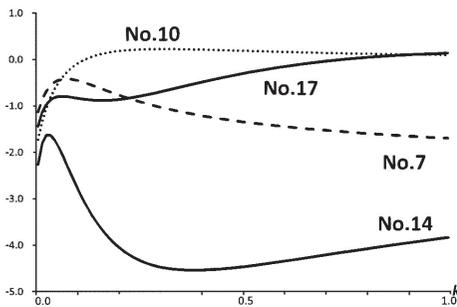


図 3.3 MI_i^* の挙動

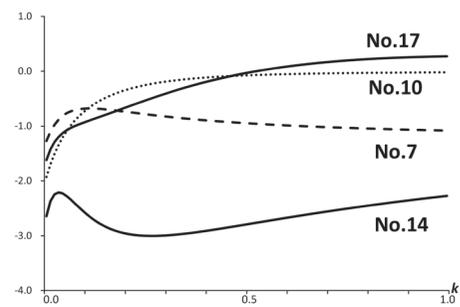


図 3.4 RMI_i の挙動

特徴的な性質として、図 3.1 および図 3.2 のグラフからは、通常の診断統計量である Cook の距離や尤度距離の場合（グラフについては竹内 [5] を参照）よりもリッジパラメータ k の値が 0 に近い領域での変化が大きいことがわかる。これはリッジパラメータの変化に着目した MSE を診断指標としたためであると思われる。通常の診断統計量は、何らかの距離規準について標準化等を行うことにより導出されているので、リッジパラメータの変化が現れにくい構造をしているものと推察できる。統計的あるいは数学的距離規準に基づく多くの診断統計量においては、何らかの標準化（あるいは調整）を行うために、比率が算出される（ベクトルや行列においては逆行列が絡む）。そのため、リッジパラメータの変化が比率の定義式における分子と分母で相殺されてしまう可能性が高まる。これに対して、MSE 規準を利用したこの補助診断統計量の場合、MSE が分散とバイアス（正確にはバイアスの 2 乗）に分けられるだけであるので、特に標準化等の調整は行われていない、リッジ推定量の大きさについての平方距離だけに基づいて影響力を評価しているため、リッジパラメータの変化による影響力の効果が直接的に現れやすいものと思われる。もちろん、リッジパラメータの直接的な影響力が現れやすいことが、必ずしも望ましい性質ということではないが、リッジパラメータ k の値を決定する際には、有力な情報を与える可能性が高いと考えられる。

また、特徴的な性質としての第二点目は、図 3.1、図 3.3 それに図 3.4 のグラフから、変化

量を示す指標である DM_i を構成する3つの要素 MI_i , MI_i^* それに RMI_i の挙動はよく似ていることがわかる. DM_i の挙動から k の値が大きい領域 (おおよそ $k \geq 0.3$ の領域) では, 変化率がほぼ0であるので, k としては0に近い狭い領域だけを調べればよいといえる. この点は Cook の距離や尤度距離の場合と同様の結論であるが, あらためてこの点も確認することができた.

以上のことから, 補助診断統計量 MI_i を利用することにより, リッジパラメータの影響力評価を含めた広い意味での観測値の影響力評価が形式的には可能になるといえる. ただ, MI_i は, 通常の距離規準に基づく診断統計量とは異なり標準化等の調整が行われていないために, リッジ推定量自体が大きく (小さく) になるとその値も連動して大きく (小さく) なるという欠点がある. この欠点があるために, 各種の診断統計量による観測値の影響力評価を補助あるいは補完する役割しか果たすことはできない. しかしながら, リッジパラメータに関する影響力評価を含めた広義の観測値の影響力評価は十分に可能であるので, データ解析に役立てる工夫を検討する余地はある.

4 まとめと今後の課題

本論文では, リッジ回帰における広義の観測値の影響力評価について, 補助診断統計量 MI_i の性質を, その導関数から検討した. 人工的なデータに基づく数値例からもわかるように, この補助診断統計量はリッジパラメータに関する影響力の効果が直接的に現れやすいことが大きな特徴であると確認できた.

