第 11 章

最小二乗法による線形回帰分析

1.4 節で 2 次元データ (x_i, y_i) $(i = 1, \cdots, n)$ で y_i を x_i で予測する単回帰直線 $y = a + bx + \varepsilon$ について論述し, 1.5 節では, 母親の身長を x_i, 女子学生の身長を y_i として, 母親の身長から女子学生の身長を予測する式を導いた. このとき, この直線の当てはまり度の尺度として決定係数があった. またその節の後半では母親のみから女子学生の身長を予測することよりも, 母親 (x_1) と父親 (x_2) から女子学生の身長 (y) を予測する重回帰直線 $y = a + b_1x_1 + b_2x_2 + \varepsilon$ について論述した. この章ではもっと変数を多くしたモデルを含む回帰直線の一般論について詳細に述べ, 変数をどこまで採るべきかの議論も行う.

11.1　線形回帰分析

1.4 節の線型単回帰モデル (1.2) を一般化した線型回帰モデル

$$Y = \theta_1 + \theta_2 x_2 + \cdots + \theta_k x_k + \varepsilon$$

を考える. 特に $k \geq 3$ のときのモデルを**線型重回帰モデル**という. Y を**被説明変数**と呼ばれる確率変数で, $k-1$ 個の要因に対応する確率変数でない変数値 $x_2, \cdots, x_k$ を**説明変数**と呼ぶ. ε は誤差と呼ばれる確率変数で平均と分散を

$$E(\varepsilon) = 0, \qquad V(\varepsilon) = \sigma^2$$

とする. このとき,

$$E(Y) = \theta_1 + \theta_2 x_2 + \cdots + \theta_k x_k, \qquad V(Y) = \sigma^2$$

である．また, Y のことを**従属変数**または**応答変数**といい, $x_2, \cdots, x_k$ を**共変量**または**共変数**ということがある．

$n > k$ とし, n 組のデータ $(Y_i, x_{i2}, \cdots, x_{ik})$ $(i = 1, \cdots, n)$ に対して, 線型回帰モデルは

$$Y_i = \theta_1 + \theta_2 x_{i2} + \cdots + \theta_k x_{ik} + \varepsilon_i \ \ (i = 1, \cdots, n) \tag{11.1}$$

であり, $\varepsilon_1, \cdots, \varepsilon_n$ は互いに独立で, 平均 0 分散 σ^2 の同一の分布に従う．便宜上ベクトルは縦ベクトルで表し, 観測ベクトル $\boldsymbol{Y} \equiv {}^t(Y_1, \cdots, Y_n)$, 係数ベクトル $\boldsymbol{\theta} \equiv {}^t(\theta_1, \cdots, \theta_k)$, 誤差ベクトル $\boldsymbol{\varepsilon} \equiv {}^t(\varepsilon_1, \cdots, \varepsilon_n)$,

$$n \times k \text{ 行列 } \boldsymbol{X} \equiv (\mathbf{1}_n, \boldsymbol{x}_2, \cdots, \boldsymbol{x}_k) \equiv \begin{pmatrix} 1 & x_{12} & \cdots & x_{1k} \\ 1 & x_{22} & \cdots & x_{2k} \\ \vdots & \vdots & \ddots & \vdots \\ 1 & x_{n2} & \cdots & x_{nk} \end{pmatrix} \tag{11.2}$$

とおくと, (11.1) はベクトルと行列で

$$\boldsymbol{Y} = \boldsymbol{X}\boldsymbol{\theta} + \boldsymbol{\varepsilon} \tag{11.3}$$

表現できる．(11.3) で $\boldsymbol{X}$ を**説明変数行列**, $\boldsymbol{\theta}$ を**回帰係数ベクトル**という．1.4 節と同じように,

誤差の二乗和:

$$\Delta^2 \equiv \|\varepsilon\|^2 = \|\boldsymbol{Y} - \boldsymbol{X}\boldsymbol{\theta}\|^2 = \sum_{i=1}^{n} (Y_i - \theta_1 - x_{i2}\theta_2 - \cdots - x_{ik}\theta_k)^2 \tag{11.4}$$

を最小にするような $\boldsymbol{\theta}$ の値を $\boldsymbol{\theta}$ の推定量とする方法を**最小二乗法**という．最小値を与える解を解くために Δ^2 を θ_j $(j = 1, \cdots, k)$ で偏微分すれば,

