KEIO UNIVERSITY CALAMVS GLADIO FORTIOR

慶應義塾大学ビジネス・スクール

回帰分析シリーズ3

— 行列 —

これまで、モデルに説明変数が２つ含まれる場合の重回帰分析を考察してきた。しかし単回帰であっても、説明変数が複数個存在する場合であっても、その基本的な仮定や仕組みは同じである。このことを理解するために、このノートでは行列や行列式を用いて回帰分析の仕組みを解説する。これは行列を用いることによって、回帰モデルの一般化が可能になるからである。さらに、多重共線性やスペック・エラー、分散の不均一性、自己相関といった様々な回帰モデルの制約や限界を考察する場合、行列を用いる方が容易に理解することができる。ただし、行列に対する知識が十分でない学生に対する配慮から、重要な証明や難解な計算については、その都度解説を行ってゆくことにする。

一般に k 個の未知数を持つ回帰モデルは、

$$Y_i = \beta_1 + \beta_2 X_{2i} + \beta_3 X_{3i} + \ldots\ldots + \beta_k X_{ki} + u_i \quad \therefore\ i = 1, 2, \ldots., n \tag{1}$$

もしくは、

$$Y_i = \sum_{g=1}^{k} \Sigma \beta_g X_{gi} + u_i \quad \therefore\ i = 1, 2, \ldots., n;\ X_{1i} = 1 \tag{2}$$

のように示すことができる。しかし行列式を用いると、このモデルは次のように記述することが可能である。

$$\mathbf{y} = \mathbf{X}\boldsymbol{\beta} + \mathbf{u} \tag{3}$$

ここで $\mathbf{y}$ と $\mathbf{u}$ は $(n \times 1)$、$\boldsymbol{\beta}$ は $(k \times 1)$ の列ベクトル、$\mathbf{X}$ は $(n \times k)$ の行列である。ただし $\mathbf{X}$ の第１列は、切片 (β_0) の係数を示すことから１になっている。

$$\mathbf{y} = \begin{pmatrix} Y_1 \\ Y_2 \\ \vdots \\ Y_n \end{pmatrix},\quad \mathbf{u} = \begin{pmatrix} u_1 \\ u_2 \\ \vdots \\ u_n \end{pmatrix},\quad \boldsymbol{\beta} = \begin{pmatrix} \beta_1 \\ \beta_2 \\ \vdots \\ \beta_k \end{pmatrix},\quad \mathbf{X} = \begin{pmatrix} 1 & X_{12} & \cdots & X_{1k} \\ 1 & X_{22} & \cdots & X_{2k} \\ \vdots & \vdots & \ddots & \vdots \\ 1 & X_{n2} & \cdots & X_{nk} \end{pmatrix}$$

そして単回帰の場合と同様に、誤差項の分布に関していくつかの仮定をおくが、それらの仮定を行列式で表せば、

(1) 誤差項の期待値はゼロである。

このノートは、慶應義塾大学ビジネス・スクールにおける補助教材として、同ビジネス・スクール教授矢作恒雄と博士課程磯辺剛彦が作成した。（1995年5月作成）

最小二乗法で推定された b_1、b_2、b_3 を用いると、被説明変数の推定値 $(\hat{Y})$ は、

$$\hat{Y} = b_1 + b_2 X_2 + b_3 X_3 \tag{7}$$

したがって (6) 式と (7) 式の差をとれば、

$$\hat{y} = b_2 x_2 + b_3 x_3 \tag{8}$$

また (7) 式と (2) 式によって、

$$Y = \hat{Y} + e \tag{9}$$

この (9) 式の偏差をとれば、

$$\begin{aligned} Y - \bar{Y} &= \left(\hat{Y} - \bar{Y}\right) + (e - \bar{e}) \\ y &= \hat{y} + e \qquad \therefore \bar{e} = 0 \end{aligned} \tag{10}$$

決定係数

(10) 式の y の平方和 (TSS) は、説明される平方和 (ESS) と残差平方和 (RSS) とに分解することができる。

$$\begin{aligned} \Sigma y^2 &= \Sigma \left(\hat{y} + e\right)^2 \\ &= \Sigma \hat{y}^2 + \Sigma e^2 + 2\, \Sigma \hat{y} e \\ &= \Sigma \hat{y}^2 + \Sigma e^2 \qquad \therefore \Sigma \hat{y} e = 0 \end{aligned} \tag{11}$$

証明：

$\Sigma \hat{y} e$ に (8) 式を代入すれば、

$$\Sigma \hat{y} e = \Sigma (b_2 x_2 + b_3 x_3) e = b_2 \Sigma x_2 e + b_3 \Sigma x_3 e = 0 \tag{12}$$

