1 部 3 章

79 頁　最小2乗推定量の導出:追加説明(初級)

(3.1)式から(3.8)式を導くには,まず人工的に設定する単回帰式

$$x_i = c_x + \omega_x z_i + \text{誤差}$$

を最小 2 乗推定する.これは 1 章で学んだ単回帰であり

$$\hat{\omega}_x = \frac{\sum_{i=1}^{n} (z_i - \bar{z}) x_i}{\sum_{i=1}^{n} (z_i - \bar{z})^2}, \qquad \hat{c}_x = \bar{x} - \hat{\omega}_x \bar{z}$$

と推定量が定まる.その結果,回帰残差として

(3.7) $$x_i^* = x_i - \hat{c}_x - \hat{\omega}_x z_i$$

がえられる.上式を用いて,(3.1)式の x_i を $(x_i^* + \hat{c}_x + \hat{\omega}_x z_i)$ で置きかえて整理すれば(3.8)式が導かれる.これは残差であるから、残差と説明変数間の直交性により、 $\sum x_i^* = 0$, $\sum x_i^* z_i = 0$, となる。

以上では「z 変数の影響を除く」ために x_i^* を定義している.この x_i^* を利用して回帰式(3.1)を書き換えれば,回帰式は

(3.8) $$y_i = \alpha^* + \beta x_i^* + \gamma^* z_i + \varepsilon_i$$

と表現できる.この変換において注意すべき点は,β係数と z 変数は(3.1)式から変化していないことである.また言うまでもなく α^* は定数項 1 の係数であるが,定数項 1 も変化していない.

ここで最小 2 乗法の原則にもどり、β の最小 2 乗推定量を導く。

$\Phi = \sum_{i=1}^{n} (y_i - \alpha^* - \beta x_i^* - \gamma^* z_i)^2$ の β に関する 1 次条件は,残差の性質 $\sum x_i^* = 0$, $\sum x_i^* z_i = 0$、により

$$0 = \sum_{i=1}^{n} (y_i - \hat{\alpha}^* - \hat{\beta} x_i^* - \hat{\gamma}^* z_i) x_i^* = \sum_{i=1}^{n} y_i x_i^* - \hat{\beta} \sum_{i=1}^{n} (x_i^*)^2$$

と簡略化され, $\hat{\beta}$ (3.9)式がもとまるのである.

x_i^* の様な変換を用いて偏回帰係数βの最小 2 乗推定量を導出するが,変換の根拠は**標本偏相関係数**である.標本偏相関係数は3個以上の変数の結び付き具合を示す指標で,例えば $r_{xy|z}$ は z 変数の影響を除いた x と y の間の標本相関を意味する.そして z 変数の影響を除くとは,

1) x 変数を z に回帰して残差を求める.

2) y 変数も同じく z に回帰して残差を求める.

という二つの操作を意味する.1)と 2)の残差間で計算した標本相関係数が標本偏相関係数 $r_{xy|z}$ である.式で表現すると

$$x_i^* = x_i - \hat{c}_x - \hat{\omega}_x z_i$$

が 1)であり, x_i^* と同じく y_i を z_i に回帰した残差

$$y_i^* = y_i - \hat{c}_y - \hat{\omega}_y z_i$$

が 2)である.そして偏相関係数とは、残差間の相関係数

$$r_{xy|z} = \frac{\sum_{i=1}^{n} x_i^* y_i^*}{\sqrt{\sum_{i=1}^{n} (x_i^*)^2 \sum_{i=1}^{n} (y_i^*)^2}}$$

である.同じく z の影響を除いた x の分散は偏分散といわれ

$$S_{xx|z}^2 = \frac{1}{n}\sum_{i=1}^{n} (x_i^*)^2$$

と定義される.z の影響を除いた x と y の共分散は偏共分散といわれ

$$S_{xy|z} = \frac{1}{n}\sum_{i=1}^{n} x_i^* y_i^* = \frac{1}{n}\sum_{i=1}^{n} x_i^* y_i$$

である.（x_i^* は回帰の残差であるので $\Sigma x_i^* y_i^* = \Sigma x_i^* y_i$ が成立する.）

偏相関係数を用いれば、$\hat{\beta}$ は

$$\hat{\beta} = r_{xy|z}\sqrt{\frac{S_{yy|z}}{S_{xx|z}}}$$

となる.偏回帰係数という用語の意味も理解できよう.**81 頁**

他の係数の推定量も全く同様に考えればよい.特に右辺の変数の順番は自由に入れ替えてよいから,z 変数が右辺の第 2 項であれば,$\hat{\beta}$ の導出は $\hat{\gamma}$ の導出に他ならない.$\hat{\alpha}$ についても同じである.

84-85 頁　自由度について:追加説明(中級)　　　9/7/2001

χ^2 確率変数の定義にもどれば自由度の意味は明白であるが,初等的に TSS,ESS,RSS の自由度を説明するのは難しい.

1.　$y_1, y_2, \cdots, y_n$ は全て自由に変動する確率変数であるから,$\sum_{i=1}^{n} y_i^2$ の自由度は n である.

2.　$y_1-\bar{y},\ y_2-\bar{y}, \cdots, y_n-\bar{y}$, の和は 0 だから,$\sum_{i=1}^{n}(y_i-\bar{y})^2$ の自由度は $(n-1)$ である.例えば $(y_1-\bar{y})$ から $(y_{n-1}-\bar{y})$ の値が決まれば,$(y_n-\bar{y})$ の値も総和が 0 になるように定まる.