今後の課題としては, リッジ回帰における観測値の影響力評価において, 通常の観測値除去法に基づく診断統計量とリッジパラメータに関する補助診断統計量をどのように連携させるのかについて十分な検討が必要となる. 数値実験のみならず理論式からもこれらの関連性あるいは補完性について検討を進めていきたい.

付録 A: DM_i の導出

以下の付録 B および竹内 [6] の付録 2 を利用して, リッジ回帰における補助診断統計量 MI_i のリッジパラメータ k についての偏微分を計算する. (2.4) 式の表現において,

$$MI_i = \frac{\hat{\sigma}^2}{k} s_i \left[\left(1 + \frac{s_i}{s_i^*} \right) \left(\frac{1}{s_i^*} - u_i^2 \right) - 2 \left(\frac{1}{s_i^{**}} - u_i u_i^* \right) \right] = \frac{\hat{\sigma}^2}{k} s_i F_i$$

つまり,

$$F_i = \left(1 + \frac{s_i}{s_i^*} \right) \left(\frac{1}{s_i^*} - u_i^2 \right) - 2 \left(\frac{1}{s_i^{**}} - u_i u_i^* \right)$$

平均2乗誤差指標によるリッジパラメータに関する影響力評価

としてから、 MI_i を k について偏微分すると、

$$DM_i \equiv \frac{\partial}{\partial k} MI_i = \hat{\sigma}^2 \left[\frac{-1}{k^2} s_i F_i + \frac{1}{k} \left\{ \frac{\partial s_i}{\partial k} F_i + s_i \frac{\partial F_i}{\partial k} \right\} \right] = \frac{\hat{\sigma}^2}{k} \left[-\frac{1}{k} s_i F_i + \frac{\partial s_i}{\partial k} F_i + s_i \frac{\partial F_i}{\partial k} \right]$$

となる。[] 内の第2項は、付録B1 (2) を利用すると

$$\frac{\partial s_i}{\partial k} F_i = \frac{s_i}{k} \left(1 - \frac{s_i}{s_i^*} \right) F_i = \frac{1}{k} s_i F_i - \frac{s_i}{s_i^*} \cdot \frac{s_i}{k} F_i$$

