第 8 回：重回帰モデルの推定

【教科書第 6 章】

北村　友宏

2020 年 11 月 20 日

本日の内容

1. 重回帰モデル

重回帰

- 定数項以外に説明変数が複数ある回帰モデルを重回帰モデル（multiple regression model）という．

定数項以外に説明変数が k 個ある場合，

$$y_i = \beta_0 + \beta_1 x_{i1} + \beta_2 x_{i2} + \cdots + \beta_k x_{ik} + u_i,$$
$$i = 1, 2, \cdots, n.$$

各観測値の式を並べると,

$$
\begin{aligned}
y_1 &= \beta_0 + \beta_1 x_{11} + \beta_2 x_{12} + \cdots + \beta_k x_{1k} + u_1,\\
y_2 &= \beta_0 + \beta_1 x_{21} + \beta_2 x_{22} + \cdots + \beta_k x_{2k} + u_2,\\
&\vdots\\
y_n &= \beta_0 + \beta_1 x_{n1} + \beta_2 x_{n2} + \cdots + \beta_k x_{nk} + u_n.
\end{aligned}
$$

ベクトル・行列を用いて表示すると,

$$
\begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{bmatrix} = \begin{bmatrix} 1 & x_{11} & x_{12} & \cdots & x_{1k} \\ 1 & x_{21} & x_{22} & \cdots & x_{2k} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & x_{n1} & x_{n2} & \cdots & x_{nk} \end{bmatrix} \begin{bmatrix} \beta_0 \\ \beta_1 \\ \beta_2 \\ \vdots \\ \beta_k \end{bmatrix} + \begin{bmatrix} u_1 \\ u_2 \\ \vdots \\ u_n \end{bmatrix}.
$$

$$\boldsymbol{y} = \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{bmatrix}, \boldsymbol{X} = \begin{bmatrix} 1 & x_{11} & x_{12} & \cdots & x_{1k} \\ 1 & x_{21} & x_{22} & \cdots & x_{2k} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & x_{n1} & x_{n2} & \cdots & x_{nk} \end{bmatrix}, \boldsymbol{\beta} = \begin{bmatrix} \beta_0 \\ \beta_1 \\ \beta_2 \\ \vdots \\ \beta_k \end{bmatrix},$$

$\boldsymbol{u} = \begin{bmatrix} u_1 \\ u_2 \\ \vdots \\ u_n \end{bmatrix}$ とすると，重回帰モデルは次のように簡潔に表すことができる．

$$\begin{aligned} \boldsymbol{y} &= \boldsymbol{X}\boldsymbol{\beta} + \boldsymbol{u}, \\ E(\boldsymbol{u} \mid \boldsymbol{X}) &= \boldsymbol{0}, \\ V(\boldsymbol{u} \mid \boldsymbol{X}) &= \sigma^2 \boldsymbol{I}_n. \end{aligned}$$

モデルを

$$\boldsymbol{y} = \boldsymbol{X}\hat{\boldsymbol{\beta}} + \boldsymbol{e},$$

と書き換え,

$$\sum_{i=1}^{n} e_i^2 = \boldsymbol{e}'\boldsymbol{e} = \left(\boldsymbol{y} - \boldsymbol{X}\hat{\boldsymbol{\beta}}\right)' \left(\boldsymbol{y} - \boldsymbol{X}\hat{\boldsymbol{\beta}}\right),$$

が最小になるように OLS 推定量を求めると,

$$\hat{\boldsymbol{\beta}} = (\boldsymbol{X}'\boldsymbol{X})^{-1}\boldsymbol{X}'\boldsymbol{y}.$$