3.　残差は,

(1)　　$$\sum_{i=1}^{n} e_i x_{ki} = 0\ , \quad k = 1, 2, \cdots, K\ ,$$

となり,説明変数と直交している.x 変数値は所与,残差を未知数と考えれば,K 式(1)に n 未知数が含まれている.ここで $(n-K)$ 個の残差,例えば e_1 から e_{n-K} の値を決めれば,残りの K 残差は K 式の解として定まる.従って自由度は $(n-K)$ である.

4.　回帰値の自由度は K である.最初の K 回帰値について

(2) $\hat{y}_i = \hat{\beta}_1 x_{1i} + \hat{\beta}_2 x_{2i} + \cdots + \hat{\beta}_K x_{Ki} \qquad i = 1, \cdots, K$

と定義される.(2)に含まれる K 個の回帰係数を K 式の未知数と見なせば, $\hat{y}_1, \cdots, \hat{y}_K$ の値を決めれば全ての回帰係数値が決まる.（x 値は所与とする.）この回帰係数値を用いて,残りの $(n-K)$ 個の回帰値も決まる.したがって自由度は K である.

5. $ESS = \sum_{i=1}^{n} (\hat{y}_i - \bar{y})^2$ は,TSS と同じく自由度は $(K-1)$ である.

93 頁 25/6/2001

「多変数回帰式における BLUE 定理」:追加説明(上級)

(3.14) $y_i = \beta_1 X_{1i} + \cdots + \beta_K X_{Ki} + \varepsilon_i, \quad i = 1, 2, \cdots, n$

と回帰式を表現し, β_1 の最小2乗推定量を本文中(3.16)式とそのすぐ下の $\hat{\beta}_1$ 式で定義します.この $\hat{\beta}_1$ が BLUE であることの証明は本文の通りです.その証明は2章5 節の証明と何ら変わるところはありません.回帰式(3.14)の右辺に含まれる K 項の順番は自由ですから, $\hat{\beta}_1$ の BLUE 性は全ての係数の BLUE 性を意味しています.

この証明は,通常の教科書で扱われるベクトル推定量の最適性の証明と同じ意味を持ちます.以下これを示します.

まずベクトル行列表示を用いて BLUE 定理を示します.回帰式は

$$\mathbf{y} = \mathbf{X}\boldsymbol{\beta} + \boldsymbol{\varepsilon}$$

となり,最小 2 乗推定量は

$$\hat{\boldsymbol{\beta}} = (\mathbf{X'X})^{-1}\mathbf{X'Y},$$

その共分散行列は

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

となります.他方,行列表示による任意の線型不偏推定量は, $\mathbf{C'}$ を $K \times n$ の定数行

列として,

$\mathbf{b}=\mathbf{C'Y}$,　　　$\mathbf{C'X}=\mathbf{I}$

となります.後者は不偏性の条件です.例えば,最小2乗推定量では,$\mathbf{C'}=(\mathbf{X'X})^{-1}\mathbf{X'}$と定義されています.b の共分散行列は

$\mathbf{V(b)}=\sigma^2\,\mathbf{C'C}$

と導かれます.$\mathbf{V}(\hat{\beta})$が最小であることの証明は

$\mathbf{Q}=\mathbf{C'C}-(\mathbf{X'X})^{-1}$

が非負定符号であることを示せば完了します.(例えば Johnstone 的な証明は,$\mathbf{C'}$と$(\mathbf{X'X})^{-1}\mathbf{X'}$の差を$\mathbf{D'}$と定義する:$\mathbf{C'}=(\mathbf{X'X})^{-1}\mathbf{X'}+\mathbf{D'}$.ところが,不偏性$\mathbf{C'X}=\mathbf{I}$により$\mathbf{D'X}=0$となる.　従って,$\mathbf{C'C}=(\mathbf{X'X})^{-1}+\mathbf{D'D}$.だから$\mathbf{D'}=0$ の時,分散最小となる.あまりおもしろくない.おもしろい証明は Wu-Hausman 検定でも用いられる分解で,不偏性により$\mathbf{COV}(\mathbf{b}-\hat{\beta},\hat{\beta})=0$だから,

$\mathbf{V}(\mathbf{b})=\mathbf{V}(\mathbf{b}-\hat{\beta}+\hat{\beta})=\mathbf{V}(\mathbf{b}-\hat{\beta})+\mathbf{V}(\hat{\beta})+\mathbf{COV}(\mathbf{b}-\hat{\beta},\hat{\beta})=\mathbf{V}(\mathbf{b}-\hat{\beta})+\mathbf{V}(\hat{\beta})$.)

以上が,ベクトル行列表示による,BLUE 定理の証明です.

以下,本書とベクトル行列分析の関係を述べます.