(12) 式がゼロになるのは、(4) 式によって $\Sigma e = 0$ と $\Sigma X_2 e = 0$、$\Sigma X_3 e = 0$ が導かれるからである。

$$\frac{\partial(\Sigma e^2)}{\partial b_1} = -2\, \Sigma (Y - b_1 - b_2 X_2 - b_3 X_3) = -2\, \Sigma e = 0$$

$$\frac{\partial(\Sigma e^2)}{\partial b_2} = -2\, \Sigma X_2 e = 0$$

$$\frac{\partial(\Sigma e^2)}{\partial b_3} = -2\, \Sigma X_3 e = 0$$

$$\begin{aligned}\mathbf{e'e} &= (\mathbf{y} - \mathbf{Xb})'(\mathbf{y} - \mathbf{Xb}) \\ &= \mathbf{y'y} - 2\,\mathbf{b'X'y} + \mathbf{b'X'Xb}\end{aligned} \tag{9}$$

ここで (6) 式を $\mathbf{b}$ で偏微分したものをゼロとすると、残差平方和を最小にする $\mathbf{b}$ を見つけることができる。

$$\frac{\partial(\mathbf{e'e})}{\partial \mathbf{b}} = -2\,\mathbf{X'y} + 2\,\mathbf{X'Xb} = \mathbf{0} \tag{10}$$

したがって、

$$(\mathbf{X'X})\,\mathbf{b} = \mathbf{X'y} \qquad \mathbf{b} = (\mathbf{X'X})^{-1}\mathbf{X'y} \tag{11}$$

また残差と $\mathbf{X}$ との相関関係は存在しないので、

$$\mathbf{X'e} = \mathbf{0} \qquad \therefore\ \mathbf{X'y} = \mathbf{X'}(\mathbf{Xb} + \mathbf{e}) = (\mathbf{X'X})\,\mathbf{b} + \mathbf{X'e} \tag{12}$$

ここで、単回帰モデルにおける傾き $b = \Sigma xy / \Sigma x^2$ と $\mathbf{b} = (\mathbf{X'X})^{-1}\mathbf{X'y}$ との関係を確認しておく。たとえば X と Y とについて、次のような5つのサンプルが与えられているものとする。

n	1	2	3	4	5	合計	平均
X	5	7	6	8	4	30	6
Y	11	16	12	19	12	70	14

$\Sigma x^2 = 10 \quad \Sigma xy = 19$　　したがって、$b = \Sigma xy \big/ \Sigma x^2 = 1.9$

また切片 a は、

$$a = \bar{Y} - b\,\bar{X} = 14 - (1.9)(6) = 2.6$$

次にこのサンプルを(11) 式の行列に代入すると、

$$\mathbf{X'X} = \begin{pmatrix} 1 & 1 & 1 & 1 & 1 \\ 5 & 7 & 6 & 8 & 4 \end{pmatrix}\begin{pmatrix} 1 & 5 \\ 1 & 7 \\ 1 & 6 \\ 1 & 8 \\ 1 & 4 \end{pmatrix} = \begin{pmatrix} 5 & 30 \\ 30 & 190 \end{pmatrix}$$

$$(\mathbf{X'X})^{-1} = \frac{1}{5 \times 190 - (30)^2}\begin{pmatrix} 190 & -30 \\ -30 & 5 \end{pmatrix}$$

$$\begin{aligned}(\mathbf{X'X})^{-1}\mathbf{X'} &= \frac{1}{50}\begin{pmatrix} 190 & -30 \\ -30 & 5 \end{pmatrix}\begin{pmatrix} 1 & 1 & 1 & 1 & 1 \\ 5 & 7 & 6 & 8 & 4 \end{pmatrix} \\ &= \begin{pmatrix} 0.8 & -0.4 & 0.2 & -1 & 1.4 \\ -0.1 & 0.1 & 0 & 0.2 & -0.2 \end{pmatrix}\end{aligned}$$

$$(\mathbf{X'X})^{-1}\mathbf{X'y} = \begin{pmatrix} 0.8 & -0.4 & 0.2 & -1 & 1.4 \\ -0.1 & 0.1 & 0 & 0.2 & -0.2 \end{pmatrix} \begin{pmatrix} 11 \\ 16 \\ 12 \\ 19 \\ 12 \end{pmatrix} = \begin{pmatrix} 2.6 \\ 1.9 \end{pmatrix} = \begin{pmatrix} a \\ b \end{pmatrix}$$

このように両者の解答は同じになる。

次に (11) 式に (3) 式を代入することにより、

$$\begin{aligned} \mathbf{b} = (\mathbf{X'X})^{-1}\mathbf{X'y} &= (\mathbf{X'X})^{-1}\mathbf{X'}(\mathbf{X}\boldsymbol{\beta} + \mathbf{u}) \\ &= \boldsymbol{\beta} + (\mathbf{X'X})^{-1}\mathbf{X'u} \end{aligned} \tag{13}$$