となる。また、[] 内の第3項は、付録B1, B3それにB4を適宜利用すると

$$\begin{aligned} \frac{\partial F_i}{\partial k} &= \left(\frac{1}{s_i^*} - u_i^2 \right) \frac{\partial}{\partial k} \left(1 + \frac{s_i}{s_i^*} \right) + \left(1 + \frac{s_i}{s_i^*} \right) \frac{\partial}{\partial k} \left(\frac{1}{s_i^*} - u_i^2 \right) - 2 \frac{\partial}{\partial k} \frac{1}{s_i^{**}} + 2u_i^* \frac{\partial u_i}{\partial k} + 2u_i \frac{\partial u_i^*}{\partial k} \\ &= \left(\frac{1}{s_i^*} - u_i^2 \right) \frac{1}{(s_i^*)^2} \left\{ s_i^* \frac{\partial s_i}{\partial k} - s_i^* \frac{\partial s_i^*}{\partial k} \right\} + \left(1 + \frac{s_i}{s_i^*} \right) \left\{ -\frac{1}{(s_i^*)^2} \frac{\partial s_i^*}{\partial k} - 2u_i \frac{\partial u_i}{\partial k} \right\} \\ &\quad - 2 \frac{-1}{(s_i^{**})^2} \frac{\partial s_i^{**}}{\partial k} + \frac{2}{k} (u_i^*)^2 + 2u_i \frac{2u_i^{**} - u_i^*}{k} \\ &= \left(\frac{1}{s_i^*} - u_i^2 \right) \frac{1}{k} \cdot \frac{s_i}{s_i^*} \left\{ 1 - \frac{s_i}{s_i^*} - 2 \left(1 - \frac{s_i^*}{s_i^{**}} \right) \right\} - \frac{2}{k} \left(1 + \frac{s_i}{s_i^*} \right) \left\{ \frac{1}{s_i^*} \left(1 - \frac{s_i^*}{s_i^{**}} \right) + u_i u_i^* \right\} \\ &\quad + \frac{6}{k} \cdot \frac{1}{s_i^{**}} \left(1 - \frac{s_i^{**}}{s_i^{***}} \right) + \frac{2}{k} (u_i^*)^2 + \frac{2}{k} u_i (2u_i^{**} - u_i^*) \\ &= \frac{1}{k} \left[\frac{s_i}{s_i^*} \left\{ 2 \frac{s_i^*}{s_i^{**}} - \left(1 + \frac{s_i}{s_i^*} \right) \right\} \left(\frac{1}{s_i^*} - u_i^2 \right) - 2 \left(1 + \frac{s_i}{s_i^*} \right) \frac{1}{s_i^*} + 2 \left(1 + \frac{s_i}{s_i^*} \right) \left(\frac{1}{s_i^{**}} - u_i u_i^* \right) \right. \\ &\quad \left. + \frac{6}{s_i^{**}} - \frac{6}{s_i^{***}} + 2(u_i^*)^2 + 4u_i u_i^{**} - 2u_i u_i^* \right] \\ &= \frac{1}{k} \left[-\frac{s_i}{s_i^*} \left\{ \left(1 + \frac{s_i}{s_i^*} \right) \left(\frac{1}{s_i^*} - u_i^2 \right) - 2 \left(\frac{1}{s_i^{**}} - u_i u_i^* \right) \right\} + 2 \frac{1}{s_i^{**}} \left(\frac{1}{s_i^*} - u_i^2 \right) - \frac{2}{s_i^*} - \frac{2s_i}{(s_i^*)^2} \right. \\ &\quad \left. + \frac{2}{s_i^{**}} - 2u_i u_i^* + \frac{6}{s_i^{**}} - \frac{6}{s_i^{***}} + 2(u_i^*)^2 + 4u_i u_i^{**} - 2u_i u_i^* \right] \\ &= -\frac{1}{k} \cdot \frac{s_i}{s_i^*} F_i \\ &\quad - \frac{2}{k} \left\{ \frac{1}{s_i^*} - \frac{4}{s_i^{**}} + \frac{3}{s_i^{***}} - \frac{s_i}{s_i^*} \cdot \frac{1}{s_i^{**}} + \frac{s_i}{(s_i^*)^2} + \frac{s_i}{s_i^{**}} u_i^2 + 2u_i u_i^* - (u_i^*)^2 - 2u_i u_i^{**} \right\} \end{aligned}$$

となる。よって、

$$\begin{aligned}
 DM_i &= \frac{\hat{\sigma}^2}{k} \left[-\frac{1}{k} s_i F_i + \frac{1}{k} s_i F_i - \frac{s_i}{s_i^*} \cdot \frac{s_i}{k} F_i - \frac{s_i}{k} \cdot \frac{s_i}{s_i^*} F_i \right. \\
 &\quad \left. - 2 \frac{s_i}{k} \left\{ \frac{1}{s_i^*} - \frac{4}{s_i^{**}} + \frac{3}{s_i^{***}} - \frac{s_i}{s_i^*} \cdot \frac{1}{s_i^{**}} + \frac{s_i}{(s_i^*)^2} + \frac{s_i}{s_i^{**}} u_i^2 + 2u_i u_i^* - (u_i^*)^2 - 2u_i u_i^{**} \right\} \right] \\
 &= -\frac{2}{k} \left[\frac{s_i}{s_i^*} MI_i + \frac{\hat{\sigma}^2}{k} s_i \left\{ \frac{1}{s_i^*} - \frac{4}{s_i^{**}} + \frac{3}{s_i^{***}} - \frac{s_i}{s_i^*} \cdot \frac{1}{s_i^{**}} + \frac{s_i}{(s_i^*)^2} \right. \right. \\
 &\quad \left. \left. + \frac{s_i}{s_i^{**}} u_i^2 + 2u_i u_i^* - (u_i^*)^2 - 2u_i u_i^{**} \right\} \right]
 \end{aligned}$$