$\mathbf{Q}$が非負定符号であることは,任意の$K\times1$の実ベクトル w について

$\mathbf{w'Qw}\geq 0$

が非負値であることを意味します.$\mathbf{V}(\hat{\beta})$や$\mathbf{V(b)}$の構成を考えれば,これは$\mathbf{w'}\beta$の推定量$\mathbf{w'}\hat{\beta}$の分散が最小である事を意味します.同じ結果を(3.14)式から示します.w のノルムは自由ですから,w の第一要素を1として$w=(1,w_2,w_3,\cdots,w_k)$と記せば,(3.14)式は整理できて

(3.14)'　$y_t=(\beta_1+\beta_2\,w_2+\cdots+\beta_k\,w_k)\,X_{1t}+\beta_2\,(X_{2t}-w_2\,X_{1t})+$

$\cdots+\beta_k(X_{kt}-w_k\,X_{1t})+\varepsilon_t$

$$=\beta_1^* X_{1t}+\beta_2 X_{2t}^*+\cdots+\beta_k X_{kt}^*+\varepsilon_t$$

となります. したがって $\mathbf{w}'\beta$ の推定量 $\mathbf{w}'\hat{\beta}$ の最適性は、β_1^* についての最小 2 乗推定量の最適性になっています. つまり, 本書で示している最適性は, ベクトル推定量の最適性と同値です. **(お終い)**

94 頁:98 頁 4 節のベクトル行列分析 22/ 9/2001

(3.47)式の証明: 回帰式 (3.43)と(3.44)を

(3.43) $y = x_a\beta_a + X\beta + \varepsilon$

(3.44) $y = X\beta + \varepsilon$

と表記する. X は $n\times(K-1)$, x_a は $n\times 1$ である. (3.43)式は,(3.45)の変換を使い

(3.43)' $y = x_a^*\beta_a + X\beta^* + \varepsilon$, $x_a^* = Mx_a$, $M = I - X(X'X)^{-1}X'$

と変型できる. x_a^* は X と直交する. 変換(3.46)式は

(3.46)' $y^* = My$

に等しい. 偏相関係数は $r_{ya|z} = \dfrac{(x_a^{*\prime}y^*)}{\sqrt{(x_a^{*\prime}x_a^*)(y^{*\prime}y^*)}}$ である. また(3.43)'に最小 2 乗法を応用すると $\hat{\beta}_a = \dfrac{x_a^{*\prime}y^*}{x_a^{*\prime}x_a^*} = r_{ya|z}\cdot\sqrt{\dfrac{y^{*\prime}y^*}{x_a^{*\prime}x_a^*}}$ となる. $\hat{\beta}^* = (X'X)^{-1}X'y$ だから,(3.43)式の推定からもたらされる残差は,

(*)
$$\begin{aligned} &y - x_a^*\hat{\beta}_a - X\hat{\beta}^* = (M - x_a^*(x_a^{*\prime}x_a^*)^{-1}x_a^{*\prime})y \\ &(= My - \frac{x_a'My}{x_a'Mx_a}Mx_a) \end{aligned},$$

である. 括弧内の表現は後で用いる。この式は,**変数追加がもたらす残差の変動式**として有用である. x_a^* の定義により $x_a^{*\prime}X = 0$, また $\hat{\beta}_a$ の定義により残差平方和 RSS は

(3.47)' $\quad \mathrm{RSS}(43) = \mathrm{RSS}(44) - y' x_a^* (x_a^{*\prime} x_a^*)^{-1} x_a^{*\prime} y$

$$= \mathrm{RSS}(44) - \hat{\beta}_a' (x_a^{*\prime} y^*)$$

となる.

(3.50)式の証明： さらに,

(3.47)'' $\quad \mathrm{RSS}(44) = (y - X\hat{\beta}^*)'(y - X\hat{\beta}) = y^{*\prime} y^*$

であることにも注意しよう.(3.47)'と(3.47)''式により(3.50)の証明は明らかで

(3.50) $$\frac{\mathrm{RSS}(44) - \mathrm{RSS}(43)}{\mathrm{RSS}(44)} = \frac{\hat{\beta}_a (x_a^{*\prime} y^*)}{y^{*\prime} y^*} = (r_{ya|z})^2$$

となる.

t 統計量： $S^2 = \mathrm{RSS}(43)/(n-K)$ により, t 統計量についても

$$t^2 = \left(\frac{\hat{\beta}_a}{\sqrt{S^2/(x_a^{*\prime} x_a^*)}} \right)^2 = \frac{\hat{\beta}_a^{\,2} (x_a^{*\prime} x_a^*)}{S^2} = \frac{\hat{\beta}_a (x_a^{*\prime} y_a^*)}{S^2}$$

であるから

(3.51) $$t^2 = \frac{\mathrm{RSS}(44) - \mathrm{RSS}(43)}{S^2} = (n-K) \frac{\mathrm{RSS}(44) - \mathrm{RSS}(43)}{\mathrm{RSS}(43)}$$

$$= (n-K) \frac{(\mathrm{RSS}(44) - \mathrm{RSS}(43))}{\mathrm{RSS}(44) - (\mathrm{RSS}(44) - \mathrm{RSS}(43))}$$

という等号が示される.分数の分母分子を RSS(44)で割れば,(3.50)式により

(3.52) $$t^2 = (n-K) \frac{r^2}{1-r^2}$$

となる.(3.53)は明らかである.

自由度調整決定係数：定義により

$$\overline{R}^2 = 1 - \frac{\mathrm{RSS}/(n-K)}{\mathrm{TSS}/(n-1)}$$

だから, $\overline{R}^2$ が増加することと $\mathrm{RSS}/(n-K)$ が減少することは等しい.したがって

(**)
$$\frac{\mathrm{RSS}(44)}{n-K+1} \geq \frac{\mathrm{RSS}(43)}{n-K}$$

であれば$\overline{R}^2$は増加する.(**)式を変型すると

$$\frac{\mathrm{RSS}(44)}{\mathrm{RSS}(43)} \geq \frac{n-K+1}{n-K} = 1 + \frac{1}{n-K}$$

であるから, $(n-K)\left(\frac{\mathrm{RSS}(44)}{\mathrm{RSS}(43)} - 1\right) \geq 1$となり,(3.51)によりこの条件は不等式

$$t^2 \geq 1$$

に一致する.