- $\boldsymbol{e} = \begin{bmatrix} e_1 \\ e_2 \\ \vdots \\ e_n \end{bmatrix}.$

重回帰モデルの OLS 推定量の導出

残差二乗和は,

$$
\begin{aligned}
\sum_{i=1}^{n} e_i^2 &= \boldsymbol{e}'\boldsymbol{e} = (\boldsymbol{y} - \boldsymbol{X}\hat{\boldsymbol{\beta}})'(\boldsymbol{y} - \boldsymbol{X}\hat{\boldsymbol{\beta}}) \\
&= \left[\boldsymbol{y}' - (\boldsymbol{X}\hat{\boldsymbol{\beta}})'\right](\boldsymbol{y} - \boldsymbol{X}\hat{\boldsymbol{\beta}}) \\
&= (\boldsymbol{y}' - \hat{\boldsymbol{\beta}}'\boldsymbol{X}')(\boldsymbol{y} - \boldsymbol{X}\hat{\boldsymbol{\beta}}) \\
&= \boldsymbol{y}'\boldsymbol{y} - \underbrace{\boldsymbol{y}'\boldsymbol{X}\hat{\boldsymbol{\beta}}}_{1\times 1\text{（スカラー）}} - \underbrace{\hat{\boldsymbol{\beta}}'\boldsymbol{X}'\boldsymbol{y}}_{1\times 1\text{（スカラー）}} + \hat{\boldsymbol{\beta}}'\boldsymbol{X}'\boldsymbol{X}\boldsymbol{\beta} \\
&= \boldsymbol{y}'\boldsymbol{y} - \boldsymbol{y}'\boldsymbol{X}\hat{\boldsymbol{\beta}} - (\hat{\boldsymbol{\beta}}'\boldsymbol{X}'\boldsymbol{y})' + \hat{\boldsymbol{\beta}}'\boldsymbol{X}'\boldsymbol{X}\boldsymbol{\beta} \\
&= \boldsymbol{y}'\boldsymbol{y} - \boldsymbol{y}'\boldsymbol{X}\hat{\boldsymbol{\beta}} - \boldsymbol{y}'(\boldsymbol{X}')'(\hat{\boldsymbol{\beta}}')' + \hat{\boldsymbol{\beta}}'\boldsymbol{X}'\boldsymbol{X}\boldsymbol{\beta} \\
&= \boldsymbol{y}'\boldsymbol{y} - \boldsymbol{y}'\boldsymbol{X}\hat{\boldsymbol{\beta}} - \boldsymbol{y}'\boldsymbol{X}\hat{\boldsymbol{\beta}} + \hat{\boldsymbol{\beta}}'\boldsymbol{X}'\boldsymbol{X}\boldsymbol{\beta} \\
&= \boldsymbol{y}'\boldsymbol{y} - 2\boldsymbol{y}'\boldsymbol{X}\hat{\boldsymbol{\beta}} + \hat{\boldsymbol{\beta}}'\boldsymbol{X}'\boldsymbol{X}\boldsymbol{\beta}.
\end{aligned}
$$

この残差二乗和

$$\sum_{i=1}^{n} e_i^2 = \boldsymbol{y}'\boldsymbol{y} - 2\boldsymbol{y}'\boldsymbol{X}\hat{\boldsymbol{\beta}} + \hat{\boldsymbol{\beta}}'\boldsymbol{X}'\boldsymbol{X}\boldsymbol{\beta}$$

が最小になるように OLS 推定量を求める．
残差二乗和最小化問題は，

$$\min_{\hat{\boldsymbol{\beta}} \in \mathbb{R}^{1+k}} \boldsymbol{y}'\boldsymbol{y} - 2\boldsymbol{y}'\boldsymbol{X}\hat{\boldsymbol{\beta}} + \hat{\boldsymbol{\beta}}'\boldsymbol{X}'\boldsymbol{X}\hat{\boldsymbol{\beta}}.$$

1 階条件は，

$$\frac{d}{d\hat{\boldsymbol{\beta}}}(\boldsymbol{y}'\boldsymbol{y} - 2\boldsymbol{y}'\boldsymbol{X}\hat{\boldsymbol{\beta}} + \hat{\boldsymbol{\beta}}'\boldsymbol{X}'\boldsymbol{X}\hat{\boldsymbol{\beta}}) = \boldsymbol{0}.$$

参考：内積の，ベクトルでの微分

$$\boldsymbol{a} = \begin{bmatrix} a_1 \\ a_2 \\ \vdots \\ a_n \end{bmatrix}, \boldsymbol{x} = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix},$$

つまり，$\boldsymbol{a}$ と $\boldsymbol{x}$ をそれぞれ $n \times 1$ ベクトルとすると，

$$\frac{d}{d\boldsymbol{x}}\boldsymbol{a}'\boldsymbol{x} = \begin{bmatrix} \frac{\partial}{\partial x_1}\boldsymbol{a}'\boldsymbol{x} \\ \frac{\partial}{\partial x_2}\boldsymbol{a}'\boldsymbol{x} \\ \vdots \\ \frac{\partial}{\partial x_n}\boldsymbol{a}'\boldsymbol{x} \end{bmatrix} = \boldsymbol{a}.$$

（証明）

$$\boldsymbol{a}'\boldsymbol{x} = a_1x_1 + a_2x_2 + \cdots + a_nx_n,$$

なので,

$$\frac{d}{d\boldsymbol{x}}\boldsymbol{a}'\boldsymbol{x} = \begin{bmatrix} \frac{\partial}{\partial x_1}\boldsymbol{a}'\boldsymbol{x} \\ \frac{\partial}{\partial x_2}\boldsymbol{a}'\boldsymbol{x} \\ \vdots \\ \frac{\partial}{\partial x_n}\boldsymbol{a}'\boldsymbol{x} \end{bmatrix} = \begin{bmatrix} \frac{\partial}{\partial x_1}(a_1x_1 + a_2x_2 + \cdots + a_nx_n) \\ \frac{\partial}{\partial x_2}(a_1x_1 + a_2x_2 + \cdots + a_nx_n) \\ \vdots \\ \frac{\partial}{\partial x_n}(a_1x_1 + a_2x_2 + \cdots + a_nx_n) \end{bmatrix}$$

$$= \begin{bmatrix} a_1 \\ a_2 \\ \vdots \\ a_n \end{bmatrix} = \boldsymbol{a}. \quad \text{（証明終）}$$

