回帰分析 (Regression Analysis)

緑川章一 *

1 回帰直線

1.1 導出

n コのデータ $(x_i,\ y_i)$ $(i = 1,\ \cdots,\ n)$ が与えられたとき、これらは直線

$$y = \theta_1 + \theta_2 x \tag{1}$$

の周りに確率的にちらばっているものとする。我々は、$\theta_1,\ \theta_2$ が存在することは知っているが、その値は分からない。データを

$$y_i = \theta_1 + \theta_2 x_i + \varepsilon_i \tag{2}$$

と表す。この時、ε_i は平均 0、分散 σ^2 の正規分布に従うとする。すなわち、

$$\mathrm{P}(y_i|\theta_1 + \theta_2 x_i) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left\{ -\frac{\big(y_i - (\theta_1 + \theta_2 x_i)\big)^2}{2\sigma^2} \right\} \tag{3}$$

となるので、

$$\begin{aligned} E(y_i) &= \int y_i \mathrm{P}(y_i|\theta_1 + \theta_2 x_i)\, dy_i = \theta_1 + \theta_2 x_i & (4) \\ V(y_i) &= E(y_i^2) - E(y_i)^2 = \sigma^2 & (5) \end{aligned}$$

である。同様に、

$$Cov(y_i, y_j) = E(y_i y_j) - E(y_i)E(y_j) = 0 \quad for \quad i \neq j \tag{6}$$

が成り立つとする。すなわち、$i \neq j$ のとき、ε_i と ε_j は無相関とする。

(1) ∼ (6) 式は、行列表示を用いて簡潔に表すことができる。まず、

$$A = \begin{bmatrix} 1 & x_1 \\ 1 & x_2 \\ \vdots & \vdots \\ 1 & x_n \end{bmatrix}, \qquad \boldsymbol{\theta} = \begin{bmatrix} \theta_1 \\ \theta_2 \end{bmatrix}, \qquad \boldsymbol{y} = \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{bmatrix}$$

*Shoichi Midorikawa

を定義する。
(3) 式は、

$$\mathrm{P}(\boldsymbol{y}|A\boldsymbol{\theta}) = \frac{1}{(\sqrt{2\pi\sigma^2})^n}\exp\left\{-\frac{(\boldsymbol{y}-A\boldsymbol{\theta})^T(\boldsymbol{y}-A\boldsymbol{\theta})}{2\sigma^2}\right\} \tag{7}$$

(4) 式は、

$$E(\boldsymbol{y}) = \int \boldsymbol{y}\,\mathrm{P}(\boldsymbol{y}|A\boldsymbol{\theta})\prod_{i=1}^{n} dy_i = A\boldsymbol{\theta} \tag{8}$$

(5), (6) 式は、

$$\begin{aligned} V(\boldsymbol{y}) &= \int (\boldsymbol{y}-E(\boldsymbol{y}))\,(\boldsymbol{y}-E(\boldsymbol{y}))^T\ \mathrm{P}(\boldsymbol{y}|A\boldsymbol{\theta})\prod_{i=1}^{n} dy_i \\ &= \int \boldsymbol{y}\boldsymbol{y}^T\,\mathrm{P}(\boldsymbol{y}|A\boldsymbol{\theta})\prod_{i=1}^{n} dy_i - E(\boldsymbol{y})E(\boldsymbol{y})^T \\ &= E(\boldsymbol{y}\boldsymbol{y}^T) - E(\boldsymbol{y})E(\boldsymbol{y})^T \\ &= \sigma^2\boldsymbol{I} \end{aligned} \tag{9}$$

となる。ここで、$\boldsymbol{I}$ は、$n \times n$ 単位行列である。

直線 (1) 式を求めるときに、普通は、

$$Q = \sum_{i=1}^{n}(\theta_1 + \theta_2 x_i - y_i)^2 \tag{10}$$

が最小になるような $\hat{\theta}_1,\ \hat{\theta}_2$ を求める。
まず、

$$\begin{aligned} \frac{\partial Q}{\partial \theta_1} &= 2\sum_{i=1}^{n}(\theta_1 + \theta_2 x_i - y_i) \\ &= 2\left\{n\theta_1 + \theta_2\sum_{i=1}^{n} x_i - \sum_{i=1}^{n} y_i\right\} = 0 \end{aligned} \tag{11}$$

より、

$$\theta_1 + \theta_2\bar{x} - \bar{y} = 0 \tag{12}$$

を得る。ここで、

$$\bar{x} = \frac{1}{n}\sum_{i=1}^{n} x_i, \qquad \bar{y} = \frac{1}{n}\sum_{i=1}^{n} y_i$$

である。
　次に、

$$\frac{\partial Q}{\partial \theta_2} = 2\left\{\theta_1 \sum_{i=1}^{n} x_i + \theta_2 \sum_{i=1}^{n} x_i^2 - \sum_{i=1}^{n} x_i y_i\right\} = 0 \tag{13}$$

を得る。
　または、(12) 式より、

$$\theta_1 = \bar{y} - \theta_2 \bar{x} \tag{14}$$

となるので、これを (10) 式に代入する。

$$Q = \sum_{i=1}^{n} \left[\theta_2(x_i - \bar{x}) - (y_i - \bar{y})\right]^2$$

ゆえに、

$$\begin{aligned}
\frac{\partial Q}{\partial \theta_2} &= 2\sum_{i=1}^{n}(x_i - \bar{x})\left[\theta_2(x_i - \bar{x}) - (y_i - \bar{y})\right] \\
&= 2\left\{\theta_2 \sum_{i=1}^{n}(x_i - \bar{x})^2 - \sum_{i=1}^{n}(x_i - \bar{x})(y_i - \bar{y})\right\} \\
&= 0
\end{aligned}$$

より、

$$\hat{\theta}_2 = \frac{\displaystyle\sum_{i=1}^{n}(x_i - \bar{x})(y_i - \bar{y})}{\displaystyle\sum_{i=1}^{n}(x_i - \bar{x})^2} \tag{15}$$

を得る。これを (14) 式に代入して、

$$\hat{\theta}_1 = \bar{y} - \hat{\theta}_2 \bar{x} \tag{16}$$

を得る。

2　射影行列

　ベクトル $\boldsymbol{b}$ の $\boldsymbol{a}$ への射影を $\boldsymbol{p}$ とする。

$$\boldsymbol{p} = k\boldsymbol{a} \tag{17}$$

さらに、$\boldsymbol{e} = \boldsymbol{p} - \boldsymbol{b}$ とおく。$\boldsymbol{e} \perp \boldsymbol{a}$ だから、

$$0 = \boldsymbol{a}^T\boldsymbol{e} = \boldsymbol{a}^T\boldsymbol{p} - \boldsymbol{a}^T\boldsymbol{b} \tag{18}$$

すなわち、

$$\boldsymbol{a}^T\boldsymbol{p} = \boldsymbol{a}^T\boldsymbol{b} \tag{19}$$

を得る。

k を求めるには、$\boldsymbol{a}$ と $\boldsymbol{p}$ の内積を取る。

$$\boldsymbol{a}^T\boldsymbol{p} = k\boldsymbol{a}^T\boldsymbol{a}$$

$$\therefore k = \frac{\boldsymbol{a}^T\boldsymbol{p}}{\boldsymbol{a}^T\boldsymbol{a}} = \frac{\boldsymbol{a}^T\boldsymbol{b}}{\boldsymbol{a}^T\boldsymbol{a}} \tag{20}$$

