2010 年度 学部 エコノメトリックス 第 13 回講義メモ (訂正版)

2010 年 11 月 17 日

3.7 多重共線性とその対策 (続)

前回スライド訂正

前回のスライド (pp.8-9) に不正確な記述があったので下記のように訂正する。

幾何的に見てゆくと、二つの説明変数ベクトル $\boldsymbol{x}_1$, $\boldsymbol{x}_2$ の相関と内積の間には、$\boldsymbol{x}_i^* = \boldsymbol{x}_i - \bar{x}_i\boldsymbol{1}$ $(i=1,2)$ とすると、

$$\mathrm{Cor}(x_1, x_2) = \frac{\frac{1}{n}(\boldsymbol{x}_1^*)'(\boldsymbol{x}_2^*)}{\sqrt{\frac{1}{n}(\boldsymbol{x}_1^*)'(\boldsymbol{x}_1^*)}\sqrt{\frac{1}{n}(\boldsymbol{x}_2^*)'(\boldsymbol{x}_2^*)}} = \frac{<\boldsymbol{x}_1^*, \boldsymbol{x}_2^*>}{\|\boldsymbol{x}_1^*\|\|\boldsymbol{x}_2^*\|} = \cos\theta$$

という関係があるので、$\boldsymbol{x}_1^*$ と $\boldsymbol{x}_2^*$ という二つのベクトルのなす角 θ が 0

か π に近いことを意味する。このことから、この二つのベクトルを含む線形空間が不安定であることが分かる。

対策

数学的な対策としては、

(a) リッジ回帰 (ridge regression)
(b) 主成分回帰

があるが、経済分析の分野では余り用いられない。ただ、どちらの対策も対称行列の**固有値** (eigenvalue) の性質を利用したものであるので、簡単に触れておくことにしよう。

まず、行列の固有値の定義を示そう。

定義 3.7.1 正方行列の固有値

いま $\boldsymbol{A}$ を $m \times m$ の行列とする。$\boldsymbol{A}$ の固有値とは、$m \times 1$ のベクトル $\boldsymbol{w} \neq \boldsymbol{0}$ に対して、$\boldsymbol{Aw} = \lambda\boldsymbol{w}$ を満たす λ のことをいう。またその λ に対応する $\boldsymbol{w}$ を固有ベクトルという。通常 $\|\boldsymbol{w}\| = 1$ として定義する。

また、行列の固有値および固有ベクトルについては次の性質がある。

定理 3.7.2

正方行列 $\boldsymbol{A}$ の相異なる固有値 λ_i, λ_j $(i \neq j)$ に対応する固有ベクトル $\boldsymbol{w}_i$, $\boldsymbol{w}_j$ は直交する。すなわち、$\boldsymbol{w}_i'\boldsymbol{w}_j = 0$ である。

定理 3.7.3

実正方行列 $\boldsymbol{A}$ が対称行列であれば、その固有値はすべて実数である。

定理 3.7.2 より、

$$\boldsymbol{w}_j'\boldsymbol{A}\boldsymbol{w}_i = \lambda_i\boldsymbol{w}_j'\boldsymbol{w}_i = \begin{cases} \lambda_i & \text{for } i = j \\ 0 & \text{otherwise} \end{cases}$$

が成立するので、次の定理が得られる。

定理 3.7.4

$\boldsymbol{\Lambda} = diag(\lambda_1, \cdots, \lambda_m)$、$\boldsymbol{W} = (\boldsymbol{w}_1, \cdots, \boldsymbol{w}_m)$ とすると $\boldsymbol{W}'\boldsymbol{A}\boldsymbol{W} = \boldsymbol{\Lambda}$ が得られるので、$\boldsymbol{A} = \boldsymbol{W}\boldsymbol{\Lambda}\boldsymbol{W}'$ と分解できる。(固有値分解)

また、$\boldsymbol{\Lambda}^{-1/2} = diag\left(\frac{1}{\sqrt{\lambda_1}}, \cdots, \frac{1}{\sqrt{\lambda_m}}\right)$ とし、$\boldsymbol{\Gamma} = \boldsymbol{W}\boldsymbol{\Lambda}^{-1/2}$ とすると、$\boldsymbol{\Gamma}'\boldsymbol{A}\boldsymbol{\Gamma} = \boldsymbol{I}$ が成立する。

(a) リッジ回帰

リッジ回帰は、正方行列の階数は非ゼロの固有値の個数に等しい、という性質に注目した手法である。定義 3.7.1 より、

$$\boldsymbol{A}\boldsymbol{w}_i = \lambda_i\boldsymbol{w}_i \qquad (i = 1, \ldots, m)$$

が成立するが、いま適当な $\kappa > 0$ をとれば、

$$(\boldsymbol{A} + \kappa\boldsymbol{I})\boldsymbol{w}_i = (\lambda_i + \kappa)\boldsymbol{w}_i$$

が成立する。したがって、$\boldsymbol{A}$ の固有値が $\lambda_i = 0$ のときでも、$\boldsymbol{A} + \kappa\boldsymbol{I}$ の固有値は正値 κ になる。

リッジ回帰はこの関係を用いて、

$$\boldsymbol{b}^* = (\boldsymbol{X}'\boldsymbol{X} + \kappa\boldsymbol{I})^{-1}\boldsymbol{X}'\boldsymbol{y}$$

を $\boldsymbol{\beta}$ の推定量としたものである。ただ、このリッジ推定量は

$$\begin{aligned}\boldsymbol{b}^* &= (\boldsymbol{X}'\boldsymbol{X} + \kappa\boldsymbol{I})^{-1}[\boldsymbol{X}'(\boldsymbol{X}\boldsymbol{\beta} + \boldsymbol{u}) + \kappa\boldsymbol{\beta} - \kappa\boldsymbol{\beta}] \\ &= \boldsymbol{\beta} + (\boldsymbol{X}'\boldsymbol{X} + \kappa\boldsymbol{I})^{-1}(-\kappa\boldsymbol{\beta} + \boldsymbol{X}'\boldsymbol{u})\end{aligned}$$

