種々の回帰推定量の TMSE 比較

地　道　正　行

I　はじめに

説明変数間に多重共線性が存在する場合に回帰係数の推定をする際、通常最小自乗（Ordinary Least Squeares ; OLS）推定量が種々の問題を引起こすことは多くの論文で指摘されてきた。(例えば、Gunst（1983）を参照。) そこで、このOLS推定量に代るものとして古くはHoerl & Kennard（1970）によって提案された通常ridge回帰（Ordinary Ridge Regression ; ORR）推定量とKendall（1957）やMarquardt（1970）によって提案された主成分回帰（Principal Component Regression ; PCR）推定量がある。

これら2つの推定量がOLS推定量の代りに使用される理論的な根拠としては、ORR推定量とPCR推定量の総平均2乗誤差（Total Mean Squared Error ; TMSE）がOLS推定量のTMSEよりも、ある種の条件のもとで小さくなるということである。(Hoerl & Kennard（1970), Theorem 4.3, Marquardt（1970), Theorem 15を参照せよ。)

また最近の話題としては、Baye & Parker（1984）によって提案された *r–k* クラス（*r–k* class）推定量がある。この推定量は、PCR推定量にridge回帰のテクニックを適用したものと見ることができ、PCR推定量とORR推定量を結合したものである。またこの推定量は、特別な場合としてOLS, ORR, PCRの各推定量を含む1つのクラスを形成することが示され、この意味で今まで多重共線性の問題で扱われてきた推定量の中で最も一般的なものであることが分か

る。

では「この推定量の存在価値はあるのか?」という疑問が起こるが、それに対しては提案者であるBaye & Parkerがある種の条件のもとでr–kクラス推定量のTMSEがPCR推定量のTMSEよりも小さくなるということを示している。(Baye & Parker (1984), Theorem 2を参照せよ。) またNomura & Ohkubo (1985) は、Baye & ParkerがPCR推定量とのTMSE比較だけにとどまっていたことを指摘し、OLS推定量、ORR推定量とのTMSE比較も行った。(Nomura & Ohkubo (1985) の一連の定理を参照せよ。)

しかしながら、r–kクラス推定量が他の推定量に対して同時にTMSE優越(TMSEが小さくなる)ということは考えられていない。そこで以下の節で、まずこれらの推定量の定義を与え、そのTMSEを求める。次に過去の文献においてOLS, ORR, PCR, r–kクラス推定量のTMSEの比較に関して重要であると思われる定理を述べる。さらに、条件を整理することによって、それらの定理を統一的に扱える定理を再構成し、最後にr–kクラス推定量が他の推定量よりもTMSEの意味でよくなることを表す定理を述べることにする。

II 種々の回帰推定量の定義とTMSE

以下モデルを考えよう:

$$\boldsymbol{Y}=\boldsymbol{X}\beta+\varepsilon$$

ここで、

1. $\boldsymbol{X}$は、$n\times p$の説明変数行列とし、$\mathrm{rank}(\boldsymbol{X})=p(\leq n)$ とする。
2. εは、$n\times 1$の誤差ベクトルとし、$\mathrm{E}(\varepsilon)=0$と$\mathrm{Var}(\varepsilon)=\sigma^2\boldsymbol{I}_n$を仮定する。

本稿で考察される回帰母数βの推定量は以下のようなものである:

OLS推定量 $\hat{\beta}\equiv(\boldsymbol{X}'\boldsymbol{X})^{-1}\boldsymbol{X}'\boldsymbol{Y}$ (1)

ORR推定量 $\hat{\beta}(\cdot,k)\equiv(\boldsymbol{X}'\boldsymbol{X}+k\boldsymbol{I}_p)^{-1}\boldsymbol{X}'\boldsymbol{Y}$ (2)

PCR推定量 $\hat{\beta}(r,\cdot)\equiv\boldsymbol{T}_r\boldsymbol{\Lambda}_r^{-1}\boldsymbol{T}_r'\boldsymbol{X}'\boldsymbol{Y}$ (3)

r–kクラス推定量 $\hat{\beta}(r,k)\equiv\boldsymbol{T}_r\boldsymbol{\Lambda}_r(k)^{-1}\boldsymbol{T}_r'\boldsymbol{X}'\boldsymbol{Y}$ (4)

ここで、PCR, r–kクラス推定量におけるrは主成分数を表わしており、ORR,

r-k クラス推定量における k は ridge 定数（ridge coefficient）とよばれる。

また、$\boldsymbol{\Lambda}_r(k)\equiv\boldsymbol{\Lambda}_r+k\boldsymbol{I}_r$ とし、行列 $\boldsymbol{T}=(\boldsymbol{t}_1,\cdots,\boldsymbol{t}_p)$ は、$\boldsymbol{T}'\boldsymbol{X}'\boldsymbol{X}\boldsymbol{T}=\boldsymbol{\Lambda}\equiv\mathrm{diag}(\lambda_1,\cdots,\lambda_p)$ とするような p 次の直交行列とする。また $\boldsymbol{T}_r\equiv(\boldsymbol{t}_1,\cdots,\boldsymbol{t}_r)$、$\boldsymbol{T}_{p\backslash r}\equiv(\boldsymbol{t}_{r+1},\cdots,\boldsymbol{t}_p)$ とし、$\boldsymbol{T}_r'\boldsymbol{X}'\boldsymbol{X}\boldsymbol{T}_r=\boldsymbol{\Lambda}_r\equiv\mathrm{diag}(\lambda_1,\cdots,\lambda_r)$、$\boldsymbol{T}_{p\backslash r}'\boldsymbol{X}'\boldsymbol{X}\boldsymbol{T}_{p\backslash r}=\boldsymbol{\Lambda}_{p\backslash r}\equiv\mathrm{diag}(\lambda_{r+1},\cdots,\lambda_p)$ と書くことにする。

このように定義された推定量において特に注意すべき点は、r-k クラス推定量が、他の3つの推定量を特別な場合として含んでいるということである。すなわち：

$$\hat\beta(p,0)=\hat\beta=(\boldsymbol{X}'\boldsymbol{X})^{-1}\boldsymbol{X}'\boldsymbol{Y}\ ;\ \text{OLS推定量、} \tag{5}$$

$$\hat\beta(p,k)=\hat\beta(\cdot,k)=(\boldsymbol{X}'\boldsymbol{X}+k\boldsymbol{I}_p)^{-1}\boldsymbol{X}'\boldsymbol{Y}\ ;\ \text{ORR推定量、} \tag{6}$$

$$\hat\beta(r,0)=\hat\beta(r,\cdot)=\boldsymbol{T}_r\boldsymbol{\Lambda}_r^{-1}\boldsymbol{T}_r'\boldsymbol{X}'\boldsymbol{Y}\ ;\ \text{PCR推定量.} \tag{7}$$