を得る。これは、

$$Q = (\boldsymbol{b} - k\boldsymbol{a})^T(\boldsymbol{b} - k\boldsymbol{a})$$

から、

$$\frac{dQ}{dk} = -2\boldsymbol{a}^T(\boldsymbol{b} - k\boldsymbol{a}) = 0$$

を解いて得られたものと一致する。

(20) 式) を (17) 式に代入すると、

$$\boldsymbol{p} = \frac{\boldsymbol{a}\boldsymbol{a}^T}{\boldsymbol{a}^T\boldsymbol{a}}\boldsymbol{b}$$

そこで、射影演算子 P は、

$$P = \frac{\boldsymbol{a}\boldsymbol{a}^T}{\boldsymbol{a}^T\boldsymbol{a}}$$

と書ける。これは、次の関係式を満たす。

$$P^2 = \frac{\boldsymbol{a}(\boldsymbol{a}^T\boldsymbol{a})\boldsymbol{a}^T}{(\boldsymbol{a}^T\boldsymbol{a})^2} = \frac{\boldsymbol{a}\boldsymbol{a}^T}{\boldsymbol{a}^T\boldsymbol{a}} = P$$

また、(18) 式より、

$$P\boldsymbol{e} = 0$$

を得る。

次に、$n \times m$ 行列の列ベクトル A と n 次元ベクトル $\boldsymbol{y}$ を定義する。

$$A = \begin{bmatrix} a_{11} & a_{12} & \dots & a_{1m} \\ a_{21} & a_{22} & \dots & a_{2m} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \dots & a_{nm} \end{bmatrix}, \qquad \boldsymbol{y} = \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{bmatrix}$$

ここで、ベクトル

$$\boldsymbol{a}_1 = \begin{bmatrix} a_{11} \\ a_{21} \\ \vdots \\ a_{n1} \end{bmatrix}, \quad \boldsymbol{a}_2 = \begin{bmatrix} a_{12} \\ a_{22} \\ \vdots \\ a_{n2} \end{bmatrix}, \quad \cdots \quad \boldsymbol{a}_m = \begin{bmatrix} a_{1m} \\ a_{2m} \\ \vdots \\ a_{nm} \end{bmatrix}$$

を定義すると、

$$A = \Big[\boldsymbol{a}_1,\ \boldsymbol{a}_2,\ \dots,\ \boldsymbol{a}_m \Big], \qquad A^T = \begin{bmatrix} \boldsymbol{a}_1^T \\ \boldsymbol{a}_2^T \\ \vdots \\ \boldsymbol{a}_m^T \end{bmatrix}$$

列ベクトル $\boldsymbol{a}_1,\ \boldsymbol{a}_2,\ \dots,\ \boldsymbol{a}_m$ が作るベクトル空間へのベクトル $\boldsymbol{y}$ の射影を $\boldsymbol{p}$ とする。この時、$\boldsymbol{y} = \boldsymbol{p} + \boldsymbol{\varepsilon}$ とすると、

$$\boldsymbol{\varepsilon} = \begin{bmatrix} \varepsilon_1 \\ \varepsilon_2 \\ \vdots \\ \varepsilon_n \end{bmatrix}$$

は、$\boldsymbol{a}_1,\ \boldsymbol{a}_2,\ \dots,\ \boldsymbol{a}_n$ と直交するので、

$$A^T \boldsymbol{\varepsilon} = 0 \quad \text{または、} \quad \boldsymbol{a}_i^T \boldsymbol{\varepsilon} = 0, \quad i = 1, 2, \cdots, m \tag{21}$$

すなわち、$\varepsilon_i\ (i = 1, 2, \cdots, n)$ はすべて独立ではなく、$\boldsymbol{\varepsilon}$ の自由度は、$n - m$ である。
(21) 式は、また、

$$A^T (\boldsymbol{p} - \boldsymbol{y}) = 0,$$

と書ける。すなわち、

$$A^T \boldsymbol{p} = A^T \boldsymbol{y} \tag{22}$$

$\boldsymbol{p}$ は、$\boldsymbol{a}_1,\ \boldsymbol{a}_2,\ \dots,\ \boldsymbol{a}_m$ が作る空間内のベクトルだから、

$$\boldsymbol{p} = A\boldsymbol{x} \tag{23}$$

となる

$$\boldsymbol{x} = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_m \end{bmatrix}$$

が存在する。これを、(22) 式に代入すると、

$$A^T A\boldsymbol{x} = A^T \boldsymbol{y} \tag{24}$$

すなわち、

$$\begin{bmatrix} \boldsymbol{a}_1^T\boldsymbol{a}_1 & \boldsymbol{a}_1^T\boldsymbol{a}_2 & \cdots & \boldsymbol{a}_1^T\boldsymbol{a}_m \\ \boldsymbol{a}_2^T\boldsymbol{a}_1 & \boldsymbol{a}_2^T\boldsymbol{a}_2 & \cdots & \boldsymbol{a}_2^T\boldsymbol{a}_m \\ \vdots & \vdots & \ddots & \vdots \\ \boldsymbol{a}_m^T\boldsymbol{a}_1 & \boldsymbol{a}_m^T\boldsymbol{a}_2 & \cdots & \boldsymbol{a}_m^T\boldsymbol{a}_m \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_m \end{bmatrix} = \begin{bmatrix} \boldsymbol{a}_1^T\boldsymbol{y} \\ \boldsymbol{a}_2^T\boldsymbol{y} \\ \vdots \\ \boldsymbol{a}_m^T\boldsymbol{y} \end{bmatrix} \tag{25}$$

を得る。

(24) 式より、

$$\boldsymbol{x} = (A^T A)^{-1} A^T \boldsymbol{y}$$

この両辺に左から A を掛けると、(23) 式より、

$$\boldsymbol{p} = A\boldsymbol{x} = A(A^T A)^{-1} A^T \boldsymbol{y}$$

を得る。そこで、射影演算子 P は、

$$P = A(A^T A)^{-1} A^T$$

となる。ゆえに、

$$P\boldsymbol{p} = \boldsymbol{p}, \qquad P\boldsymbol{\varepsilon} = 0$$

が成り立つ。

ここで、

$$\begin{aligned} Q &= ||A\boldsymbol{x} - \boldsymbol{y}||^2 \\ &= (\boldsymbol{a}_1 x_1 + \boldsymbol{a}_2 x_2 + \cdots + \boldsymbol{a}_m x_m - \boldsymbol{y})^T (\boldsymbol{a}_1 x_1 + \boldsymbol{a}_2 x_2 + \cdots + \boldsymbol{a}_m x_m - \boldsymbol{y}) \end{aligned}$$

とおいて、

$$\frac{\partial Q}{\partial x_i} = 2\boldsymbol{a}_i^T (\boldsymbol{a}_1 x_1 + \boldsymbol{a}_2 x_2 + \cdots + \boldsymbol{a}_m x_m - \boldsymbol{y}) = 0, \quad (i = 1, 2, \cdots, m)$$

