複数のパラメータ（係数）の同時検定について

2001 年 7 月 11 日

1　複数のパラメータ（係数）の同時検定問題

線形回帰モデル $\boldsymbol{y} = \boldsymbol{X}\boldsymbol{\beta} + \boldsymbol{\epsilon}$ を考えることにしよう。（ただし、$\boldsymbol{y}$, $\boldsymbol{\epsilon}$ は $n \times 1$ ベクトル、$\boldsymbol{\beta}$ は $k \times 1$ ベクトル、$\boldsymbol{X}$ は $n \times k$ 行列をあらわし、また $\boldsymbol{\epsilon}$ は古典的仮定をみたしているものとするが、説明の簡単化のために $\boldsymbol{\epsilon} \sim \mathrm{N}(\boldsymbol{0}, \sigma^2\boldsymbol{I})$ と強い仮定をおこう。）

授業で解説したように、一つのパラメータ（係数）を検定するには t 検定を用いるが、複数のパラメータ（係数）を同時に検定するには、別な検定方法を用いる。いま、同時に検定したい問題が次の式で表現できるものとしよう。

$$\begin{cases} H_0: & \boldsymbol{Q}'\boldsymbol{\beta} = \boldsymbol{c} \\ H_1: & \boldsymbol{Q}'\boldsymbol{\beta} \neq \boldsymbol{c} \end{cases}$$

ここで、$\boldsymbol{Q}$ は $m \times k$ 行列であり、$m\,(m \leq k)$ 個の係数間制約があることを表している。

係数 $\boldsymbol{\beta}$ の最小 2 乗推定量 $\hat{\boldsymbol{\beta}}$ の分布は

$$(\hat{\boldsymbol{\beta}} - \boldsymbol{\beta}) \sim \mathrm{N}\left(\boldsymbol{0}, \sigma^2(\boldsymbol{X}'\boldsymbol{X})^{-1}\right)$$

であるから、2 次形式は

$$\frac{(\hat{\boldsymbol{\beta}} - \boldsymbol{\beta})'(\boldsymbol{X}'\boldsymbol{X})(\hat{\boldsymbol{\beta}} - \boldsymbol{\beta})}{\sigma^2} \sim \chi^2(k)$$

さらに、σ^2 を不偏推定量 $\hat{\sigma^2} \equiv \boldsymbol{e}'\boldsymbol{e}/(n-k)$ で置き換え（ちなみに $\boldsymbol{e} \equiv \boldsymbol{y} - \hat{\boldsymbol{y}} = \boldsymbol{y} - \boldsymbol{X}\hat{\boldsymbol{\beta}}$）、$k$ で割ると、

$$\frac{(\hat{\boldsymbol{\beta}} - \boldsymbol{\beta})'(\boldsymbol{X}'\boldsymbol{X})(\hat{\boldsymbol{\beta}} - \boldsymbol{\beta})/k}{\hat{\sigma^2}} \sim F(k, n-k)$$

という結果がえられる。この結果を使うと、$\boldsymbol{Q}'\hat{\boldsymbol{\beta}}$ の分布は

$$\boldsymbol{Q}'(\hat{\boldsymbol{\beta}} - \boldsymbol{\beta}) = (\boldsymbol{Q}'\hat{\boldsymbol{\beta}} - \boldsymbol{c}) \sim \mathrm{N}\left(\boldsymbol{0}, \sigma^2\boldsymbol{Q}'(\boldsymbol{X}'\boldsymbol{X})^{-1}\boldsymbol{Q}\right)$$

であるので、

$$f = \frac{(\boldsymbol{Q}'\hat{\boldsymbol{\beta}} - \boldsymbol{c})'(\boldsymbol{Q}'(\boldsymbol{X}'\boldsymbol{X})^{-1}\boldsymbol{Q})^{-1}(\boldsymbol{Q}'\hat{\boldsymbol{\beta}} - \boldsymbol{c})/k}{\hat{\sigma^2}} \sim F(k, n-k) \tag{1}$$

これから F 検定によって仮説を検定できることがわかる。(1) 式の計算は大変なように見えるが、実は通常の最小 2 乗推定と、$\boldsymbol{Q}'\boldsymbol{\beta} = \boldsymbol{c}$ という制約があるときの最小 2 乗推定という、２つの推定を行い、それぞれの推定式の残差 2 乗和を用いて、簡単に求めることができる。そこで次節では、係数間に制約があるときの最小 2 乗推定量について考察することにする。