参考：2 次形式の，ベクトルでの微分

$$\boldsymbol{x} = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix}, \boldsymbol{A} = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{12} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{1n} & a_{2n} & \cdots & a_{nn} \end{bmatrix},$$

つまり，$\boldsymbol{x}$ を $n \times 1$ ベクトル，$\boldsymbol{A}$ を $n \times n$ 対称行列とすると，

$$\frac{d}{d\boldsymbol{x}} \boldsymbol{x}' \boldsymbol{A} \boldsymbol{x} = \begin{bmatrix} \frac{\partial}{\partial x_1} \boldsymbol{x}' \boldsymbol{A} \boldsymbol{x} \\ \frac{\partial}{\partial x_2} \boldsymbol{x}' \boldsymbol{A} \boldsymbol{x} \\ \vdots \\ \frac{\partial}{\partial x_n} \boldsymbol{x}' \boldsymbol{A} \boldsymbol{x} \end{bmatrix} = 2\boldsymbol{A}\boldsymbol{x}.$$

※ $\boldsymbol{A}$ が対称でなければ，$\frac{d}{d\boldsymbol{x}} \boldsymbol{x}' \boldsymbol{A} \boldsymbol{x} = (\boldsymbol{A} + \boldsymbol{A}')\boldsymbol{x}$.

（証明【$\boldsymbol{A}$ が対称の場合】）簡単化のため,

$$\boldsymbol{x} = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}, \boldsymbol{A} = \begin{bmatrix} a_{11} & a_{12} \\ a_{12} & a_{22} \end{bmatrix},$$

とすると,

$$\begin{aligned} \boldsymbol{x}'\boldsymbol{A}\boldsymbol{x} &= \begin{bmatrix} x_1 & x_2 \end{bmatrix} \begin{bmatrix} a_{11} & a_{12} \\ a_{12} & a_{22} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} \\ &= \begin{bmatrix} a_{11}x_1 + a_{12}x_2 & a_{12}x_1 + a_{22}x_2 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} \\ &= (a_{11}x_1 + a_{12}x_2)x_1 + (a_{12}x_1 + a_{22}x_2)x_2 \\ &= a_{11}x_1^2 + a_{12}x_1x_2 + a_{12}x_1x_2 + a_{22}x_2^2 \\ &= a_{11}x_1^2 + 2a_{12}x_1x_2 + a_{22}x_2^2, \end{aligned}$$

なので,

$$
\begin{aligned}
\frac{d}{d\boldsymbol{x}}\boldsymbol{x}'\boldsymbol{A}\boldsymbol{x} &= \begin{bmatrix} \frac{\partial}{\partial x_1}\boldsymbol{x}'\boldsymbol{A}\boldsymbol{x} \\ \frac{\partial}{\partial x_2}\boldsymbol{x}'\boldsymbol{A}\boldsymbol{x} \end{bmatrix} \\
&= \begin{bmatrix} \frac{\partial}{\partial x_1}(a_{11}x_1^2 + 2a_{12}x_1x_2 + a_{22}x_2^2) \\ \frac{\partial}{\partial x_2}(a_{11}x_1^2 + 2a_{12}x_1x_2 + a_{22}x_2^2) \end{bmatrix} \\
&= \begin{bmatrix} 2a_{11}x_1 + 2a_{12}x_2 \\ 2a_{12}x_1 + 2a_{22}x_2 \end{bmatrix} = \begin{bmatrix} 2(a_{11}x_1 + a_{12}x_2) \\ 2(a_{12}x_1 + a_{22}x_2) \end{bmatrix} \\
&= 2\begin{bmatrix} a_{11}x_1 + a_{12}x_2 \\ a_{12}x_1 + a_{22}x_2 \end{bmatrix} \\
&= 2\begin{bmatrix} a_{11} & a_{12} \\ a_{12} & a_{22} \end{bmatrix}\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} \\
&= 2\boldsymbol{A}\boldsymbol{x}. \quad \text{（証明終）}
\end{aligned}
$$