となる。他方、

$$\begin{aligned}
 MI_i - MI_i^* - RMI_i &= \frac{\hat{\sigma}^2}{k} s_i \left[\left(1 + \frac{s_i}{s_i^*} \right) \left(\frac{1}{s_i^*} - u_i^2 \right) - 2 \left(\frac{1}{s_i^{**}} - u_i u_i^* \right) - \left(1 + \frac{s_i}{s_i^*} \right) \left(\frac{1}{s_i^{**}} - u_i^2 \right) \right. \\
 &\quad \left. + 2 \left(\frac{1}{s_i^{***}} - u_i u_i^{**} \right) - \frac{1}{s_i^{**}} (1 - s_i u_i^2) + \frac{1}{s_i^{***}} \{ 1 - s_i^{***} (u_i^*)^2 \} \right] \\
 &= \frac{\hat{\sigma}^2}{k} s_i \left\{ \left(1 + \frac{s_i}{s_i^*} \right) \left(\frac{1}{s_i^*} - \frac{1}{s_i^{**}} \right) - \frac{3}{s_i^{**}} + \frac{3}{s_i^{***}} + 2u_i u_i^* - 2u_i u_i^{**} + \frac{s_i}{s_i^{**}} u_i^2 - (u_i^*)^2 \right\} \\
 &= \frac{\hat{\sigma}^2}{k} s_i \left\{ \frac{1}{s_i^*} - \frac{4}{s_i^{**}} + \frac{3}{s_i^{***}} - \frac{s_i}{s_i^*} \cdot \frac{1}{s_i^{**}} + \frac{s_i}{(s_i^*)^2} + \frac{s_i}{s_i^{**}} u_i^2 + 2u_i u_i^* - (u_i^*)^2 - 2u_i u_i^{**} \right\}
 \end{aligned}$$

となり、上記 DM_i の [] 内の第2項 ({ } 部分) と一致する。

したがって、

$$DM_i = -\frac{2}{k} \left[\frac{s_i}{s_i^*} MI_i + MI_i - MI_i^* - RMI_i \right] = -\frac{2}{k} \left[\left(1 + \frac{s_i}{s_i^*} \right) MI_i - MI_i^* - RMI_i \right]$$

となり、(3.1) 式の導関数を得ることができる。

付録 B: 基本要素の偏微分

上記の付録 A において利用した基本要素のリッジパラメータ k についての偏微分について、それらの結果を個々に示す。

平均2乗誤差指標によるリッジパラメータに関する影響力評価

付録 B1: h_{ii}^{R2} , s_i , u_i および u_i^* の偏微分

以下の4つの偏微分については、竹内 [6] の付録2において算出されているので、結果のみ示しておく。

$$(1) \frac{\partial h_{ii}^{R2}}{\partial k} = \frac{2(h_{ii}^{R3} - h_{ii}^{R2})}{k} \quad (2) \frac{\partial s_i}{\partial k} = \frac{s_i}{k} \left(1 - \frac{s_i}{s_i^*}\right) \quad (3) \frac{\partial u_i}{\partial k} = \frac{u_i^*}{k} \quad (4) \frac{\partial u_i^*}{\partial k} = \frac{2u_i^{**} - u_i^*}{k}$$

付録 B2: h_{ii}^{R3} の偏微分

竹内 [6] 付録2.1 の h_{ii}^{R2} の偏微分と同様に、 h_{ii}^{R3} を $X'X$ の固有値および固有ベクトルの成分によって表してから、 k について偏微分をする。説明変数行列 X を以下のように特異値分解する。

$$X = LG\frac{1}{2}U'$$

ただし、 G は $p \times p$ の対角行列であり、その第 j 対角成分は $X'X$ の第 j 固有値 $\delta_j (> 0)$ である。その第 j 固有値 δ_j に対応する固有ベクトルを第 j 列にもつ行列が U であり、 $p \times p$ の正方行列になり $U'U = UU' = I_p$ を満たす。また、 L は $n \times p$ の行列であり $L'L = I_p$ を満たす。 ℓ_i は L の第 i 行 (ベクトル) であり、 ℓ_{ij} は L の第 (i, j) 成分、つまり ℓ_i の第 j 成分を表す。

h_{ii}^{R3} は H_{ii}^{R3} の第 i 対角成分であり、竹内 [6] 付録2.1における (A.1) 式のように固有値 δ_j および固有ベクトルの成分 ℓ_{ij} によって

$$h_{ii}^{R3} = \sum_{j=1}^p \left(\frac{\delta_j}{\delta_j + k} \right)^3 \ell_{ij}^2$$