係数推定量間の相関係数：(3.54)式の証明

回帰式

$$y = X\beta + x_a\beta_a + x_b\beta_b + \varepsilon$$

において, x_aとx_bは共に$n \times 1$のベクトルで,以下ではβ_aとβ_bの最小 2 乗推定量間の相関係数を導く.まず$\hat{\beta}_a$を導くが,これはx_aを X とx_bに回帰した残差をx_a^*とすれば

$$\hat{\beta}_a = \frac{x_a^{*}{}'y}{x_a^{*}{}'x_a^*}$$

であり,(3.54)の最初の等号は容易に示すことができる.これは(3.38)式と同じだが,証明したいのは第 2 の等号である. x_a^*はx_aを X とx_bに回帰した残差だから, $M = 1 - X(X'X)^{-1}X'$ と定義して,本解説 6 頁の(*)式の括弧内の表現により

$$x_a^* = M x_a - \frac{x_b' M x_a}{x_b' M x_b} M x_b,$$

となる.同様にx_bを X とx_aに回帰した残差は

$$x_b^* = M x_b - \frac{x_a' M x_b}{x_a' M x_a} M x_a$$

となる.以上の結果を用いて計算すれば

$$x_a^{*}{}'x_a^{*} = x_a{}'M\,x_a(1-r_{ab|x}^2)\,,\; x_b^{*}{}'x_b^{*} = x_b{}'M\,x_b(1-r_{ab|x}^2)$$

さらに

$$x_a^{*}{}'x_b^{*} = x_a{}'M\,x_b(1-r_{ab|x}^2)$$

となる.したがって(3.54)式の第 s 二の等号が証明される.

練習問題解答

問 5　　任意の線型推定量を $b=\sum_{i=1}^{n} d_i y_i$ と定義する.この線型推定量の期待値を求めれば, $E(b)=\sum_{i=1}^{n} d_i E(y_i)=\sum_{i=1}^{n} d_i(\sum_{k=1}^{K}\beta_k x_{ki})=\sum_{k=1}^{K}\beta_k(\sum_{i=1}^{n} d_i x_{ki})=\beta_1$　となるから,不偏性の条件により 2 章 3 節と同様に $\sum_{i=1}^{n} d_i x_{1i}=1\,,\ \sum_{i=1}^{n} d_i x_{2i}=0\,,\cdots,\sum_{i=1}^{n} d_i x_{Ki}=0$ が満たされないといけない.(つまり, x_{2i} から x_{Ki} については d_i と直交している.)したがって,問題中の式は証明できる.

最小 2 乗推定量では重みが　$w_i=\dfrac{x_{1i}^{*}}{\Sigma(x_{1i}^{*})^2}$ と定義される.また最小二乗推定量は不偏な線形推定量であるから,この重み w_i は先の条件を満たす. x_{1i}^{*} と,任意の線形不偏推定量の重み d_i の積和については,問題中の式により, $\Sigma\, d_i x^{*}_{1i} = \Sigma\, d_i x_{1i} = 1$,となる.したがって,

$$(*)\quad \sum_{i=1}^{n} d_i w_i = \frac{\Sigma d_i x_{1i}^{*}}{\Sigma(x_{1i}^{*})^2} = \frac{1}{\Sigma(x_{1i}^{*})^2} = \sum_{i=1}^{n}(w_i)^2$$

という等式が満たされる.他方で,任意の線形不偏推定量の分散は $V(b-\beta_1)^2=\sigma^2\Sigma(d_i)^2$ と表現できるが,この式の右辺は

$\sum_{i=1}^{n}[(d_i-w_{1i})+w_{1i}]^2=\Sigma(d_i-w_{1i})^2+2\Sigma(d_i-w_{1i})w_{1i}+\Sigma w_{1i}{}^2$,と分解できる.ここ

で,*式により,$\Sigma(d_i - w_{1i})w_{1i} = 0$であるから，第一項が0になる時に線形不偏推定量の分散は最小になる.つまり,$d_i = w_{1i}$のときBLUE.

問8　(3.44)式では変数aが落ちている.そのためRSS(43)とRSS(44)の間に問題中の関係式が成立する.(3.44)式からさらに変数bを落とすと,

$$RSS(44)=RSS(a,b)(1-r^2_{yb|c,d,\cdots,K})$$

となる.ただし,RSS(a,b)は変数a,bを除いた回帰式の残差変動和である.問題中の式と合わせると,

$$RSS(43)=RSS(a,b)\ (1-r^2_{ya|b,c,\cdots,K})(1-r^2_{yb|c,d,\cdots,K})$$

となる.この操作を$k-1$番目まで繰り返すと,

$$RSS(43)=RSS(a,b,\ldots,K-1)\ (1-r^2_{ya|b,c,\cdots,K})(1-r^2_{yb|c,d,\cdots,K})\cdots(1-r^2_{y,K-1|K})$$

となる.ここで,RSS(a,b,…,K-1)は変数aから変数K-1までを全て除くから,残っているのは変数Kつまり定数項だけである.さらに,yを定数項に回帰した場合のRSSは,全変動に一致する.式を使えば，$x_{K-1,i} = 1, \hat{\beta}_K = \bar{y}$なので

$$\Sigma(y_i - \hat{\beta}_K x_{K-1,i})^2 = \Sigma y_i{}^2 - n\bar{y} = TSS$$

となる.