では次に各推定量の TMSE をもとめよう。

$$\begin{aligned}\mathrm{TMSE}(\hat\beta)&=\mathrm{E}\{(\hat\beta-\beta)'(\hat\beta-\beta)\}\\&=\mathrm{trVar}(\hat\beta)\\&=\sigma^2\mathrm{tr}(\boldsymbol{X}'\boldsymbol{X})^{-1}\\&=\sigma^2\ \mathrm{tr}\ \boldsymbol{T}\boldsymbol{\Lambda}^{-1}\boldsymbol{T}'\\&=\sigma^2\ \mathrm{tr}\ \boldsymbol{\Lambda}^{-1}\\&=\sigma^2\sum_{i\in N_p}\frac{1}{\lambda_i}\end{aligned}\tag{8}$$

$$\begin{aligned}\mathrm{TMSE}(\hat\beta(\cdot,k))&=\mathrm{E}\{(\hat\beta(\cdot,k)-\beta)'(\hat\beta(\cdot,k)-\beta)\}\\&=\mathrm{trVar}(\hat\beta(\cdot,k))+\|\mathrm{bias}(\hat\beta(\cdot,k))\|^2\\&=\sigma^2\ \mathrm{tr}(\boldsymbol{X}'\boldsymbol{X}+k\boldsymbol{I}_p)^{-1}\boldsymbol{X}'\boldsymbol{X}(\boldsymbol{X}'\boldsymbol{X}+k\boldsymbol{I}_p)^{-1}\\&\quad+\|\{(\boldsymbol{X}'\boldsymbol{X}+k\boldsymbol{I}_p)^{-1}\boldsymbol{X}'\boldsymbol{X}-\boldsymbol{I}_p\}\beta\|^2\\&=\sigma^2\ \mathrm{tr}\ \boldsymbol{T}(\boldsymbol{\Lambda}+k\boldsymbol{I}_p)^{-1}\boldsymbol{\Lambda}(\boldsymbol{\Lambda}+k\boldsymbol{I}_p)^{-1}\boldsymbol{T}'\\&\quad+\|\boldsymbol{T}\{(\boldsymbol{\Lambda}+k\boldsymbol{I}_p)^{-1}-\boldsymbol{I}_p\}\boldsymbol{T}'\beta\|^2\\&=\sigma^2\ \mathrm{tr}(\boldsymbol{\Lambda}+k\boldsymbol{I}_p)^{-1}\boldsymbol{\Lambda}(\boldsymbol{\Lambda}+k\boldsymbol{I}_p)^{-1}\\&\quad+\|\boldsymbol{T}\{(\boldsymbol{\Lambda}+k\boldsymbol{I}_p)^{-1}\boldsymbol{\Lambda}^{-1}-\boldsymbol{I}_p\}\boldsymbol{\alpha}\|^2\\&=\sigma^2\sum_{i\in N_p}\frac{\lambda_i}{(\lambda_i+k)^2}+\sum_{i\in N_p}\frac{k^2\alpha_i^2}{(\lambda_i+k)^2}\end{aligned}\tag{9}$$

$$
\begin{aligned}
\mathrm{TMSE}(\hat{\beta}(r,\bullet)) &= \mathrm{E}\{(\hat{\beta}(r,\bullet)-\beta)'(\hat{\beta}(r,\bullet)-\beta)\} \\
&= \mathrm{trVar}(\hat{\beta}(r,\bullet)) + \|\mathrm{bias}(\hat{\beta}r,\bullet))\|^2 \\
&= \sigma^2\,\mathrm{tr}\,\boldsymbol{T}_r\boldsymbol{\Lambda}_r^{-1}\boldsymbol{T}_r' + \|-(\boldsymbol{I}_p-\boldsymbol{T}_r\boldsymbol{T}_r')\beta\|^2 \\
&= \sigma^2\,\mathrm{tr}\,\boldsymbol{\Lambda}_r^{-1} + \beta'(\boldsymbol{I}_p-\boldsymbol{T}_r\boldsymbol{T}_r')\beta \\
&= \sigma^2\,\mathrm{tr}\,\boldsymbol{\Lambda}_r^{-1} + \beta'\beta - \beta'\boldsymbol{T}_r\boldsymbol{T}_r'\beta \\
&= \sigma^2\,\mathrm{tr}\,\boldsymbol{\Lambda}_r^{-1} + \boldsymbol{\alpha}'\boldsymbol{\alpha} - \boldsymbol{\alpha}'_r\boldsymbol{\alpha}_r \\
&= \sigma^2\sum_{i\in N_r}\frac{1}{\lambda_i} + \sum_{i\in N_{p\backslash r}}\alpha_i^2 \qquad (10)
\end{aligned}
$$

$$
\begin{aligned}
\mathrm{TMSE}(\hat{\beta}(r,k)) &= \mathrm{E}\{(\hat{\beta}(r,k)-\beta)'(\hat{\beta}(r,k)-\beta)\} \\
&= \mathrm{trVar}(\hat{\beta}(r,k)) + \|\mathrm{bins}(\hat{\beta}(r,k))\|^2 \\
&= \sigma^2\,\mathrm{tr}\,\boldsymbol{T}_r\boldsymbol{\Lambda}_r(k)^{-1}\boldsymbol{T}_r'\boldsymbol{X}'\boldsymbol{X}\boldsymbol{T}_r\boldsymbol{\Lambda}_r(k)^{-1}\boldsymbol{T}_r' + \|(\boldsymbol{T}_r\boldsymbol{\Lambda}_r(k)^{-1}\boldsymbol{\Lambda}_r\boldsymbol{T}_r' - \boldsymbol{I}_p)\beta\|^2 \\
&= \sigma^2\,\mathrm{tr}\,\boldsymbol{\Lambda}_r(k)^{-1}\boldsymbol{\Lambda}_r\boldsymbol{\Lambda}_r(k)^{-1} \\
&\quad + \beta'(\boldsymbol{T}_r\boldsymbol{\Lambda}_r\boldsymbol{\Lambda}_r(k)^{-1}\boldsymbol{T}_r' - \boldsymbol{I}_p)(\boldsymbol{T}_r\boldsymbol{\Lambda}_r(k)^{-1}\boldsymbol{\Lambda}_r\boldsymbol{T}_r' - \boldsymbol{I}_p)\beta \\
&= \sigma^2\,\mathrm{tr}\,\boldsymbol{\Lambda}_r(k)^{-1}\boldsymbol{\Lambda}_r\boldsymbol{\Lambda}_r(k)^{-1} \\
&\quad + \beta'\boldsymbol{T}_r\boldsymbol{\Lambda}_r\boldsymbol{\Lambda}_r(k)^{-2}\boldsymbol{\Lambda}_r\boldsymbol{T}_r'\beta - 2\beta'\boldsymbol{T}_r\boldsymbol{\Lambda}_r\boldsymbol{\Lambda}_r(k)^{-1}\boldsymbol{T}_r'\beta + \beta'\beta \\
&= \sigma^2\mathrm{tr}\,\boldsymbol{\Lambda}_r(k)^{-1}\boldsymbol{\Lambda}_r\boldsymbol{\Lambda}_r(k)^{-1} \\
&\quad + \boldsymbol{\alpha}_r'\boldsymbol{\Lambda}_r\boldsymbol{\Lambda}_r(k)^{-2}\boldsymbol{\Lambda}_r\boldsymbol{\alpha}_r - 2\boldsymbol{\alpha}_r'\boldsymbol{\Lambda}_r\boldsymbol{\Lambda}_r(k)^{-1}\boldsymbol{\alpha}_r + \boldsymbol{\alpha}_r'\boldsymbol{\alpha}_r + \boldsymbol{\alpha}_{p\backslash r}'\boldsymbol{\alpha}_{p\backslash r}' \\
&= \sigma^2\sum_{i\in N_r}\frac{\lambda_i}{(\lambda_i+k)^2} + \left\{\sum_{i\in N_r}\frac{k^2\alpha_i^2}{(\lambda_i+k)^2} + \sum_{i\in N_{p\backslash r}}\alpha_i^2\right\} \qquad (11)
\end{aligned}
$$