ここで $\mathbf{b}$ の期待値をとると、

$$E(\mathbf{b}) = \boldsymbol{\beta} \qquad \therefore E\left[(\mathbf{X'X})^{-1}\mathbf{X'u}\right] = (\mathbf{X'X})^{-1}\mathbf{X'}E(\mathbf{u}) = \mathbf{0} \tag{14}$$

また $\mathbf{b}$ の分散・共分散行列については、

$\mathbf{b} - E(\mathbf{b}) = \mathbf{b} - \boldsymbol{\beta} = (\mathbf{X'X})^{-1}\mathbf{X'u}$　と仮定 (2) より、

$$\begin{aligned} Var(\mathbf{b}) &= E\left[(\mathbf{b} - \boldsymbol{\beta})^2\right] \\ &= E\left[(\mathbf{X'X})^{-1}\mathbf{X'uu'X}(\mathbf{X'X})^{-1}\right] \\ &= (\mathbf{X'X})^{-1}\mathbf{X'}\,\sigma^2\mathbf{I}\,\mathbf{X}(\mathbf{X'X})^{-1} \\ &= \sigma^2(\mathbf{X'X})^{-1} \end{aligned} \tag{15}$$

したがって、回帰モデルにおける $\mathbf{b}$ は次のような平均と分散をもつ正規分布にしたがうことが分かる。

$$\mathbf{b} = N\left(\boldsymbol{\beta},\ \sigma^2(\mathbf{X'X})^{-1}\right) \tag{16}$$

<u>行列式の偏差型</u>

今までの単回帰、重回帰分析ノートでは、偏差型（たとえば、$x = X - \bar{X}$）を用いて解説してきた。このような偏差型は行列式においても適用できる。そのような偏差転換行列を $\mathbf{A}$ とすると、

$$\mathbf{A} = \mathbf{I} - \frac{1}{n}\mathbf{ii'} \tag{17}$$

によって表すことができる。たとえば被説明変数 $\mathbf{Y}$ の偏差型は、

$$\mathbf{Ay} = \left(\mathbf{I} - \frac{1}{n}\mathbf{ii'}\right)\mathbf{y} = \begin{pmatrix} 1 & 0 & \cdots & 0 \\ 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 1 \end{pmatrix} \begin{pmatrix} Y_1 \\ Y_2 \\ \vdots \\ Y_n \end{pmatrix} - \frac{1}{n}\left[\begin{pmatrix} 1 \\ 1 \\ \vdots \\ 1 \end{pmatrix} \begin{pmatrix} 1 & 1 & \cdots & 1 \end{pmatrix}\right] \begin{pmatrix} Y_1 \\ Y_2 \\ \vdots \\ Y_n \end{pmatrix}$$

$$= \begin{pmatrix} Y_1 \\ Y_2 \\ \vdots \\ Y_n \end{pmatrix} - \frac{1}{n} \begin{pmatrix} 1 & 1 & \cdots & 1 \\ 1 & 1 & \cdots & 1 \\ \vdots & \vdots & \ddots & \vdots \\ 1 & 1 & \cdots & 1 \end{pmatrix} \begin{pmatrix} Y_1 \\ Y_2 \\ \vdots \\ Y_n \end{pmatrix} = \begin{pmatrix} Y_1 - \bar{Y} \\ Y_2 - \bar{Y} \\ \vdots \\ Y_n - \bar{Y} \end{pmatrix}$$

より簡単に示すと、

$$\mathbf{Ay} = \mathbf{y} - \mathbf{i}\bar{Y}$$

このような偏差変換行列 **A** はいくつかの特性をもつ。

(1)　$\mathbf{Ai} = \mathbf{0}$　(18)

$$\left(\mathbf{I} - \frac{1}{n}\mathbf{ii'}\right)\mathbf{i} = \mathbf{i} - \frac{1}{n}\mathbf{ii'i} = \mathbf{0}$$

$$\therefore \mathbf{ii'i} = \underset{(n \times 1)}{\begin{pmatrix} 1 \\ 1 \\ \vdots \\ 1 \end{pmatrix}} \underset{(1 \times n)}{\begin{pmatrix} 1 & 1 & \cdots & 1 \end{pmatrix}} \underset{(n \times 1)}{\begin{pmatrix} 1 \\ 1 \\ \vdots \\ 1 \end{pmatrix}} = \begin{pmatrix} n \\ n \\ \vdots \\ n \end{pmatrix}, \quad \mathbf{Ii} = \underset{(n \times n)}{\begin{pmatrix} 1 & & & \\ & 1 & & \\ & & \ddots & \\ & & & 1 \end{pmatrix}} \underset{(n \times 1)}{\begin{pmatrix} 1 \\ 1 \\ \vdots \\ 1 \end{pmatrix}} = \mathbf{i}$$