1部4章

127頁

26/6/2001

R^2を使ってF統計量を計算する方法:追加説明

(4.33)式の分母分子をTSSで割ると

（追1）$f = \dfrac{(RSS(32)/TSS) - (RSS(31)/TSS)}{RSS(31)/TSS}\dfrac{n-K}{m}$

となります.$R^2 = 1-(RSS/TSS)$　ですから,このfは

（追 2）　$f=\frac{(1-R^2(32)-(1-R^2(31))}{1-R^2(31)}\frac{n-K}{m}=\frac{R^2(31)-R^2(32)}{1-R^2(31)}\frac{n-K}{m}$

となり,二個の回帰式の R^2 値を使って F 値を計算できることがわかります.

127頁　定数項は必要か:追加説明

F 検定を使う場合は(4.31)式ならびに(4.32)式に定数項が含まれる必要はありません.全ての説明変数が変数であってよいわけです.重要なのは(4.32)式は(4.31)式の一部の説明変数を用いるという点です.

しかし回帰式に定数項が含まれていないと,TSS=RSS+ESS の分解式が成立しないために, R^2 を使って F を表現することはできません.（(追 2)式や 128 頁の f 統計量を R^2 で表現した式が導けません）理由は以下のようになります.

TSS$\sum_{i=1}^{n}(y_i-\bar{y})^2$ の分解は

$$\sum_{i=1}^{n}(y_i-\bar{y})^2=\sum_{i=1}^{n}(y_i-\hat{y}_i)^2+\sum_{i=1}^{n}(\hat{y}_i-\hat{y}\text{の平均})^2$$

となります.しかし,定数項がないために,回帰値 $\hat{y}_i$ の平均は $\bar{y}$ になりません.(21 頁から 22 頁にかけての説明を読む.) $\Sigma(y_i-\hat{y}_i)$ も 0 にはなりません.　そのため,TSS を分解しても,積の項が残ってしまいます.（他方 $\Sigma(y_i-\hat{y}_i)\hat{y}_i=0$ は成立します.）

129頁　LM検定法 （上級）　　7/2001

(4.32)　$\mathbf{y}=\mathbf{X}\beta+\mathbf{Z}\gamma+\varepsilon$

ここで X は $n\times K$,Z は $n\times P$ とする..　検定の帰無仮説は $H_0:\gamma=0$ である.ラグランジ乗数検定統計量は、

(4.35) $\mathrm{LM} = n\dfrac{\mathrm{RSS(X)} - \mathrm{RSS(X,Z)}}{\mathrm{RSS(X)}}$.

と定義されるが,RSS(X) と RSS(X,Z) は帰無および対立仮説の下で計算される残差変動である．LM の漸近分布は $\chi^2(\mathrm{P})$ となる．おなじ検定統計量を,残差を用いた回帰で表す． その手順は以下のようになる.

1. 帰無仮説の下での回帰

(4.32) $y = X\beta + v$

の最小２乗残差ベクトルを $\hat{\varepsilon}$ とする.

2. 回帰式

(A1) $\hat{\varepsilon} = X\beta + Z\delta + \mathrm{error} = (X, Z)\begin{pmatrix} \beta \\ \delta \end{pmatrix} + \mathrm{error}$

の nR^2 を求める． nR^2 **と LM は等しい．** ただし、

$$nR^2 = n\frac{\mathrm{ESS}}{\mathrm{TSS}} = n\frac{1}{\hat{\varepsilon}'\hat{\varepsilon}}\hat{\varepsilon}'\begin{pmatrix} X & Z \end{pmatrix}\begin{pmatrix} X'X & X'Z \\ Z'X & Z'Z \end{pmatrix}^{-1}\begin{pmatrix} X' \\ Z' \end{pmatrix}\hat{\varepsilon}$$

である。証明は以下の順に行う.

1.1 $\hat{\varepsilon} = M_X y \equiv (I - X(X'X)^{-1}X')y$ ． これは明らかである.

1.2 $\mathbf{W} = (\mathbf{X}, \mathbf{Z})$ と定義する.(4.35) の分子は次の２次形式になる.（明らか）

(A2) $\mathbf{y}'\left\{\mathbf{W}(\mathbf{W}'\mathbf{W})^{-1}\mathbf{W}' - \mathbf{X}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\right\}\mathbf{y}$.

1.3 (A2) は次の公式により $\mathbf{y}'\mathbf{Z}^*(\mathbf{Z}^{*\prime}\mathbf{Z}^*)^{-1}\mathbf{Z}^{*\prime}\mathbf{y}$ と同等になる．ただし $\mathbf{Z}^* = \mathbf{M_X Z}$ ． だから

$$\mathrm{LM} = n\frac{y'Z^*(Z^{*\prime}Z^*)^{-1}Z^{*\prime}y}{y'M_X y} = n\frac{\hat{\varepsilon}'Z^*(Z^{*\prime}Z^*)^{-1}Z^{*\prime}\hat{\varepsilon}}{\hat{\varepsilon}'\hat{\varepsilon}}$$

他方、

$$\hat{\varepsilon}'\begin{pmatrix}X & Z\end{pmatrix}\begin{pmatrix}X'X & X'Z\\ Z'X & Z'Z\end{pmatrix}^{-1}\begin{pmatrix}X'\\ Z'\end{pmatrix}\hat{\varepsilon} = (0 \quad \hat{\varepsilon}'Z)\begin{pmatrix}X'X & X'Z\\ Z'X & Z'Z\end{pmatrix}^{-1}\begin{pmatrix}0\\ Z'\hat{\varepsilon}\end{pmatrix}$$

$$= \hat{\varepsilon}'Z^*(Z^{*\prime}Z^*)^{-1}Z^{*\prime}\hat{\varepsilon}.$$

したがって、両者は一致する。

FORMULA $\mathbf{W}(\mathbf{W}'\mathbf{W})^{-1}\mathbf{W}' = \mathbf{X}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}' + \mathbf{Z}^*(\mathbf{Z}^{*\prime}\mathbf{Z}^*)^{-1}\mathbf{Z}^{*\prime}$.