ここで $N_r\equiv\{1,\cdots,r\}$, $N_{p\backslash r}\equiv\{r+1,\cdots,p\}$, $r=1,\cdots,p$ とし、$N_{p\backslash p}\equiv\emptyset$ と定義する。また、

$$
\boldsymbol{\alpha}=(\alpha_1,\cdots,\alpha_r,\alpha_{r+1},\cdots,\alpha_p)' = \begin{bmatrix}\boldsymbol{\alpha}_r\\ \boldsymbol{\alpha}_{p\backslash r}\end{bmatrix} \equiv \begin{bmatrix}\boldsymbol{T}_r'\beta\\ \boldsymbol{T}_{p\backslash r}'\beta\end{bmatrix} = \begin{bmatrix}\boldsymbol{T}_r'\\ \boldsymbol{T}_{p\backslash r}'\end{bmatrix}\beta = \boldsymbol{T}'\beta
$$

とする。

Ⅲ　TMSE の差に関する定理と系

この節では、前節で述べた回帰推定量間の TMSE の比較に関する定理を述べる。その際、関連する論文から引用された定理は明示的にそれらの著者名を

書くことにする。

Ⅲ. 1 ORR 推定量と OLS 推定量の TMSE 比較

以下の定理は、Hoerl & Kennard（1970）によって与えられた。

Theorem 1（Hoerl & Kennard）

$$\exists k(>0)\ s.t.\ \mathrm{TMSE}(\hat{\beta}(\cdot, k)) < \mathrm{TMSE}(\hat{\beta})$$

上の定理の証明については、後に述べる Theorem 3 において $r=p$ とおくことによってえられるのでここでは省略する[1]。

この定理は、ORR 推定量が OLS 推定量よりも TMSE の意味でよくなる k の存在を示しているが、k の範囲が明示的には与えられていない。そこで、以下の定理を与える：

Theorem 2 任意の $i \in N_p$ に対して $\alpha_i^2 - \dfrac{\sigma^2}{\lambda_i} \leq 0$ である時、

(i) $\mathrm{TMSE}(\hat{\beta}(\cdot, k)) < \mathrm{TMSE}(\hat{\beta})$ *for* $k>0$

$\alpha_i^2 - \dfrac{\sigma^2}{\lambda_i} > 0$ となる $i \in N_p$ が存在する時、

(ii) $\mathrm{TMSE}(\hat{\beta}(\cdot, k)) < \mathrm{TMSE}(\hat{\beta})$ *for* $0<k<K_1$

ここで、

$$K_1 \equiv \frac{2\sigma^2}{\max_{i \in N_p}\left(\alpha_i^2 - \dfrac{\sigma^2}{\lambda_i}\right)} \tag{12}$$

である。

Proof：TMSE の差を求めると、

$$\begin{aligned}\mathrm{TMSE}(\hat{\beta}(\cdot, k)) - \mathrm{TMSE}(\hat{\beta}) &= \sum_{i \in N_p} \frac{\sigma^2 \lambda_i}{(\lambda_i + k)^2} + \sum_{i \in N_p} \frac{\alpha_i^2 k^2}{(\lambda_i + k)^2} - \sigma^2 \sum_{i \in N_p} \frac{1}{\lambda_i} \\ &= \sum_{i \in N_p} \left\{ \frac{\sigma^2 \lambda_i}{(\lambda_i + k)^2} - \frac{\sigma^2}{\lambda_i} + \frac{\alpha_i^2 k^2}{(\lambda_i + k)^2} \right\}\end{aligned}$$

1）この定理の証明の詳細は、Hoerl & Kennard（1970）を参照。

$$= \sum_{i \in N_p} \frac{1}{(\lambda_i + k)^2} \left\{ \sigma^2 \lambda_i - \frac{(\lambda_i + k)^2 \sigma^2}{\lambda_i} + \alpha_i^2 k^2 \right\}$$

$$= \sum_{i \in N_p} \frac{k}{(\lambda_i + k)^2} \left\{ \left(\sigma_i^2 - \frac{\sigma^2}{\lambda_i} \right) k - 2\sigma^2 \right\} \tag{13}$$

となる。

$\boxed{\text{(i)の証明}}$

仮定より $\sigma_i^2 - \dfrac{\sigma^2}{\lambda_i} \leq 0$ が任意の $i \in N_p$ に対して成り立つので、

$$\frac{k}{(\lambda_i + k)^2} \left\{ \left(\sigma_i^2 - \frac{\sigma^2}{\lambda_i} \right) k - 2\sigma^2 \right\} < 0 \quad \text{for} \quad k > 0 \tag{14}$$

となり、これより(13)式は $k>0$ で負となることから(i)が示された。

$\boxed{\text{(ii)の証明}}$

$\sigma_i^2 - \dfrac{\sigma^2}{\lambda_i} > 0$ となる $i \in N_p$ に対して

$$\frac{k}{(\lambda_i + k)^2} \left\{ \left(\sigma_i^2 - \frac{\sigma^2}{\lambda_i} \right) k - 2\sigma^2 \right\} < 0 \quad \text{for} \quad 0 < k < \frac{2\sigma^2}{\alpha_i^2 - \dfrac{\sigma^2}{\lambda_i}}$$

が成り立ち、$\alpha_i^2 - \dfrac{\sigma^2}{\lambda_i} \leq 0$ となる i に対しては(14)式が成り立つ。

これらのことから $0 < k < \dfrac{2\sigma^2}{\max_{i \in N_r} \left(\alpha_i^2 - \dfrac{\sigma^2}{\lambda_i} \right)}$ という範囲では(13)式が負となる。よって(ii)が示された。□

さらに、$\sigma_i^2 - \dfrac{\sigma^2}{\lambda_i}, (i \in N_p)$ の符号によらず以下の系が成り立つ：

Corollary 1

$$\mathrm{TMSE}(\hat{\beta}(\cdot, k)) < \mathrm{TMSE}(\hat{\beta}) \quad \text{for} \quad 0 < k \leq K_2$$

ここで、

$$K_2 \equiv \frac{2\sigma^2}{\|\beta\|^2} \tag{15}$$

である。

Proof : $\alpha_i^2-\dfrac{\sigma^2}{\lambda_i}>0$ となる i （$\in N_p$）が存在する時、

$$\frac{2\sigma^2}{\sum_{i\in N_p}\alpha_i^2}<\frac{2\sigma^2}{\max_{i\in N_p}\left(\alpha_i^2-\dfrac{\sigma^2}{\lambda_i}\right)}$$

が成り立つ。このことから Theorem 2(ii)を使うと、

$$\mathrm{TMSE}(\hat{\beta}(\bullet,k))<\mathrm{TMSE}(\hat{\beta})\quad \text{for}\quad 0<k\leq\frac{2\sigma^2}{\sum_{i\in N_p}\alpha_i^2}=\frac{2\sigma^2}{||\beta||^2}$$

がいえる。また任意の i に対して $\alpha_i^2-\dfrac{\sigma^2}{\lambda_i}\leq 0$ が成り立っている時、Theorem 2(i)より、

$$\mathrm{TMSE}(\hat{\beta}(\bullet,k))<\mathrm{TMSE}(\hat{\beta})\quad \text{for}\quad k>0$$

がいえるので、結局

$$0<k\leq\frac{2\sigma^2}{||\beta||^2}$$

に対して $\mathrm{TMSE}(\hat{\beta}(\bullet,k))<\mathrm{TMSE}(\hat{\beta})$ が成り立つ。□

よって Corollary 1 より、ORR 推定量が OLS 推定量よりも TMSE の意味でよくなる k の範囲が明示的に表された。

Ⅲ.2 *r–k* クラス推定量と PCR 推定量の TMSE 比較

以下の定理は Baye & Parker（1984）によって与えられた。

Theorem 3（Baye & Parker）

$$\forall r\in N_p,\ \exists k(>0)\ s.t.\ \mathrm{TMSE}(\hat{\beta}(r,k))<\mathrm{TMSE}(\hat{\beta}(r,\ \bullet))$$