$$
\begin{aligned}
\frac{\partial}{\partial \boldsymbol{\theta}}\Delta^2 &= \begin{pmatrix} \dfrac{\partial \Delta^2}{\partial \theta_1} \\ \dfrac{\partial \Delta^2}{\partial \theta_2} \\ \vdots \\ \dfrac{\partial \Delta^2}{\partial \theta_k} \end{pmatrix} = \begin{pmatrix} 2\sum\limits_{i=1}^{n}(-1)(Y_i - \theta_1 - x_{i2}\theta_2 - \cdots - x_{ik}\theta_k) \\ 2\sum\limits_{i=1}^{n}(-x_{i2})(Y_i - \theta_1 - x_{i2}\theta_2 - \cdots - x_{ik}\theta_k) \\ \vdots \\ 2\sum\limits_{i=1}^{n}(-x_{ik})(Y_i - \theta_1 - x_{i2}\theta_2 - \cdots - x_{ik}\theta_k) \end{pmatrix} \\
&= -2({}^t\boldsymbol{X}\boldsymbol{Y} - {}^t\boldsymbol{X}\boldsymbol{X}\boldsymbol{\theta})
\end{aligned}
$$

を得るから, 最小二乗法の正規方程式は

$$
{}^t\boldsymbol{X}\boldsymbol{X}\boldsymbol{\theta} = {}^t\boldsymbol{X}\boldsymbol{Y}
$$

となる．この章では今後 ${}^t\boldsymbol{X}\boldsymbol{X}$ が正則行列の場合を考える．このとき正規方程式より得られる最小二乗解は,

$$
\hat{\boldsymbol{\theta}} \equiv ({}^t\boldsymbol{X}\boldsymbol{X})^{-1}\,{}^t\boldsymbol{X}\boldsymbol{Y}
$$

である．そのとき,

$$
\hat{\boldsymbol{Y}} \equiv {}^t(\hat{Y}_1, \cdots, \hat{Y}_n) \equiv \boldsymbol{X}\hat{\boldsymbol{\theta}} = \boldsymbol{X}({}^t\boldsymbol{X}\boldsymbol{X})^{-1}\,{}^t\boldsymbol{X}\boldsymbol{Y}
$$

を**回帰ベクトル**といい, $\hat{\boldsymbol{Y}}$ によって $\boldsymbol{Y}$ を近似する．

$$
\boldsymbol{Y} - \hat{\boldsymbol{Y}} = \boldsymbol{Y} - \boldsymbol{X}\hat{\boldsymbol{\theta}} = \{\boldsymbol{I}_n - \boldsymbol{X}({}^t\boldsymbol{X}\boldsymbol{X})^{-1}\,{}^t\boldsymbol{X}\}\boldsymbol{Y}
$$

を**残差ベクトル**といい,

$$
RSS \equiv \|\boldsymbol{Y} - \hat{\boldsymbol{Y}}\|^2 = \|\boldsymbol{Y} - \boldsymbol{X}\hat{\boldsymbol{\theta}}\|^2 = \sum_{i=1}^{n}(Y_i - \hat{\theta}_1 - x_{i2}\hat{\theta}_2 - \cdots - x_{ik}\hat{\theta}_k)^2
$$

を**残差平方和** (Residual Sum of Squares) といい, 1.4 節の単回帰の場合と同様に $\bar{Y} \equiv \dfrac{1}{n}\sum\limits_{i=1}^{n} Y_i$ とおき

$$
CD \equiv 1 - \frac{RSS}{\sum\limits_{i=1}^{n}(Y_i - \bar{Y})^2}
$$

を**寄与率**または**決定係数** (Coefficient of Determination) という．Y_i と $\hat{Y}_i$ $(i =$

$1, \cdots, n)$ の標本相関係数 ((7.1) で定義) を $\tilde{\rho}_{Y\hat{Y}}$ とすれば, 寄与率と標本相関係数の間のつぎの興味ある結果を得る.

定理 11.1 寄与率 CD は標本相関係数 $\tilde{\rho}_{Y\hat{Y}}$ の二乗に等しい. すなわち,

$$0 \leq CD = \tilde{\rho}^2_{Y\hat{Y}} = \frac{\{\sum_{i=1}^{n}(Y_i - \bar{Y})(\hat{Y}_i - \bar{\hat{Y}})\}^2}{\sum_{i=1}^{n}(Y_i - \bar{Y})^2 \sum_{i=1}^{n}(\hat{Y}_i - \bar{\hat{Y}})^2} \leq 1$$

が成り立つ, ただし $\bar{\hat{Y}}$ は $\hat{Y}_1, \cdots, \hat{Y}_n$ の標本平均とする.