2　制約付最小２乗法

制約条件があるときの最小 2 乗推定量を考えることにしよう。このときの最小化問題は、

$$\max_{\boldsymbol{\beta}} \boldsymbol{\epsilon}\boldsymbol{\epsilon}' \quad \text{subject to } (\boldsymbol{Q}'\boldsymbol{\beta} - \boldsymbol{c}) = \boldsymbol{0}$$

となるので、ラグランジェ乗数法を使って解くことができる。いま $S = \boldsymbol{\epsilon}'\boldsymbol{\epsilon}$ とすると、目的関数

$$S^* = S + \boldsymbol{\lambda}'(\boldsymbol{Q}'\boldsymbol{\beta} - \boldsymbol{c})$$

が最小となる $\boldsymbol{\beta}$ を求めることになる。最小化のための 1 階の条件より、

$$\frac{\partial S^*}{\partial \boldsymbol{\beta}} = \frac{\partial S}{\partial \boldsymbol{\beta}} + \boldsymbol{Q}\boldsymbol{\lambda} = \boldsymbol{0}$$

ところで、$\frac{\partial S}{\partial \boldsymbol{\beta}} = -2\boldsymbol{X}'(\boldsymbol{y} - \boldsymbol{X}\boldsymbol{\beta})$ であるから、

$$-2\boldsymbol{X}'\boldsymbol{y} - 2(\boldsymbol{X}'\boldsymbol{X})\boldsymbol{\beta} + \boldsymbol{Q}\boldsymbol{\lambda} = \boldsymbol{0} \tag{2}$$

上式の左から $\frac{1}{2}(\boldsymbol{X}'\boldsymbol{X})^{-1}$ をかけると、求める制約付最小 2 乗推定量 $\hat{\boldsymbol{\beta}^+}$ は、

$$\hat{\boldsymbol{\beta}^+} = \hat{\boldsymbol{\beta}} - \frac{1}{2}(\boldsymbol{X}'\boldsymbol{X})^{-1}\boldsymbol{Q}\boldsymbol{\lambda} \tag{3}$$

となる。ところで、(2) の左から $\boldsymbol{Q}'(\boldsymbol{X}'\boldsymbol{X})^{-1}$ をかけて整理すると、

$$\boldsymbol{\lambda} = 2[\boldsymbol{Q}'(\boldsymbol{X}'\boldsymbol{X})^{-1}\boldsymbol{Q}]^{-1}\boldsymbol{Q}'(\hat{\boldsymbol{\beta}} - \boldsymbol{\beta}) \tag{4}$$

となるので、(4) を (3) に代入すると、

$$\hat{\boldsymbol{\beta}^+} = \hat{\boldsymbol{\beta}} - (\boldsymbol{X}'\boldsymbol{X})^{-1}\boldsymbol{Q}[\boldsymbol{Q}'(\boldsymbol{X}'\boldsymbol{X})^{-1}\boldsymbol{Q}]^{-1}\boldsymbol{Q}'(\hat{\boldsymbol{\beta}} - \boldsymbol{\beta}) \tag{5}$$

となる。両辺から $\boldsymbol{\beta}$ を引くと、

$$\begin{aligned}
\hat{\boldsymbol{\beta}^+} - \boldsymbol{\beta} &= (\hat{\boldsymbol{\beta}} - \boldsymbol{\beta}) - (\boldsymbol{X}'\boldsymbol{X})^{-1}\boldsymbol{Q}[\boldsymbol{Q}'(\boldsymbol{X}'\boldsymbol{X})^{-1}\boldsymbol{Q}]^{-1}\boldsymbol{Q}'(\hat{\boldsymbol{\beta}} - \boldsymbol{\beta}) \\
&= [\boldsymbol{I} - (\boldsymbol{X}'\boldsymbol{X})^{-1}\boldsymbol{Q}[\boldsymbol{Q}'(\boldsymbol{X}'\boldsymbol{X})^{-1}\boldsymbol{Q}]^{-1}\boldsymbol{Q}'](\hat{\boldsymbol{\beta}} - \boldsymbol{\beta})
\end{aligned}$$