Proof : $\hat{\beta}(r,0)=\hat{\beta}(r,\bullet)$ であることから、$\mathrm{TMSE}(\hat{\beta}(r,k))$ が $k=0$ の近傍で減

少していることを示せば、$\mathrm{TMSE}(\hat{\beta}(r,k))$が$k$の連続関数であるので減少している範囲で$\mathrm{TMSE}(\hat{\beta}(r,k))<\mathrm{TMSE}(\hat{\beta}(r,\ \bullet))$が示されたことになる。まず$\mathrm{TMSE}(\hat{\beta}(r,k))$の$k$に関する微分を求めると、

$$\frac{d\mathrm{TMSE}(\hat{\beta}(r,k))}{dk}=\frac{d}{dk}\left[\sigma^2\sum_{i\in N_r}\frac{\lambda_i}{(\lambda_i+k)^2}+\left\{\sum_{i\in N_r}\frac{k^2\alpha_i^2}{(\lambda_i+k)^2}+\sum_{i\in N_{p\setminus r}}\alpha_i^2\right\}\right]$$
$$=-2\sigma^2\sum_{i\in N_r}\frac{\lambda_i}{(\lambda_i+k)^3}+\sum_{i\in N_r}\frac{2k\alpha_i(\lambda_i+k)-2k^2\alpha_i^2}{(\lambda_i+k)^3}$$
$$=-2\sigma^2\sum_{i\in N_r}\frac{\lambda_i}{(\lambda_i+k)^3}+2k\sum_{i\in N_r}\frac{\lambda_i}{(\lambda_i+k)^3}\alpha_i^2$$

となる。次に、この1次導関数の原点の（右側）近傍における符号を調べると、

$$\lim_{k\to+0}\frac{d\mathrm{TMSE}(\hat{\beta}(r,k))}{dk}=\lim_{k\to+0}\left\{-2\sigma^2\sum_{i\in N_r}\frac{\lambda_i}{(\lambda_i+k)^3}+2k\sum_{i\in N_r}\frac{\lambda_i}{(\lambda_i+k)^3}\alpha_i^2\right\}$$
$$=-2\sigma^2\sum_{i\in N_r}\frac{1}{\lambda_i^2}<0$$

となることから、$\mathrm{TMSE}(\hat{\beta}(r,k))$ は$k=0$の近傍で減少することがわかる。

さらに$\dfrac{d\mathrm{TMSE}(\hat{\beta}(r,k))}{dk}=0$とおくことによって、

$$k=\frac{\sigma^2\sum_{i\in N_r}\frac{\lambda_i}{(\lambda_i+k)^3}}{\sum_{i\in N_r}\frac{\lambda_i}{(\lambda_i+k)^3}\alpha_i^2}$$

となり、この等式をみたすk（k^*とおく）が$\mathrm{TMSE}(\hat{\beta}(r,k))$の最小値を与えることがわかる。この$k^*$は、

$$k^*>\frac{\sigma^2}{\|\beta\|^2}(>0)$$

という範囲にある。なぜなら、

$$w_i\equiv\frac{\frac{\lambda_i}{(\lambda_i+k^*)^3}}{\sum_{i\in N_r}\frac{\lambda_i}{(\lambda_i+k^*)^3}}$$

とおくと、（但し、$\sum_{i\in N_r}w_i=1,\ 0<w_i<1$となることに注意）

$$k^*=\frac{\sigma^2\sum_{i\in N_r}\frac{\lambda_i}{(\lambda_i+k^*)^3}}{\sum_{i\in N_r}\frac{\lambda_i}{(\lambda_i+k^*)^3}\alpha_i^2}=\frac{\sigma^2}{\frac{\sum_{i\in N_r}\frac{\lambda_i}{(\lambda_i+k^*)^3}\alpha_i^2}{\sum_{i\in N_i}\frac{\lambda_i}{(\lambda_i+k^*)}}}=\frac{\sigma^2}{\sum_{i\in N_r}w_i\alpha_i^2}$$

となり、分母は α_i^2 の重付けされた平均を表す。よって α_i^2 を最大値 $\max_{i\in N_r}\alpha_i^2$ で置き換えると、

$$k^*\geq\frac{\sigma^2}{\max_{i\in N_r}\alpha_i^2}>\frac{\sigma^2}{\|\alpha\|^2}=\frac{\sigma^2}{\|\beta\|^2}>0$$

となるからである。□

この定理から、任意に固定された $r\in N_r$ に対して、うまく k を選んでやると、r-k クラス推定量が PCR 推定量よりも TMSE の意味で良くなることはわかるが、Theorem 1 と同様に、k がどのような範囲で良くなっているかが明示的に表されていない。そこで以下のような定理を与える：

Theorem 4 $r\in N_p$ とすると以下のことが成り立つ。

任意の $i\in N_r$ に対して $\alpha_i^2-\frac{\sigma^2}{\lambda_i}\leq 0$ である時、

(i) $\mathrm{TMSE}(\hat{\beta}(r,k))<\mathrm{TMSE}(\hat{\beta}(r,\cdot))$ for $k<0$

$\alpha_i^2-\frac{\sigma^2}{\lambda_i}>0$ となる $i\in N_r$ が存在する時、

(ii) $\mathrm{TMSE}(\hat{\beta}(r,k))<\mathrm{TMSE}(\hat{\beta}(r,\cdot))$ for $0<k<K_3$

ここで、

$$K_3\equiv\frac{2\sigma^2}{\max_{i\in N_r}\left(\alpha_i^2-\frac{\sigma^2}{\lambda_i}\right)} \tag{16}$$

である。

Proof : TMSEの差を求めると、

$$\mathrm{TMSE}(\hat{\beta}(r,k))-\mathrm{TMSE}(\hat{\beta}(r,\bullet))$$

$$=\sum_{i\in N_r}\frac{\sigma^2\lambda_i}{(\lambda_i+k)^2}+\left\{\sum_{i\in N_r}\frac{\alpha_i^2k^2}{(\lambda_i+k)^2}+\sum_{i\in N_{p\backslash r}}\alpha_i^2\right\}-\left\{\sigma^2\sum_{i\in N_r}\frac{1}{\lambda_i}+\sum_{i\in N_{p\backslash r}}\alpha_i^2\right\}$$

$$=\sum_{i\in N_r}\left\{\frac{\sigma^2\lambda_i}{(\lambda_i+k)^2}-\frac{\sigma^2}{\lambda_i}+\frac{\alpha_i^2k^2}{(\lambda_i+k)^2}\right\}$$

$$=\sum_{i\in N_r}\frac{1}{(\lambda_i+k)^2}\left\{\sigma^2\lambda_i-\frac{(\lambda_i+k)^2\sigma^2}{\lambda_i}+\alpha_i^2k^2\right\}$$

$$=\sum_{i\in N_r}\frac{k}{(\lambda_i+k)^2}\left\{\left(\alpha_i^2-\frac{\sigma^2}{\lambda_i}\right)k-2\sigma^2\right\} \tag{17}$$

となる。

$\boxed{\text{(i)の証明}}$

仮定より $\alpha_i^2-\dfrac{\sigma^2}{\lambda_i}\leq 0$ が任意の $i\in N_r$ に対して成り立つので、

$$\frac{k}{(\lambda_i+k)^2}\left\{\left(\alpha_i^2-\frac{\sigma^2}{\lambda_i}\right)k-2\sigma^2\right\}<0 \quad \text{for} \quad k>0 \tag{18}$$

