

2010 年度 学部 エコノメトリックス 第 10 回講義メモ

2010 年 11 月 9 日

2

3.3 線形重回帰モデルの行列表現

3.2 節と同様にして、線形重回帰モデルを行列・ベクトルを使って表現してみよう。

線形重回帰モデルは、一般に k 個の説明変数の一次式

$$Y_i = \beta_1 X_{1,i} + \beta_2 X_{2,i} + \cdots + \beta_k X_{k,i} + \epsilon_i, \quad (i = 1, \dots, n) \quad (3.3.1)$$

で表現されるモデルである。定数項を含む場合には $X_{1,i} = 1$, ($i = 1, \dots, n$) であるとする。

3

いま、 $n \times 1$ のベクトル $\mathbf{y}$, $\mathbf{1}$, $\mathbf{x}_j$, $\boldsymbol{\epsilon}$ を以下のように定義する。

$$\mathbf{y} = \begin{pmatrix} Y_1 \\ \vdots \\ Y_n \end{pmatrix}, \quad \mathbf{1} = \begin{pmatrix} 1 \\ \vdots \\ 1 \end{pmatrix}, \quad \mathbf{x}_j = \begin{pmatrix} X_{j,1} \\ \vdots \\ X_{j,n} \end{pmatrix}, \quad \boldsymbol{\epsilon} = \begin{pmatrix} \epsilon_1 \\ \vdots \\ \epsilon_n \end{pmatrix}$$

このとき、定数項 β_1 を含む線形重回帰モデルは、

$$\mathbf{y} = \beta_0 \mathbf{1} + \beta_2 \mathbf{x}_2 + \cdots + \beta_k \mathbf{x}_k + \boldsymbol{\epsilon} \quad (3.3.2)$$

として表すことができる。さらに、

$$\mathbf{X} = (\mathbf{1} \quad \mathbf{x}_2 \quad \cdots \quad \mathbf{x}_k) = \begin{pmatrix} 1 & X_{2,1} & \cdots & X_{k,1} \\ \vdots & \vdots & & \vdots \\ 1 & X_{2,n} & \cdots & X_{k,n} \end{pmatrix}, \quad \boldsymbol{\beta} = \begin{pmatrix} \beta_1 \\ \beta_2 \\ \vdots \\ \beta_k \end{pmatrix}$$

4

とすると、(3.3.2) 式は

$$\mathbf{y} = \mathbf{X}\boldsymbol{\beta} + \boldsymbol{\epsilon} \quad (3.3.3)$$

になる。なお、攪乱項に関する仮定が

仮定 2' $\epsilon_i \sim iid(0, \sigma^2)$ の場合には、 $\boldsymbol{\epsilon} \sim IID(\mathbf{0}, \sigma^2 \mathbf{I}_n)$

仮定 5 $\epsilon_i \sim N(0, \sigma^2)$ の場合には、 $\boldsymbol{\epsilon} \sim N(\mathbf{0}, \sigma^2 \mathbf{I}_n)$

と表記することにする。

これより、 $\boldsymbol{\beta} = (\beta_1, \beta_2, \dots, \beta_k)'$ の最小 2 乗推定量 $\mathbf{b} = (b_1, b_2, \dots, b_k)'$ は、 $k \times k$ 行列である $\mathbf{X}'\mathbf{X}$ に逆行列 $(\mathbf{X}'\mathbf{X})^{-1}$ が存在すれば、

$$\mathbf{b} = (\mathbf{X}'\mathbf{X})^{-1} \mathbf{X}'\mathbf{y} \quad (3.3.4)$$

と表現できる。

5

行列の階数と逆行列

$n \times k$ の行列 $\mathbf{X}$ の階数は、 $k \leq n$ であることから $\text{rank}(\mathbf{X}) \leq k$ となる。等号が成立するのは、 $\mathbf{X}$ を構成するベクトル $\mathbf{x}_j$ ($j = 1, \dots, k$) が一次独立である、すなわち、 $\alpha_1 = \cdots = \alpha_k = 0$ であるときのみに限って $\alpha_1 \mathbf{x}_1 + \cdots + \alpha_k \mathbf{x}_k = \mathbf{0}$ が成立する、ときに限られる。

ところで、 $\mathbf{X}'\mathbf{X}$ が逆行列を持つ必要十分条件は、full rank であること、すなわち $\text{rank}(\mathbf{X}'\mathbf{X}) = k$ であり、また

$$\text{rank}(\mathbf{X}) = \text{rank}(\mathbf{X}') = \text{rank}(\mathbf{X}'\mathbf{X})$$

という関係があるので、 $\mathbf{x}_j$ ($j = 1, \dots, k$) が一次独立であるときに限って $\mathbf{X}'\mathbf{X}$ が逆行列を持つ。

さて、残差ベクトルを $\mathbf{e} = (e_1, \dots, e_n)'$ で表すと、

$$\mathbf{e} = \mathbf{y} - \mathbf{X}\mathbf{b}$$

6

となるが、

$$E\left(\sum_{i=1}^n e_i^2\right) = E(\mathbf{e}'\mathbf{e}) = (n - k)\sigma^2$$

となることから、攪乱項 ϵ_i の分散 σ^2 の不偏推定量は

$$\hat{\sigma}^2 = \frac{1}{n - k} \mathbf{e}'\mathbf{e}$$

によって求めることができる。

7

3.4 OLS 推定量の幾何的理解

射影行列

いま、二つの行列

$$P_X \equiv X(X'X)^{-1}X',$$

$$M_X \equiv I - P_X = I - X(X'X)^{-1}X'$$

を考えよう。この行列を用いると、 y の推定量 $\hat{y}$ および残差ベクトル e はそれぞれ、

$$\hat{y} = Xb = P_X y$$

$$e = y - \hat{y} = (y - P_X y) = M_X y$$

で表される。

8

実は P_X は y を $Xb = \sum_{j=1}^k b_j X_j$ で表現される線形空間へ射影する変換行列、 M_X は y を Xb に直交する空間へ射影する変換行列になっており、射影行列と呼ばれている。

定理 3.4.1

P_X, M_X には次の性質がある。

- (1) $(P_X)' = P_X, (M_X)' = M_X$ (対称)
- (2) $(P_X)^2 = P_X, (M_X)^2 = M_X$ (べき等)
- (3) $P_X X = X, X' P_X = X'$
- (4) $M_X X = 0, X' M_X = 0$

さらに、定理 3.4.1 より次のことが得られる。

9

系 3.4.2

- (1) $X'e = 0$
- (2) $\hat{y}'e = 0$
- (3) $P_X \hat{y} = \hat{y}$
- (4) $M_X e = e$

系 3.4.1 の (1) より、残差ベクトル e は、線形空間 Xb の法線ベクトルになっていることが分かる。

10

図による理解

$k = 1$ のケース (図 3.4.a)

$$y = b_1 x_1 + e$$

$k = 2$ のケース (図 3.4.b)

$$y = b_1 x_1 + b_2 x_2 + e$$

もし、 $\alpha_1 x_1 + \alpha_2 x_2 = 0$ となる $\alpha = (\alpha_1, \alpha_2)'$ が存在すれば、 $\text{rank}(X'X) = 1$ となり、 $b = (b_1, b_2)'$ が一意に定まらない。この状況を **多重共線性** (multicollinearity) と呼ぶ。

図 3.4.a

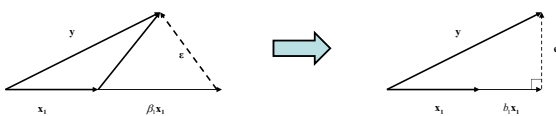


図 3.4.b