と表されるので、

$$\frac{\partial h_{ii}^{R3}}{\partial k} = - \sum_{j=1}^p \frac{3\delta_j^3}{(\delta_j + k)^4} \ell_{ij}^2 = \frac{3}{k} \sum_{j=1}^p \left[\left(\frac{\delta_j}{\delta_j + k} \right)^4 - \left(\frac{\delta_j}{\delta_j + k} \right)^3 \right] \ell_{ij}^2$$

となる。ここでも竹内 [6] 付録2.1と同じく、 $H^{R4} = (H^R)^4 = L(I_p + kG^{-1})^{-4}L'$ であるので、この第 i 対角成分は

$$h_{ii}^{R4} = \sum_{j=1}^p \left(\frac{\delta_j}{\delta_j + k} \right)^4 \ell_{ij}^2$$

と表すことができる。よって、

$$\frac{\partial h_{ii}^{R3}}{\partial k} = \frac{3(h_{ii}^{R4} - h_{ii}^{R3})}{k}$$

となる。

付録 B3: s_i^* の偏微分

s_i^* の k についての偏微分は、

$$\frac{\partial s_i^*}{\partial k} = \frac{\partial}{\partial k} \frac{1-h_{ii}}{1-h_{ii}^{R2}} = (1-h_{ii}) \frac{\partial}{\partial k} \frac{1}{1-h_{ii}^{R2}} = (1-h_{ii}) \frac{1}{(1-h_{ii}^{R2})^2} \frac{\partial h_{ii}^{R2}}{\partial k}$$

となるので、上記付録 B1 (1) を利用すると、

$$\begin{aligned} \frac{\partial s_i^*}{\partial k} &= \frac{1-h_{ii}}{(1-h_{ii}^{R2})^2} \cdot \frac{2(h_{ii}^{R3}-h_{ii}^{R2})}{k} = \frac{1}{k} \cdot \frac{1-h_{ii}}{1-h_{ii}^{R2}} \cdot \frac{2\{(1-h_{ii}^{R2})-(1-h_{ii}^{R3})\}}{1-h_{ii}^{R2}} \\ &= \frac{2s_i^*}{k} \left(1 - \frac{1-h_{ii}}{1-h_{ii}^{R2}} \right) = 2 \frac{s_i^*}{k} \left(1 - \frac{s_i^*}{s_i^{**}} \right) \end{aligned}$$

と変形することができる。

付録 B4 : s_i^{**} 偏微分

s_i^{**} の k についての偏微分は、

$$\frac{\partial s_i^{**}}{\partial k} = \frac{\partial}{\partial k} \frac{1-h_{ii}}{1-h_{ii}^{R3}} = (1-h_{ii}) \frac{\partial}{\partial k} \frac{1}{1-h_{ii}^{R3}} (1-h_{ii}) \frac{1}{(1-h_{ii}^{R3})^2} \frac{\partial h_{ii}^{R3}}{\partial k}$$

となるので、上記付録 B2 を利用して、

$$\begin{aligned} \frac{\partial s_i^{**}}{\partial k} &= \frac{1-h_{ii}}{(1-h_{ii}^{R3})^2} \cdot \frac{3(h_{ii}^{R4}-h_{ii}^{R3})}{k} = \frac{1}{k} \cdot \frac{1-h_{ii}}{1-h_{ii}^{R3}} \cdot \frac{3\{(1-h_{ii}^{R3})-(1-h_{ii}^{R4})\}}{1-h_{ii}^{R3}} \\ &= \frac{3s_i^{**}}{k} \left(1 - \frac{1-h_{ii}}{1-h_{ii}^{R3}} \right) = 3 \frac{s_i^{**}}{k} \left(1 - \frac{s_i^{**}}{s_i^{***}} \right) \end{aligned}$$

と変形することができる。

付録 C: 全観測値の挙動を示すグラフ

$0 < k \leq 1$ の領域について、(2.4) 式の MI_i 、(3.1) 式の DM_i 、(3.2) 式の MI_i^* それに (3.3) 式の RMI_i の挙動を全観測値に関してグラフ化した結果が、それぞれ図 C.1、図 C.2、図 C.3 それに図 C.4 である。