となり、これより(17)式は $k>0$ で負となることから(i)が示された。

$\boxed{\text{(ii)の証明}}$

$\alpha_i^2-\dfrac{\sigma^2}{\lambda_i}>0$ となる $i\in N_r$ に対して

$$\frac{k}{(\lambda_i+k)^2}\left\{\left(\alpha_i^2-\frac{\sigma^2}{\lambda_i}\right)k-2\sigma^2\right\}<0 \quad \text{for} \quad 0<k<\frac{2\sigma^2}{\alpha_i^2-\dfrac{\sigma^2}{\lambda_i}}$$

が成り立ち、$\alpha_i^2-\dfrac{\sigma^2}{\lambda_i}\leq 0$ となる i に対しては(18)式が成り立つ。

これらのことから $0<k<\dfrac{2\sigma^2}{\max_{i\in N_r}\left(\alpha_i^2-\dfrac{\sigma^2}{\lambda_i}\right)}$ という範囲では(17)が負となる。

よって(ii)が示された。□

さらに、$\alpha_i^2-\dfrac{\sigma^2}{\lambda_i}$, $(i\in N_r)$ の符号によらず以下の系が成り立つ：

Corollary 2 $r \in N_p$ とすると以下のことが成り立つ

$$\mathrm{TMSE}(\hat{\beta}(r,k)) < \mathrm{TMSE}(\hat{\beta}(r,\cdot)) \text{ for } 0<k \leq K_4$$

ここで、

$$K_4 \equiv \frac{2\sigma^2}{\|\alpha_r\|^2}. \tag{19}$$

証明は、Corollary 1 と同様にできる。

よって Corollary 2 より、r-k クラス推定量が PCR 推定量よりも TMSE の意味でよくなる k の範囲が明示的に表された。

III. 3　PCR 推定量と OLS 推定量の TMSE 比較

以下の定理は Marquardt（1970）によって与えられた。

Theorem 5（**Marquardt**）$r \in N_{p-1}$ のとき以下のことが成り立つ：

$$\sum_{i \in N_{p\setminus r}} \frac{\sigma^2}{\lambda_i} > \|\beta\|^2 \Rightarrow \mathrm{TMSE}(\hat{\beta}(r,\cdot)) < \mathrm{TMSE}(\hat{\beta})$$

Proof: TMSE の差を求めると、

$$\begin{aligned}\mathrm{TMSE}(\hat{\beta}(r,\cdot)) - \mathrm{TMSE}(\hat{\beta}) &= \sigma^2 \sum_{i \in N_r} \frac{1}{\lambda_i} + \sum_{i \in N_{p\setminus r}} \alpha_i^2 - \sigma^2 \sum_{i \in N_p} \frac{1}{\lambda_i} \\ &= \sum_{i \in N_{p\setminus r}} \left(\alpha_i^2 - \frac{\sigma^2}{\lambda_i}\right)\end{aligned} \tag{20}$$

この差が負となる 1 つの十分条件は、

$$\forall i \in N_{p\setminus r}, \ \alpha_i^2 - \frac{\sigma^2}{\lambda_i} < 0 \tag{21}$$

である。しかしながら(21)式が成り立つことをチェックすることは、実際上困難であると思えるので他の十分条件を考えると、

$$\|\beta\|^2 = \|\alpha\|^2 > \|\alpha_{p\setminus r}\|^2 = \sum_{i \in N_{p\setminus r}} \alpha_i^2 \tag{22}$$

となることから、

$$\sum_{i \in N_{p\setminus r}} \frac{\sigma^2}{\lambda_i} > \|\beta\|^2 \Rightarrow \sum_{i \in N_{p\setminus r}} \frac{\sigma^2}{\lambda_i} > \sum_{i \in N_{p\setminus r}} \alpha_i^2 \Longleftrightarrow \sum_{i \in N_{p\setminus r}} \left(\alpha_i^2 - \frac{\sigma^2}{\lambda_i}\right) < 0$$

となり、

$$\sum_{i \in N_{p \setminus r}} \frac{\sigma^2}{\lambda_i} > \|\beta\|^2 \tag{23}$$

が TMSE の差が負となるための十分条件となる。この条件は、厳しさという点では欠けるが、条件(21)よりもチェックしやすいという点で便利であろう。□

この定理の証明において Marquardt がいうように、条件(21)が成り立つことをチェックすることは難しいかもしれないが、他の定理の条件との比較も考えて以下のような定理を用意する。

Theorem 6 $r \in N_{p-1}$ のとき以下のことが成り立つ：

$$\forall i \in N_{p \setminus r}, \alpha_i^2 - \frac{\sigma^2}{\lambda_i} < 0 \Rightarrow \mathrm{TMSE}(\hat{\beta}(r, \bullet)) < \mathrm{TMSE}(\hat{\beta})$$

Ⅲ.4 *r-k* クラス推定量と ORR 推定量の TMSE 比較

以下の定理と系は Nomura & Ohkubo (1985) によって与えられた。

Theorem 7 (**Nomura & Ohkubo**) $i \in N_{p-1}$ とする、以下のことが成り立つ：

(i)

$$\alpha_i^2 - \frac{\sigma^2}{\lambda_i} < 0 \Rightarrow \begin{cases} \mathrm{TMSE}(\hat{\beta}(i, k)) > \mathrm{TMSE}(\hat{\beta}(i-1, k)) \text{ for } 0 \le k < \dfrac{\lambda_i}{2\alpha_i^2}\left(\dfrac{\sigma^2}{\lambda_i} - \alpha_i^2\right) \\ \mathrm{TMSE}(\hat{\beta}(r, k)) \le \mathrm{TMSE}(\hat{\beta}(i-1, k)) \text{ for } k \ge \dfrac{\lambda_i}{2\alpha_i^2}\left(\dfrac{\sigma^2}{\lambda_i} - \alpha_i^2\right) \end{cases}$$

等号成立は$k = \dfrac{\lambda_i}{2\alpha_i^2}\left(\dfrac{\sigma^2}{\lambda_i} - \alpha_i^2\right)$の時である。

(ii) $$\alpha_i^2 - \frac{\sigma^2}{\lambda_i} \ge 0 \Rightarrow \mathrm{TMSE}(\hat{\beta}(i, k)) \le \mathrm{TMSE}(\hat{\beta}(i-1, k)) \quad \text{for} \quad k \ge 0$$

等号成立は$\dfrac{\sigma^2}{\lambda_i} - \alpha_i^2 = 0$かつ$k = 0$の時である。

Proof : $\hat{\beta}(i,k)$と$\hat{\beta}(i-1,k)$の TMSE の差は以下のようになる：

$$\begin{aligned}
&\mathrm{TMSE}(\hat{\beta}(i,k))-\mathrm{TMSE}(\hat{\beta}(i-1,k))\\
&=\sigma^2\frac{\lambda_i}{(\lambda_i+k)^2}+\frac{k^2\alpha_i^2}{(\lambda_i+k)^2}-\alpha_i^2\\
&=\frac{\lambda_i^2}{(\lambda_i+k)^2}\left\{\frac{\sigma^2}{\lambda_i}+\frac{\alpha_i^2}{\lambda_i^2}k^2-\frac{(\lambda_i+k)^2\alpha_i^2}{\lambda_i^2}\right\}\\
&=\frac{\lambda_i^2}{(\lambda_i+k)^2}\left(\frac{\sigma^2}{\lambda_i}-\alpha_i^2-2\frac{\alpha_i^2}{\lambda_i}k\right)\\
&=-\frac{2\alpha_i^2\lambda_i}{(\lambda_i+k)^2}\left\{k-\frac{\lambda_i}{2\alpha_i^2}\left(\frac{\sigma^2}{\lambda_i}-\alpha_i^2\right)\right\} \qquad (24)
\end{aligned}$$

