多重共線性について

井　上　勝　雄

§0

経済諸量の関係式を推定する際，単純最小二乗法（OLS）が統計的に望ましい特性をもつためにいくつかのモデルの前提が必要である．その中には説明変数間に完全な相関がないという仮定がある．これに反して，多重共線性というのは，説明変数間に相関があること，あるいは，説明変数間に線型関係が存在することを示すのに用いられる．

本稿で，この多重共線性の問題のいくつかを単一方程式モデルの場合について考察する．

§1では，多重共線性をどの様な統計量で測定したらよいかと，多重共線性が推定に及ぼす影響としてしばしば指摘されることについて検討を加える．**§2**では，多重共線性の定義を与える統計的仮説検定の方法を，基本的には Farrar-Glauber のそれに従って考察する．最後に，**§3**において，多重共線性がある標本の場合，単純最小二乗推定よりも好ましい推定法についてを，予測の問題の中で考察する．

§1

1.1　説明変数$X_1, \cdots\cdots X_k$と被説明変数Yとの関係式

$$Y=\beta_0+\beta_1X_1+\cdots\cdots+\beta_kX_k+u$$

を考察するモデルとしよう．

いま，変数Y，$X_1, \cdots\cdots$，X_kについての標本がT個得られたとして，そのすべての変量を標本平均からの偏差（y_t，$x_{1t}, \cdots\cdots$，x_{kt}）（$t=1, \cdots\cdots T$）で表わすならば，考察しようとするモデルを

$$y_t=\beta_1x_{1t}+\cdots\cdots\beta_kx_{kt}+u_t \quad (t=1, 2, \cdots, \ T)$$

で示せる．さらに，これを

(1.1)　$y=X\beta+u$

$$y=\begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_T \end{pmatrix},\quad X=\begin{pmatrix} x_{11}\cdots\cdots x_{k1} \\ x_{12}\cdots\cdots x_{k2} \\ \vdots \qquad\quad \vdots \\ x_{1T}\cdots\cdots x_{kT} \end{pmatrix},\quad \beta=\begin{pmatrix} \beta_1 \\ \vdots \\ \beta_k \end{pmatrix},\quad u=\begin{pmatrix} u_1 \\ u_2 \\ \vdots \\ u_T \end{pmatrix}$$

で表わそう．(1.1)において確率変数ベクトルuについて

(1.2)　$u\sim N(0,\ \sigma_u^2 I)$

を前提しておこう．以下で，われわれはモデル(1.1)，(1.2)を考察する．

1.2　周知の様に，未知パラメータβのOLS推定量$\hat{\beta}$は，

(1.3)　$\hat{\beta}=(X'X)^{-1}X'y$

であり，$\hat{\beta}$の分散・共分散行列$V(\hat{\beta})$は，

(1.4)　$V(\hat{\beta})=\sigma_u^2(X'X)^{-1}$

である．また，確率変数u_tの分散の推定量$\hat{\sigma}_u^2$は，

(1.5)　$\hat{\sigma}_u^2=\dfrac{1}{T-k}y'[I-X(X'X)^{-1}X']y$

となり，したがって，$\hat{\beta}$の分散・共分散行列$V(\hat{\beta})$の推定量$\hat{V}(\beta)$は，

(1.6)　$\hat{V}(\beta)=\hat{\sigma}_u^2(X'X)^{-1}$

となることもよく知られている．また，βの OLS 推定量$\hat{\beta}$の著しい統計学的特徴は，(1.2)の下では$\hat{\beta}$が最良線型不偏推定量となることも周知に属する．

さて，一般に(1.1)のβを推定する場合，(1.2)の他に，

(1.7)　$rank\ \ X=k$

を前提する．(1.3)に示されるように推定量$\hat{\beta}$が求められるためには，

(1.8)　$rank\ \ X'X=k$

でなければならないが，(1.7)は(1.8)を保証するのである．つまり，OLS 推定量$\hat{\beta}$が決定されるためには，(1.7) に示される様にXのk個の列が一次独立でなければならない．

(1.7)を満たさないケースを考えると，説明変数の標本の間に，たとえば，

(1.9)　$x_{1t}=\alpha x_{2t}\qquad (t=1, 2, \cdots T)$

の関係があるか，あるいはもっと一般的に，

(1.10)　$\alpha_1 x_{1t}+\alpha_2 x_{2t}+\cdots\cdots+\alpha_k x_{kt}=0 \quad (t=1,2,\cdots T)$

なる関係が存在する場合である．(1.9) を満たすような場合には，本質的には変量X_1とX_2とは等しい変量であるから，元来，説明変量X_1とX_2のそれぞれをYの変動因としては識別不可能であるし，(1.1) のモデルを分析しようとすることが良くないと言えるだろう[1]．したがって，OLSの推定量$\hat{\beta}$は未決定のままである．同じことが(1.10)のケースについても言える．

しかし，実際に得られる標本が(1.9) や (1.10)を満たすことは，ほとんど有り得ないと言えるだろう．つまり，説明変数の標本間に(1.9)あるいは(1.10)に示されるような完全な線型関係があるということ，したがって (1.7)を満たさないということはないと考えてよい．換言すれば，標本は一般的には (1.7) を満たすであろうし，それと同義であるが，説明変数の標本間に完全な線型関係があること，つまり，完全な多重共線性があるということは現実的な問題ではないということになる．

