2010年度 エコノメトリックスII & 上級エコノメトリックスII 第3回講義メモ

2010年10月8日

§2.2 線形回帰モデルの幾何的理解

2.2.1 準備：ベクトル空間

いま、すべての要素が実数値をとる $n \times 1$ ベクトル $\boldsymbol{x}$ を考えることにしよう。すなわち

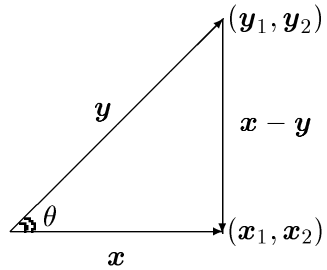
$$\boldsymbol{x} \in \boldsymbol{R}^n$$

である。n 次元の実数空間 $\boldsymbol{R}^n$ に距離と内積を導入する。厳密に言うと、n 次元のベクトル空間は $\boldsymbol{R}^n$ を単なる集合でなく $\boldsymbol{R}^n$ の演算まで入れて考えたものを指している。

$\boldsymbol{x}$, $\boldsymbol{y} \in \boldsymbol{R}^n$ としたとき、距離 $d(\boldsymbol{x}, \boldsymbol{y})$、内積 $< \boldsymbol{x}, \boldsymbol{y} >$ を次のように定義する。

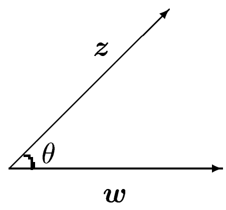
$$\begin{aligned} d(\boldsymbol{x}, \boldsymbol{y}) &\equiv \sqrt{(\boldsymbol{x} - \boldsymbol{y})'(\boldsymbol{x} - \boldsymbol{y})} \\ < \boldsymbol{x}, \boldsymbol{y} > &\equiv \|\boldsymbol{x}\| \|\boldsymbol{y}\| \cos\theta = \boldsymbol{x}'\boldsymbol{y} \\ &\quad \text{where} \quad \|\boldsymbol{x}\| = \sqrt{\boldsymbol{x}'\boldsymbol{x}} \quad (\text{ノルム}) \end{aligned}$$

$\boldsymbol{R}^n$ に距離と内積を導入した空間を n 次元ユークリッド空間といい $\boldsymbol{E}^n$ で表す。



Note.

$\boldsymbol{w}, \boldsymbol{z} \in \boldsymbol{E}^2$ とする。いま $\boldsymbol{w} = (1,0)'$ とする。ここで $\|\boldsymbol{w}\| = 1, \|\boldsymbol{z}\| = 1$ であるものとする。このとき、$\|\boldsymbol{z}\|^2 = \cos^2\theta + \sin^2\theta = 1$ であるから、$\boldsymbol{z} = (\cos\theta, \sin\theta)'$ と表すことができる。これより、$<\boldsymbol{w}, \boldsymbol{z}> = \boldsymbol{w}'\boldsymbol{z} = \cos\theta$ が得られる。



一般化して $\boldsymbol{x} = \alpha\boldsymbol{w}, \boldsymbol{y} = \gamma\boldsymbol{z} (\alpha, \gamma > 0)$ とすると、

$$\begin{aligned} <\boldsymbol{x}, \boldsymbol{y}> &= \alpha\gamma <\boldsymbol{w}, \boldsymbol{z}> \\ &= \|\boldsymbol{x}\|\|\boldsymbol{y}\| \cos\theta \end{aligned}$$

がえられる。ところで、$-1 \leq \cos\theta \leq 1$ であるから、

$$|\ \boldsymbol{x}'\boldsymbol{y}\ | = \|\boldsymbol{x}\|\|\boldsymbol{y}\|\ |\cos\theta\ | \leq \|\boldsymbol{x}\|\|\boldsymbol{y}\|$$

という不等式が得られる。なお、$|\ \boldsymbol{x}'\boldsymbol{y}\ |^2 \leq \|\boldsymbol{x}\|^2\|\boldsymbol{y}\|^2$ を Cauchy-Schwartz の不等式と言う。

線形部分空間 (subspace)

$\boldsymbol{R}^n$ の空でない部分集合 $\mathcal{W}$ が以下の条件を満たすとき、$\mathcal{W}$ は $\boldsymbol{R}^n$ の線型部分空間（linear Subspace）であるという。

a) if $\boldsymbol{x}, \boldsymbol{y} \in \mathcal{W}$ then $\boldsymbol{x} + \boldsymbol{y} \in \mathcal{W}$

b) if $\boldsymbol{x} \in \mathcal{W},\ a \in \boldsymbol{R}$ then $a\boldsymbol{x} \in \mathcal{W}$