FORMULA の証明

$$\mathbf{W}(\mathbf{W}'\mathbf{W})^{-1}\mathbf{W}' = \mathbf{WQ}(\mathbf{Q}'\mathbf{W}'\mathbf{WQ})^{-1}\mathbf{Q}'\mathbf{W}'$$

ここで Q の定義は

$$\mathbf{Q} = \begin{pmatrix}\mathbf{I} & -(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{Z}\\ 0 & \mathbf{I}\end{pmatrix}, \quad \mathbf{WQ} = \left(\mathbf{X}, \mathbf{M_X Z}\right).$$

したがって公式は明らかである

1.4 　R^2 の分子は次のようになる.

$$(\mathbf{X}\hat{\beta} + \mathbf{z}\,\hat{\gamma})'(\mathbf{X}\hat{\beta} + \mathbf{z}\,\hat{\gamma}) = (\hat{\beta}' \;\; \hat{\gamma}')\mathbf{W}'\mathbf{W}\begin{pmatrix}\hat{\beta}\\ \hat{\gamma}\end{pmatrix} = \hat{\varepsilon}'\mathbf{W}(\mathbf{W}'\mathbf{W})^{-1}\mathbf{W}'\hat{\varepsilon} = \mathbf{y}'\mathbf{M_X}\mathbf{W}(\mathbf{W}'\mathbf{W})^{-1}\mathbf{W}'\mathbf{M_X}\mathbf{y}$$

$$= \mathbf{y}'\mathbf{M_X}\mathbf{Z}(\mathbf{Z}'\mathbf{M_X}\mathbf{Z})^{-1}\mathrm{Z}'\mathbf{M_X}\mathbf{y} = \mathbf{y}'\mathbf{Z}^*(\mathbf{Z}^{*\prime}\mathbf{Z}^*)^{-1}\mathbf{Z}^{*\prime}\mathbf{y}$$

(LM検定法終わり)

8/2001

132-133 頁　予測のはずれ具合検定 (predictive failure test)　(上級)

m を 2 として,(4.39)式から(4.40)式を導いてみる.

(4.39) 　$y_i = \beta_1 x_{1i} + \beta_2 x_{2i} + \beta_3 x_{3i} + \beta_4 x_{4i} + r_3 z_{3i} + r_4 z_{4i} + \varepsilon_i$, $i = 1, \cdots, n, n+1, n+2$

とする.つまり m が 2 であるので, x_3 と x_4 変数を「適当」に選んだ**人工変数**とし, z_3 と z_4 を定義するのである.ところがこの z_3 と z_4 は最初の n 期では 0, $(n+1)$ と $(n+2)$ 期のみ 1 である.そこで

(A1)　$(n+1)$ 期　$r_3 z_{3n+1} + r_4 z_{4n+1} = r_3^* D_{1,n+1}$

(A2)　$(n+2)$ 期　$r_3 z_{3n+2} + r_4 z_{4n+2} = r_4^* D_{2,n+2}$

と理解すれば（4.40）式が導かれる．$(m=2)$　D_1 は $(n+1)$ 期では 1，他の全ての時点において 0 である．D_2 についても同様とする．また(A1)式,(A2)式より，r_3^* と r_4^* が推定できれば,逆に元の r_3 と r_4 も推定できることがわかる.一般の m についても説明は変わらない.

予測はずれという用語は最小２乗法基準からくる.これは

(A3)　$$\sum_{i=1}^{n+m}(y_i - \beta_1 x_{1i} - \beta_2 x_{2i} - \beta_3 x_{3i} - \beta_4 x_{4i} - r_1 D_{1i} - \cdots - r_m D_{mi})^2$$

を最小にするという基準だが,(A3)はダミー変数の性質により

(A4)　$$\sum_{i=1}^{n}(y_i - \beta_1 x_{1,i} - \beta_2 x_{2,i} - \beta_3 x_{3,i} - \beta_4 x_{4,i})^2 + (y_{n+1} - \beta_1 x_{1,n+1} - \beta_2 x_{2,n+1} - \beta_3 x_{3,n+1} - \beta_4 x_{4,n+1} - r_1)^2$$
$$+ \cdots + (y_{n+m} - \beta_1 x_{1,n+m} - \beta_2 x_{2,n+m} - \beta_3 x_{3,n+m} - \beta_4 x_{4,n+m} - r_1)^2$$

と分解できる.第一項は n 期までの偏差の２乗和だが,β 係数はすべて第一項の最小化から導くことができる.第２項より

(A5)　$$\hat{r}_1 = y_{n+1} - \beta_1 x_{1,n+1} - \beta_2 x_{2,n+1} - \beta_3 x_{3,n+1} - \beta_4 x_{4,n+1}$$