(2)　$\mathbf{Ae} = \mathbf{e}$　(19)

$$\left(\mathbf{I} - \frac{1}{n}\mathbf{ii'}\right)\mathbf{e} = \mathbf{e} - \frac{1}{n}\Sigma e = \mathbf{e}$$

$$\therefore \mathbf{ii'e} = \begin{pmatrix} 1 \\ 1 \\ \vdots \\ 1 \end{pmatrix} \begin{pmatrix} 1 & 1 & \cdots & 1 \end{pmatrix} \begin{pmatrix} e_1 \\ e_2 \\ \vdots \\ e_n \end{pmatrix} = \begin{pmatrix} \Sigma e \\ \Sigma e \\ \vdots \\ \Sigma e \end{pmatrix}, \quad \frac{1}{n}\Sigma e = \begin{pmatrix} \dot{e} \\ \dot{e} \\ \vdots \\ \dot{e} \end{pmatrix} = \mathbf{0}$$

(3)　$\mathbf{A} = \mathbf{A}^2 = \mathbf{A}^3 = \ldots\ldots$　(20)

A は巾（べき）等行列 (idempotent matrix) である。巾等行列とは、いくら自分自身の積を作っていっても不変な行列のことをいい、その必要条件は (21) 式のように、行列を倒置しても元の行列と同じになることである。

$$\mathbf{A} = \mathbf{A'} \tag{21}$$

次に、この偏差転換行列を用いて回帰モデルの構造を分析する。まずここでは、(7) 式の **X** を切片と傾きとに分解すると、

$$\mathbf{y} = \mathbf{x}_1\mathbf{b}_1 + \mathbf{x}_2\mathbf{b}_2 + \mathbf{e} \tag{22}$$

ここで $\mathbf{x}_1$ は $(1\times n)$、$\mathbf{x}_2$ は $(n\times k-1)$ である。

$$\mathbf{X} = [\mathbf{x}_1 \quad \mathbf{x}_2] = \left[\begin{array}{c:ccc} x_{11} & x_{21} & \cdots & x_{k1} \\ x_{12} & x_{22} & \cdots & x_{k1} \\ \vdots & \vdots & \ddots & \vdots \\ x_{1n} & x_{2n} & \cdots & x_{kn} \end{array}\right]$$
$$(1\times n) \qquad (n\times k-1)$$

(22) 式の両辺に $\mathbf{A}$ を乗じると、

$$\mathbf{Ay} = \mathbf{Ax}_1\mathbf{b}_1 + \mathbf{Ax}_2\mathbf{b}_2 + \mathbf{Ae} \tag{23}$$

ここで $\mathbf{b}_1$ は切片なので、$\mathbf{x}_1 = \mathbf{i}$。さらに $\mathbf{Ai} = \mathbf{0}$、$\mathbf{Ae} = \mathbf{e}$ により、(23) 式は次のように変換できる。

$$\mathbf{Ay} = \mathbf{Ax}_2\mathbf{b}_2 + \mathbf{e} \tag{24}$$

さらにこの (24) 式に $\mathbf{y}'$ を乗じると、

$$\begin{aligned} \mathbf{y}'\mathbf{Ay} &= \mathbf{y}'\mathbf{Ax}_2\mathbf{b}_2 + \mathbf{y}'\mathbf{e} \\ &= (\mathbf{x}_2\mathbf{b}_2 + \mathbf{e})'\mathbf{Ax}_2\mathbf{b}_2 + (\mathbf{x}_2\mathbf{b}_2 + \mathbf{e})'\mathbf{e} \\ &= \mathbf{b}_2'\mathbf{x}_2'\mathbf{Ax}_2\mathbf{b}_2 + \mathbf{e}'\mathbf{e} \qquad \therefore \mathbf{x}_2'\mathbf{e} = \mathbf{0} \end{aligned} \tag{25}$$
$$(\mathrm{TSS}) = \quad (\mathrm{ESS}) \quad + \quad (\mathrm{RSS})$$

したがって決定係数は、

$$R^2 = \frac{\mathrm{ESS}}{\mathrm{TSS}} = 1 - \frac{\mathbf{e}'\mathbf{e}}{\mathbf{y}'\mathbf{Ay}} = \frac{\mathbf{b}_2'\mathbf{x}_2'\mathbf{Ax}_2\mathbf{b}_2}{\mathbf{y}'\mathbf{Ay}} = \frac{\mathbf{b}_2'\mathbf{x}_2'\mathbf{Ay}}{\mathbf{y}'\mathbf{Ay}} \tag{26}$$