1.3　前述のように，説明変数間の完全な多重共線性がほとんど問題ではないとすると，計量経済学の領域において，問題とされる多重共線性は何であろうか．しかし，問題とされるべき多重共線性についての厳密な定義は現在ではなされていない．直観的には(1.9)あるいは，(1.10) の関係に近い状態にある変量間の関係として把握されている．ここで，(1.9) あるいは(1.10)に近い状態ということを明示的に考えてみよう．

一般に，説明変量の標本間には，

$$\alpha_1' x_{1t}+\alpha_2' x_{2t}+\cdots\cdots+\alpha_k' x_{kt}+v_t'=0 \quad (t=1,2,\cdots,\ T)$$

なる関係が想定できる．ここでα_i' $(i=1,\cdots k)$ は定数であり，v_t' は計算上必然的な残差である．いま，$\alpha_1' \neq 0$ とできるならば，

1)　β_2 は他の変数の値を一定にし x_2 を微小変化させたときの y の変化量を表わすが，(1.9)ならば，少なくともx_1を一定に保てないのである．この意味で(1.1)を分析することはできない．

$$x_{1t}=\hat{\alpha}_2 x_{2t}+\cdots+\hat{\alpha}_k x_{kt}+\hat{v}_t \quad (t=1, 2, \cdots, T)$$

の関係が規定できる．$\hat{\alpha}_i(i=2, \cdots, k)$ は定係数であり，$\hat{v}_t$ は残差である．上式を

(1.11) $x_1=X_2\hat{\alpha}+\hat{v}$

$$x_1=\begin{pmatrix} x_{11} \\ x_{12} \\ \vdots \\ x_{1T} \end{pmatrix}, \quad X_2=\begin{pmatrix} x_{21} \cdots\cdots x_{k1} \\ x_{22} \cdots\cdots x_{k1} \\ \vdots \qquad \vdots \\ x_{2T} \cdots\cdots x_{kT} \end{pmatrix}, \quad \hat{\alpha}=\begin{pmatrix} \hat{\alpha}_2 \\ \vdots \\ \hat{\alpha}_k \end{pmatrix}, \quad \hat{v}=\begin{pmatrix} \hat{v}_1 \\ \hat{v}_2 \\ \vdots \\ \hat{v}_T \end{pmatrix}$$

で表現しよう．さらに残差 $\hat{v}$ について，

(1.12) $\sum\hat{v}_t=0$

$X_2'\hat{v}=0$

であるとしてよい．この規定は，(1.11)を，変量 x_1 の$(k-1)$の変量 X_2 への回帰式と見做すならば，$\hat{\alpha}$，$\hat{v}$ は最小二乗法で求めることができ，(1.12) が満たされることは明らかである[1]．

(1.11)のように説明変数間の関係を規定すると，

(i) 完全な多重共線性が存在するとき，

$\hat{v}=0$であり，したがって $\hat{v}'\hat{v}=0$ で，しかも$\hat{\alpha}\neq 0$である．

(ii) 全く多重共線性が存在しないとき，$\hat{\alpha}=0$であって，したがって，$\hat{v}'\hat{v}=x_1'x_1$ である．

これから，一般的には，

(1.13) $0\leqq\hat{v}'\hat{v}\leqq x_1'x_1$

が導かれる．通常の説明変数の標本については，$\hat{v}'\hat{v}=0$ あるいは $\hat{v}'\hat{v}=x_1'x_1$ となることはないと考えてよい．したがって $\hat{v}'\hat{v}$ の値は，(1.13)において厳密な不等号を成立させると考えられる．換言すれば，(1.9) あるいは(1.10)の状態に近いという場合，つまり完全な多重共線性に近いときには，$\hat{v}'\hat{v}$ の値が 0 に近い値をとることが予想されるし，$\hat{v}'\hat{v}$ の値が$x_1'x_1$の値に近い値をとる場合，多重共線性がほとんどないと言えるであろう．

1) この§ではその必要がないが，以下では$\hat{\alpha}=(X_2'X_2)^{-1}X_2'x_1$，$\hat{v}=x_1-X_2\hat{\alpha}$である．

1.4 **1.3** の考察を**§2**に続ける前に，多重共線性が推定に及ぼす影響として，しばしば指摘されることをこの節と**1.5**で若干考察しておこう．

モデル(1.1)において，説明変数Xの間に多重共線性が存在する場合，

（i） 係数推定値$\hat{\beta}$が不安定になる．

つまり，モデル(1.1)を推定する場合，標本数が1，2個増減することによって係数の推定値が大きく変化する．また説明変数を1，2個増減させることによって，言い換えればモデル(1.1)の定式化を少し変更するだけで係数推定値が大きく変化する．

(ii) 係数推定値の標本誤差，あるいは分散が一般に大きくなる．このことから係数の区間推定を導出するとき，その信頼区間が大きくなる．さらに，推定値の標本誤差が大きくなることから，その係数の有意性が低下し，関係式からその変数を落すことになりやすい．したがって誤まったモデルを定式化することになりやすい．