とすると、(25) 式を得ることに注意しよう。

3 最小二乗法の幾何学的意味

方程式

$$A\boldsymbol{\theta} = \boldsymbol{y}$$

は、$\mathrm{rank}\ A < \mathrm{rank}[A|\boldsymbol{\theta}]$ のとき、解を持たない。このとき、行列 A の列ベクトルがつくるベクトル空間に $\boldsymbol{y}$ を射影してできるベクトルを $\hat{\boldsymbol{y}}$ とすると、方程式

$$A\boldsymbol{x} = \hat{\boldsymbol{y}}, \qquad \hat{\boldsymbol{y}} = P\boldsymbol{y} = A(A^TA)^{-1}A^T\boldsymbol{y} \tag{26}$$

を、今度は解くことができる。

もしも n コの対の観測データ $(x_i,\ y_i),\ (n = 1, 2, \cdots, n)$ 間に線形関係 $y_i = \hat{\theta}_1 + \hat{\theta}_2 x_i$ が成り立つならば、

$$A\hat{\boldsymbol{\theta}} = \boldsymbol{y}$$

と書くことができる。ここで、

$$A = \begin{bmatrix} 1 & x_1 \\ 1 & x_2 \\ \vdots & \vdots \\ 1 & x_n \end{bmatrix}, \qquad \hat{\boldsymbol{\theta}} = \begin{bmatrix} \hat{\theta}_1 \\ \hat{\theta}_2 \end{bmatrix}, \qquad \boldsymbol{y} = \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{bmatrix}$$

とおく。さらに、

$$\boldsymbol{a}_1 = \begin{bmatrix} 1 \\ 1 \\ \vdots \\ 1 \end{bmatrix}, \qquad \boldsymbol{a}_2 = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix}$$

とおくと、

$$A = \begin{bmatrix} \boldsymbol{a}_1,\ \boldsymbol{a}_2 \end{bmatrix}, \qquad A^T = \begin{bmatrix} \boldsymbol{a}_1^T \\ \boldsymbol{a}_2^T \end{bmatrix}$$

$\boldsymbol{y}$ に誤差があるとき、すなわち、

$$\boldsymbol{y} = A\hat{\boldsymbol{\theta}} + \hat{\boldsymbol{\varepsilon}}$$

の場合には、(26) 式が成り立つ。

$$A\hat{\boldsymbol{\theta}} = A(A^TA)^{-1}A^T\boldsymbol{y}$$

この両辺に A^T を左から掛けると、

$$A^T A\hat{\boldsymbol{\theta}} = A^T \boldsymbol{y} \tag{27}$$

が得られる。

$$A^T A = \begin{bmatrix} \boldsymbol{a}_1^T \boldsymbol{a}_1 & \boldsymbol{a}_1^T \boldsymbol{a}_2 \\ \boldsymbol{a}_2^T \boldsymbol{a}_1 & \boldsymbol{a}_2^T \boldsymbol{a}_2 \end{bmatrix} = \begin{bmatrix} n & \sum_i x_i \\ \sum_i x_i & \sum_i x_i^2 \end{bmatrix} \tag{28}$$

$$A^T \boldsymbol{y} = \begin{bmatrix} \boldsymbol{a}_1^T \boldsymbol{y} \\ \boldsymbol{a}_2^T \boldsymbol{y} \end{bmatrix} = \begin{bmatrix} \sum_i y_i \\ \sum_i x_i y_i \end{bmatrix}$$

なので、(27) 式は、

$$\begin{bmatrix} n & \sum_i x_i \\ \sum_i x_i & \sum_i x_i^2 \end{bmatrix} \begin{bmatrix} \hat{\theta}_1 \\ \hat{\theta}_2 \end{bmatrix} = \begin{bmatrix} \sum_i y_i \\ \sum_i x_i y_i \end{bmatrix} \tag{29}$$

を得る。これは、(11) 式と (13) 式の行列表示である。

(10) 式のベクトル表示である

$$\begin{aligned} Q(\theta_1, \theta_2) &= ||A\boldsymbol{\theta} - \boldsymbol{y}||^2 \\ &= (\theta_1 \boldsymbol{a}_1 + \theta_2 \boldsymbol{a}_2 - \boldsymbol{y})^T (\theta_1 \boldsymbol{a}_1 + \theta_2 \boldsymbol{a}_2 - \boldsymbol{y}) \end{aligned}$$

から

$$\frac{\partial Q(\hat{\theta}_1, \hat{\theta}_2)}{\partial \theta_1} = 0, \qquad \frac{\partial Q(\hat{\theta}_1, \hat{\theta}_2)}{\partial \theta_2} = 0$$

としたものに等しい。

(29) 式を解くと、

$$\begin{aligned} \begin{bmatrix} \hat{\theta}_1 \\ \hat{\theta}_2 \end{bmatrix} &= (A^T A)^{-1} \begin{bmatrix} \sum_i y_i \\ \sum_i x_i y_i \end{bmatrix} \\ &= \frac{1}{n \sum_i x_i^2 - (\sum_i x_i)^2} \begin{bmatrix} \sum_i x_i^2 & -\sum_i x_i \\ -\sum_i x_i & n \end{bmatrix} \begin{bmatrix} \sum_i y_i \\ \sum_i x_i y_i \end{bmatrix} \\ &= \frac{1}{\sum_i x_i^2 - n\bar{x}^2} \begin{bmatrix} \frac{1}{n} \sum_i x_i^2 & -\bar{x} \\ -\bar{x} & 1 \end{bmatrix} \begin{bmatrix} n\bar{y} \\ \sum_i x_i y_i \end{bmatrix} \\ &= \frac{1}{\sum_i x_i^2 - n\bar{x}^2} \begin{bmatrix} \bar{y} \sum_i x_i^2 - \bar{x} \sum_i x_i y_i \\ -n\bar{x}\bar{y} + \sum_i x_i y_i \end{bmatrix} \\ &= \frac{1}{\sum_i (x_i - \bar{x})^2} \begin{bmatrix} \bar{y} \sum_i (x_i - \bar{x})^2 - \bar{x} \sum_i (x_i - \bar{x})(y_i - \bar{y}) \\ \displaystyle\sum_i (x_i - \bar{x})(y_i - \bar{y}) \end{bmatrix} \end{aligned}$$

ゆえに、

$$\begin{bmatrix}\hat{\theta}_1\\ \hat{\theta}_2\end{bmatrix}=\begin{bmatrix}1-\dfrac{\bar{x}\sum_i(x_i-\bar{x})(y_i-\bar{y})}{\sum_i(x_i-\bar{x})^2}\\ \\ \dfrac{\sum_i(x_i-\bar{x})(y_i-\bar{y})}{\sum_i(x_i-\bar{x})^2}\end{bmatrix} \tag{30}$$

を得る。

また、

$$P\hat{\boldsymbol{\varepsilon}}=A^T\hat{\boldsymbol{\varepsilon}}=0 \tag{31}$$

より、

$$\boldsymbol{a}_1^T\hat{\boldsymbol{\varepsilon}}=\boldsymbol{a}_2^T\hat{\boldsymbol{\varepsilon}}=0$$