証明 ${}^t\boldsymbol{X}\{\boldsymbol{I}_n - \boldsymbol{X}({}^t\boldsymbol{X}\boldsymbol{X})^{-1}{}^t\boldsymbol{X}\} = \boldsymbol{O}$ より,

$$\sum_{i=1}^{n}(Y_i - \hat{Y}_i) = {}^t\boldsymbol{1}_n(\boldsymbol{Y} - \hat{\boldsymbol{Y}}) = {}^t\boldsymbol{1}_n\{\boldsymbol{I}_n - \boldsymbol{X}({}^t\boldsymbol{X}\boldsymbol{X})^{-1}{}^t\boldsymbol{X}\}\boldsymbol{Y} = 0$$

かつ

$$\begin{aligned}\sum_{i=1}^{n}(Y_i - \hat{Y}_i)\hat{Y}_i &= {}^t(\boldsymbol{Y} - \hat{\boldsymbol{Y}})\hat{\boldsymbol{Y}} \\ &= {}^t\boldsymbol{Y}\{\boldsymbol{I}_n - \boldsymbol{X}({}^t\boldsymbol{X}\boldsymbol{X})^{-1}{}^t\boldsymbol{X}\}\boldsymbol{X}({}^t\boldsymbol{X}\boldsymbol{X})^{-1}{}^t\boldsymbol{X}\boldsymbol{Y} = 0\end{aligned}$$

を得る. この 2 つの関係式より,

$$\bar{Y} = \bar{\hat{Y}} \qquad \text{かつ} \qquad \sum_{i=1}^{n}(Y_i - \hat{Y}_i)(\hat{Y}_i - \bar{\hat{Y}}) = 0. \tag{11.5}$$

(11.5) より

$$\begin{aligned} CD &= \frac{\sum_{i=1}^{n}(Y_i - \hat{Y}_i + \hat{Y}_i - \bar{\hat{Y}})^2 - \sum_{i=1}^{n}(Y_i - \hat{Y}_i)^2}{\sum_{i=1}^{n}(Y_i - \bar{Y})^2} \\ &= \frac{\sum_{i=1}^{n}(\hat{Y}_i - \bar{\hat{Y}})^2}{\sum_{i=1}^{n}(Y_i - \bar{Y})^2}\end{aligned}$$

$$
\begin{aligned}
&= \frac{\{\sum_{i=1}^{n}(\hat{Y}_i - \bar{\hat{Y}})^2\}^2}{\sum_{i=1}^{n}(Y_i - \bar{Y})^2 \sum_{i=1}^{n}(\hat{Y}_i - \bar{\hat{Y}})^2} \\
&= \frac{\{\sum_{i=1}^{n}(Y_i - \hat{Y}_i + \hat{Y}_i - \bar{\hat{Y}})(\hat{Y}_i - \bar{\hat{Y}})\}^2}{\sum_{i=1}^{n}(Y_i - \bar{Y})^2 \sum_{i=1}^{n}(\hat{Y}_i - \bar{\hat{Y}})^2} \\
&= \tilde{\rho}^2_{Y\hat{Y}}.
\end{aligned}
$$

この値が 0 以上 1 以下であることは標本相関係数の定義より解る．これらにより結論を得た．　□

$$E(\hat{\boldsymbol{\theta}}) = \boldsymbol{\theta},\ \ E(\hat{\boldsymbol{Y}}) = \boldsymbol{X}\boldsymbol{\theta},\ \ E(\boldsymbol{Y} - \hat{\boldsymbol{Y}}) = \boldsymbol{0}$$

より, $\hat{\boldsymbol{\theta}}$, $\hat{\boldsymbol{Y}}$, $\boldsymbol{Y} - \hat{\boldsymbol{Y}}$ はそれぞれ $\boldsymbol{\theta}$, $\boldsymbol{X}\boldsymbol{\theta}$, $\boldsymbol{0}$ の不偏推定量である．

$\boldsymbol{P} \equiv \boldsymbol{X}({}^t\boldsymbol{X}\boldsymbol{X})^{-1\,t}\boldsymbol{X}$ と $\boldsymbol{I}_n - \boldsymbol{P}$ はともに対称な巾等行列でそれぞれ階数が $\mathrm{tr}(\boldsymbol{P}) = k$ と $\mathrm{tr}(\boldsymbol{I}_n - \boldsymbol{P}) = n - k$ である．また積が $\boldsymbol{O}$ である, すなわち

$$\boldsymbol{P} \cdot (\boldsymbol{I}_n - \boldsymbol{P}) = (\boldsymbol{I}_n - \boldsymbol{P}) \cdot \boldsymbol{P} = \boldsymbol{O}.$$

$\boldsymbol{\lambda} = {}^t(\lambda_1, \cdots, \lambda_k)$ の線形関数

$${}^t\boldsymbol{\lambda}\boldsymbol{\theta} = \lambda_1\theta_1 + \cdots + \lambda_k\theta_k$$

の最小二乗法による推定量

$${}^t\boldsymbol{\lambda}\hat{\boldsymbol{\theta}} = {}^t\boldsymbol{\lambda}({}^t\boldsymbol{X}\boldsymbol{X})^{-1\,t}\boldsymbol{X}\boldsymbol{Y}$$

を得る．また,

定理 11.2

$$\hat{\sigma}^2 \equiv \frac{||\boldsymbol{Y} - \boldsymbol{X}\hat{\boldsymbol{\theta}}||^2}{n-k} = \frac{RSS}{n-k} \tag{11.6}$$