上に挙げた多重共線性の結果としてしばしば指摘される特徴[1)]について，以下で若干の考察を加えておく．

モデル(1.1)の係数βのOLS推定量$\hat{\beta}$を導出する過程で一般に正規方程式

(1.18)　$X'X\hat{\beta}=X'y$

を導く．標本個数の相異による係数推定値$\hat{\beta}$の差異は，この正規方程式の各係数の変化（標本個数の相異による$X'X$，$X'y$の相異）に依存することは明らかである．

結論的に言えることは，係数間に大差ない場合に推定値$\hat{\beta}$に大きな差異をもたらすことも，それほど大きな差異が生じない場合もあり得る，ということである．このことを一般的な形で示すことは困難であるので，例示することにしたい．

---

1) 多くの著書で指摘されている．たとえば，J. Johnston, *Econometric Methods*, (McGraw-Hill, 1972), p. 160. 参照．誤まった定式化に関する誤差については本稿ではとり挙げない．

いままでの定式化から

$$X=[x_1\ \ X_2]$$

であるから上式に(1.11)を代入し，(1.12)に注意して(1.18)を書き換えるならば，

$$(1.19)\quad \begin{bmatrix} \hat{\alpha}'X_2'X_2\hat{\alpha}+\hat{v}'\hat{v} & \hat{\alpha}'X_2'X_2 \\ X_2'X_2\hat{\alpha} & X_2'X_2 \end{bmatrix}\hat{\beta}=\begin{bmatrix} \hat{\alpha}'X_2'y+\hat{v}'y \\ X_2'y \end{bmatrix}$$

が導出できる．いま，$\hat{\beta}$の第1要素$\hat{\beta}_1$に注目すると，

$$\hat{\beta}_1=\frac{\hat{v}'y\cdot|X_2'X_2|}{\hat{v}'\hat{v}}$$

である．問題を単純化して，標本個数の違いがあるが，$\frac{1}{T}X_2'X_2$は変化しないで，$\frac{1}{T}\hat{v}'y$，$\frac{1}{T}\hat{v}'\hat{v}$が微少に変化したとする．$\hat{v}'y$が0に近いときには，標本個数の違いがその符号を変える場合があり得る．このことは，別な見方をするならば，yとx_1との偏相関が0に近いときに，標本を1，2個増減させると符号を変えてしまい，それが推定値$\hat{\beta}_1$の符号を変化させることになる．これは，後述するように，x_1がX_2の影響を受け，つまりx_1がX_2と多重共線関係にありそのためにyの説明にはX_2だけで充分説明力があるときに起こりやすいであろう．

他方，$\hat{v}'y$の値に比して$\hat{v}'\hat{v}$が充分0に近いとき，標本個数の違いによって$\hat{\beta}_1$の分母にある$\hat{v}'\hat{v}$の微少変化が$\hat{\beta}_1$の値そのものを大きく変化させる．これはx_1がyに対しても充分説明力をもっているが，それのX_2との多重共線関係の程度が非常に強い場合に起こるであろう．

標本個数による係数推定値の差異は，単に多重共線性だけに依存するのではなく，モデル(1.1)において，共線関係にあるx_1がyに対してもつ説明力にも依存する．そうして，x_1のyに対する説明力と多重共線関係とは必然的な関係にあるのではないから，多重共線性が係数推定値に与える影響は一義的ではないのである．

上のことを纏めると，説明変数間に多重共線性のあるとき，標本個数の違い

によって係数推定値が大きく変化する傾向があるとしても，必然的にそうなるとは言えないのである．

1.5 次に推定係数の分散について考察しよう．モデル(1.1)の未知パラメータ β の OLS 推定量 $\hat{\beta}$ の分散・共分散は先述のように(1.4)で与えられる． $V(\hat{\beta})$ の 1-1 要素，つまり $\hat{\beta}_1$ の分散 $var(\hat{\beta}_1)$は，

$$(1.20)\quad var(\hat{\beta}_1)=\sigma_u^2\frac{|X_2'X_2|}{|X'X|}$$

である．ここで(1.11)より

$$(1.21)\quad |X'X|=\begin{vmatrix}\hat{\alpha}'X_2'X_2\hat{\alpha}+\hat{v}'\hat{v} & \hat{\alpha}'X_2'X_2\\ X_2'X_2\hat{\alpha} & X_2'X_2\end{vmatrix}$$

$$=|X_2'X_2|\,|\hat{\alpha}'X_2'X_2\hat{\alpha}+\hat{v}'\hat{v}-\hat{\alpha}'X_2'X_2(X_2'X_2)^{-1}X_2'X_2\hat{\alpha}|$$

$$=|X_2'X_2|\,|\hat{v}'\hat{v}|$$

なることに注意して，

$$(1.22)\quad var(\hat{\beta}_1)=\frac{\sigma_u^2}{|\hat{v}'\hat{v}|}$$

が導出される．したがって **1.3** で示したように多重共線性の程度を(1.11)の形式から $\hat{v}'\hat{v}$ の大きさで測定すると，完全な多重共線性に近い程 $\hat{\beta}_1$ の分散は大きくなる．換言すれば，推定値 $\hat{\beta}_1$と真値β_1との乖離は多重共線性の程度が強くなるに従って大きくなる傾向があり，推定値の精度は落ちることになる．