になるので、

$$\mathrm{E}(\boldsymbol{b}^* \mid \boldsymbol{X}) = \boldsymbol{\beta} - \kappa(\boldsymbol{X}'\boldsymbol{X} + \kappa\boldsymbol{I})^{-1}\boldsymbol{\beta} \neq \boldsymbol{\beta}$$

となり、不偏推定量ではない。

また、平均 2 乗誤差は、

$$\begin{aligned}\mathrm{MSE}(\boldsymbol{b}^* \mid \boldsymbol{X}) = \sigma^2&(\boldsymbol{X}'\boldsymbol{X} + \kappa\boldsymbol{I})^{-1}\boldsymbol{X}'\boldsymbol{X}(\boldsymbol{X}'\boldsymbol{X} + \kappa\boldsymbol{I})^{-1} \\ &+\kappa^2(\boldsymbol{X}'\boldsymbol{X} + \kappa\boldsymbol{I})^{-1}\boldsymbol{\beta}\boldsymbol{\beta}'(\boldsymbol{X}'\boldsymbol{X} + \kappa\boldsymbol{I})^{-1}\end{aligned}$$

となる。

実際には、κ は平均 2 乗誤差を小さくするように選ぶ。しかし、経済分析では、不偏推定量でないためか、リッジ回帰はあまり用いられない。

(b) 主成分回帰

主成分回帰は、主成分分析を用いて説明変数 $\boldsymbol{X}$ を主成分変数 $\boldsymbol{V}$ に変換し、$\boldsymbol{V}$ を新たな説明変数として回帰分析を行うものである。以下で概略を述べよう。

定数項付の重回帰モデルの推定結果は、

$$\boldsymbol{y} = \boldsymbol{X}\boldsymbol{b} + \hat{\boldsymbol{\epsilon}} = \mathbf{1}b_1 + \boldsymbol{X}_2\boldsymbol{b}_2 + \hat{\boldsymbol{\epsilon}} \tag{3.7.1}$$

で表される。ここで、

$$\boldsymbol{X}_2 = (\boldsymbol{x}_2, \cdots, \boldsymbol{x}_k), \quad \boldsymbol{\beta}_2 = \begin{pmatrix} \beta_2 \\ \vdots \\ \beta_k \end{pmatrix}$$

である。

いま、

$$\bar{\boldsymbol{y}} = \boldsymbol{y} - \mathbf{1}\bar{y}, \;\; \bar{\boldsymbol{x}}_j = \boldsymbol{x}_j - \mathbf{1}\bar{x}_j \; (j = 2, \ldots, k), \; \bar{\boldsymbol{X}}_2 = (\bar{\boldsymbol{x}}_2, \cdots, \bar{\boldsymbol{x}}_k)$$

とすると、(3.7.1) 式は

$$\bar{\boldsymbol{y}} = \bar{\boldsymbol{X}}_2\boldsymbol{b}_2 + residual \tag{3.7.2}$$

と書きなおせる。定理 3.7.4 により、$\bar{\boldsymbol{X}}_2'\bar{\boldsymbol{X}}_2$ を

$$\bar{\boldsymbol{X}}_2'\bar{\boldsymbol{X}}_2 = \boldsymbol{W}\boldsymbol{\Lambda}\boldsymbol{W}'$$

に固有値分解できるから、$(k-1)\times(k-1)$ の直交行列 $\boldsymbol{W}$ を用いて $\bar{\boldsymbol{X}}_2$ を線形変換して、

$$\boldsymbol{V} = \bar{\boldsymbol{X}}_2\boldsymbol{W}$$

を得ることができる。$n\times(k-1)$ 行列 $\boldsymbol{V} = (\boldsymbol{v}_2, \cdots, \boldsymbol{v}_k)$ を構成する列ベクトル $\boldsymbol{v}_j \; (j = 2, \ldots, k)$ には、

$$\boldsymbol{v}_i'\boldsymbol{v}_j = \begin{cases} \lambda_j & \text{for } i = j \\ 0 & \text{for } i \neq j \end{cases}$$

という性質があるので、$\frac{1}{n}\boldsymbol{V}'\boldsymbol{V} = \frac{1}{n}\boldsymbol{\Lambda}$ が成立し、$\boldsymbol{v}_j$ は（算術）平均 0、分散 λ_j/n の変数に変換されており、かつ $\boldsymbol{v}_i$ と $\boldsymbol{v}_j$ は互いに直交している。

$\boldsymbol{\Lambda} = diag(\lambda_2, \cdots, \lambda_k)$ を $\lambda_2 \geq \cdots \geq \lambda_k$ のようにとるとき、$\boldsymbol{v}_j$ $(j = 2, \ldots, k)$ を第 $(j-1)$ 主成分と呼ぶ。

$\bar{\boldsymbol{y}}$ を主成分 $\boldsymbol{V}$ について回帰したモデルを主成分回帰モデルとよぶ。その係数推定量を $\boldsymbol{c}_2$ で表すと、

$$\bar{\boldsymbol{y}} = \boldsymbol{V}\boldsymbol{c}_2 + residual$$

となるが、$\boldsymbol{V}\boldsymbol{c}_2 = \boldsymbol{X}\boldsymbol{W}\boldsymbol{c}_2$ であるので、

$$\boldsymbol{b}_2 = \boldsymbol{W}\boldsymbol{c}_2$$

として求めることができる。