平均2乗誤差指標によるリッジパラメータに関する影響力評価

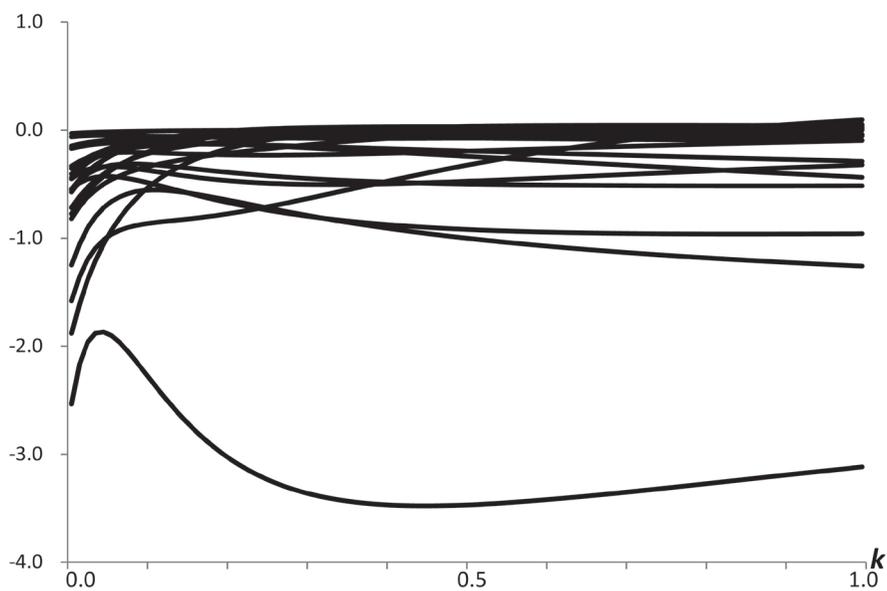


図 C.1 全観測値に関する MI_i の挙動

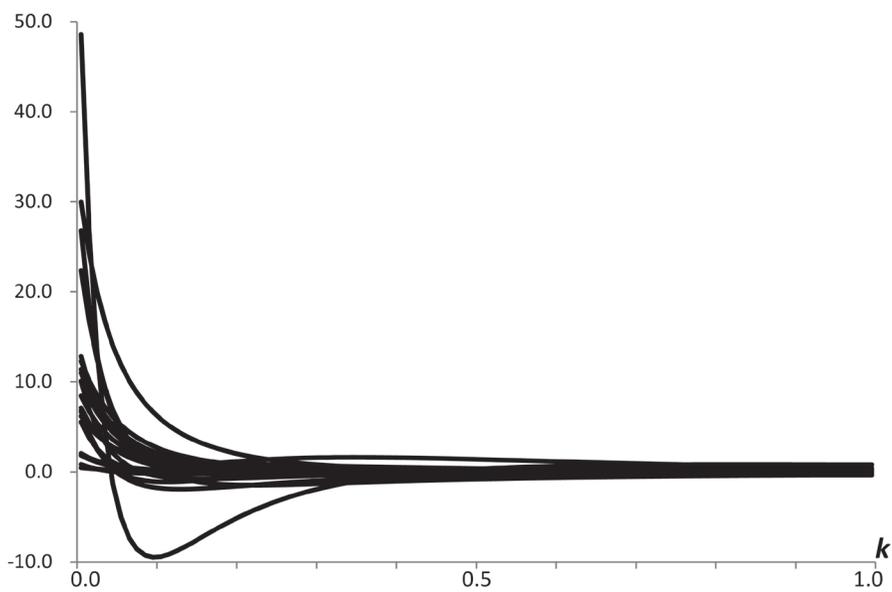


図 C.2 全観測値に関する DM_i の挙動

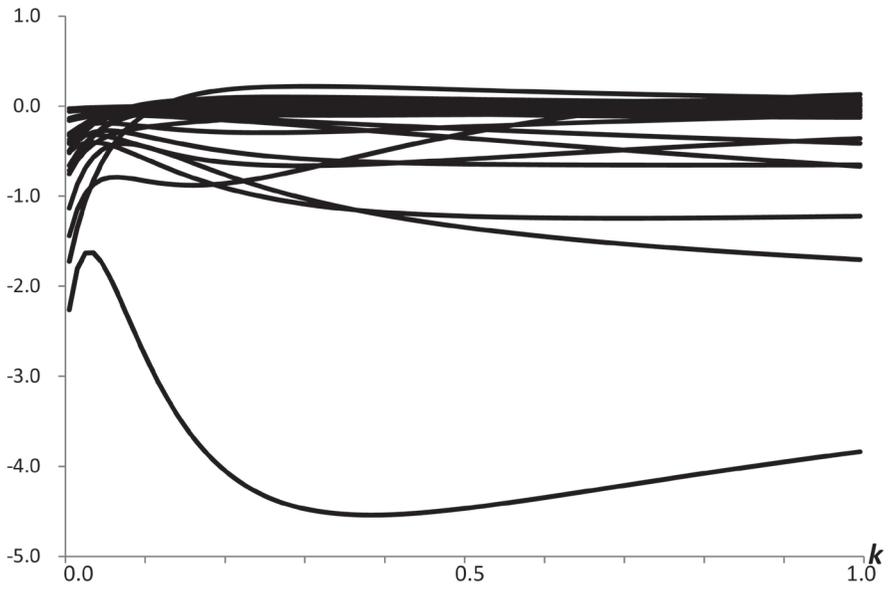


図 C.3 全観測値に関する MI_i^* の挙動

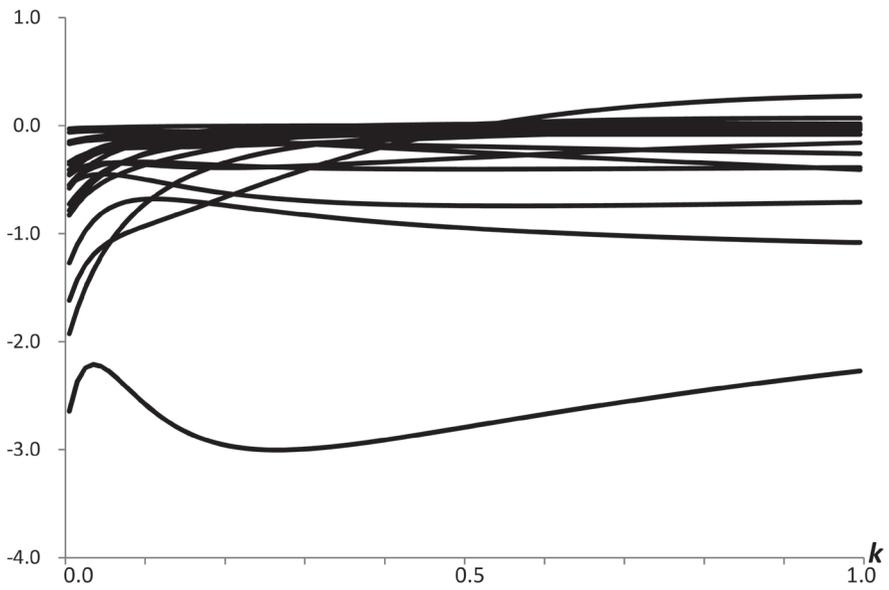


図 C.4 全観測値に関する RMI_i の挙動

参 考 文 献

- 1) Cook, R. D. and Weisberg, S. (1982) , *Residuals and Influence in Regression*, New York: Chapman and Hall.
- 2) Groß, J. (2003) , *Linear Regression*, Berlin: Springer.
- 3) Hoerl, A. E. and Kennard, R. W. (1970) , Ridge regression: biased estimation for nonorthogonal problems, *Technometrics*, 12, 55-67.
- 4) Takeuchi, H. (1994) , Sensitivity analysis with an extension of Cook's distance in ridge regression, *Journal of the Japan Statistical Society*, 24, 221-236.
- 5) 竹内秀一 (2007) , リッジ回帰における尤度距離による影響力評価, 人文自然科学論集, 123号, 3-16.
- 6) 竹内秀一 (2008) , 尤度距離におけるリッジパラメータに関する影響力, 人文自然科学論集, 125号, 57-71.
- 7) 竹内秀一 (2010) , 尤度距離におけるリッジパラメータの推定量の選定方法, 人文自然科学論集, 129号, 133-147.