しかし，$\hat{\beta}_1$の標本分散，つまり $var(\hat{\beta}_1)$ の推定値が多重共線性によって大きくなるということは(1.22)では不明である．それは(1.22)には未知パラメータ σ_u^2 を含んでいるからであり，$\hat{V}(\beta)$の推定値を考察するには (1.6) に示されたように，σ_u^2 の推定値 $\hat{\sigma}_u^2$ を検討しなければならない．しかし，(1.5) に示される σ_u^2 の推定量 $\hat{\sigma}_u^2$ が，$\hat{v}'\hat{v}$ の値が 0 に近ずくことによってその値がどのようになるかに一義的な判断はできない．換言すれば，推定係数の分散の推定値が多重共線性によって影響される方向については[1)]，明確な結論はなされ得ない．

---

1) σ_u^2 の推定量 $\hat{\sigma}_u^2$ は，説明変数x_1， X_2 の y に対してもつ説明力と関係があって，それと多重共線性とに一義的な関係がないということは **1.4** の議論と同様である．

§2

多重共線性の定義を説明変数間の統計的仮説検定によってなそうとする試みがある．Farrar-Glauber[1)]による検定方式がそれであるが，その検定方式を概観してみよう．

2.1 **1.2** でわれわれは説明変数間の完全な多重共線性について考察したが，この完全な多重共線性を1つの極端なケースとするならば，これと反対の極は，説明変数Xの直交性である．つまり，説明変数のいずれの2変数も全く統計的に独立であるケースである．[2)]

多重共線性に関する第1の検定方式の帰無仮説は，

H_0； 説明変数は直交している．

である．この帰無仮説が統計学的に有意に棄却できるならば，説明変数間に多重共線性が存在すると判定するのである．

一方，説明変数の相関係数行列Rを，

$$(2.1)\qquad R=\begin{pmatrix} 1 & r_{12} & r_{13}\cdots\cdots r_{1k} \\ r_{21} & 1 & r_{23}\cdots\cdots r_{2k} \\ \vdots & & \ddots \quad \vdots \\ r_{k1}\cdots & \cdots\cdots & \cdots\cdots 1 \end{pmatrix}$$

とすると，一般に，

$$(2.2)\qquad 0\leq|R|\leq 1$$

であって，完全な多重共線性が存在する場合は，$|R|=0$ であり，諸説明変数が直交している場合，つまり帰無仮説が成立するならば，

$$(2.3)\qquad H_0\,;\ |R|=1$$

である．**§1**で議論したように，現実に得られる標本については完全な多重共線

---

1) D. E. Farrar & R. R. Glauber, "Multicollinearity in Regression Analysis," *The Review of Economics and Statistics*, Vol. 49, pp. 92—107.

2) 直交しているということは，

$$\sum_t x_{it}\ x_{jt}=0 \quad (i\neq j)$$

であって，したがって，$X'X$が対角行列となることである．

また，この§では，説明変数が正規確率変数であり，Xはその実現値行列である．**§3** の **3.1** 参照．

はないであろうし，あるいは直交性を満すことはないであろうから，標本相関係数行列$\hat{R}$が，$|\hat{R}|=0$ あるいは，$|\hat{R}|=1$ を満たすことは，ほとんどない．

さて，標本相関係数行列$\hat{R}$の行列式に関して，帰無仮説 (2.3) が成立するとき，近似的に，

$$(2.4)\quad \hat{\chi}^2=-\left[T-1-\frac{1}{6}(2k+5)\right]log|\hat{R}|\sim\chi^2\left(\frac{1}{2}k(k-1)\right)$$

が得られる.[1] したがって，自由度$\frac{1}{2}k(k-1)$ の χ^2 分布をする変量 χ^2 について

$$Pr\{\chi^2\geq\chi_\alpha{}^2\}=\alpha$$

を満たす $\chi_\alpha{}^2$ の値に対して，

$$(2.4)\quad \hat{\chi}^2\leq\chi_\alpha{}^2$$

ならば，帰無仮説を受容する．逆に，

$$(2.5)\quad \chi_\alpha{}^2\leq\hat{\chi}^2$$

であれば，帰無仮説を棄却する．この検定方式によって，説明変数間の標本相関係数行列の行列式$|\hat{R}|$が (2.5) を満たす場合多重共線性が存在すると判定するのである．他方，(2.4) が満たされるときには，説明変数間に有意な多重共線性はないと判断するのである．

2.2　**2.1**で考察した多重共線性の検出に関する検定方法は，個々の説明変数や，あるいはその一部の共線関係の有無を検出するのではなく，説明変数全体の中に共線性が有るか否かを検出しようとするものである．

一方，**1.3** で考察したような，説明変数間に 1 つの説明変数が他の説明変数の影響を受けるという多重共線関係を検出することも考察さなければならない．

たとえば，説明変数間に(1.11)の関係を規定し，(1.13)に示したように，全く多重共線関係がない場合は，$\hat{v}'\hat{v}=x_1'x_1$ であり，完全な多重共線関係がある場合は $\hat{v}'\hat{v}=0$ である．この **1.3** での考察の結果から，上述の多重共線関係の検出をする接近として以下のことが考えられる．

1) *ibid.*, p. 101.