は共通分散 σ^2 の不偏推定量である．

証明　$E(\boldsymbol{Y} - \boldsymbol{X}\hat{\boldsymbol{\theta}}) = \boldsymbol{0}$ より一般性を失うことなく $\boldsymbol{\theta} = \boldsymbol{0}$ と仮定する．定理

A.2 より, $\boldsymbol{I}_n - \boldsymbol{P} = {}^t\boldsymbol{U}\mathrm{diag}(\overbrace{1,\cdots,1}^{\text{n-k 個}},0,\cdots,0)\boldsymbol{U}$ なる直交行列 $\boldsymbol{U}$ が存在する. $\boldsymbol{Z} \equiv {}^t(Z_1,\cdots,Z_n) \equiv \boldsymbol{UY}$ とおくと $E(\boldsymbol{Z}) = \boldsymbol{0}$, $V(\boldsymbol{Z}) = \sigma^2\boldsymbol{I}_n$. ここで,

$$E(\hat{\sigma}^2) = \frac{1}{n-k}E\{{}^t\boldsymbol{Y}(\boldsymbol{I}_n - \boldsymbol{P})\boldsymbol{Y}\} = \frac{1}{n-k}E\{\sum_{i=1}^{n-k} Z_i^2\} = \sigma^2 \quad \square$$

$\varepsilon_1,\cdots,\varepsilon_n$ は互いに独立で同一の $N(0,\sigma^2)$ に従うとき, (11.1) のモデルを**正規線形モデル**といい, 記号 $GMN(\boldsymbol{Y};\boldsymbol{X\theta},\sigma^2\boldsymbol{I}_n)$ で表す. このとき, $\boldsymbol{Y} \sim N_n(\boldsymbol{X\theta},\sigma^2\boldsymbol{I}_n)$ である.

定理 11.3 $GMN(\boldsymbol{Y};\boldsymbol{X\theta},\sigma^2\boldsymbol{I}_n)$ のとき,

$$T = \frac{({}^t\boldsymbol{\lambda}\hat{\boldsymbol{\theta}} - {}^t\boldsymbol{\lambda\theta})}{\hat{\sigma}\sqrt{{}^t\boldsymbol{\lambda}({}^t\boldsymbol{XX})^{-1}\boldsymbol{\lambda}}} \sim t_{n-k}.$$

証明 一般性を失うことなく, $\boldsymbol{\theta} = \boldsymbol{0}$, $\sigma^2 = 1$ とする. このとき, 系 A.2 より $\hat{\sigma}T \sim N(0,1)$ は容易に示される. 定理 11.2 の証明で定義した $\boldsymbol{Z} \equiv {}^t(Z_1,\cdots,Z_n)$ は $N_n(\boldsymbol{0},\boldsymbol{I}_n)$ に従う. これにより $(n-k)\hat{\sigma}^2 = \sum_{i=1}^{n-k} Z_i^2 \sim \chi^2_{n-k}$. また ${}^t\boldsymbol{X}\cdot(\boldsymbol{I}_n - \boldsymbol{P}) = \boldsymbol{O}$ より $\hat{\sigma}T$ と $(n-k)\hat{\sigma}^2$ の独立性は示せる. $\square$

したがって, ${}^t\boldsymbol{\lambda\theta}$ の $100(1-\alpha)$ パーセント信頼区間は

$$({}^t\boldsymbol{\lambda}\hat{\boldsymbol{\theta}} - t_{n-k,\frac{\alpha}{2}}\hat{\sigma}\sqrt{{}^t\boldsymbol{\lambda}({}^t\boldsymbol{XX})^{-1}\boldsymbol{\lambda}},\ {}^t\boldsymbol{\lambda}\hat{\boldsymbol{\theta}} + t_{n-k,\frac{\alpha}{2}}\hat{\sigma}\sqrt{{}^t\boldsymbol{\lambda}({}^t\boldsymbol{XX})^{-1}\boldsymbol{\lambda}}) \quad (11.7)$$

となる.

次に

$$\boldsymbol{A} \equiv \begin{pmatrix} a_{11} & \cdots & a_{1k} \\ \vdots & & \vdots \\ a_{r1} & \cdots & a_{rk} \end{pmatrix}, \qquad \mathrm{rank}(\boldsymbol{A}) = r \ \ (r < k)$$

とし, 回帰係数ベクトルの線形制約 $\boldsymbol{\Theta}_0 \equiv \{\boldsymbol{\theta} \in R^k : \boldsymbol{A\theta} = \boldsymbol{0}\}$ を考える. このとき, $\boldsymbol{\theta}$ が線型制約を満たすかどうかの検定仮説は, 以下のようになる.