により求めることができる。このような結果は、説明変数 $\mathbf{X}$ を分解しなくとも得ることができる。(7) 式を二乗すると、

$$\begin{aligned} \mathbf{y}'\mathbf{y} &= (\mathbf{Xb} + \mathbf{e})'(\mathbf{Xb} + \mathbf{e}) \\ &= \mathbf{b}'\mathbf{X}'\mathbf{Xb} + 2\mathbf{X}'\mathbf{be} + \mathbf{e}'\mathbf{e} \\ &= \mathbf{b}'\mathbf{X}'\mathbf{Xb} + \mathbf{e}'\mathbf{e} \qquad \therefore \mathbf{X}'\mathbf{e} = \mathbf{0} \end{aligned} \tag{27}$$

また (27) 式に(11) 式の $\mathbf{b} = (\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{y}$ を代入すると、

$$\begin{aligned} \mathbf{y}'\mathbf{y} &= \mathbf{y}'\mathbf{X}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{X}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{y} + \mathbf{e}'\mathbf{e} \\ &= \mathbf{y}'\mathbf{X}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{y} + \mathbf{e}'\mathbf{e} \qquad \therefore \mathbf{X}'\mathbf{X}(\mathbf{X}'\mathbf{X})^{-1} = \mathbf{I} \end{aligned} \tag{28}$$

ここで、$\mathbf{y}'\mathbf{y} = \Sigma Y^2$ および、

$$
\begin{aligned}
\mathbf{y'Ay} &= \Sigma\left(Y-\bar{Y}\right)^2 \\
&= \Sigma Y^2 + 2\,\bar{Y}\Sigma Y + n\,\bar{Y}^2 \\
&= \Sigma Y^2 - n\,\bar{Y}^2 \qquad \therefore\ \bar{Y}\Sigma Y = \bar{Y}\left(n\,\bar{Y}\right) \\
&= \mathbf{y'y} - n\,\bar{Y}^2
\end{aligned}
\tag{29}
$$

(27) 式の両辺から $n\,\bar{Y}^2$ を引くと、

$$
\begin{aligned}
&\mathbf{y'y} - n\,\bar{Y}^2 = \left(\mathbf{b'x'xb} - n\,\bar{Y}^2\right) + \mathbf{e'e} \\
&\quad (\mathrm{TSS}) \quad = \quad (\mathrm{ESS}) \quad + \quad (\mathrm{RSS})
\end{aligned}
\tag{30}
$$

したがって決定係数は、

$$
R^2 = 1 - \frac{\mathbf{e'e}\,/\,n}{\mathbf{y'Ay}\,/\,n}
\tag{31}
$$

となり、(26) 式と同じであることが分かる。また自由度修正済み決定係数については、

$$
\bar{\mathrm{R}}^2 = 1 - \frac{\mathbf{e'e}\,/\,(n-k)}{\mathbf{y'Ay}\,/\,(n-1)}
\tag{32}
$$

誤差項の分散

未知数である誤差項の分散 σ^2 を推定するためには、(7) 式を用いるとよい。

$$
\begin{aligned}
\mathbf{e} &= \mathbf{y} - \mathbf{Xb} \\
&= \mathbf{y} - \mathbf{X(X'X)^{-1}X'y} && \therefore\ (11)\text{式} \\
&= \left[\mathbf{I} - \mathbf{X(X'X)^{-1}X'}\right]\mathbf{y} \\
&= \mathbf{My} && \therefore\ \mathbf{M} = \mathbf{I} - \mathbf{X(X'X)^{-1}X'} \\
&= \mathbf{M(X\boldsymbol{\beta} + u)} && \therefore\ (3)\text{式} \\
&= \mathbf{Mu} && \therefore\ \mathbf{MX} = 0
\end{aligned}
\tag{33}
$$

ここでの行列 $\mathbf{M}$ は巾等行列である。なぜなら、

$$
\begin{aligned}
\mathbf{MM} &= \left[\mathbf{I} - \mathbf{X(X'X)^{-1}X'}\right]\left[\mathbf{I} - \mathbf{X(X'X)^{-1}X'}\right] \\
&= \mathbf{I} - \mathbf{X(X'X)^{-1}X'} - \mathbf{X(X'X)^{-1}X'} + \mathbf{X(X'X)^{-1}X'} \\
&= \mathbf{I} - \mathbf{X(X'X)^{-1}X'} \\
&= \mathbf{M}
\end{aligned}
\tag{34}
$$

また $\mathbf{M}$ がゼロになるのは、

$$\mathbf{MX} = \left[\mathbf{I} - \mathbf{X}(\mathbf{X'X})^{-1}\mathbf{X'}\right]\mathbf{X} = \mathbf{X} - \mathbf{X}(\mathbf{X'X})^{-1}\mathbf{X'X} = \mathbf{0} \tag{35}$$

残差平方和を求めるには (33) 式を二乗すればよいので、

$$\begin{aligned} \mathbf{e'e} &= \mathbf{u'M'Mu} \\ &= \mathbf{u'Mu} \end{aligned} \tag{36}$$