説明変量間の真の関係として，

(2.6)　$x_1=\alpha_2x_2+\alpha_3x_3+\cdots\cdots\alpha_kx_k+v$

を想定し，帰無仮説 H_0 を

(2.7)　$H_0\,;\,v'v=x_1'\,x_1$

とする．つまり，帰無仮説は (2.6) に示される多重共線関係が全く存在しないということを表現しており，もし得られた標本によってこの帰無仮説が有意に棄却されると多重共線関係が存在すると判断するのである．

一方，帰無仮説(2.7)は，真のパラメータ $\alpha_2, \alpha_3, \cdots, \alpha_k$ が，

(2.8)　$\alpha_2=\alpha_3=\cdots=\alpha_k=0$

であることを意味しており，(2.8)が成立するもとに，標本から得られる $\hat{v}'\hat{v}$ に関して，

(2.9)　$\hat{F}=\dfrac{(x_1'x_1-\hat{v}'\hat{v})/(k-1)}{\hat{v}'\hat{v}/(T-k)}\sim F(k-1,\ T-k)$

が導出されることは周知の通りである.[1)]

自由度 $k-1$，$T-k$ の F 分布をする変量 F に対して，

$$Pr\{F_\alpha\leq F\}=\alpha$$

となる F_α を得て，有意水準 α の帰無仮説(2.7)の検定方式は，

(2.10)　$F_\alpha\leq\hat{F}$

ならば仮説を棄却する．つまり，説明変数の標本から得られる統計量 $\hat{F}$ が (2.10)を満たすならば，x_1 がそれ以外の説明変数に影響されるという多重共線関係が存在すると判断するのである．

逆に，標本が

(2.11)　$\hat{F}<F_\alpha$

---

1)　(1.11)における重相関係数 $\hat{R}^2x_1.x_2$ と $\hat{v}'\,\hat{v}$ の関係は，

$$R^2x_1.x_2=1-\hat{v}'\,\hat{v}/x_1'x_1$$

である．したがって，帰無仮説(2.7)あるいは(2.8)は，

$$H_0\,;\,R^2x_1.x_2=0$$

と同等である．

を満たすならば，帰無仮説が受容され，有意な多重共線関係は存在しないと判定するのである．

ここで注意されなければならないのは，上述の考察はx_1と他の変数間との多重共線関係の検出を前提にしている．したがって，実際には (2.6) の左辺の変量を順次 x_2，x_3，……x_k と変更し，それぞれのケースについて上述と同様の検定を試みる必要がある．

2.3　**2.2** で考察した検定の帰無仮説は(2.8)となることはすでに述べたが，いまこの検定で帰無仮説が棄却され，説明変数間に (2.6) で示すような多重共線関係が存在すると判定されたとしよう．他方，たとえば，

$$x_k=\gamma_1 x_1+\gamma_2 x_2+\cdots\cdots\gamma_{k-1}x_{k-1}+v$$

なるパターンの多重共線関係はないと判断されたとき，(2.6)の型式において，真の関係としては，$\alpha_k=0$ であると考えられるであろう．つまり，(2.6) において，各々の α_i $(i=2,\cdots,k)$ が有意に0であるか否かが問題となる．換言すれば，(2.8)に示される仮説が棄却されても，その多重共線関係が(2.6)の型式ではなく，上述の例では，実は，

$$x_1=\alpha_2 x_2+\cdots\cdots+\alpha_{k-1}x_{k-1}+v$$

がその多重共線関係のパターンであるケースがある．これより，(2.6) における各々の係数 α_i $(i=2,\cdots k)$が0と有意に異なるか否かをも検討しなければならないのである．

多重共線関係の第3の検定方式における帰無仮説は，(2.6) の各々の係数$\alpha_i(i=2,\cdots k)$について，

(2.12)　H_i；　$\alpha_i=0$

となる.[1)]

1)　ここで α_i と表記しているのは，厳密な偏回帰係数の書き方に従うならば

$$\alpha_{1i.2\cdots\cdots i-1\ i+1\cdots k}$$

あるいは，

$$\alpha_{1i.2\cdots\cdots)i(\cdots\cdots k}$$

とすべきである.

さらに，帰無仮説(2.12)は **Farrar-Glauber** の提案している偏相関の検定と基本的には同等である.

2. 2 で考察した多重共線関係に関する第2の検定は，上述の帰無仮説 (2. 12) の検定と同時に採用する必要がある．換言すれば，第2の検定方式と第3のそれとは併行的に採用して多重共線関係の検出を試みるものと見做される．

さて，周知の様に，$(X_2'X_2)^{-1}$ の第 $i-i$ 要素を$(X_2'X_2)^{-1}{}_{ii}$で表わすならば，帰無仮説(2. 12)のもとに，

$$(2.13)\quad \hat{t}_i=\frac{\hat{\alpha}_i}{\sqrt{(X_2'X_2)^{-1}{}_{ii}}}\bigg/\sqrt{\frac{\hat{v}'\,\hat{v}}{T-k+1}}\sim t(T-k+1)$$