$$\text{帰無仮説 } H_0 : \boldsymbol{A\theta} = \boldsymbol{0} \text{ と 対立仮説 } H_1 : \boldsymbol{A\theta} \neq \boldsymbol{0}. \qquad (11.8)$$

線形制約 $\boldsymbol{\Theta}_0 = \left\{\boldsymbol{\theta} \in R^k : \boldsymbol{A\theta} = \mathbf{0}\right\}$ の下での制限最小二乗法を解くために, ラグランジュ乗数法を使う.

$$g(\boldsymbol{\theta}, \boldsymbol{\nu}) \equiv \Delta^2(\boldsymbol{\theta}) + 2{}^t\boldsymbol{\nu}\boldsymbol{A\theta}, \quad \boldsymbol{\nu} = {}^t(\nu_1, \cdots, \nu_r)$$

とおき, それぞれのパラメータで偏微分して $\mathbf{0}$ とおくと,

$$\begin{aligned} \frac{\partial}{\partial \boldsymbol{\theta}} g(\boldsymbol{\theta}, \boldsymbol{\nu}) &= -2({}^t\boldsymbol{XY} - {}^t\boldsymbol{XX\theta}) + 2{}^t\boldsymbol{A\nu} = \mathbf{0}, \\ \frac{\partial}{\partial \boldsymbol{\nu}} g(\boldsymbol{\theta}, \boldsymbol{\nu}) &= 2\boldsymbol{A\theta} = \mathbf{0} \end{aligned}$$

となり, 制限正規方程式は

$$\begin{cases} {}^t\boldsymbol{XX\theta} + {}^t\boldsymbol{A\nu} = {}^t\boldsymbol{XY} \\ \boldsymbol{A\theta} = \mathbf{0} \end{cases}$$

$$\begin{pmatrix} {}^t\boldsymbol{XX} & {}^t\boldsymbol{A} \\ \boldsymbol{A} & \boldsymbol{O} \end{pmatrix} \begin{pmatrix} \boldsymbol{\theta} \\ \boldsymbol{\nu} \end{pmatrix} = \begin{pmatrix} {}^t\boldsymbol{XY} \\ \mathbf{0} \end{pmatrix}$$

となる. $W \equiv ({}^t\boldsymbol{XX})^{-1}$ とおき逆行列を求めると,

$$\begin{pmatrix} {}^t\boldsymbol{XX} & {}^t\boldsymbol{A} \\ \boldsymbol{A} & \boldsymbol{O} \end{pmatrix}^{-1} = \begin{pmatrix} \boldsymbol{W} - \boldsymbol{W}{}^t\boldsymbol{A}(\boldsymbol{AW}{}^t\boldsymbol{A})^{-1}\boldsymbol{AW} & \boldsymbol{W}{}^t\boldsymbol{A}(\boldsymbol{AW}{}^t\boldsymbol{A})^{-1} \\ (\boldsymbol{AW}{}^t\boldsymbol{A})^{-1}\boldsymbol{AW} & -(\boldsymbol{AW}{}^t\boldsymbol{A})^{-1} \end{pmatrix}$$

となる. ここでこの逆行列を

$$\begin{pmatrix} \boldsymbol{C}_1 & {}^t\boldsymbol{C}_2 \\ \boldsymbol{C}_2 & \boldsymbol{C}_3 \end{pmatrix}$$

とおく. 制限正規方程式を解くと,

$$\begin{pmatrix} \hat{\boldsymbol{\theta}}^* \\ \hat{\boldsymbol{\nu}}^* \end{pmatrix} = \begin{pmatrix} \boldsymbol{C}_1{}^t\boldsymbol{XY} \\ \boldsymbol{C}_2{}^t\boldsymbol{XY} \end{pmatrix}$$

となり, ここで, 線形制約 $\boldsymbol{\Theta}_0 = \left\{\boldsymbol{\theta} \in R^k : \boldsymbol{A\theta} = \mathbf{0}\right\}$ の下での $\boldsymbol{\theta}$ の制限最小

二乗推定量は

$$\hat{\boldsymbol{\theta}}^* = \boldsymbol{C}_1{}^t\boldsymbol{X}\boldsymbol{Y} \tag{11.9}$$

であり, 制限回帰ベクトルは

$$\boldsymbol{X}\hat{\boldsymbol{\theta}}^* = \boldsymbol{X}\boldsymbol{C}_1{}^t\boldsymbol{X}\boldsymbol{Y}.$$