と r_1 の推定量が求まるが,これは第 $(n+1)$ 期の予測誤差である.以下 $\hat{r}_m$ まで同様に計算する.(A5)などの推定により第２項以下は 0 となり,(A3)の最小化は,(A4)第一項の最小化と一致する.F 統計量は

$$f = \frac{RSS(4.39) - RSS(4.40)}{RSS(4.40)} \ \frac{(n+m)-(K+m)}{m}$$

$$= \frac{RSS(4.39) - RSS(4.36)}{RSS(4.36)} \ \frac{n-K}{m}$$

となる. **(predictive failure test 終わり)**

二回帰式の係数に関する検定（中級）　　(9/4/2002)

被説明変数ｙと説明変数 x に関する観測値を２つの集団から得たとする.ただしｘは K 次元ベクトルとし,便宜的に定数項を含むとする.（難しければ,各変数は平均が０になるように計測されており,ｘは一次元であると考えてよい.）各集団の観測個数は n_1,n_2 とする.各グループに関する回帰式を

$$y_i = x_i'\beta_i + u_i \quad i=1,2$$

と表し,両回帰式の**誤差分散が共通**であるとして,係数全体,β_1 と β_2 が等しいという帰無仮説　$H_0:\beta_1=\beta_2$　の検定方法を説明しよう.

変換

$$y_2 = x_2'\beta_1 + x_2'(\beta_2-\beta_1) + u_2 = x_2'\beta_1 + x_2'\gamma + u_2$$

を使えば,両式は対立仮説のもとで

$$\begin{pmatrix} y_1 \\ y_2 \end{pmatrix} = \begin{pmatrix} x_1' \\ 0 \end{pmatrix}\beta_1 + \begin{pmatrix} 0 \\ x_2' \end{pmatrix}\beta_2 + \begin{pmatrix} u_1 \\ u_2 \end{pmatrix} = \begin{pmatrix} x_1' \\ x_2' \end{pmatrix}\beta_1 + \begin{pmatrix} 0 \\ x_2' \end{pmatrix}\gamma + \begin{pmatrix} u_1 \\ u_2 \end{pmatrix}$$

となる.この変換により,検定の帰無仮説は　$H_0:\gamma=0$　となるが,新しい帰無仮説の下で,帰無仮説下の回帰式が対立仮説下の回帰式に含まれることが明らかになる.従って,通常の F 検定が利用できる.

誤差分散が異なるとき

誤り：グループごとに推定をして残差分散を求める.そしてグループごとの回帰式を残差分散の平方根（標準偏差）で割って分散を（近似的に）均一にする.この分散均一になるように変換された式について上述の手法を用いればよいと考えられるかもしれないが,これは誤りである.F 検定は使えない.このことを式を使って説明しよう.まずグループごとに最小２乗推定を行うが,その推定量を $\hat\beta_1,\hat\beta_2$ とすれば,誤差分散は

$$\hat s_i^{\,2} = \frac{1}{n_i-K}\mathrm{RSS}(\text{第}i\text{グループの回帰式}), \qquad i=1,2$$

となる.各式を標準偏差で割ると,対立仮説の下での回帰式は

$$(*) \qquad \begin{pmatrix} \frac{1}{\hat{s}_1} y_1 \\ \frac{1}{\hat{s}_2} y_2 \end{pmatrix} = \begin{pmatrix} \frac{1}{\hat{s}_1} x_1' \\ 0 \end{pmatrix} \beta_1 + \begin{pmatrix} 0 \\ \frac{1}{\hat{s}_2} x_2' \end{pmatrix} \beta_2 + \begin{pmatrix} \frac{1}{\hat{s}_1} u_1 \\ \frac{1}{\hat{s}_2} u_2 \end{pmatrix}$$

となり,近似的に均一分散である.しかし,この式を最小２乗法で推定しても,求まる両係数の推定量は元の$\hat{\beta}_1$と$\hat{\beta}_2$に戻ってしまう.なぜなら,(*)式はグループごとに分解されて,分解された式は

$$\frac{1}{s_i} y_i = \frac{1}{s_i} x_i' \beta_i + \frac{1}{s_i} u_i, \qquad i=1,2$$

となり,(*)式の最小２乗推定はグループごとの式の最小２乗推定と同じになるからである.各式を標準偏差で割ろうが割るまいが,推定結果は同じである.この結果,(*)式の残差平方和は

$$\frac{1}{\hat{s}_1^2} \text{RSS(第一グループの回帰式)} + \frac{1}{\hat{s}_2^2} \text{RSS(第二グループの回帰式)}$$

となるが,これは誤差分散の定義により定数になってしまう.

正解:一般的にはWald法を使う.Kが１の場合を説明しよう.かつ各変数からは平均が引かれているとする.良く知られているように,この場合例えば

$$\hat{\beta}_1 = \frac{\sum_{t=1,n1} x_{1t} y_{1t}}{\sum_{t=1,n1} x_{1t}^2}$$

となるが,この分散は $V(\hat{\beta}_1) = \dfrac{\sigma_1^2}{\sum_{t=1,n1} x_{1t}^2}$ である.さらに,グループの独立性により,

$$V(\hat{\beta}_1 - \hat{\beta}_2) = \frac{\sigma_1^2}{\sum_{t=1,n1} x_{1t}^2} + \frac{\sigma_2^2}{\sum_{t=1,n1} x_{2t}^2}$$

となる.この結果により,帰無仮説の下では

$$z = \frac{\hat{\beta}_1 - \hat{\beta}_2}{\sqrt{\dfrac{\sigma_1^{\ 2}}{\sum_{t=1,n1} x_{1t}^{\ 2}} + \dfrac{\sigma_2^{\ 2}}{\sum_{t=1,n1} x_{2t}^{\ 2}}}}$$

が,漸近的に標準正規分布にしたがう.（誤差項が正規分布にしたがうと仮定しておけば,z の分布は厳密に標準正規である.）したがって,z に含まれる未知の誤差分散を　$\hat{s}_i^{\ 2} = \dfrac{1}{n_i - K}$RSS(第iグループの回帰式）　と推定すれば,z は帰無仮説の下で漸近的に標準正規になり,検定に用いることができる.