テキストでは、$\boldsymbol{x}_i \in \mathcal{W}, \quad b_i \in \boldsymbol{R}$ について $(i = 1, \ldots, n)$

$$\sum_{i=1}^{n} b_i \boldsymbol{x}_i \in \mathcal{W}$$

として定義。そこで、再び定義し直すと、次のようになる。

部分空間（線形）

$\{\boldsymbol{x}_1, \ldots, \boldsymbol{x}_n\};\ \boldsymbol{x}_i \in \boldsymbol{E}^n$ を基底ベクトル (basis vector) とする。さらに、$k(< n)$ 次元の $\boldsymbol{E}^n$ の部分空間 $\mathcal{S}(\boldsymbol{x}_1, \ldots, \boldsymbol{x}_k)$ を次のように定義する。

$$\mathcal{S}(\boldsymbol{x}_1, \ldots, \boldsymbol{x}_k) \equiv \{\boldsymbol{z} \in \boldsymbol{E}^n \mid \boldsymbol{z} = \sum_{i=1}^{k} b_i \boldsymbol{x}_i, b_i \in \boldsymbol{R}\}$$

$\mathcal{S}(\boldsymbol{x}_1, \ldots, \boldsymbol{x}_k)$ は $\boldsymbol{x}_1, \ldots, \boldsymbol{x}_k$ が張る空間という。

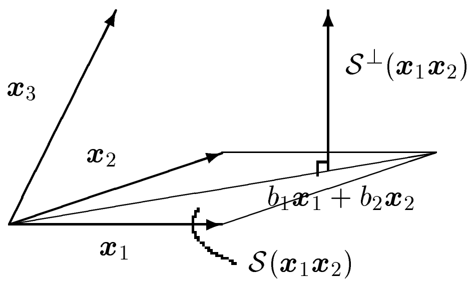
部分空間 $\mathcal{S}(\boldsymbol{x}_1, \cdots, \boldsymbol{x}_k)$ の直交補空間 $\mathcal{S}^{\perp}(\boldsymbol{x}_1, \cdots, \boldsymbol{x}_k)$ は次のように定義される。

$$\mathcal{S}^{\perp}(\boldsymbol{x}_1, \cdots, \boldsymbol{x}_k) \equiv \{\boldsymbol{w} \in \boldsymbol{E}^n \mid \boldsymbol{w}'\boldsymbol{z} = 0 \text{ for all } \boldsymbol{z} \in \mathcal{S}(\boldsymbol{x}_1, \ldots, \boldsymbol{x}_k)\}$$

直交補空間の次元は $n - k$ となる。

（例）$k = 3$ のケース

$\boldsymbol{x}_1, \boldsymbol{x}_2, \boldsymbol{x}_3 \in \boldsymbol{E}^n$ としよう。この時 $\mathcal{S}(\boldsymbol{x}_1, \boldsymbol{x}_2)$ は $\boldsymbol{x}_1$ と $\boldsymbol{x}_2$ の線形結合の作る空間全体となり、$\mathcal{S}^{\perp}(\boldsymbol{x}_1, \boldsymbol{x}_2)$ はその空間に直交する空間となる。



$\mathcal{S}(\boldsymbol{x}_1, \ldots, \boldsymbol{x}_k) \cap \mathcal{S}^{\perp}(\boldsymbol{x}_1, \ldots, \boldsymbol{x}_k) = \{\boldsymbol{0}\}$ 点：ゼロベクトルからなる集合

$\boldsymbol{E}^n = \mathcal{S}(\boldsymbol{x}_1, \ldots, \boldsymbol{x}_k) \oplus \mathcal{S}^{\perp}(\boldsymbol{x}_1, \ldots, \boldsymbol{x}_k)$