線形制約の下での残差平方和は,

$$RSS^* = ||\boldsymbol{Y} - \boldsymbol{X}\hat{\boldsymbol{\theta}}^*||^2.$$

正規線形モデル $GMN(\boldsymbol{Y}; \boldsymbol{X}\boldsymbol{\theta}, \sigma^2\boldsymbol{I}_n)$ において, 回帰係数が線形制約を満たすかどうかの 帰無仮説 $H_0 : \boldsymbol{A}\boldsymbol{\theta} = \boldsymbol{0}$ と 対立仮説 $H_1 : \boldsymbol{A}\boldsymbol{\theta} \neq \boldsymbol{0}$ を検定するための統計量は

$$T_S = \frac{(RSS^* - RSS)/r}{RSS/(n-k)} = \frac{(||\boldsymbol{Y} - \boldsymbol{X}\hat{\boldsymbol{\theta}}^*||^2 - ||\boldsymbol{Y} - \boldsymbol{X}\hat{\boldsymbol{\theta}}||^2)/r}{||\boldsymbol{Y} - \boldsymbol{X}\hat{\boldsymbol{\theta}}||^2/(n-k)} \tag{11.10}$$

である.

$\boldsymbol{Q} \equiv \boldsymbol{X}\boldsymbol{C}_1{}^t\boldsymbol{X}$ とおくと, 次が成り立つ.

補題 11.4 3 つの等式

$$\boldsymbol{Q} = {}^t\boldsymbol{U}\mathrm{diag}(\overbrace{1, \cdots, 1}^{\text{(k-r) 個}}, \overbrace{0, \cdots, 0}^{\text{(n-k+r) 個}})\boldsymbol{U} \tag{11.11}$$

$$\boldsymbol{P} - \boldsymbol{Q} = {}^t\boldsymbol{U}\mathrm{diag}(\overbrace{0, \cdots, 0}^{\text{(k-r) 個}}, \overbrace{1, \cdots, 1}^{\text{r 個}}, \overbrace{0, \cdots, 0}^{\text{(n-k) 個}})\boldsymbol{U} \tag{11.12}$$

$$\boldsymbol{I}_n - \boldsymbol{P} = {}^t\boldsymbol{U}\mathrm{diag}(\overbrace{0, \cdots, 0}^{\text{k 個}}, \overbrace{1, \cdots, 1}^{\text{(n-k) 個}})\boldsymbol{U} \tag{11.13}$$

を同時に満たすある直交行列 $\boldsymbol{U}$ が存在する.

証明 $\boldsymbol{Q}$, $\boldsymbol{P} - \boldsymbol{Q}$, $\boldsymbol{I}_n - \boldsymbol{P}$ はいずれも対称な巾等行列より, 階数はそれぞれ $\mathrm{tr}(\boldsymbol{Q}) = k - r$, $\mathrm{tr}(\boldsymbol{P} - \boldsymbol{Q}) = r$, $\mathrm{tr}(\boldsymbol{I}_n - \boldsymbol{P}) = n - k$ である. 定理 A.2(1) より, $\boldsymbol{Q}$ の固有値は 1 か 0 で 1 に対する $\boldsymbol{Q}$ の固有ベクトルとして互いに直交し長さが 1 のものを $k - r$ 個選ぶことができ, それらを $\boldsymbol{u}_1, \cdots, \boldsymbol{u}_{k-r}$ とする. 同様に $\boldsymbol{P} - \boldsymbol{Q}$ の固有値 1 対する $\boldsymbol{P} - \boldsymbol{Q}$ の固有ベクトルとして互いに直

交し長さが 1 のものを r 個選ぶことができ, それらを $\boldsymbol{u}_1^*, \cdots, \boldsymbol{u}_r^*$, $\boldsymbol{I}_n - \boldsymbol{P}$ の固有値 1 対する $\boldsymbol{I}_n - \boldsymbol{P}$ の固有ベクトルとして互いに直交し長さが 1 のものを $n-k$ 個選ぶことができ, それらを $\boldsymbol{u}_1^\sharp, \cdots, \boldsymbol{u}_{n-k}^\sharp$ とする.

また $\boldsymbol{Q}$, $\boldsymbol{P}-\boldsymbol{Q}$, $\boldsymbol{I}_n - \boldsymbol{P}$ の間の積は $\boldsymbol{O}$ である, すなわち

$$\boldsymbol{Q}(\boldsymbol{P}-\boldsymbol{Q}) = \boldsymbol{Q}(\boldsymbol{I}_n - \boldsymbol{P}) = \boldsymbol{O}. \tag{11.14}$$