が成り立つ。すなわち、$\boldsymbol{\varepsilon}$ の自由度は、$n-2$ である。

4　最小 2 乗推定量の平均と分散

4.1　平均

$$\boldsymbol{\theta}=\begin{pmatrix}\theta_1\\ \theta_2\end{pmatrix} \tag{32}$$

とおくと、

$$\boldsymbol{y}=A\boldsymbol{\theta}+\boldsymbol{\varepsilon} \tag{33}$$

である。ところが、(4) 式より、

$$E(\boldsymbol{y})=A\boldsymbol{\theta}$$

(27) 式より、

$$A^TAE(\hat{\boldsymbol{\theta}})=A^TE(\boldsymbol{y}) \tag{34}$$

となるので、

$$E(\hat{\boldsymbol{\theta}})=\boldsymbol{\theta} \tag{35}$$

を得る。すなわち、

$$E(\hat{\theta}_1)=\theta_1,\qquad E(\hat{\theta}_2)=\theta_2 \tag{36}$$

を得る。

4.2 分散

(27) 式より、

$$\begin{bmatrix} \hat{\theta}_1^2 & \hat{\theta}_1\hat{\theta}_2 \\ \hat{\theta}_1\hat{\theta}_2 & \hat{\theta}_2^2 \end{bmatrix} = \hat{\boldsymbol{\theta}}\hat{\boldsymbol{\theta}}^T = (A^TA)^{-1}A^T\boldsymbol{y}\boldsymbol{y}^TA\left((A^TA)^{-1}\right)^T \tag{37}$$

ここで、

$$\boldsymbol{y}\boldsymbol{y}^T = \begin{bmatrix} y_1^2 & y_1y_2 & \cdots & y_1y_n \\ y_2y_1 & y_2^2 & \cdots & y_2y_n \\ \vdots & \vdots & \ddots & \vdots \\ y_ny_1 & y_ny_2 & \cdots & y_n^2 \end{bmatrix}$$

となるので、(4)、 (5)、 (6) 式を用いて、

$$\begin{aligned} E(\boldsymbol{y}\boldsymbol{y}^T) - E(\boldsymbol{y})E(\boldsymbol{y}^T) &= \begin{bmatrix} V(y_1) & Cov(y_1, y_2) & \cdots & Cov(y_1, y_n) \\ Cov(y_2, y_1) & V(y_2) & \cdots & Cov(y_2, y_n) \\ \vdots & \vdots & \ddots & \vdots \\ Cov(y_n, y_1) & Cov(y_n, y_2) & \cdots & V(y_n) \end{bmatrix} \\ &= \begin{bmatrix} \sigma^2 & 0 & \cdots & 0 \\ 0 & \sigma^2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \sigma^2 \end{bmatrix} \\ &= \sigma^2\boldsymbol{I} \end{aligned} \tag{38}$$

を得る。

(37) 式と (38) 式より、

$$\begin{aligned} \boldsymbol{\Sigma} &= \begin{bmatrix} V(\hat{\theta}_1) & Cov(\hat{\theta}_1, \hat{\theta}_2) \\ Cov(\hat{\theta}_1, \hat{\theta}_2) & V(\hat{\theta}_2) \end{bmatrix} \\ &= (A^TA)^{-1}A^T\sigma^2\boldsymbol{I}A\left((A^TA)^{-1}\right)^T \\ &= \sigma^2\left((A^TA)^{-1}\right)^T \\ &= \sigma^2(A^TA)^{-1} \end{aligned} \tag{39}$$

最後の変形では、(39) 式が対称行列であることを用いた。この式を (28) 式を用いて書き直すと、

$$\begin{bmatrix} V(\hat{\theta}_1) & Cov(\hat{\theta}_1, \hat{\theta}_2) \\ Cov(\hat{\theta}_1, \hat{\theta}_2) & V(\hat{\theta}_2) \end{bmatrix} = \frac{\sigma^2}{\sum_i(x_i - \bar{x})^2}\begin{pmatrix} \frac{1}{n}\sum_i x_i^2 & -\bar{x} \\ -\bar{x} & 1 \end{pmatrix} \tag{40}$$

を得る。すなわち、

$$\hat{\theta}_1\text{は、}\quad N\left(\theta_1, \frac{\sigma^2\sum_i x_i^2}{n\sum_i(x_i-\bar{x})^2}\right) \tag{41}$$

$$\hat{\theta}_2\text{は、}\quad N\left(\theta_2, \frac{\sigma^2}{\sum_i(x_i-\bar{x})^2}\right) \tag{42}$$

に従う。

再び、$\mathbf{\Sigma}=\sigma^2(A^TA)^{-1}$ を (28) 式を用いて書き直すと、

$$\begin{aligned}\mathbf{\Sigma} &= \begin{pmatrix}\sigma_{11}^2 & \sigma_{12}^2\\ \sigma_{12}^2 & \sigma_{22}^2\end{pmatrix}\\ &= \frac{\sigma^2}{|\boldsymbol{a}_1|^2|\boldsymbol{a}_2|^2-(\boldsymbol{a}_1^T\boldsymbol{a}_2)^2}\begin{pmatrix}|\boldsymbol{a}_2|^2 & -\boldsymbol{a}_1^T\boldsymbol{a}_2\\ \boldsymbol{a}_1^T\boldsymbol{a}_2 & |\boldsymbol{a}_1|^2\end{pmatrix}\\ &= \frac{\sigma^2}{|A^TA|}\begin{pmatrix}|\boldsymbol{a}_2|^2 & -\boldsymbol{a}_1^T\boldsymbol{a}_2\\ \boldsymbol{a}_1^T\boldsymbol{a}_2 & |\boldsymbol{a}_1|^2\end{pmatrix} \qquad (43)\end{aligned}$$

5 誤差分散の推定

残差平方和

$$\begin{aligned}S(\hat{\boldsymbol{\varepsilon}}) &= \sum_{i=1}^n \hat{\varepsilon}_i^2 \qquad (44)\\ &= \hat{\boldsymbol{\varepsilon}}^T\hat{\boldsymbol{\varepsilon}} \qquad (45)\\ &= \{(1-P)\boldsymbol{y}\}^T(1-P)\boldsymbol{y} \qquad (46)\end{aligned}$$

ここで、$E\left(S(\hat{\boldsymbol{\varepsilon}})\right)$ を求めたいが、そのためには、上の式を直接に使わないで、

$$\begin{aligned}\mathcal{M}_\varepsilon &= \hat{\boldsymbol{\varepsilon}}\hat{\boldsymbol{\varepsilon}}^T\\ &= \begin{bmatrix}\hat{\varepsilon}_1^2 & \hat{\varepsilon}_1\hat{\varepsilon}_2 & \cdots & \hat{\varepsilon}_1\hat{\varepsilon}_n\\ \hat{\varepsilon}_2\hat{\varepsilon}_1 & \hat{\varepsilon}_2^2 & \cdots & \hat{\varepsilon}_1\hat{\varepsilon}_n\\ \vdots & \vdots & \ddots & \vdots\\ \hat{\varepsilon}_n\hat{\varepsilon}_1 & \hat{\varepsilon}_n\hat{\varepsilon}_2 & \cdots & \hat{\varepsilon}_n^2\end{bmatrix} \qquad (47)\end{aligned}$$

を定義して、

$$S(\hat{\boldsymbol{\varepsilon}}) = \mathrm{Tr}\left[\mathcal{M}_\varepsilon\right] \tag{48}$$