ここで、$\oplus$ は直和 (direct sum) を表す記号である。

Note: 直和

$\mathcal{W}, \mathcal{W}_1, \mathcal{W}_2$ を $\mathcal{V}$ (いまの例では $\boldsymbol{R}^n$) の部分空間としよう。このとき、$\mathcal{W}_1 \cap \mathcal{W}_2 = \{\boldsymbol{0}\}$ であって、

$$\mathcal{W} = \mathcal{W}_1 + \mathcal{W}_2$$

かつ、$u \in \mathcal{W}$, $u_i \in \mathcal{W}_i$ $(i = 1, 2)$ について

$$u = u_1 + u_2$$

の形に一意に表現できるとき、$\mathcal{W}$, は、$\mathcal{W}_1$, $\mathcal{W}_2$ の直和であるといい、

$$\mathcal{W} = \mathcal{W}_1, \oplus \mathcal{W}_2$$

で表す。

定義 2.2.1 一次独立

$\boldsymbol{x}_1, \cdots, \boldsymbol{x}_k$ が一次独立 (linearly independent) であるとは、次と同値である。

$$\sum_{i=1}^{k} b_i \boldsymbol{x}_i = 0, \quad \text{iff} \ \ b_i = 0 \ \forall_i$$

定義 2.2.2 一次従属

$\boldsymbol{x}_1, \cdots, \boldsymbol{x}_k$ が一次従属 (linearly dependent) であるとは、次と同値である。

$$\sum_{i=1}^{k} c_i \boldsymbol{x}_i = 0 \; \exists c_i \neq 0$$

命題 2.2.3

いま、$\boldsymbol{X} = (\boldsymbol{x}_1, \ldots, \boldsymbol{x}_k)$ であるものとする。もし、$\boldsymbol{x}_1, \ldots, \boldsymbol{x}_k$ が一次独立でない、すなわち一次従属であるならば、$\boldsymbol{X}'\boldsymbol{X}$ は可逆 (invertible) ではない。
[証明] text p.53 参照。もし、$\boldsymbol{x}_1, \ldots, \boldsymbol{x}_k$ が一次従属であるならば、$\boldsymbol{X}\boldsymbol{b} = \boldsymbol{0}$ を満たす $\boldsymbol{b} \neq \boldsymbol{0}$ が存在する。

$$\boldsymbol{X}\boldsymbol{b} = \boldsymbol{0} \text{であるから、} \boldsymbol{X}'\boldsymbol{X}\boldsymbol{b} = \boldsymbol{0}$$

もし $\boldsymbol{X}'\boldsymbol{X}$ が可逆であるなら、

$$(\boldsymbol{X}'\boldsymbol{X})^{-1}(\boldsymbol{X}'\boldsymbol{X}) = \boldsymbol{I}$$
$$\boldsymbol{b} = \boldsymbol{I}\boldsymbol{b} = (\boldsymbol{X}'\boldsymbol{X})^{-1}(\boldsymbol{X}'\boldsymbol{X})\boldsymbol{b} = \boldsymbol{0}$$

を満たす $(\boldsymbol{X}'\boldsymbol{X})^{-1}$ が存在することになる。これは $\boldsymbol{b} \neq \boldsymbol{0}$ であることに反する。

部分空間の次元

いま、行列 $\boldsymbol{X} = (\boldsymbol{x}_1, \cdots, \boldsymbol{x}_k)$ を考えよう。$\boldsymbol{X}$ の列ベクトルのうち、一次独立であるものの（最大）個数を $\boldsymbol{X}$ の階数 (rank) といい、$\mathrm{rank}(\boldsymbol{X})$ で表す。このとき、

$$\mathrm{rank}(\boldsymbol{X}) = \dim \mathcal{S}(\boldsymbol{X})$$

という関係が成立する。

定理 2.2.4

もし、$\dim \mathcal{S}(\boldsymbol{X}) = k',\; k' < k$ であるならば、

$$\mathcal{S}\underbrace{(\boldsymbol{X})}_{n \times k} = \mathcal{S}\underbrace{(\boldsymbol{X}^*)}_{n \times k'}$$

が成立する。ここで、$\boldsymbol{X}^*$ は $\boldsymbol{X}$ のうち、一次独立な列ベクトルからなる行列である。

§2.2.2 OLSE の幾何的理解

さて、前節のフレームワークを使って、OLS 推定量の幾何的理解を試みよう。

線形回帰モデル $\boldsymbol{y} = \boldsymbol{X\beta} + \boldsymbol{u}$ の OLSE は、$\hat{\boldsymbol{\beta}} = (\boldsymbol{X}'\boldsymbol{X})\boldsymbol{X}'\boldsymbol{y}$ で与えられる。いま、残差ベクトル $\hat{\boldsymbol{u}}$ を

$$\hat{\boldsymbol{u}} = \boldsymbol{y} - \boldsymbol{X}\hat{\boldsymbol{\beta}}$$