(i)の証明

$\dfrac{\sigma^2}{\lambda_i}-\alpha_i^2>0$ より k が $0\leq k<\dfrac{\lambda_i}{2\alpha_i^2}\left(\dfrac{\sigma^2}{\lambda_i}-\alpha_i^2\right)$という範囲にあれば、

$$\frac{2\alpha_i^2\lambda_i}{(\lambda_i+k)^2}\left\{k-\frac{\lambda_i}{2\alpha_i^2}\left(\frac{\sigma^2}{\lambda_i}-\alpha_i^2\right)\right\}<0$$

が成り立ち、(24)式の右辺は正となる。故に TMSE$(\hat{\beta}(r,k))>$TMSE$(\hat{\beta}(i-1,k))$が示された。

また k が $k\geq\dfrac{\lambda_i}{2\alpha_i^2}\left(\dfrac{\sigma^2}{\lambda_i}-\alpha_i^2\right)$ (>0) という範囲にあれば、

$$\frac{2\alpha_i^2\lambda_i}{(\lambda_i+k)^2}\left\{k-\frac{\lambda_i}{2\alpha_i^2}\left(\frac{\sigma^2}{\lambda_i}-\alpha_i^2\right)\right\}\geq 0$$

が成り立ち、(24)式の右辺は非正となる。よって(i)が示された。

(ii)の証明

$\dfrac{\sigma^2}{\lambda_i}-\alpha_i^2\leq 0$ より、$\forall k\geq 0$ に対して、

$$\frac{2\alpha_i^2\lambda_i}{(\lambda_i+k)^2}\left\{k-\frac{\lambda_i}{2\alpha_i^2}\left(\frac{\sigma^2}{\lambda_i}-\alpha_i^2\right)\right\}\geq 0$$

が成り立ち、(24)式の右辺は非正となる。よって(ii)が示された。□

Corollary 3 (Nomura & Ohkubo) $r\in N_{p-1}$ とすると以下のことが成り立つ：
任意の $i\in N_{p\backslash r}$ に対して $\alpha_i^2-\dfrac{\sigma^2}{\lambda_i}<0$ であるとき、

(i)　$\mathrm{TMSE}(\hat{\beta}(r,k))<\mathrm{TMSE}(\hat{\beta}(r+1,k))<\cdots<\mathrm{TMSE}(\hat{\beta}(\cdot,k))$　for　$0\leq k<K_5$

となるような $K_5(>0)$ が存在する。

　また $\alpha_i^2-\dfrac{\sigma^2}{\lambda_i}\geq 0$ となるような $i\in N_{p\backslash r}$ が存在するとき

(ii)　$\mathrm{TMSE}(\hat{\beta}(\cdot,k))<\mathrm{TMSE}(\hat{\beta}(p-1,k))<\cdots<\mathrm{TMSE}(\hat{\beta}(r,k))$ for $k>K_6$

となるような $K_6(\geq 0)$ が存在する。

　ここで、K_5、K_6 は、

$$K_5\equiv\min_{i\in N_{p\backslash r}}\left\{\frac{\lambda_i}{2\alpha_i^2}\left(\frac{\sigma^2}{\lambda_i}-\alpha_i^2\right)\right\},\tag{25}$$

$$K_6\equiv\max\left[0,\ \max_{i\in N_{p\backslash r}}\left\{\frac{\lambda_i}{2\alpha_i^2}\left(\frac{\sigma^2}{\lambda_i}-\alpha_i^2\right)\right\}\right]\tag{26}$$

ととればよい。

Proof : $\boxed{\text{(i)の証明}}$

　すべての $i\in N_{p\backslash r}$ に対して $\dfrac{\sigma^2}{\lambda_i}-\alpha_i^2>0$ が成り立っており、$N_{p\backslash r}\subseteq N_{p\backslash 1}$ であるので、Theorem 7 の(i)を各 i に対して適用すると、

$$\mathrm{TMSE}(\hat{\beta}r,k))<\mathrm{TMSE}(\hat{\beta}(r+1,k))\quad\text{for}\quad 0\leq k<\frac{\lambda_{r+1}}{2\alpha_{r+1}^2}\left(\frac{\sigma^2}{\lambda_{r+1}}-\alpha_{r+1}^2\right)$$

$$\mathrm{TMSE}(\hat{\beta}(r+1,k))<\mathrm{TMSE}(\hat{\beta}(r+2,k))\quad\text{for}\quad 0\leq k<\frac{\lambda_{r+2}}{2\alpha_{r+2}^2}\left(\frac{\sigma^2}{\lambda_{r+2}}-\alpha_{r+2}^2\right)$$

$$\vdots$$

$$\mathrm{TMSE}(\hat{\beta}(p-1,k))<\mathrm{TMSE}(\hat{\beta}(p,k))\quad\text{for}\quad 0\leq k<\frac{\lambda_p}{2\alpha_p^2}\left(\frac{\sigma^2}{\lambda_p}-\alpha_p^2\right)$$

となる。これより、$K_5=\displaystyle\min_{i\in N_{p\backslash r}}\left\{\frac{\lambda_i}{2\alpha_i^2}\left(\frac{\sigma^2}{\lambda_i}-\alpha_i^2\right)\right\}$ とおくと、$0\leq k<K_5$ という範囲で、

$$\mathrm{TMSE}(\hat{\beta}(r,k))<\cdots<\mathrm{TMSE}(\hat{\beta}(p-1,k))<\mathrm{TMSE}(\hat{\beta}(\cdot,k))$$

が成り立つ。

(ii)の証明

$\dfrac{\sigma^2}{\lambda_i}-\alpha_i^2>0$ となる $i\in N_{p\backslash r}$ に対して Theorem 7(i)より

$$\mathrm{TMSE}(\hat{\beta}(i,k))<\mathrm{TMSE}(\hat{\beta}(i-1,k)) \quad \text{for} \quad k>\frac{\lambda_i}{2\alpha_i^2}\left(\frac{\sigma^2}{\lambda_i}-\alpha_i^2\right)$$

が成り立つ。また $\dfrac{\sigma^2}{\lambda_i}-\alpha_i^2\leq 0$ となる i に対して、Theorem 7(ii)より

$$\mathrm{TMSE}(\hat{\beta}(i,k))<\mathrm{TMSE}(\hat{\beta}(i-1,k)) \quad \text{for} \quad k>0$$

が成り立つ。よって $K_5=\max\left[0,\ \max\limits_{i\in N_{p\backslash r}}\left\{\dfrac{\lambda_i}{2\alpha_i^2}\left(\dfrac{\sigma^2}{\lambda_i}-\alpha_i^2\right)\right\}\right]$ とおくと、$k>K_6$ という範囲で

$$\mathrm{TMSE}(\hat{\beta}(\cdot,k))<\mathrm{TMSE}(\hat{\beta}(p-1,k))<\cdots<\mathrm{TMSE}(\hat{\beta}(r,k))$$

が成り立つ。□

このCorollary 3(i)から、任意に固定された r（但し、$r\in N_{p-1}$）に関してうまく k を選んでやると、r-k クラス推定量は ORR 推定量よりも TMSE の意味で良くなることが分かる。