を用いて計算する。

$$\begin{aligned}\mathrm{Tr}\left[\mathcal{M}_\varepsilon\right] &= \mathrm{Tr}\left[(1-P)\boldsymbol{y}\left\{(1-P)\boldsymbol{y}\right\}^T\right]\\ &= \mathrm{Tr}\left[(1-P)\boldsymbol{y}\boldsymbol{y}^T(1-P)\right]\\ &= \mathrm{Tr}\left[(1-P)\boldsymbol{y}\boldsymbol{y}^T\right] \qquad (49)\end{aligned}$$

ここで (38) 式を用いると、

$$\begin{aligned} S\left(E(\hat{\boldsymbol{\varepsilon}})\right) &= \mathrm{Tr}\left[E(\mathcal{M}_\varepsilon)\right] \\ &= \sigma^2 \mathrm{Tr}\left[(1-P)\boldsymbol{I}\right] \\ &= \sigma^2 \mathrm{Tr}\left[\boldsymbol{I}-P\right] \\ &= \sigma^2 \big(\mathrm{Tr}[\boldsymbol{I}] - \mathrm{Tr}[P]\big) \end{aligned} \tag{50}$$

ここに、$\boldsymbol{I}$ は $n \times n$ の単位行列だから、

$$\mathrm{Tr}[\boldsymbol{I}] = n$$

また、射影演算子 $P = A(A^TA)^{-1}A^T$ も $n \times n$ 行列で、

$$\mathrm{Tr}[P] = \mathrm{Tr}[A(A^TA)^{-1}A^T] = \mathrm{Tr}[A^TA(A^TA)^{-1}]$$

ところが (28) 式からも分かるように、A^TA も、その逆行列である $(A^TA)^{-1}$ も 2×2 行列であるので、

$$A^TA(A^TA)^{-1} = \boldsymbol{I}$$

は、2×2 の単位行列となる。したがって、

$$\mathrm{Tr}[P] = \mathrm{Tr}[\boldsymbol{I}] = 2$$

である。これらの結果を (50) 式に代入して、

$$S\left(E(\hat{\boldsymbol{\varepsilon}})\right) = (n-2)\sigma^2 \tag{51}$$

を得る。よって、

$$\hat{\sigma}^2 = \frac{S(\hat{\boldsymbol{\varepsilon}})}{n-2} = \frac{\sum_{i=1}^n \left(y_i - (\hat{\theta}_1 + \hat{\theta}_2 x_i)\right)^2}{n-2} \tag{52}$$

が不偏推定量を与える。

6 最尤推定量の分布

(7) 式の指数の肩にある $S(\boldsymbol{\varepsilon}) = (\boldsymbol{y} - A\boldsymbol{\theta})^T(\boldsymbol{y} - A\boldsymbol{\theta})$ を $\hat{\boldsymbol{\theta}}$ の周りで展開すると、

$$\begin{aligned} S(\boldsymbol{\varepsilon}) &= (\boldsymbol{y} - A\boldsymbol{\theta})^T(\boldsymbol{y} - A\boldsymbol{\theta}) && (53) \\ &= \left(\boldsymbol{y} - A\hat{\boldsymbol{\theta}} + A(\hat{\boldsymbol{\theta}} - \boldsymbol{\theta})\right)^T \left(\boldsymbol{y} - A\hat{\boldsymbol{\theta}} + A(\hat{\boldsymbol{\theta}} - \boldsymbol{\theta})\right) \\ &= \left(\hat{\boldsymbol{\varepsilon}} + A(\hat{\boldsymbol{\theta}} - \boldsymbol{\theta})\right)^T \left(\hat{\boldsymbol{\varepsilon}} + A(\hat{\boldsymbol{\theta}} - \boldsymbol{\theta})\right) \\ &= \hat{\boldsymbol{\varepsilon}}^T\hat{\boldsymbol{\varepsilon}} + 2(\hat{\boldsymbol{\theta}} - \boldsymbol{\theta})^T A^T \hat{\boldsymbol{\varepsilon}} + (\hat{\boldsymbol{\theta}} - \boldsymbol{\theta})^T A^T A (\hat{\boldsymbol{\theta}} - \boldsymbol{\theta}) && (54) \end{aligned}$$

ところが、(31) 式より、$A^T\hat{\boldsymbol{\varepsilon}} = 0$ だから、

$$S(\boldsymbol{\varepsilon}) = S(\hat{\boldsymbol{\varepsilon}}) + (\hat{\boldsymbol{\theta}} - \boldsymbol{\theta})^T A^T A(\hat{\boldsymbol{\theta}} - \boldsymbol{\theta}) \tag{55}$$

である。ところが、(39) 式より、$A^T A = \sigma^2 \boldsymbol{\Sigma}^{-1}$ だから、

$$S(\boldsymbol{\varepsilon}) = S(\hat{\boldsymbol{\varepsilon}}) + \sigma^2(\hat{\boldsymbol{\theta}} - \boldsymbol{\theta})^T \boldsymbol{\Sigma}^{-1}(\hat{\boldsymbol{\theta}} - \boldsymbol{\theta})$$

を得る。これを (7) 式に代入して、尤度推定量 $\hat{\boldsymbol{\theta}}$ の分布関数とみなすと、

$$\mathrm{P}(\hat{\boldsymbol{\theta}}|\boldsymbol{\theta}, \boldsymbol{\Sigma}) = C \exp\left\{-\frac{1}{2}(\hat{\boldsymbol{\theta}} - \boldsymbol{\theta})^T \boldsymbol{\Sigma}^{-1}(\hat{\boldsymbol{\theta}} - \boldsymbol{\theta})\right\} \tag{56}$$

を得る。

C の計算

C を計算するのに、(56) 式の代わりに、$\boldsymbol{\Sigma}^{-1}$ を $A^T A/\sigma^2$ で置き換えた

$$\mathrm{P}(\hat{\boldsymbol{\theta}}|\boldsymbol{\theta}, \boldsymbol{\Sigma}) = C \exp\left\{-\frac{1}{2\sigma^2}(\hat{\boldsymbol{\theta}} - \boldsymbol{\theta})^T A^T A(\hat{\boldsymbol{\theta}} - \boldsymbol{\theta})\right\} \tag{57}$$

を用いる。

$\boldsymbol{\Theta} = \hat{\boldsymbol{\theta}} - \boldsymbol{\theta}$ とおくと、

$$A(\hat{\boldsymbol{\theta}} - \boldsymbol{\theta}) = A\boldsymbol{\Theta} = \Theta_1 \boldsymbol{a}_1 + \Theta_2 \boldsymbol{a}_2$$

さらに、

$$\boldsymbol{e}_1 = \left(\frac{\boldsymbol{a}_1}{|\boldsymbol{a}_1|} + \frac{\boldsymbol{a}_2}{|\boldsymbol{a}_2|}\right) \tag{58}$$

$$\boldsymbol{e}_2 = \left(\frac{\boldsymbol{a}_1}{|\boldsymbol{a}_1|} - \frac{\boldsymbol{a}_2}{|\boldsymbol{a}_2|}\right) \tag{59}$$