(11.15) より ${}^t\boldsymbol{u}_i\boldsymbol{u}_j^* = {}^t\boldsymbol{u}_i\boldsymbol{Q}(\boldsymbol{P}-\boldsymbol{Q})\boldsymbol{u}_j^* = 0$. 同様にして, $\boldsymbol{u}_1, \cdots, \boldsymbol{u}_{k-r}, \boldsymbol{u}_1^*, \cdots, \boldsymbol{u}_r^*$, $\boldsymbol{u}_1^\sharp, \cdots, \boldsymbol{u}_{n-k}^\sharp$ は互いに直交することが解る. さらに, $\boldsymbol{Q}\boldsymbol{u}_j^* = \boldsymbol{Q}(\boldsymbol{P}-\boldsymbol{Q})\boldsymbol{u}_j^* = \boldsymbol{0}$ より, $\boldsymbol{u}_1^*, \cdots, \boldsymbol{u}_r^*$ は固有値 0 に対する $\boldsymbol{Q}$ の固有ベクトルである. 同様に, $\boldsymbol{u}_1^\sharp, \cdots, \boldsymbol{u}_{n-k}^\sharp$ は固有値 0 に対する $\boldsymbol{Q}$ の固有ベクトル, $\boldsymbol{u}_1, \cdots, \boldsymbol{u}_{k-r}, \boldsymbol{u}_1^\sharp, \cdots, \boldsymbol{u}_{n-k}^\sharp$ は固有値 0 に対する $\boldsymbol{P}-\boldsymbol{Q}$ の固有ベクトル, $\boldsymbol{u}_1, \cdots, \boldsymbol{u}_{k-r}, \boldsymbol{u}_1^*, \cdots, \boldsymbol{u}_r^*$ は固有値 0 に対する $\boldsymbol{I}_n - \boldsymbol{P}$ の固有ベクトルである. 以上により ${}^t\boldsymbol{U} = (\boldsymbol{u}_1, \cdots, \boldsymbol{u}_{k-r}, \boldsymbol{u}_1^*, \cdots,$
$\boldsymbol{u}_r^*, \boldsymbol{u}_1^\sharp, \cdots, \boldsymbol{u}_{n-k}^\sharp)$ とおくことにより,

$$\begin{aligned}
\boldsymbol{Q}{}^t\boldsymbol{U} &= (\boldsymbol{Q}\boldsymbol{u}_1, \cdots, \boldsymbol{Q}\boldsymbol{u}_{k-r}, \boldsymbol{Q}\boldsymbol{u}_1^*, \cdots, \boldsymbol{Q}\boldsymbol{u}_r^*, \boldsymbol{Q}\boldsymbol{u}_1^\sharp, \cdots, \boldsymbol{Q}\boldsymbol{u}_{n-k}^\sharp) \\
&= (\boldsymbol{u}_1, \cdots, \boldsymbol{u}_{k-r}, \overbrace{0, \cdots, 0}^{\text{(n-k+r) 個}}) \\
&= {}^t\boldsymbol{U}\mathrm{diag}(\overbrace{1, \cdots, 1}^{\text{(k-r) 個}}, \overbrace{0, \cdots, 0}^{\text{(n-k+r) 個}})
\end{aligned}$$

が示せ, (11.12) が解る. (11.13), (11.14) も同様に示せる. □

定理 11.4　正規線形モデル $GMN(\boldsymbol{Y}; \boldsymbol{X}\boldsymbol{\theta}, \sigma^2\boldsymbol{I}_n)$ において, 帰無仮説 H_0 の下で $T_S \sim F_{n-k}^r$.

証明　H_0 の下で $E\{\boldsymbol{Y} - \boldsymbol{X}\hat{\boldsymbol{\theta}}^*\} = E\{\boldsymbol{Y} - \boldsymbol{X}\hat{\boldsymbol{\theta}}\} = \boldsymbol{0}$ より一般性を失うことなく $\boldsymbol{\theta} = \boldsymbol{0}$ と仮定してよく, (11.11) 式の分母と分子両方を σ^2 で割ることにより一般性を失うことなく $\sigma^2 = 1$ と仮定してよい. 補題 11.4 の直交行列 $\boldsymbol{U}$ を使って, $\boldsymbol{Y}$ の直交変換を考えると,

$$\boldsymbol{Z} \equiv {}^t(Z_1, \cdots, Z_n) \equiv \boldsymbol{U}\boldsymbol{Y} \sim N_n(\boldsymbol{0}, \boldsymbol{I}_n).$$

ここで 補題 11.4 を使うことにより,

$$\begin{aligned}||\boldsymbol{Y}-\boldsymbol{X}\hat{\boldsymbol{\theta}}^*||^2-||\boldsymbol{Y}-\boldsymbol{X}\hat{\boldsymbol{\theta}}||^2 &= {}^t\boldsymbol{Y}(\boldsymbol{I}_n-\boldsymbol{Q})\boldsymbol{Y}-{}^t\boldsymbol{Y}(\boldsymbol{I}_n-\boldsymbol{P})\boldsymbol{Y} \\ &= {}^t\boldsymbol{Y}(\boldsymbol{P}-\boldsymbol{Q})\boldsymbol{Y} \\ &= Z_{k-r+1}^2+\cdots+Z_k^2 \qquad (11.15) \\ &\sim \chi_r^2\end{aligned}$$