となるから、$\mathrm{E}(\hat{\boldsymbol{\beta}^+}) = \boldsymbol{\beta}$。よって不偏推定量である。また、この推定量の分散は、

$$\begin{aligned}
\mathrm{Var}(\hat{\boldsymbol{\beta}^+}) &= \sigma^2[\boldsymbol{I} - (\boldsymbol{X}'\boldsymbol{X})^{-1}\boldsymbol{Q}[\boldsymbol{Q}'(\boldsymbol{X}'\boldsymbol{X})^{-1}\boldsymbol{Q}]^{-1}\boldsymbol{Q}'](\boldsymbol{X}'\boldsymbol{X})^{-1} \\
&\quad \times[\boldsymbol{I} - \boldsymbol{Q}[\boldsymbol{Q}'(\boldsymbol{X}'\boldsymbol{X})^{-1}\boldsymbol{Q}]^{-1}\boldsymbol{Q}'(\boldsymbol{X}'\boldsymbol{X})^{-1}] \\
&= \sigma^2[(\boldsymbol{X}'\boldsymbol{X})^{-1} - (\boldsymbol{X}'\boldsymbol{X})^{-1}\boldsymbol{Q}[\boldsymbol{Q}'(\boldsymbol{X}'\boldsymbol{X})^{-1}\boldsymbol{Q}]^{-1}\boldsymbol{Q}'(\boldsymbol{X}'\boldsymbol{X})^{-1}]
\end{aligned}$$

であることがわかる。

次節では、本節で得られた結果を用いて、(1) 式を導出することにしよう。

3 f の導出

制約付き最小 2 乗法で推定した回帰式の残差 $\boldsymbol{e}^+$ は、(5) 式から

$$\begin{aligned} \boldsymbol{e}^+ &\equiv \boldsymbol{y} - \boldsymbol{X}\hat{\boldsymbol{\beta}}^+ \\ &= (\boldsymbol{y} - \boldsymbol{X}\hat{\boldsymbol{\beta}}) - \boldsymbol{X}(\boldsymbol{X}'\boldsymbol{X})^{-1}\boldsymbol{Q}[\boldsymbol{Q}'(\boldsymbol{X}'\boldsymbol{X})^{-1}\boldsymbol{Q}]^{-1}\boldsymbol{Q}'(\hat{\boldsymbol{\beta}} - \boldsymbol{\beta}) \\ &= \boldsymbol{e} - \boldsymbol{X}(\boldsymbol{X}'\boldsymbol{X})^{-1}\boldsymbol{Q}[\boldsymbol{Q}'(\boldsymbol{X}'\boldsymbol{X})^{-1}\boldsymbol{Q}]^{-1}\boldsymbol{Q}'(\hat{\boldsymbol{\beta}} - \boldsymbol{\beta}) \end{aligned} \tag{6}$$

で表すことができる。よって、その残差 2 乗和は、

$$\begin{aligned} \boldsymbol{e}^{+\prime}\boldsymbol{e}^+ &= \boldsymbol{e}'\boldsymbol{e} - \boldsymbol{e}'\boldsymbol{X}(\boldsymbol{X}'\boldsymbol{X})^{-1}\boldsymbol{Q}[\boldsymbol{Q}'(\boldsymbol{X}'\boldsymbol{X})^{-1}\boldsymbol{Q}]^{-1}\boldsymbol{Q}'(\hat{\boldsymbol{\beta}} - \boldsymbol{\beta}) \\ &\quad -(\hat{\boldsymbol{\beta}} - \boldsymbol{\beta})'\boldsymbol{Q}[\boldsymbol{Q}'(\boldsymbol{X}'\boldsymbol{X})^{-1}\boldsymbol{Q}]^{-1}\boldsymbol{Q}'(\boldsymbol{X}'\boldsymbol{X})^{-1}\boldsymbol{X}'\boldsymbol{e} \\ &\quad +(\hat{\boldsymbol{\beta}} - \boldsymbol{\beta})'\boldsymbol{Q}[\boldsymbol{Q}'(\boldsymbol{X}'\boldsymbol{X})^{-1}\boldsymbol{Q}]^{-1}\boldsymbol{Q}'(\boldsymbol{X}'\boldsymbol{X})^{-1}\boldsymbol{X}'\boldsymbol{X}(\boldsymbol{X}'\boldsymbol{X})^{-1}\boldsymbol{Q}[\boldsymbol{Q}'(\boldsymbol{X}'\boldsymbol{X})^{-1}\boldsymbol{Q}]^{-1} \\ &\quad \times\boldsymbol{Q}'(\hat{\boldsymbol{\beta}} - \boldsymbol{\beta}) \\ &= \boldsymbol{e}'\boldsymbol{e} + (\hat{\boldsymbol{\beta}} - \boldsymbol{\beta})'\boldsymbol{Q}[\boldsymbol{Q}'(\boldsymbol{X}'\boldsymbol{X})^{-1}\boldsymbol{Q}]^{-1}\boldsymbol{Q}'(\hat{\boldsymbol{\beta}} - \boldsymbol{\beta}) \end{aligned} \tag{7}$$