ここで $\mathbf{u'Mu}$ は行列ではなく、ある一定の数値になる。

$$\underset{(1\times n)}{\mathbf{u'Mu} = \begin{pmatrix} u_1 & u_2 & \cdots & u_n \end{pmatrix}} \underset{(n\times n)}{\begin{pmatrix} M_{11} & M_{12} & \cdots & M_{1n} \\ M_{21} & M_{22} & \cdots & M_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ M_{n1} & M_{n2} & \cdots & M_{nn} \end{pmatrix}} \underset{(n\times 1)}{\begin{pmatrix} u_1 \\ u_2 \\ \vdots \\ u_n \end{pmatrix}} \underset{(1\times 1)}{= \lambda} \tag{37}$$

(36) 式の期待値をとると、

$$\begin{aligned} E\,(\mathbf{e'e}) &= E\,(\mathbf{u'Mu}) \\ &= E\,[tr\,(\mathbf{u'Mu})] \qquad \therefore (37)\text{式} \\ &= E\,[tr\,(\mathbf{Muu'})] \\ &= \sigma^2\, tr\, \mathbf{M} \qquad \therefore (6)\text{式} \end{aligned} \tag{38}$$

ここで $tr\,\mathbf{A}$（**A** のトレースと呼ぶ）とは、正方行列 **A** の対角線上のセルを合計したものをいう。たとえば **A** と **B** が、それぞれ $(m\times n)$ と $(n\times m)$ の行列だとすると、その積は、

$$\mathbf{AB} = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{pmatrix} \begin{pmatrix} b_{11} & b_{12} & \cdots & b_{1m} \\ b_{21} & b_{22} & \cdots & b_{2m} \\ \vdots & \vdots & \ddots & \vdots \\ b_{n1} & b_{n2} & \cdots & b_{nm} \end{pmatrix} = \begin{pmatrix} \sum_{j=1}^{n} a_{1j} b_{j1} & & \\ & \sum_{j=1}^{n} a_{ij} b_{ji} & \\ & & \sum_{j=1}^{n} a_{mj} b_{jm} \end{pmatrix}$$

したがって、

$tr\,(\mathbf{AB}) = \sum_{i=1}^{m}\sum_{j=1}^{n} a_{ij} b_{ji}$ および、$tr\,(\mathbf{BA}) = \sum_{j=1}^{n}\sum_{i=1}^{m} b_{ji} a_{ij}$ より、

$tr\,(\mathbf{AB}) = tr\,(\mathbf{BA})$

この結果を拡張すると、

$$tr\,(\mathbf{ABC}) = tr\,(\mathbf{CAB}) = tr\,(\mathbf{BCA})$$

この法則を利用すれば、(38) 式は次のように展開できる。

$$\begin{aligned} tr\,(\mathbf{M}) &= tr\left[\mathbf{I} - \mathbf{X}(\mathbf{X'X})^{-1}\mathbf{X'}\right] \\ &= tr\,(\mathbf{I}) - tr\left[\left(\mathbf{X}(\mathbf{X'X})^{-1}\mathbf{X'}\right)\right] \\ &= tr\,(\mathbf{I}) - tr\left[\left((\mathbf{X'X})^{-1}\mathbf{X'X}\right)\right] \end{aligned} \tag{39}$$

ここで、

$$\mathbf{I} = \begin{pmatrix} 1 & & & \\ & 1 & & \\ & & \ddots & \\ & & & 1 \end{pmatrix} \quad \text{より、} tr\,(\mathbf{I}) = n$$

$$tr\left[\left((\mathbf{X'X})^{-1}\mathbf{X'X}\right)\right] = tr\,(\mathbf{I_k}) = k$$

より、

$$\begin{aligned} E\,(\mathbf{e'e}) &= \sigma^2 E\,[tr\,(\mathbf{M})] \\ &= \sigma^2\,(n-k) \end{aligned} \tag{40}$$

したがって、(40) 式の未知数 σ^2 を推定値 s^2 に置き換えると、

$$\frac{\mathbf{e'e}}{n-k} = s^2 \tag{41}$$

通常最小二乗法 (ordinary least squares ; OLS) による推定

これまでの結果を用いて、ＯＬＳによる推定問題について考えてみる。まずβの仮説検証において、$\mathbf{R}$ と $\mathbf{r}$ という２つの行列を導入する。たとえば (3) 式のような回帰モデルの傾き$\boldsymbol{\beta}$（ただし、切片 β_1 は除く）がゼロであるという仮説を検証する場合、$\mathbf{R}$ と $\mathbf{r}$ とはそれぞれ次のような行列になる。

$$\underset{(k-1)\times k}{\mathbf{R} = \begin{pmatrix} 0 & 1 & 0 & \cdots & 1 \\ 0 & 0 & 1 & \cdots & 1 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & 1 \end{pmatrix}}, \quad \underset{(k-1)\times 1}{\mathbf{r} = \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 0 \end{pmatrix}}$$