III. 5 r-k クラス推定量と OLS 推定量の TMSE 比較

以下の定理は、Nomura & Ohkubo (1985) によって与えられた：

Theorem 8 (Nomura & Ohkubo) $r\in N_p$ とすると、以下のことが成り立つ：

$\sum\limits_{i\in N_{p\backslash r}}\left(\alpha_i^2-\dfrac{\sigma^2}{\lambda_i}\right)\leq 0$ であり、任意の $i\in N_r$ に対して $\alpha_i^2-\dfrac{\sigma^2}{\lambda_i}\leq 0$ である時、

(i) $\mathrm{TMSE}(\hat{\beta}(r,k))<\mathrm{TMSE}(\hat{\beta})$ for $k>0.$

$\sum\limits_{i\in N_{p\backslash r}}\left(\alpha_i^2-\dfrac{\sigma^2}{\lambda_i}\right)\leq 0$ であり、$\alpha_i^2-\dfrac{\sigma^2}{\lambda_i}>0$ となる $i\in N_r$ が存在する時、

(ii) $\mathrm{TMSE}(\hat{\beta}(r,k))<\mathrm{TMSE}(\hat{\beta})$ for $0<k<K_3.$

ここで，K_3 は(16)式で定義されたものである。

Proof : $\hat{\beta}(r,k)$ と $\hat{\beta}$ の TMSE の差は以下のようになる：

$$
\begin{aligned}
&\mathrm{TMSE}(\hat{\beta}(r,k))-\mathrm{TMSE}(\hat{\beta})\\
&=\sigma^2\sum_{i\in N_r}\frac{\lambda_i}{(\lambda_i+k)^2}+\left\{\sum_{i\in N_r}\frac{k^2\alpha_i^2}{(\lambda_i+k)^2}+\sum_{i\in N_{p\backslash r}}\alpha_i^2\right\}-\sigma^2\sum_{i\in N_p}\frac{1}{\lambda_i}\\
&=\sum_{i\in N_r}\left\{\frac{\sigma^2\lambda_i+k^2\alpha_i^2}{(\lambda_i+k)^2}-\frac{\sigma^2}{\lambda_i}\right\}+\sum_{i\in N_{p\backslash r}}\left(\alpha_i^2-\frac{\sigma^2}{\lambda_i}\right)\\
&=\sum_{i\in N_r}\left\{\frac{\sigma^2\lambda_i^2+\lambda_i k^2\alpha_i^2-\sigma^2(\lambda_i+k)^2}{\lambda_i(\lambda_i+k)^2}\right\}+\sum_{i\in N_{p\backslash r}}\left(\alpha_i^2-\frac{\sigma^2}{\lambda_i}\right)\\
&=\sum_{i\in N_r}\frac{1}{\lambda_i(\lambda_i+k)^2}\ (\lambda_i\alpha_i^2k^2-2\sigma^2\lambda_i k-\sigma^2k^2)\ +\sum_{i\in N_{p\backslash r}}\left(\alpha_i^2-\frac{\sigma^2}{\lambda_i}\right)\\
&=\sum_{i\in N_r}\frac{k}{(\lambda_i+k)^2}\left(\alpha_i^2k-2\sigma^2-\frac{\sigma^2}{\lambda_i}k\right)+\sum_{i\in N_{p\backslash r}}\left(\alpha_i^2-\frac{\sigma^2}{\lambda_i}\right)\\
&=\sum_{i\in N_r}\frac{k}{(\lambda_i+k)^2}\left\{\left(\alpha_i^2-\frac{\sigma^2}{\lambda_i}\right)k-2\sigma^2\right\}+\sum_{i\in N_{p\backslash r}}\left(\alpha_i^2-\frac{\sigma^2}{\lambda_i}\right). \qquad (27)
\end{aligned}
$$

となる。ここで(27)式の第2項は≦0であると仮定されていたので、第1項の符号を調べればよい。

(i)の証明

仮定より $\alpha_i^2-\dfrac{\sigma^2}{\lambda_i}\leq 0$ が任意の $i\in N_r$ に対して成り立つので、

$$
\frac{k}{(\lambda_i+k)^2}\left\{\left(\alpha_i^2-\frac{\sigma^2}{\lambda_i}\right)k-2\sigma^2\right\}<0 \quad \text{for} \quad k>0 \qquad (28)
$$

となり、これより(27)式の第1項は $k>0$ で負となることから(i)が示された。

(ii)の証明

$\alpha_i^2-\dfrac{\sigma^2}{\lambda_i}>0$ となる $i\in N_r$ に対して、

$$
\frac{k}{(\lambda_i+k)^2}\left\{\left(\alpha_i^2-\frac{\sigma^2}{\lambda_i}\right)k-2\sigma^2\right\}<0 \quad \text{for} \quad 0<k<\frac{2\sigma^2}{\alpha_i^2-\dfrac{\sigma^2}{\lambda_i}}
$$

が成り立ち、$\alpha_i^2-\dfrac{\sigma^2}{\lambda_i}\leq 0$ となる i に対しては(28)式が成り立つ。

これからのことから $0<k<\dfrac{2\sigma^2}{\max_{i\in N_r}\left(\alpha_i^2-\dfrac{\sigma^2}{\lambda_i}\right)}$ という範囲では(27)の第1項が負となる。よって(ii)が示された。□

さらに、Nomura & Ohkubo は、$\alpha_i^2-\dfrac{\sigma^2}{\lambda_i}$, $(i\in N_r)$ の符号によらずに以下の系が成り立つことを示した。

Corollary 4（Nomura & Ohkubo） $\sum_{i\in N_{p\backslash r}}\left(\alpha_i^2-\dfrac{\sigma^2}{\lambda_i}\right)\leq 0$ である時、

$$\mathrm{TMSE}(\hat{\beta}(r,k))<\mathrm{TMSE}(\hat{\beta})\quad \text{for}\quad 0<k\leq K_4$$

ここで、K_4 は、(19)式で定義されたものである。

Proof : $\alpha_i^2-\dfrac{\sigma^2}{\lambda_i}>0$ となる $i(\in N_r)$ が存在する時、

$$\frac{2\sigma^2}{\sum_{i\in N_r}\alpha_i^2}<\frac{2\sigma^2}{\max_{i\in N_r}\left(\alpha_i^2-\dfrac{\sigma^2}{\lambda_i}\right)}$$

が成り立つ。このことと $\sum_{i\in N_{p\backslash r}}\left(\alpha_i^2-\dfrac{\sigma^2}{\lambda_i}\right)\leq 0$ という仮定から Theorem 8(ii)を使うと、

$$\mathrm{TMSE}(\hat{\beta}(r,k))<\mathrm{TMSE}(\hat{\beta})\quad \text{for}\quad 0<k\leq\frac{2\sigma^2}{\sum_{i\in N_r}\alpha_i^2}$$

がいえる。また任意の i に対して $\alpha_i^2-\dfrac{\sigma^2}{\lambda_i}\leq 0$ が成り立っている時、Theorem 8(i)より、

$$\mathrm{TMSE}(\hat{\beta}(r,k))<\mathrm{TMSE}(\hat{\beta})\quad \text{for}\quad k>0$$

がいえるので、結局 $\sum_{i\in N_{p\backslash r}}\left(\alpha_i^2-\dfrac{\sigma^2}{\lambda_i}\right)\leq 0$ が成り立つ時、

$$0<k\leq\frac{2\sigma^2}{\sum_{i\in N_r}\alpha_i^2}$$

に対して $\mathrm{TMSE}(\hat{\beta}(r,k))<\mathrm{TMSE}(\hat{\beta})$ が成り立つ。□

このCorollary 4より、任意に選ばれた $r(\in N_p)$ に対して、$\sum_{i\in N_{p\backslash r}}\left(\alpha_i^2-\frac{\sigma^2}{\lambda_i}\right)\leq 0$ ならば、うまく k を選ぶと、r-k クラス推定量はOLS推定量よりもTMSEの意味でよくなることが分かる。