と定義するとき、

$$\boldsymbol{X}'\hat{\boldsymbol{u}} = \boldsymbol{0} \tag{2.2.2.a}$$

が得られる。なぜなら

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

であるから、

$$\begin{aligned}
\boldsymbol{X}'\hat{\boldsymbol{u}} &= \boldsymbol{X}'(\boldsymbol{I} - \boldsymbol{X}(\boldsymbol{X}'\boldsymbol{X})^{-1}\boldsymbol{X}')\boldsymbol{y} \\
&= \boldsymbol{O}\boldsymbol{y} = \boldsymbol{0}
\end{aligned}$$

が得られる。

言い換えると、

$$\boldsymbol{X}'\hat{\boldsymbol{u}} = \boldsymbol{0} \qquad \Longleftrightarrow \qquad \bar{\boldsymbol{x}}_i'\hat{\boldsymbol{u}} =< \bar{\boldsymbol{x}}_i'\hat{\boldsymbol{u}} >= 0 \quad \forall i$$

すなわち、$\bar{\boldsymbol{x}}_i,\ (i = 1, \ldots, k)$ と $\hat{\boldsymbol{u}}$ は直交する。ここで、$\bar{\boldsymbol{x}}_i$ は説明変数行列 $\boldsymbol{X}$ の i 番目の列ベクトルを指し、この授業の記法では

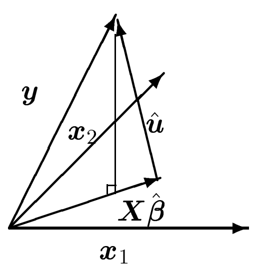
$$\boldsymbol{X} = (\bar{\boldsymbol{x}}_1, \bar{\boldsymbol{x}}_2, \cdots, \bar{\boldsymbol{x}}_k) = \begin{pmatrix} \boldsymbol{x}_1' \\ \boldsymbol{x}_2' \\ \vdots \\ \boldsymbol{x}_n' \end{pmatrix}$$

という関係になっている。

さて、$\hat{\boldsymbol{y}} \equiv \boldsymbol{X}\hat{\boldsymbol{\beta}}$ とすると、

$$\hat{\boldsymbol{y}} \in \mathcal{S}(\boldsymbol{X}), \quad \hat{\boldsymbol{u}} \in \mathcal{S}\perp(\boldsymbol{X})$$

という性質が導かれる。



上図より、$\|\ \hat{\boldsymbol{u}}\ \|$ が最小となるのは $\boldsymbol{X}\hat{\boldsymbol{\beta}}$ と $\hat{\boldsymbol{u}}$ が直交するときのみであることが分かる。ここから、モーメント法推定量 (MME) が最小 2 乗推定量 (OLSE) と等しくなることが分かる。

さらに、ピタゴラスの定理より

$$\begin{aligned}\|\, \boldsymbol{y} \,\|^2 &= \|\, \boldsymbol{X}\hat{\boldsymbol{\beta}} \,\|^2 + \|\, \hat{\boldsymbol{u}} \,\|^2 \\ \underbrace{\boldsymbol{y}'\boldsymbol{y}}_{TSS} &= \underbrace{(\boldsymbol{X}\hat{\boldsymbol{\beta}})'\boldsymbol{X}\hat{\boldsymbol{\beta}}}_{ESS} + \underbrace{(\boldsymbol{y}-\boldsymbol{X}\hat{\boldsymbol{\beta}})'(\boldsymbol{y}-\boldsymbol{X}\boldsymbol{\beta})}_{RSS}\end{aligned}$$

直交射影

$\boldsymbol{y} \in \boldsymbol{E}^n$ において

$$\boldsymbol{y} = \boldsymbol{z} + \boldsymbol{w} \qquad \boldsymbol{z} \in \mathcal{S}(\boldsymbol{X}),\ \ \boldsymbol{w} \in \mathcal{S}^{\perp}(\boldsymbol{X})$$

と表されるとき、$\boldsymbol{z}$ を $\boldsymbol{y}$ の $\mathcal{S}(\boldsymbol{X})$ の上への直交射影という。この直交射影を変換行列（＝射影行列）を使って考えるとわかりやすい。いまの例では、

$$\begin{aligned}\boldsymbol{z} &= \boldsymbol{P}_{\boldsymbol{X}}\boldsymbol{y} \\ \boldsymbol{w} &= (\boldsymbol{I} - \boldsymbol{P}_{\boldsymbol{X}})\boldsymbol{y}\end{aligned}$$

になる。

[memo]