とおくと、

$$\boldsymbol{e}_1^T \boldsymbol{e}_1 = 2\left(\frac{|\boldsymbol{a}_1||\boldsymbol{a}_2| + \boldsymbol{a}_1^T \boldsymbol{a}_2}{|\boldsymbol{a}_1||\boldsymbol{a}_2|}\right) \tag{60}$$

$$\boldsymbol{e}_2^T \boldsymbol{e}_2 = 2\left(\frac{|\boldsymbol{a}_1||\boldsymbol{a}_2| - \boldsymbol{a}_1^T \boldsymbol{a}_2}{|\boldsymbol{a}_1||\boldsymbol{a}_2|}\right) \tag{61}$$

$$\boldsymbol{e}_1^T \boldsymbol{e}_2 = 0 \tag{62}$$

となることに注意しよう。さらに、

$$\boldsymbol{i}_1 = \frac{\boldsymbol{e}_1}{|\boldsymbol{e}_1|} \tag{63}$$

$$\boldsymbol{i}_2 = \frac{\boldsymbol{e}_2}{|\boldsymbol{e}_2|} \tag{64}$$

(65)

と置くと、

$$A\boldsymbol{\Theta} = \frac{|\boldsymbol{e}_1|}{2}\big(\Theta_1|\boldsymbol{a}_1| + \Theta_2|\boldsymbol{a}_2|\big)\boldsymbol{i}_1 + \frac{|\boldsymbol{e}_2|}{2}\big(\Theta_1|\boldsymbol{a}_1| - \Theta_2|\boldsymbol{a}_2|\big)\boldsymbol{i}_2$$

ここで、

$$\begin{aligned} u_1\sigma &= \frac{|\boldsymbol{e}_1|}{2}(\Theta_1|\boldsymbol{a}_1| + \Theta_2|\boldsymbol{a}_2|) \\ u_2\sigma &= \frac{|\boldsymbol{e}_2|}{2}(\Theta_1|\boldsymbol{a}_1| - \Theta_2|\boldsymbol{a}_2|) \end{aligned}$$

と置くと、

$$A\boldsymbol{\Theta} = u_1\sigma\boldsymbol{i}_1 + u_2\sigma\boldsymbol{i}_2 \tag{66}$$

だから、

$$\frac{1}{2\sigma^2}\boldsymbol{\Theta}^T A^T A\boldsymbol{\Theta} = \frac{u_1^2}{2} + \frac{u_2^2}{2} \tag{67}$$

また、

$$\begin{aligned} \frac{\partial(u_1, u_2)}{\partial(\Theta_1, \Theta_2)} &= \begin{vmatrix} \dfrac{|\boldsymbol{e}_1||\boldsymbol{a}_1|}{2} & \dfrac{|\boldsymbol{e}_1||\boldsymbol{a}_2|}{2} \\ \\ \dfrac{|\boldsymbol{e}_2||\boldsymbol{a}_1|}{2} & -\dfrac{|\boldsymbol{e}_2||\boldsymbol{a}_2|}{2} \end{vmatrix} & (68) \\ &= -\frac{1}{2\sigma^2}|\boldsymbol{e}_1||\boldsymbol{e}_2||\boldsymbol{a}_1||\boldsymbol{a}_2| & (69) \\ &= -\sqrt{\frac{|\boldsymbol{a}_1|^2|\boldsymbol{a}_2|^2 - (\boldsymbol{a}_1^T\boldsymbol{a}_2)^2}{\sigma^4}} & (70) \\ &= -\sqrt{\frac{|A^TA|}{\sigma^4}} & (71) \end{aligned}$$

となるので、結局、

$$\frac{\partial(u_1, u_2)}{\partial(\Theta_1, \Theta_2)} = -\frac{1}{\sqrt{|\boldsymbol{\Sigma}|}} \tag{72}$$

を得る。ここで、$\boldsymbol{\Sigma} = \dfrac{\sigma^2}{A^TA}$ は、2×2 行列である。従って、$|\boldsymbol{\Sigma}| = \dfrac{\sigma^4}{|A^TA|}$ となることを用いた。

ゆえに、

$$\begin{aligned} \int \mathrm{P}(\hat{\boldsymbol{\theta}}|\boldsymbol{\theta}, \boldsymbol{\Sigma})\, d\Theta_1 d\Theta_2 &= \left|\frac{\partial(\Theta_1, \Theta_2)}{\partial(u_1, u_2)}\right| C \int \exp\left\{-\frac{u_1^2}{2} - \frac{u_2^2}{2}\right\} du_1\, du_2 \\ &= \sqrt{|\boldsymbol{\Sigma}|}C \int \exp\left\{-\frac{u_1^2}{2} - \frac{u_2^2}{2}\right\} du_1\, du_2 \\ &= 2\pi\sqrt{|\boldsymbol{\Sigma}|}\, C \\ &= 1 \end{aligned}$$

$$\therefore\ C = \frac{1}{2\pi\sqrt{|\mathbf{\Sigma}|}}$$

以上より、

$$\begin{aligned} \mathrm{P}(\hat{\boldsymbol{\theta}}|\boldsymbol{\theta},\mathbf{\Sigma}) &= \frac{1}{2\pi}\sqrt{\frac{|A^TA|}{\sigma^4}}\exp\left\{-\frac{1}{2\sigma^2}(\hat{\boldsymbol{\theta}}-\boldsymbol{\theta})^TA^TA(\hat{\boldsymbol{\theta}}-\boldsymbol{\theta})\right\} && (73) \\ &= \frac{1}{2\pi\sqrt{|\mathbf{\Sigma}|}}\exp\left\{-\frac{1}{2}(\hat{\boldsymbol{\theta}}-\boldsymbol{\theta})^T\mathbf{\Sigma}^{-1}(\hat{\boldsymbol{\theta}}-\boldsymbol{\theta})\right\} && (74) \end{aligned}$$

を得る。

7　周辺分布と共分散

(73) 式より、

$$\mathrm{P}(\hat{\mathbf{\Theta}}|\mathbf{\Sigma}) = \frac{1}{2\pi}\sqrt{\frac{|A^TA|}{\sigma^4}}\exp\left\{-\frac{1}{2\sigma^2}\hat{\mathbf{\Theta}}^TA^TA\hat{\mathbf{\Theta}}\right\} \qquad (75)$$

ここで、

$$\begin{aligned} \hat{\mathbf{\Theta}}^TA^TA\hat{\mathbf{\Theta}} &= \boldsymbol{a}_1^T\boldsymbol{a}_1\hat{\Theta}_1^2 + 2\boldsymbol{a}_1^T\boldsymbol{a}_2\hat{\Theta}_1\hat{\Theta}_2 + \boldsymbol{a}_2^T\boldsymbol{a}_2\hat{\Theta}_2^2 && (76) \\ &= \left(|\boldsymbol{a}_1|^2 - \frac{(\boldsymbol{a}_1^T\boldsymbol{a}_2)^2}{|\boldsymbol{a}_2|^2}\right)\hat{\Theta}_1^2 + |\boldsymbol{a}_2|^2\left(\hat{\Theta}_2 + \frac{\boldsymbol{a}_1^T\boldsymbol{a}_2}{|\boldsymbol{a}_2|^2}\hat{\Theta}_1\right)^2 && (77) \\ &= |\boldsymbol{a}_1|^2\left(\hat{\Theta}_1 + \frac{\boldsymbol{a}_1^T\boldsymbol{a}_2}{|\boldsymbol{a}_1|^2}\hat{\Theta}_2\right)^2 + \left(|\boldsymbol{a}_2|^2 - \frac{(\boldsymbol{a}_1^T\boldsymbol{a}_2)^2}{|\boldsymbol{a}_1|^2}\right)\hat{\Theta}_2^2 && (78) \end{aligned}$$