が得られる．このことから，帰無仮説(2. 12)の検定が可能になる．つまり，自由度 $T-k+1$ の t 分布をする t に関して

$$Pr\{t_\alpha\leq|\,t\,|\}=\alpha$$

なる t_α の値を用いて，有意水準 α での帰無仮説 (2. 12) の検定方式は次の様に結論する．

もし，(2. 13)の統計量 $\hat{t}_i$ が，

$$t_\alpha\leq|\hat{t}_i|$$

ならば，帰無仮説を棄却し，x_1とx_i は有意に多重共線関係にあると判定できるが，

$$|\hat{t}_i|<t_\alpha$$

を満すならば，x_1 と x_i とは有意な多重共線関係にはないと判定するのである．

2. 4 上に考察した多重共線性の検定方式によるその定義は，基本的には Farrar-Glauber の提案しているものである．これは標準的な回帰分析の手法を説明変数間に適用させるという点で採用されやすいであろう．これに反して，Frisch の線束図による分析[1)]は，その適用に際して，多くの計算と図示を要し，又それによっては，多重共線性に関しての決定的な結論を導くことが困難な場合も多

1) 線束図分析 (*Bunch Map Analysis*) を本稿ではとり挙げなかったが，
Beach, E. F., *Economic Models—An Exposition*—,(John Wiley & Sons, 1957), pp. 197-208,
Valavanis, S., *Econometrics* (McGraw-Hill, 1959) pp. 175-180.
の解説がわかりやすい．

く有り得る．しかし，ある変数を説明変数として関係式に加えた場合，その変数の他の説明変数の推定値やその標準誤差に与える効果を測定し，その変数の導入の良否を判定するという点は長所といえるであろう．この点を考慮するならば，**2.2**及び**2.3**に述べたF検定，t検定は，説明変数を1変数ずつ増加させていく過程でそれぞれ採用するという仕方がよいと言えるであろう．

また，上述の考察では説明変数間の多重共線性のみに注意して本来推定したい関係式における被説明変数yと説明変数Xとの関係に多重共線性が与える影響を無視している傾向がある．一般的に説明変数間の多重共線性が被説明変数に与える影響も少なくない場合もあろう．

たとえば，いま被説明変数yに対して説明変数となり得る変数x_1，x_2，x_3がある場合，$R^2_{y\cdot123}$と$R^2_{y\cdot12}$を比較して，この差が微少なものであるならば変数x_3が，ほとんど被説明変数yに対して説明力をもたない変数であるかもしれない．しかし，変数x_3は他の説明変数x_1，x_2とに多重共線関係にあるために生ずる結果とも考えられる．したがって，説明変数のみの検定だけでなく，説明変数の被説明変数への効果をも多重共線性の検出やその定義を与える検定に際しても考慮に入れる必要があるであろう．

§3

3.1 **§1**で既に指摘したように，多重共線性の定義について必らずしも決定的な統一的な理論はないが，モデル(1.1)の分析に関して，多重共線性の問題の取り扱いについてもいくつかの異なる方向があることを明確にしなければならない．

まず，モデル(1.1)において，通常なされるように攪乱項と説明変数とは独立であると前提しても，説明変数自体が定数（*fixed variables*）とするか，確率変数と考えられるかによって取り扱い方は異なる．**§2**で考察した検定方式は，説明変数を定数とした場合は形式的な適用にすぎないのであって，本来，説明変数が確率変数である場合にのみ統計学的に意味があるのである．

第2に，§2の考察は説明変数間の多重共線性の有無の判定と同時にその関係の検出を意図するが，実際に共線性が在ると判断せられたときに，モデル(1.1)の分析にどの様に対処するかが問題になる.[1] ①全変数の標本のいくつかを，もしくはある変数の標本を全部捨てるか，②多重共線性を回避するために標本を追加することができるのか，③他の分析の仕方によって多重共線性の問題を回避できるか，④多重共線性が存在しても，現在の標本からモデル分析しなければならないのか，等を考えなければならない.

最後に，上述の④に述べた様な，たとえ多重共線性が在ると判定されても，現在の標本によって(1.1)のモデルを考察しなければならない場合，被説明変数に与える説明変数の個々の影響を推定するのか，あるいは，被説明変数の予測のために(1.1)を分析しているのかの差異がある．たとえば，後者の場合，つまり予測のためにモデルを分析するときには個々の係数推定よりも予測値を導出することに主眼がおかれるだろう.

われわれが，多重共線性の問題を考える際，上述の分析の目的を充分明確にする必要がある.

3.2　この節では，モデル(1.1)によって被説明変数の予測値が考察の対象になっている場合を取り挙げる．モデル(1.1)によって被説明変数の予測をするということは，現在のデータとは別に，説明変数のデータが先決的に，

(3.1)　$\bar{x}'=(\bar{x}_1,\ \bar{x}_2, \cdots\cdots, \bar{x}_k)$

と与えられたとき，それに対応する未知の被説明変数

(3.2)　$y_f=\bar{x}'\beta+u_f$

の期待値

(3.3)　$y_e=E(y_f)=\bar{x}'\beta$

1)　多重共線性のある場合の具体的な対処の方法については，たとえば，Koutsoyiannis, A., *Theory of Econometrics* (Macmillan, 1973), pp. 241-244 に整理されている.