となる（なぜなら、$\boldsymbol{X}'\boldsymbol{e} = \boldsymbol{0}$)。(3) 式を整理すると、

$$\boldsymbol{e}^{+\prime}\boldsymbol{e}^+ - \boldsymbol{e}'\boldsymbol{e} = (\boldsymbol{Q}'\hat{\boldsymbol{\beta}} - \boldsymbol{c})'[\boldsymbol{Q}'(\boldsymbol{X}'\boldsymbol{X})^{-1}\boldsymbol{Q}]^{-1}(\boldsymbol{Q}'\hat{\boldsymbol{\beta}} - \boldsymbol{c}) \tag{8}$$

が得られるので、(1) 式に (8) 式を代入すると、

$$f = \frac{(\boldsymbol{e}^{+\prime}\boldsymbol{e}^+ - \boldsymbol{e}'\boldsymbol{e})/k}{\boldsymbol{e}'\boldsymbol{e}/(n-k)} \tag{9}$$

のように、通常の最小 2 乗推定と、制約付き最小 2 乗推定の、２つの推定式の残差 2 乗和を用いても f 統計量を求められることが確認できる。

4 尤度比検定との関係

前節までに考察した f 検定統計量は、漸近的には（＝標本数 n が大きくなると）尤度比検定統計量と一致するという性質がある。授業では、尤度について殆ど触れていないので、結果だけを簡単に説明することにしよう。

$\boldsymbol{\epsilon} \sim \mathrm{N}(\boldsymbol{0}, \sigma^2\boldsymbol{I})$ という仮定のもとで、通常の最小 2 乗推定量の尤度は、

$$L(\hat{\boldsymbol{\beta}}|\boldsymbol{X}, \boldsymbol{y}) = (2\pi)^{-n/2}(\sigma^2)^{-n/2}\exp\left(-\frac{(\boldsymbol{y} - \boldsymbol{X}\hat{\boldsymbol{\beta}})'(\boldsymbol{y} - \boldsymbol{X}\hat{\boldsymbol{\beta}})}{2\sigma^2}\right)$$

で、また制約付き最小 2 乗推定量の尤度は、

$$L(\hat{\boldsymbol{\beta}}^+|\boldsymbol{X}, \boldsymbol{y}) = 2\pi^{-n/2}(\sigma^2)^{-n/2}\exp\left(-\frac{(\boldsymbol{y} - \boldsymbol{X}\hat{\boldsymbol{\beta}}^+)'(\boldsymbol{y} - \boldsymbol{X}\hat{\boldsymbol{\beta}}^+)}{2\sigma^2}\right)$$

で表される。対数尤度比を 2 倍したものは帰無仮説 H_0 の下で

$$\begin{aligned} 2\log\left(\frac{L(\hat{\boldsymbol{\beta}}|\boldsymbol{X},\boldsymbol{y})}{L(\hat{\boldsymbol{\beta}}^{+}|\boldsymbol{X},\boldsymbol{y})}\right) &= 2\left[\left(-\frac{\boldsymbol{e}'\boldsymbol{e}}{2\sigma^2}\right)-\left(-\frac{\boldsymbol{e}^{+\prime}\boldsymbol{e}^{+}}{2\sigma^2}\right)\right] \\ &= \frac{\boldsymbol{e}^{+\prime}\boldsymbol{e}^{+}-\boldsymbol{e}'\boldsymbol{e}}{\sigma^2}\sim\chi^2(k) \end{aligned} \tag{10}$$

にしたがう。この結果を用いて検定を行なうのが、尤度比検定である。

(1) 式や (9) 式で用いる σ^2 の推定量 $\hat{\sigma^2}=\boldsymbol{e}'\boldsymbol{e}/(n-k)$ は一致推定量でもあるから、標本数が大きいとき、f 検定統計量（(1)、(9) 式）に基づく検定は、(10) 式の尤度比検定と一致することがわかる。