ここでの帰無仮説は、

$$\mathrm{H}_0: \ \beta_2 = \beta_2 = \ldots\ldots = \beta_k = 0$$

つまり、

$$\mathbf{R\beta} = E\,(\mathbf{Rb}) = \mathbf{r} = \mathbf{0} \tag{42}$$

ここで **Rb** の分散は、

$$\begin{aligned} Var\,(\mathbf{Rb}) &= E\,[\mathbf{R}\,(\mathbf{b}-\boldsymbol{\beta})\,(\mathbf{b}-\boldsymbol{\beta})'\,\mathbf{R}'] \\ &= \sigma^2\,\mathbf{R}\,(\mathbf{X'X})^{-1}\mathbf{R}' \qquad \therefore (15)\text{式} \end{aligned} \tag{43}$$

b は多変量正規分布 (multivariate normal distribution) にしたがうので、

$$\mathbf{Rb} = N\left(\mathbf{R\beta}\ ,\ \ \sigma^2\,\mathbf{R}\,(\mathbf{X'X})^{-1}\mathbf{R}'\right) \tag{44}$$

また (42) 式の仮説、つまり $\mathbf{Rb} = \mathbf{r}$ が真であるならば、次のような正規分布に変換できる。

$$(\mathbf{Rb}-\mathbf{r}) = N\left(\mathbf{0}\ ,\ \ \sigma^2\,\mathbf{R}\,(\mathbf{X'X})^{-1}\mathbf{R}'\right) \tag{45}$$

したがって $(\mathbf{Rb}-\mathbf{r})$ の二乗を分散で割った数値は、自由度 $(k-1)$ のカイ二乗分布にしたがう。

$$(\mathbf{Rb}-\mathbf{r})'\left[\sigma^2\,\mathbf{R}\,(\mathbf{X'X})^{-1}\mathbf{R}'\right]^{-1}(\mathbf{Rb}-\mathbf{r}) \sim \chi^2\,(k-1) \tag{46}$$

さらに (41) 式により、残差平方和を分散で割った数値も自由度 $(n-k)$ のカイ二乗分布にしたがう。

$$\frac{\mathbf{e'e}}{\sigma^2} \sim \chi^2\,(n-k) \tag{47}$$

以上の (46) 式と (47) 式とを組み合わせることにより、未知の分散 σ^2 を取り除くことができる。そして、もし帰無仮説 $\mathbf{Rb} = \mathbf{r}$ が真であるならば、(46) 式を (47) 式で割った数値は、自由度が $k-1$ と $n-k$ の F 分布にしたがう。この F 検定は、モデルに含まれる説明変数の連結有意性 (joint significance) を検証するものである。

$$\frac{(\mathbf{Rb}-\mathbf{r})'\,[\mathbf{R}\,(\mathbf{X'X})^{-1}\mathbf{R}']^{-1}\,(\mathbf{Rb}-\mathbf{r})\big/(k-1)}{\mathbf{e'e}\big/(n-k)} \sim F\,(k-1\ ,\ \ n-k) \tag{48}$$

次に、重回帰に含まれている特定の傾き β_j が、ゼロであるという仮説の検証を考えてみる。この場合には、**R** と **r** はそれぞれ次のような行列になる。

$$\mathrm{H}_0\ :\ \beta_j = 0$$

つまり、

$$\mathbf{R} = \begin{array}{c} \begin{array}{cccccccc} 1 & & (j-1) & j & & & \end{array} \\ \left(\begin{array}{cccccccc} 0 & \cdots & 0 & 1 & 0 & \cdots & 0 \end{array}\right) \end{array}\text{、および}\quad \mathbf{r} = \mathbf{0} \tag{49}$$

(49) 式の$\mathbf{R}$と$\mathbf{r}$とを (48) 式に代入すると、$\mathbf{Rb}=b_j$、また$\mathbf{R}(\mathbf{X'X})^{-1}\mathbf{R'}=c_{jj}$（一定値）なので、

$$F=\frac{b_j^2}{s^2c_{jj}}\sim F(1,\ n-k) \qquad \therefore s^2=\frac{\mathbf{e'e}}{n-k}$$

また自由度 $(1\ ,\ n-k)$ の F 分布にしたがうのであれば、この平方根は自由度 $(n-k)$ の t 分布にしたがうことから、

$$t=\frac{b_j}{s\sqrt{c_{jj}}}\sim t(n-k)$$

例題

これらの結果を例題を用いて確認する。いま、次のように *X2* から *X4* の３つの説明変数によって、被説明変数 (*Y*) を説明するサンプル数１０の重回帰モデルを分析してゆく。

n	X2	X3	X4	Y
1	25	111	123	16
2	27	118	126	16
3	9	107	109	15
4	-4	105	96	14
5	-15	93	85	12
6	-7	90	93	12
7	1	112	101	14
8	2	109	102	13
9	15	123	119	13
10	18	113	126	12