だから、

$$\begin{aligned} \mathrm{P}(\hat{\Theta}_1|\mathbf{\Sigma}) &= \int \mathrm{P}(\hat{\mathbf{\Theta}}|\mathbf{\Sigma})\,d\Theta_2 \\ &= \sqrt{\frac{|A^TA|}{2\pi\sigma^2|\boldsymbol{a}_2|^2}}\exp\left\{-\frac{1}{2\sigma^2}\left(|\boldsymbol{a}_1|^2 - \frac{(\boldsymbol{a}_1^T\boldsymbol{a}_2)^2}{|\boldsymbol{a}_2|^2}\right)\hat{\Theta}_1^2\right\} \\ &= \sqrt{\frac{|A^TA|}{2\pi\sigma^2|\boldsymbol{a}_2|^2}}\exp\left\{-\frac{|A^TA|}{2\sigma^2|\boldsymbol{a}_2|^2}\hat{\Theta}_1^2\right\} && (79) \end{aligned}$$

$$\begin{aligned} \mathrm{P}(\hat{\Theta}_2|\mathbf{\Sigma}) &= \int \mathrm{P}(\hat{\mathbf{\Theta}}|\mathbf{\Sigma})\,d\Theta_1 \\ &= \sqrt{\frac{|A^TA|}{2\pi\sigma^2|\boldsymbol{a}_1|^2}}\exp\left\{-\frac{1}{2\sigma^2}\left(|\boldsymbol{a}_2|^2 - \frac{(\boldsymbol{a}_1^T\boldsymbol{a}_2)^2}{|\boldsymbol{a}_1|^2}\right)\hat{\Theta}_1^2\right\} \\ &= \sqrt{\frac{|A^TA|}{2\pi\sigma^2|\boldsymbol{a}_1|^2}}\exp\left\{-\frac{|A^TA|}{2\sigma^2|\boldsymbol{a}_1|^2}\hat{\Theta}_1^2\right\} && (80) \end{aligned}$$

を得る。ここで、(43) 式を用いると、

$$
\begin{aligned}
\mathrm{P}(\hat{\Theta}_1|\boldsymbol{\Sigma}) &= \frac{1}{\sqrt{2\pi\sigma_{11}^2}} \exp\left(-\frac{\hat{\Theta}_1^2}{2\sigma_{11}^2}\right) \\
&= \sqrt{\frac{n\sum_i (x_i-\bar{x})^2}{2\pi\sigma^2\sum_i x_i^2}} \exp\left(-\frac{n\sum_i (x_i-\bar{x})^2}{2\sigma^2\sum_i x_i^2}\hat{\Theta}_1^2\right) \qquad (81) \\
\mathrm{P}(\hat{\Theta}_2|\boldsymbol{\Sigma}) &= \frac{1}{\sqrt{2\pi\sigma_{22}^2}} \exp\left(-\frac{\hat{\Theta}_2^2}{2\sigma_{22}^2}\right) \\
&= \sqrt{\frac{\sum_i (x_i-\bar{x})^2}{2\pi\sigma^2}} \exp\left(-\frac{\sum_i (x_i-\bar{x})^2}{2\sigma^2}\hat{\Theta}_2^2\right) \qquad (82) \\
& \qquad (83)
\end{aligned}
$$

を得る。

また、

$$
\begin{aligned}
\sigma_{12}^2 &= \frac{1}{2\pi}\sqrt{\frac{|A^TA|}{\sigma^4}} \int \hat{\Theta}_1\hat{\Theta}_2\ \exp\left\{-\frac{1}{2\sigma^2}\hat{\boldsymbol{\Theta}}^T A^T A \hat{\boldsymbol{\Theta}}\right\} d\hat{\Theta}_1 d\hat{\Theta}_2 \\
&= \frac{1}{2\pi}\sqrt{\frac{|A^TA|}{\sigma^4}} \int d\hat{\Theta}_1 d\hat{\Theta}_2\, \hat{\Theta}_1\hat{\Theta}_2 \\
&\quad \exp\left\{-\frac{1}{2\sigma^2}\left[|\boldsymbol{a}_1|^2\left(\hat{\Theta}_1 + \frac{\boldsymbol{a}_1^T\boldsymbol{a}_2}{|\boldsymbol{a}_1|^2}\hat{\Theta}_2\right)^2 + \frac{|A^TA|}{|\boldsymbol{a}_1|^2}\hat{\Theta}_2^2\right]\right\} \\
&= \frac{1}{2\pi}\sqrt{\frac{|A^TA|}{\sigma^4}}\left(-\frac{\boldsymbol{a}_1^T\boldsymbol{a}_2}{|\boldsymbol{a}_1|^2}\right)\sqrt{\frac{2\pi\sigma^2}{|\boldsymbol{a}_1|^2}} \int d\hat{\Theta}_2\, \hat{\Theta}_2^2 \exp\left\{-\frac{|A^TA|}{2\sigma^2|\boldsymbol{a}_1|^2}\hat{\Theta}_2^2\right\} \\
&= \frac{1}{2\pi}\sqrt{\frac{|A^TA|}{\sigma^4}}\left(-\frac{\boldsymbol{a}_1^T\boldsymbol{a}_2}{|\boldsymbol{a}_1|^2}\right)\sqrt{\frac{2\pi\sigma^2}{|\boldsymbol{a}_1|^2}}\,\frac{\sqrt{\pi}}{2}\,\frac{(2\sigma^2|\boldsymbol{a}_1|)^{3/2}}{|A^TA|^{3/2}} \\
&= -\frac{\sigma^2\boldsymbol{a}_1^T\boldsymbol{a}_2}{|A^TA|} \qquad (84)
\end{aligned}
$$

これは、(43) 式の結果と一致している。

8 t 分布

8.1 一変量 t 分布

(81) 式より、$\hat{\Theta}_1$ は、$N\left(0, \sigma_{11}^2\right)$ の正規分布に従う。そこで、変数変換

$$
\begin{aligned}
\mathrm{P}(u_1') &= \int \delta\left(u_1' - \frac{\sqrt{|A^TA|}}{\sigma|\boldsymbol{a}_2|}\hat{\Theta}_1\right) \mathrm{P}(\hat{\Theta}_1|\boldsymbol{\Sigma})\, d\hat{\Theta}_1 \\
&= \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{u_1'^2}{2}\right) \qquad (85)
\end{aligned}
$$

を行うと、

$$u'_1 = \sqrt{\frac{n\sum_i(x_i-\bar{x})^2}{\sigma^2\sum_i x_i^2}}\,(\hat{\theta}_1-\theta_1) \tag{86}$$

は、$N(0,1)$ の正規分布に従う。
一方、

$$v = \frac{S(\hat{\boldsymbol{\varepsilon}})}{\sigma^2}$$

は、自由度 $n-2$ の χ^2 分布に従う。
(52) 式より、不偏分散は、

$$\hat{\sigma}^2 = \frac{S(\hat{\boldsymbol{\varepsilon}})}{n-2} = \frac{\left(\boldsymbol{y}-A\hat{\boldsymbol{\theta}}\right)^2}{n-2}$$