を推定することである.[1]

このとき，y_e の推定量として，最小二乗予測量 $\hat{y}_e$ は，

(3.4)　$\hat{y}_e=\bar{x}'\hat{\beta}$

である．(3.4)における $\hat{\beta}$ はもちろん(1.3)に与えられた $\hat{\beta}$ である．$\hat{\beta}$ が β の不偏推定量となっているのと同様に，

(3.5)　$E(\hat{y}_e)=E(\bar{x}'\hat{\beta})=\bar{x}'E(\hat{\beta})=\bar{x}'\beta$

であるから y_e の予測量 $\hat{y}_e$ も不偏性を保持している．OLS 推定量 $\hat{\beta}$ と同様に OLS 予測量 $\hat{y}_e$ の不偏性はその統計学的に顕著な特性といえる．

一方，平均平方誤差基準に照らすとき，特に説明変数間に多重共線性が存在するならば，OLS 予測量よりも好ましい予測量があり得る．

変数 x_1 がそれ以外の説明変数 X_2 と多重共線関係にある場合，y を X_2 に回帰させて，

(3.6)　$\hat{\gamma}=(X_2'X_2)^{-1}X_2'y$

を得，予測量 $\tilde{y}_e$ を

(3.7)　$\tilde{y}_e=\bar{x}_2'\hat{\gamma}$

とする．(3.7)で $\bar{x}_2'$ は $\bar{x}'$ の第2要素以下を含む $\bar{x}'$ の$(k-1)$次部分ベクトルである．

さて，二つの y_e の予測量 $\hat{y}_e$ と $\tilde{y}_e$ を平均平方誤差基準に照らして比較しようとするとき，次の定理が有用である．

[定理3.1][2] もし，

(3.8)　$\beta_1^2\leqq V(\hat{\beta}_1)$

ならば，$\hat{y}_e$ の平均平方誤差 $MSE(\hat{y}_e)$ と $\tilde{y}_e$ の平均平方誤差 $MSE(\tilde{y}_e)$ との間に，

---

1)　もちろん，y_f の推定量 $\hat{y}_f$ を求めることもあるが一般的には $E(u_f)=0$ であるかぎり，$\hat{y}_f=\hat{y}_e$ であろう．しかし $\hat{y}_f$，と $\hat{y}_e$ の分散は異なるから，推定区間は一般的には異なるだろう．

2)　説明変数が2個の場合については，竹内啓，「計量経済学の研究」，(東洋経済新報社，昭和47年9月) pp. 23-24にある．

(3.9)　$MSE(\tilde{y}_p) \leqq MSE(\hat{y}_p)$

の関係がある.

定理3.1によって,(3.8)が成立するとき,つまり,

(3.10)　$\dfrac{\beta_1^2}{V(\hat{\beta}_1)} \leqq 1$

のときには,平均平方誤差基準では y_e の予測量 $\tilde{y}_e$ の方が,OLS 予測量 $\hat{y}_e$ よりも好ましいと結論できる.さらに,§1 で与えた $\hat{v}'$ を用いて

(3.11)　$V(\hat{\beta}_1) = \sigma_u^2 / \hat{v}'\hat{v}$

が得られるから,$\hat{v}'\hat{v}$ が小さいときには,換言すれば,多重共線関係が強くなればなるだけ,変量 x_1 を含めない予測量 $\tilde{y}_e$ の方が OLS 予測量 $\hat{y}_e$ より好ましい.

一方,(3.10),(3.11)より判るように,その判定には未知パラメータβ_1,σ_u^2 を含んでいるので,それらの推定量

(3.12)　$\hat{t}^2 = \hat{\beta}_1^2 \Big/ \dfrac{\hat{\sigma}_u^2}{\hat{v}'\hat{v}}$

を平均平方誤差基準での予測量の判定統計量となし得るであろう.

3.3　この節は **3.2** で与えた定理 3.1 の証明とそれと関連する考察にあてる.

モデル(1.1)を

(3.12)　$y = [x_1\ X_2]\begin{pmatrix}\beta_1\\ \beta_2\end{pmatrix} + u$

の様に表現し,

(3.13)　$\hat{\alpha} = (X_2'X_2)^{-1}X_2'x_1$

(3.14)　$\hat{v} = x_1 - X_2(X_2'X_2)^{-1}X_2'x_1$

を用いて,(3.12)をそれと同値な関係

(3.15)　$y = \hat{v}\beta_1 + X_2(\hat{\alpha}\beta_1 + \beta_2) + u$

で表わす.さらに,

(3.16)　$Z = [\hat{v}\ \ X_2]$

(3.17)　$\delta=\begin{pmatrix}\delta_1\\ \delta_2\end{pmatrix}=\begin{pmatrix}\beta_1\\ \hat{\alpha}\beta_1+\beta_2\end{pmatrix}$

を定義すると，モデル(1.1)は同等のモデル

(3.18)　$y=Z\delta+u$

で表現できる.

(3.18)における δ の OLS 推定量 $\hat{\delta}$ は

(3.19)　$\hat{\delta}=\begin{pmatrix}\hat{\delta}_1\\ \hat{\delta}_2\end{pmatrix}=(Z'Z)^{-1}Z'y$

であって，(1.1)における β の OLS 推定量 $\hat{\beta}$

(3.20)　$\hat{\beta}=\begin{pmatrix}\hat{\beta}_1\\ \hat{\beta}_2\end{pmatrix}=(X'X)^{-1}X'y$

との間に，真のパラメータでの関係(3.17)と同様な関係

(3.21)　$\hat{\delta}_1=\hat{\beta}_1$

$\hat{\delta}_2=\hat{\alpha}\hat{\beta}_1+\hat{\beta}_2$

が導出される．これは次の様に示される.