まず、ここでの回帰モデルが全体としてどの程度説明力をもつのかについては、(48) 式の各構成要素に数値を代入する必要がある。この場合の $\mathbf{R}$ は、(42) 式と同様なものになる。

$$\mathbf{X}=\begin{pmatrix} 1 & 25 & 111 & 123 \\ 1 & 27 & 118 & 126 \\ 1 & 9 & 107 & 109 \\ 1 & -4 & 105 & 96 \\ 1 & -15 & 93 & 85 \\ 1 & -7 & 90 & 93 \\ 1 & 1 & 112 & 101 \\ 1 & 2 & 109 & 102 \\ 1 & 15 & 123 & 119 \\ 1 & 18 & 113 & 126 \end{pmatrix} \qquad \mathbf{R}=\begin{pmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}$$

まず分子については、それぞれ

$$\left[\mathbf{R}(\mathbf{X'X})^{-1}\mathbf{R'}\right]^{-1} = \begin{pmatrix} 1{,}774.9 & 1{,}012.9 & 1{,}838.0 \\ & 934.9 & 1{,}096.0 \\ & & 1{,}978.0 \end{pmatrix}$$

$$\mathbf{Rb} = \begin{pmatrix} 0.458 \\ 0.035 \\ -0.394 \end{pmatrix} = \begin{pmatrix} b_2 \\ b_3 \\ b_4 \end{pmatrix}$$

により、

$$(\mathbf{Rb})'\left[\mathbf{R}(\mathbf{X'X})^{-1}\mathbf{R'}\right]^{-1}\mathbf{Rb} = 19.554$$

また分母の残差平方和については、推定値 **b** を導き出すことによって求めることができる。**b** の各推定値は、

$$(\mathbf{X'X})^{-1}\mathbf{X'y} = \begin{pmatrix} 49.215 \\ 0.458 \\ 0.035 \\ -0.394 \end{pmatrix} = \begin{pmatrix} b_1 \\ b_2 \\ b_3 \\ b_4 \end{pmatrix}$$

これらの推定値を用いることによって、総平方和を説明される平方和と残差平方和に分解でき、さらに決定係数も測定される。

$$\mathrm{TSS} = 22.100、\mathrm{ESS} = 19.554、\mathrm{RSS} = 2.546、R^2 = 0.885$$

したがって例題の重回帰モデルは、自由度が3と6の F 分布にしたがう。この場合の5% 水準の F 値は4.76であり、モデル全体としての有意性が究めて高いことが理解できる。

$$F = \frac{19.554/3}{2.546/(10-4)} = 15.361$$

次に、モデル全体ではなく、ある特定の傾き b_3 がどの程度の説明力をもつかについて考察する。このような分析目的においても、行列 **R** の形状だけを操作することによって容易に行うことができる。(49) 式により、ここでの **R** は次のようなものになる。

$$\mathbf{R} = \begin{pmatrix} 0 & 0 & 1 & 0 \end{pmatrix}$$

計算を進めてゆくと、

$$\mathbf{Rb} = b_3 = 0.035$$

$$(\mathbf{Rb})'\left[\mathbf{R}(\mathbf{X'X})^{-1}\mathbf{R'}\right]^{-1}\mathbf{Rb} = 0.390$$

などが導かれ、F 値は 0.920 になる。ただしこの場合には、分子の自由度は３から１に変わることに注意する必要がある。また自由度が１、６の F 分布は、自由度６の t 分布にしたがうので、F 値の平方根によって t 検定を行うことも可能である。また表１に、例題についての統計パッケージによる結果も併せて掲載している。

$$F = \frac{0.390/1}{2.546/(10-4)} = 0.920$$

$$t = \sqrt{F} = \sqrt{0.920} = 0.959$$

（表１）統計パッケージによる計算結果

DEP VAR: Y N: 10
MULTIPLE R: 0.941
SQUARED MULTIPLE R: 0.885
ADJUSTED SQUARED MULTIPLE R: 0.827
STANDARD ERROR OF ESTIMATE: 0.651

VARIABLE	COEFFICIENT	STD ERROR	STD COEF	T	P(2 TAIL)
CONSTANT	49.215	7.731	0.000	6.366	.71E-03
X2	0.458	0.080	4.108	5.756	0.001
X3	0.035	0.036	0.225	0.959	0.375
X4	-0.394	0.079	-3.723	-4.998	0.002

ANALYSIS OF VARIANCE

SOURCE	SUM-OF-SQUARES	DF	MEAN-SQUARE	F-RATIO	P
REGRESSION	19.554	3	6.518	15.361	0.003
RESIDUAL	2.546	6	0.424		