IV　r-k クラス推定量が他の推定量をTMSE優越する定理

前節において述べられた定理と系において使用される条件の主なものは、以下のような形式に整理することができる：

$$(C1)\ :\ 0<k\leq\frac{2\sigma^2}{\|\beta\|^2},$$

$$(C2)\ :\ \frac{\sigma^2}{\lambda_i}-\alpha_i^2>0,\quad \text{for } \forall_i\in N_{p\backslash r},$$

$$(C3)\ :\ 0<k\leq\frac{2\sigma^2}{\|\alpha_r\|^2},$$

$$(C4)\ :\ 0\leq k<\min_{i\in N_{p\backslash r}}\left\{\frac{\lambda_i}{2\alpha_i^2}\left(\frac{\sigma^2}{\lambda_i}-\alpha_i^2\right)\right\},\quad (\text{and } (C2))$$

これらの条件（$C1$）～（$C4$）を使うと、前節の定理、系は以下のように再構成できる。まずCorollary 2は以下の定理になる：

Theorem 9　任意に固定された $r\in N_p$ に対して、条件（$C3$）のもとで r-k クラス推定量のTMSEはPCR推定量のTMSEよりも小さくなる。すなわち、

$$\mathrm{TMSE}(\hat{\beta}(r,k))<\mathrm{TMSE}(\hat{\beta}(r,\bullet)).$$

この定理において，$r=p$ とおくことによって、以下の系が与えられる。

Corollary 5　条件（$C1$）のもとでORR推定量のTMSEはOLS推定量のTMSEよりも小さくなる。すなわち、

$$\mathrm{TMSE}(\hat{\beta}(\bullet,k))<\mathrm{TMSE}(\hat{\beta}).$$

これは Corollary 1 に他ならない。

一方、Corollary 3(i)は以下のようになる：

Theorem 10 任意に固定された $r \in N_{p-1}$ に対して、条件（$C4$）のもとで r–k クラス推定量の TMSE は ORR 推定量の TMSE よりも小さくなる。すなわち、

$$\mathrm{TMSE}(\hat{\beta}(r, k)) < \mathrm{TMSE}(\hat{\beta}(\cdot, k)).$$

ここで $k=0$ とおくことによって、以下の系が与えられる：

Corollary 6 任意に固定された $r \in N_{p-1}$ に対して、条件（$C2$）のもとで PCR 推定量の TMSE は OLS 推定量の TMSE よりも小さくなる。すなわち、

$$\mathrm{TMSE}(\hat{\beta}(r, \cdot)) < \mathrm{TMSE}(\hat{\beta}).$$

これは Theorem 6 に他ならない。

以上の定理と系から、r–k クラス推定量が他の推定量を TMSE 優越する以下の系を得ることができる。(図 1 も参照)：

図 1　TMSE 比較に関するダイアグラム

$$\begin{array}{ccc}
 & (C1) & \\
\mathrm{TMSE}(\hat{\beta}(\cdot, k)) & < & \mathrm{TMSE}(\hat{\beta}) \\
(C4)\ \vee & & \vee\ (C2) \\
\mathrm{TMSE}(\hat{\beta}(r, k) & < & \mathrm{TMSE}(\hat{\beta}(r, \cdot)) \\
 & (C3) &
\end{array}$$

Corollary 7 任意に固定された $r \in N_{p-1}$ に対して、条件（$C2$）と（$C1$）$\wedge$（$C4$）のもとで

$$\mathrm{TMSE}(\hat{\beta}(r, k)) < \mathrm{TMSE}(\text{the other estimators})$$

Proof：$r \in N_{p-1}$ に注意しよう。

$$\frac{2\sigma^2}{\|\beta\|^2} < \frac{2\sigma^2}{\|\alpha_r\|^2}$$

が成り立つことから、Theorem 9、Corollary 5 より、条件（$C1$）のもとで、

$$\mathrm{TMSE}(\hat{\beta}(\cdot, k)) < \mathrm{TMSE}(\hat{\beta}) \ \& \ \mathrm{TMSE}(\hat{\beta}(r, k)) < \mathrm{TMSE}(\hat{\beta}(r, \cdot)) \tag{29}$$

が成り立つ。

次に、条件（$C2$）、（$C4$）のもとで Theorem 10、Corollary 6 より、

$$\mathrm{TMSE}(\hat{\beta}(r, \cdot)) < \mathrm{TMSE}(\hat{\beta}) \ \& \ \mathrm{TMSE}(\hat{\beta}(r, k)) < \mathrm{TMSE}(\hat{\beta}(\cdot, k)) \tag{30}$$

が成り立つ。この２つの結果より、条件（$C2$）と（$C1$）∧（$C4$）、すなわち $0 < k < \min(K_2, K_5)$ のもとで、

$$\begin{cases} \mathrm{TMSE}(\hat{\beta}(r, k)) < \mathrm{TMSE}(\hat{\beta}(r, \cdot)) < \mathrm{TMSE}(\hat{\beta}) \\ \mathrm{TMSE}(\hat{\beta}(r, k)) < \mathrm{TMSE}(\hat{\beta}(\cdot, k)) < \mathrm{TMSE}(\hat{\beta}) \end{cases}$$

となる。このことは、*r–k* クラス推定量が他の推定量（OLS，ORR，PCR 推定量）よりも TMSE が小さいことを表している。□

V　おわりに

本稿では、多重共線性が説明変数間に存在する場合の回帰係数の代表的な推定量として OLS，PCR，ORR，*r–k* クラス推定量をとりあげ、過去の著名な文献で述べられている TMSE の比較をもとに、条件を整備することによって、その比較をより明らかなものにすることを定理の形でまとめた。（Theorem 9, 10，Corollary 5，6 参照。）この試みから、*r–k* クラス推定量がその他の推定量を TMSE 優越する条件も得られた。（Corollary 7 を参照。）これらの定理と系は、実際に *r–k* クラス推定量を使用する際に問題となる主成分数 r と ridge 定数 k の選択問題に役立つであろう。また、これらの定理と系は、Jimichi & Inagaki（1996）において補助変量を含む場合の回帰係数の推定問題に拡張されている。

（筆者は関西学院大学商学部専任講師）

参考文献

[1] Baye, M. R. and D. F. Parker (1984), "Combining ridge and principal component regression : A money demand illustration," *Communication in Statistics. Theor. and Meth.*, vol. 13, pp. 197–205.

[2] Gunst, R. F. (1983), "Regression analysis with multicollinear predictor variables : Definition, detection, and effects," *Communication in Statistics. Theor. and Meth.*, vol. 12, pp. 2217–2260.

[3] Hoerl, A. E. and R. W. Kennard (1970), "Ridge regression : biased estimation for nonorthogonal problems," *Technometrics*, vol. 12, pp. 55–67.

[4] Jimichi, M. and N. Inagaki (1996) "*r–k* class estimation in regression model with concomitant variables,", *Annals of Institute Statistical Mathematics*, in printing.

[5] Kendall, M. G. (1957), *A Course in Multivariate Analysis*, London : Griffin.

[6] Marquardt, D. W. (1970), "Generalized inverses, ridge regression biased estimation, and nonlinear estimation," *Technometrics*, vol. 12, pp. 591–612.

[7] Nomura, M. and T. Ohkubo (1985), "A note on combining ridge and principal component regression," *Communication in Statistics. Theor. and Meth.*, vol. 14, pp. 2489–2493.