残差二乗和最小化問題の１階条件は,

$$
\begin{aligned}
&\frac{d}{d\hat{\boldsymbol{\beta}}}(\boldsymbol{y}'\boldsymbol{y} - 2\boldsymbol{y}'\boldsymbol{X}\hat{\boldsymbol{\beta}} + \hat{\boldsymbol{\beta}}'\boldsymbol{X}'\boldsymbol{X}\hat{\boldsymbol{\beta}}) = \mathbf{0} \\
\Leftrightarrow\ &\underbrace{\frac{d}{d\hat{\boldsymbol{\beta}}}\boldsymbol{y}'\boldsymbol{y}}_{=\mathbf{0}} + \frac{d}{d\hat{\boldsymbol{\beta}}}(-2\boldsymbol{y}'\boldsymbol{X}\hat{\boldsymbol{\beta}}) + \frac{d}{d\hat{\boldsymbol{\beta}}}\hat{\boldsymbol{\beta}}'\boldsymbol{X}'\boldsymbol{X}\hat{\boldsymbol{\beta}} = \mathbf{0} \\
\Leftrightarrow\ &-2\cdot\frac{d}{d\hat{\boldsymbol{\beta}}}\boldsymbol{y}'\boldsymbol{X}\hat{\boldsymbol{\beta}} + \frac{d}{d\hat{\boldsymbol{\beta}}}\hat{\boldsymbol{\beta}}'\boldsymbol{X}'\boldsymbol{X}\hat{\boldsymbol{\beta}} = \mathbf{0}.
\end{aligned}
$$

- $\boldsymbol{y}'\boldsymbol{X}\hat{\boldsymbol{\beta}}$ の中で,
    - $\boldsymbol{y}'\boldsymbol{X}$（$1 \times (1+k)$ ベクトル）が $\boldsymbol{a}'\boldsymbol{x}$ の $\boldsymbol{a}'$ に相当
    - $\hat{\boldsymbol{\beta}}$（$(1+k) \times 1$ ベクトル）が $\boldsymbol{a}'\boldsymbol{x}$ の $\boldsymbol{x}$ に相当
- $\hat{\boldsymbol{\beta}}'\boldsymbol{X}'\boldsymbol{X}\hat{\boldsymbol{\beta}}$ の中で,
    - $\hat{\boldsymbol{\beta}}'$（$1 \times (1+k)$ ベクトル）が $\boldsymbol{x}'\boldsymbol{A}\boldsymbol{x}$ の $\boldsymbol{x}'$ に相当
    - $\boldsymbol{X}'\boldsymbol{X}$（$(1+k) \times (1+k)$ 対称行列）が $\boldsymbol{x}'\boldsymbol{A}\boldsymbol{x}$ の $\boldsymbol{A}$ に相当
    - $\hat{\boldsymbol{\beta}}$（$(1+k) \times 1$ ベクトル）が $\boldsymbol{x}'\boldsymbol{A}\boldsymbol{x}$ の $\boldsymbol{x}$ に相当

のように考えれば,

$$
\begin{aligned}
& -2 \cdot \frac{d}{d\hat{\boldsymbol{\beta}}} \boldsymbol{y}' \boldsymbol{X} \hat{\boldsymbol{\beta}} + \frac{d}{d\hat{\boldsymbol{\beta}}} \hat{\boldsymbol{\beta}}' \boldsymbol{X}' \boldsymbol{X} \hat{\boldsymbol{\beta}} = \mathbf{0} \\
\Leftrightarrow & -2(\boldsymbol{y}'\boldsymbol{X})' + 2\boldsymbol{X}'\boldsymbol{X}\hat{\boldsymbol{\beta}} = \mathbf{0} \\
\Leftrightarrow & -2\boldsymbol{X}'(\boldsymbol{y}')' + 2\boldsymbol{X}'\boldsymbol{X}\hat{\boldsymbol{\beta}} = \mathbf{0} \\
\Leftrightarrow & 2\boldsymbol{X}'\boldsymbol{X}\hat{\boldsymbol{\beta}} = 2\boldsymbol{X}'\boldsymbol{y} \\
\Leftrightarrow & \boldsymbol{X}'\boldsymbol{X}\hat{\boldsymbol{\beta}} = \boldsymbol{X}'\boldsymbol{y} \\
\Leftrightarrow & \underbrace{(\boldsymbol{X}'\boldsymbol{X})^{-1}\boldsymbol{X}'\boldsymbol{X}}_{=\boldsymbol{I}_{1+k}}\hat{\boldsymbol{\beta}} = (\boldsymbol{X}'\boldsymbol{X})^{-1}\boldsymbol{X}'\boldsymbol{y} \\
\Leftrightarrow & \hat{\boldsymbol{\beta}} = (\boldsymbol{X}'\boldsymbol{X})^{-1}\boldsymbol{X}'\boldsymbol{y}.
\end{aligned}
$$