だから、

$$v = \frac{(n-2)\hat{\sigma}^2}{\sigma^2} \tag{87}$$

と書ける。
(86) 式の σ^2 は未知なので、(52) 式の $\hat{\sigma}_2$ で置き換えた、

$$t'_1 = \sqrt{\frac{n\sum_i(x_i-\bar{x})^2}{\hat{\sigma}^2\sum_i x_i^2}}\,(\hat{\theta}_1-\theta_1) = \sqrt{\frac{n-2}{v}}u'_1 \tag{88}$$

は、自由度 $n-2$ の t 分布に従う。
同様にして、

$$t'_2 = \sqrt{\frac{\sum_i(x_i-\bar{x})^2}{\hat{\sigma}^2}}\,(\hat{\theta}_2-\theta_2) \tag{89}$$

もまた自由度 $n-2$ の t 分布に従う。

8.2 二変量 t 分布

(67) 式より、

$$A\hat{\boldsymbol{\Theta}} = A\hat{\boldsymbol{\theta}} - A\boldsymbol{\theta}$$

は $N(\mathbf{0},\sigma^2\mathbf{1})$ に従う。
また、(66) 式より、

$$\boldsymbol{u} = \frac{A\hat{\boldsymbol{\theta}}-A\boldsymbol{\theta}}{\sigma} = u_1\boldsymbol{i}_1 + u_2\boldsymbol{i}_2 \tag{90}$$

は、$N(\mathbf{0},\mathbf{1})$ に従う。
そこで、

$$\begin{aligned} \boldsymbol{t} &= \frac{A\hat{\boldsymbol{\theta}} - A\boldsymbol{\theta}}{\hat{\sigma}} & (91) \\ &= \frac{(A\hat{\boldsymbol{\theta}} - A\boldsymbol{\theta})}{\sigma} \frac{\sigma}{\hat{\sigma}} & (92) \\ &= \sqrt{\frac{n-2}{v}}\, \boldsymbol{u} & (93) \end{aligned}$$

と置くと、これは自由度 $n-2$ の 2 変量 t 分布に従う。
その分布は、

$$\begin{aligned} f(\boldsymbol{t}) &= \int du_1\, du_2\, dv\ \delta\left(\boldsymbol{t} - \sqrt{\frac{n-2}{v}}\, \boldsymbol{u}\right) \times \\ &\qquad \frac{1}{2\pi} \exp\left\{-\frac{1}{2}\boldsymbol{u}^T\boldsymbol{u}\right\} \frac{1}{2^{(n-2)/2}\Gamma\left(\frac{n-2}{2}\right)} v^{(n-4)/2} e^{-v/2} & (94) \\ &= \frac{1}{\pi(n-2)2^{n/2}\,\Gamma\left(\frac{n-2}{2}\right)} \int dv\, v^{(n-2)/2} \exp\left\{-\left(1 + \frac{\boldsymbol{t}^T\boldsymbol{t}}{n-2}\right)\frac{v}{2}\right\} & (95) \end{aligned}$$

となる。ここで、

$$x = \left(1 + \frac{\boldsymbol{t}^T\boldsymbol{t}}{n-2}\right)\frac{v}{2}$$

とおいて、(95) 式の積分を行うと、

$$\begin{aligned} &\int dv\ v^{(n-2)/2} \exp\left\{-\left(1 + \frac{\boldsymbol{t}^T\boldsymbol{t}}{n-2}\right)\frac{v}{2}\right\} \\ &= 2^{n/2}\left(1 + \frac{\boldsymbol{t}^T\boldsymbol{t}}{n-2}\right)^{-n/2} \int_0^\infty x^{(n-2)/2} e^{-x}\, dx \\ &= 2^{n/2}\Gamma\left(\frac{n}{2}\right)\left(1 + \frac{\boldsymbol{t}^T\boldsymbol{t}}{n-2}\right)^{-n/2} \end{aligned}$$

となるので、

$$f(\boldsymbol{t}) = \frac{\Gamma\left(\frac{n}{2}\right)}{(n-2)\pi\Gamma\left(\frac{n-2}{2}\right)} \left(1 + \frac{\boldsymbol{t}^T\boldsymbol{t}}{(n-2)}\right)^{-n/2} \qquad (96)$$

を得る。また、ガンマ関数の性質

$$\Gamma\left(\frac{n}{2}\right) = \left(\frac{n-2}{2}\right)\Gamma\left(\frac{n-2}{2}\right)$$

より、(96) 式は、

$$f(\boldsymbol{t}) = \frac{1}{2\pi}\left(1 + \frac{\boldsymbol{t}^T\boldsymbol{t}}{(n-2)}\right)^{-n/2} \tag{97}$$

となる。

(90) 式と同様に (91) 式は、

$$\boldsymbol{t} = t_1\boldsymbol{i}_1 + t_2\boldsymbol{i}_2$$

と書けるので、

$$\boldsymbol{t}^T\boldsymbol{t} = t_1^2 + t_2^2$$

となる。そこで、(96) 式も、また、

$$f(t_1, t_2) = \frac{1}{2\pi}\left(1 + \frac{t_1^2 + t_2^2}{n-2}\right)^{-n/2}$$

と表すことができる。

(91) 式より、

$$\boldsymbol{t}^T\boldsymbol{t} = \frac{1}{\hat{\sigma}^2}(\hat{\boldsymbol{\theta}} - \boldsymbol{\theta})^T A^T A(\hat{\boldsymbol{\theta}} - \boldsymbol{\theta})$$

ここで、

$$\hat{\boldsymbol{\Sigma}} = \frac{\hat{\sigma}^2}{A^T A}$$

と置くと、

$$\boldsymbol{t}^T\boldsymbol{t} = (\hat{\boldsymbol{\theta}} - \boldsymbol{\theta})^T\hat{\boldsymbol{\Sigma}}^{-1}(\hat{\boldsymbol{\theta}} - \boldsymbol{\theta})$$

また、(72) 式と類似の関係式、

$$\frac{\partial(t_1, t_2)}{\partial(\Theta_1, \Theta_2)} = -\frac{1}{\sqrt{|\hat{\boldsymbol{\Sigma}}|}} \tag{98}$$

が成り立つので、(97) 式を $\hat{\boldsymbol{\theta}}$ を用いて書き直すと、

$$f_\theta(\boldsymbol{\theta}) = \frac{1}{\sqrt{|\hat{\boldsymbol{\Sigma}}|}}f\left(\frac{A\hat{\boldsymbol{\theta}} - A\boldsymbol{\theta}}{\hat{\sigma}}\right) \tag{99}$$

より、

$$f_\theta(\boldsymbol{\theta}) = \frac{1}{2\pi\sqrt{|\hat{\boldsymbol{\Sigma}}|}}\left(1 + \frac{(\hat{\boldsymbol{\theta}} - \boldsymbol{\theta})^T\hat{\boldsymbol{\Sigma}}^{-1}(\hat{\boldsymbol{\theta}} - \boldsymbol{\theta})}{(n-2)}\right)^{-n/2} \tag{100}$$

と表すことができる。