つまり(3.20)を導出する正規方程式より

(i)　$x_1'x_1\hat{\beta}_1+x_1'X_2\hat{\beta}_2=x_1'y$

(ii)　$X_2'x_1\hat{\beta}_1+X_2'X_2\hat{\beta}_2=X_2'y$

が得られる．また(3.19)を導出する正規方程式より，

$$Z'Z=\begin{bmatrix}\hat{v}'\hat{v} & 0\\ 0 & X_2'X_2\end{bmatrix}$$

なることに注意して，

(iii)′　$\hat{v}'\hat{v}\hat{\delta}_1 \qquad\qquad =\hat{v}'y$

(iv)　$\qquad\qquad X_2'X_2\hat{\delta}_2=X_2'y$

が得られる．(iii)′を書き換えると，

(iii)　$x_1'[I-X_2(X_2'X_2)^{-1}X_2']x_1\hat{\delta}_1=[I-X_2(X_2'X_2)^{-1}X_2']x_1'y$

である.

(ii)より $\hat{\beta}_2$ を導出し，(i)に代入すれば，

$$x_1'x_1\hat{\beta}_1+x_1'X_2(X_2'X_2)^{-1}[X_2'y-X_2'X_1\hat{\beta}_1]=x_1'y$$

が与えられ，これと(iii)とを比較して

$$\hat{\delta}_1=\hat{\beta}_1$$

が導出される．一方，(ii)より，

$$(X_2'X_2)^{-1}X_2'x_1\hat{\beta}_1+\hat{\beta}_2=(X_2'X_2)^{-1}X_2'y$$

であり，(3.13)に注意して，

$$\hat{\alpha}\hat{\beta}_1+\hat{\beta}_2=(X_2'X_2)^{-1}X_2'y$$

他方，(iv)より

$$\hat{\delta}_2=(X_2'X_2)^{-1}X_2'y$$

が得られ，したがって(3.21)の第二式が導出された．

さて，$\hat{\delta}$ について

$$\hat{\delta}=(Z'Z)^{-1}Z'y=\delta+(Z'Z)^{-1}Z'u$$

であるから，

(3.22) $E(\hat{\delta})=\delta$

(3.23) $V(\hat{\delta})=\sigma_u^2(Z'Z)^{-1}=\sigma_u^2\begin{pmatrix}\hat{v}'\hat{v}^{-1} & 0\\ 0 & (X_2'X_2)^{-1}\end{pmatrix}$

が得られる．

一方，$\bar{x}'$ について上述のように

$$\bar{x}'=(\bar{x}_1\ \ \bar{x}_2')$$

であって，いま，

$$\bar{z}'=(\bar{v},\ \ \bar{x}_2')$$

$$\bar{v}=\bar{x}_1-\bar{x}_2'\hat{\alpha}$$

とすると，(3.17)の関係を利用して，

(3.24) $y_e=\bar{x}'\beta=\bar{z}'\delta$

が，さらに(3.21)を利用して，

(3.25) $\hat{y}_e=\bar{x}'\hat{\beta}=\bar{z}'\hat{\delta}$

が導ける．したがって，$\hat{y}_e$ の平均平方誤差 $MSE(\hat{y}_e)$は，

$$MSE(\hat{y}_e)=MSE(\bar{z}'\hat{\delta})=E(\bar{z}'\hat{\delta}-\bar{z}'\delta)^2=z'V(\hat{\delta})z$$

である．上式に(3.23)を代入すると，

(3.26) $$MSE(\hat{y}_e)=\sigma_u^2\left(\frac{\bar{v}^2}{\hat{v}'\hat{v}}+\bar{x}_2'(X_2'X_2)^{-1}\bar{x}_2\right)$$

が導ける．

他方，(3.6) で与えられた $\hat{\gamma}$ に対して，

$$\hat{\gamma}=(X_2'X_2)^{-1}X_2'y=\delta_2+(X_2'X_2)^{-1}X_2'u$$

であるから，

(3.27) $E(\hat{\gamma})=\delta_2$

(3.28) $V(\hat{\gamma})=\sigma_u^2(X_2'X_2)^{-1}$

が明らかになる．(3.27)より，y_eの予測量 $\tilde{y}_e$ は

$$E(\tilde{y}_e)=E(\bar{x}_2'\hat{\gamma})=\bar{x}_2'\delta_2$$

となり，一般的には，

$$\bar{x}_2'\delta_2\neq\bar{v}\delta_1+\bar{x}_2'\delta_2=z'\delta=y_e$$

であるから，$\tilde{y}_e$ は y_e の不偏予測量でないことがわかる．

さて，(3.27)，(3.28)を利用して，

(3.29) $$MSE(\tilde{y}_e)=E(\bar{x}_2\hat{\gamma}-\bar{z}'\delta)^2=E(\bar{x}_2\hat{\gamma}-x_2'\delta_2-\bar{v}\beta_1)^2$$
$$=\bar{x}_2'V(\hat{\gamma})\bar{x}_2+(\bar{v}\beta_1)^2=\sigma_u^2\bar{x}_2'(X_2'X_2)^{-1}\bar{x}_2+(\bar{v}\beta_1)^2$$

が導出でき，(3.26)と(3.29)を比較して定理 3.1 が明らかとなる．