- 部分空間に垂線をおろすイメージで捉えると理解しやすい。
- invariant sub-supace 不変部分空間　＝　部分空間への射影が部分空間で一定。

OLS では、次の 2 つの射影行列

$$\begin{aligned}\boldsymbol{P}_{\boldsymbol{X}} &= \boldsymbol{X}(\boldsymbol{X}'\boldsymbol{X})^{-1}\boldsymbol{X}' \\ \boldsymbol{M}_{\boldsymbol{X}} &= \boldsymbol{I} - \boldsymbol{X}(\boldsymbol{X}'\boldsymbol{X})^{-1}\boldsymbol{X}' = \boldsymbol{I} - \boldsymbol{P}_{\boldsymbol{X}}\end{aligned}$$

が重要である。なぜなら、

$$\begin{aligned}\hat{\boldsymbol{y}} = \boldsymbol{P}_{\boldsymbol{X}}\boldsymbol{y} &\in \mathcal{S}(\boldsymbol{X}) \\ \hat{\boldsymbol{u}} = \boldsymbol{M}_{\boldsymbol{X}}\boldsymbol{y} &\in \mathcal{S}^{\perp}(\boldsymbol{X})\end{aligned}$$

この直交性は $\boldsymbol{P}_{\boldsymbol{X}}$ が対称 (symmetric) であることに由来している。テキスト p.59 後半を見よ。
また、$\boldsymbol{M}_{\boldsymbol{X}}, \boldsymbol{P}_{\boldsymbol{X}}$ は冪等（べき等：idempotent）と呼ばれる行列であり、

$$\boldsymbol{M}_{\boldsymbol{X}}\boldsymbol{M}_{\boldsymbol{X}} = (\boldsymbol{M}_{\boldsymbol{X}})^{\ell} = \boldsymbol{M}_{\boldsymbol{X}}, \qquad (\ell = 1, 2, \ldots)$$

$$\boldsymbol{P}_{\boldsymbol{X}}\boldsymbol{P}_{\boldsymbol{X}} = (\boldsymbol{P}_{\boldsymbol{X}})^{\ell} = \boldsymbol{P}_{\boldsymbol{X}}, \qquad (\ell = 1, 2, \ldots)$$

という性質を持つ。
次に $\boldsymbol{P}_{\boldsymbol{X}}$、$\boldsymbol{M}_{\boldsymbol{X}}$ が onto であることを示そう。

$$\begin{aligned}\boldsymbol{P}_{\boldsymbol{X}}\boldsymbol{X}\boldsymbol{b} &= \boldsymbol{X}\boldsymbol{b} \\ \boldsymbol{M}_{\boldsymbol{X}}\hat{\boldsymbol{u}} &= \hat{\boldsymbol{u}} - \boldsymbol{X}(\boldsymbol{X}'\boldsymbol{X})^{-1}\boldsymbol{X}'\hat{\boldsymbol{u}} \ = \ \hat{\boldsymbol{u}}\end{aligned}$$

であるが、さらに $\boldsymbol{M_X X} = \boldsymbol{O}$ より、

$$\hat{\boldsymbol{u}} = \boldsymbol{M_X y} = \boldsymbol{M_X X\beta} + \boldsymbol{M_X u} = \boldsymbol{M_X u}$$

であることにも注意。

一般に、

$$\boldsymbol{I} = \boldsymbol{P_X} + \boldsymbol{M_X} \Rightarrow \quad \begin{cases} \boldsymbol{y} & = \quad \boldsymbol{P_X y} + \boldsymbol{M_X y} \\ & = \quad \hat{\boldsymbol{y}} + \hat{\boldsymbol{u}} \end{cases}$$

と分解できるが、$\boldsymbol{X}'\hat{\boldsymbol{u}} = \boldsymbol{0}$ とすると、$\boldsymbol{P_X M_X} = \boldsymbol{O}$ が得られる。これを annihilate（アナイアレイト）という。

ピタゴラスの定理より

$$\| \boldsymbol{P_X y} \|^2 \leq \| \boldsymbol{y} \|^2$$

Non Singlar transformation (非特異変換)

いま $\boldsymbol{A}$ を $k \times k$ の full rank の行列とする。このとき、

$$\begin{aligned} \mathcal{S}(\boldsymbol{X}) &= \mathcal{S}(\boldsymbol{XA}) \\ \boldsymbol{P_X} &= \boldsymbol{P_{XA}} \end{aligned}$$

が成立し、invariance residual の話がここから導き出せる。