共通項を持つ場合： 異分散でも両回帰式に共通項があるとする.つまり

$$y_1 = z_1'\delta + x_1'\beta_1 + u_1 ,\ y_2 = z_2'\delta + x_2'\beta_2 + u_2$$

となっているとする.この場合は,対立仮説の下での統合された式

$$\begin{pmatrix} y_1 \\ y_2 \end{pmatrix} = \begin{pmatrix} z_1' \\ z_2' \end{pmatrix}\delta + \begin{pmatrix} x_1' \\ 0 \end{pmatrix}\beta_1 + \begin{pmatrix} 0 \\ x_2' \end{pmatrix}\beta_2 + \begin{pmatrix} u_1 \\ u_2 \end{pmatrix}$$

を最小２乗法で推定し,各グループの残差平方和を用いて,各式を標準化し,F 検定を応用する事ができる.先の問題は,二段階目の推定が,元の推定に戻ってしまうことにある.一般的には,**一般化最小2乗推定量が,単純最小2乗推定量に等しくなる**時の問題といわれる（終わり）

練習問題解答

問１　(11)式において $n = m$ ならば, $t = \left(2S^2 / n\right)^{-1/2}(\overline{X}_M - \overline{X}_F)$.ただし,

$S^2 = \{\sum_{i=1}^{n} (X_{Mi} - \overline{X}_M)^2 + \sum_{i=n+1}^{2n} (X_{Fi} - \overline{X}_F)^2\}/\{2(n-1)\}$ となっている.一方(16)式において, $n = m$ ならば, $t = \{(S_M^2 + S_F^2)/n\}^{-1/2}(\overline{X}_M - \overline{X}_F)$.また, $S^2 = (S_M^2 + S_F^2)/2$ である.なぜなら, $S_M^2 = \sum_{i=1}^{n} (X_{Mi} - \overline{X}_M)^2/(n-1)$ などより

$$\Sigma(X_{Mi} - \overline{X}_M)^2 + \Sigma(X_{Fi} - \overline{X}_F)^2 = (n-1)(S_M^2 + S_F^2)$$

となる.

問 5　　正規方程式は, $(n+m)\hat{\mu}_M + m\hat{\delta} = n\overline{X}_M + m\overline{X}_F$, $m\hat{\mu}_M + m\hat{\delta} = m\overline{X}_F$ となる. だから,正規方程式を解いて, $\hat{\mu}_M = \overline{X}_M$, さらに $\hat{\delta} = \overline{X}_F - \overline{X}_M$. t 統計量については,(9)式より $V(\hat{\delta}) = (\frac{1}{n} + \frac{1}{m})\sigma^2$ なので, $\hat{\delta}$ を標準偏差で割り, σ を S に置き換えると,(11)式となる.残差分散が(10)に一致することは明らかだろう.($D_{Fi}+D_{Mi}=1$ となる.)

問題中のコメントは, $\hat{\delta}$ は回帰係数の最小 2 乗推定量であるので,3 章の結果から t 統計量を求めると言うことである.つまり,89 頁(3.35)式が,(11)式に一致することを示すことが問題となる.係数推定量は既に求まっている.回帰の RSS を求めることも容易であろう.問題は, x_t^* を定義し, $\sum (x_t^*)^2$ を計算すること. x_t は何か, x_t を他の説明変数に回帰して求まる残差は何かを考える.

x_t は D_{Fi} であり,他の変数は定数しかない.従って x_t^* は D_{Fi} を定数に回帰した残差になる.回帰式は, $D_{Fi}=c+error$, という形になる.ここで c の最小 2 乗推定量を求めると,それは D_{Fi} の平均であるから, $m/(n+m)$ になる.さらに RSS は

$$RSS=\sum_{i=1}^{n+m}(D_{Fi}-\frac{m}{n+m})^2=\sum_{i=1}^{n+m}D_{Fi}{}^2-(n+m)(\frac{m}{n+m})^2=m-\frac{m^2}{n+m}=\frac{nm}{n+m}$$

となるから,後は代入して整理すればよい.

問 7　　$\Phi = \sum_{i=1}^{n+m}(y_i - \beta_1 x_{1i} - \ldots - \beta_4 x_{4i} - r_1 D_{1i} - \ldots - r_m D_{mi})^2$ と置いて, D_m は $n+m$ で 1, 他は 0 である. $\partial\Phi/\partial r_m = -2(y_{n+m} - \beta_1 x_{1n+m} - \ldots - \beta_4 x_{4n+m} - r_m) = 0$ より, $\hat{r}_m = y_{n+m} - \beta_1 x_{1n+m} - \ldots - \beta_4 x_{4n+m}$ となり, $n+m$ 期の残差は 0 と分かる.この点は重要である.

問8　ln(Q/L)=b0 + b1 ln(K/L) + b3 ln(L) + b3 ln(PM/P) + b4 year

を推定すればよい.b3 の推定値が有意であれば帰無仮説（収穫不変）は棄却される.

	係数	t
切片	-10.56	-7.25
lkh	0.69	12.09
lpm(-1)	0.0002	0.0084
year	-0.02	-9.06
l(L)	0.53	4.25