であり,

$$\begin{aligned}||\boldsymbol{Y}-\boldsymbol{X}\hat{\boldsymbol{\theta}}||^2 &= {}^t\boldsymbol{Y}(\boldsymbol{I}_n-\boldsymbol{P})\boldsymbol{Y} \\ &= Z_{k+1}^2+\cdots+Z_n^2 \qquad (11.16) \\ &\sim \chi_{n-k}^2\end{aligned}$$

(11.16), (11.17) より (11.11) 式の分母と分子の独立性が示せ結論を得る. □

この定理により, $F^r_{n-k,\alpha}$ を自由度 $(r, n-k)$ のエフ分布の上側 α 点とすると, $P_0(T_S > F^r_{n-k,\alpha}) = \alpha$. ゆえに有意水準 α の検定方式は,

$$\phi(\boldsymbol{X}) = \begin{cases} 1 & (T_S > F^r_{n-k,\alpha}\text{のとき}) \\ 0 & (T_S < F^r_{n-k,\alpha}\text{のとき}) \end{cases}$$

と表現される.

例 11.1 正規線形モデル $GMN(\boldsymbol{Y}; \boldsymbol{X\theta}, \sigma^2\boldsymbol{I}_n)$ において, 帰無仮説 H_i : $\theta_i = 0$ と 対立仮説 H_i' : $\theta_i \neq 0$ を検定したいならば,

$$\boldsymbol{A} \equiv (\overbrace{0,\cdots,0}^{i-1}, 1, \overbrace{0,\cdots,0}^{k-i})$$

とおき, (11.11) の T_S を計算する. 有意水準 α の検定方式は,

$$\phi(\boldsymbol{X}) = \begin{cases} 1 & (T_S > F^1_{n-k,\alpha}\text{のとき}) \\ 0 & (T_S < F^1_{n-k,\alpha}\text{のとき}) \end{cases}$$

となる. □

11.2　回帰係数の次元の決定

k 次元縦ベクトル $\boldsymbol{\theta}_p$ を $\boldsymbol{\theta}_p = {}^t(\theta_1, \cdots, \theta_p, 0, \cdots, 0)$ $(p = 1, \cdots, k)$ とおき, 正規線形モデル $GMN(\boldsymbol{Y}; \boldsymbol{X\theta}, \sigma^2 \boldsymbol{I}_n)$ において母数 $\boldsymbol{\theta}$ の次元 p $(1 \leq p \leq k)$ に関する

$$\begin{cases} \text{帰無仮説 } H_p : \boldsymbol{\theta} = \boldsymbol{\theta}_p \ \ (i = p+1, \cdots, k \text{ について} \theta_i = 0) \\ \text{対立仮説 } H_k : \boldsymbol{\theta} = \boldsymbol{\theta}_k \ \ (p+1 \leq i \leq k \text{ なるある } i \text{ について } \theta_i \neq 0), \end{cases}$$

の検定について考える.

$$\boldsymbol{A} \equiv (\mathbf{O}_{(k-p)\times p}, \boldsymbol{I}_{k-p})$$

とおくことにより, 帰無仮説と対立仮説は (11.9) で表現できる.

$$\boldsymbol{X}_p \equiv (\mathbf{1}_n, \boldsymbol{x}_2, \cdots, \boldsymbol{x}_p) \equiv \begin{pmatrix} 1 & x_{12} & \cdots & x_{1p} \\ \vdots & \vdots & & \vdots \\ 1 & x_{n2} & \cdots & x_{np} \end{pmatrix},$$

とおけば, 帰無仮説 H_p の下で (11.3) は

$$\boldsymbol{Y} = \boldsymbol{X}_p {}^t(\theta_1, \cdots, \theta_p) + \varepsilon \qquad (11.17)$$

となる. モデル (11.18) の下での残差平方和を

$$RSS_p \equiv ||\boldsymbol{Y} - \boldsymbol{X}_p({}^t\boldsymbol{X}_p\boldsymbol{X}_p)^{-1t}\boldsymbol{X}_p\boldsymbol{Y}||^2$$

とおくと, $\boldsymbol{X}_k = \boldsymbol{X}$ より, モデル (11.3) の下での残差平方和 RSS は便宜上 RSS_k で表される. 回帰係数が線形制約を満たすかどうかの仮説 (帰無仮説 H_p と対立仮説 H_k) を検定するための統計量は (11.11) より

$$T_{Sp} = \frac{(RSS_p - RSS_k)/(k-p)}{RSS_k/(n-k)}$$

となり, 定理 11.4 より帰無仮説 H_p の下で $T_{Sp} \sim F_{n-k}^{k-p}$ である. このことから, 統計量 T_{Sp} によって仮説 H_p と H_k のどちらが成り立っているかを調べることができる. しかし真の母数の次元 r $(1 \leq r \leq k)$ が未知であって, データ

によって決定されなければならない場合, 有意確率が異なり, 有意水準の決め方に大きく依存する. そこで通常は, $\hat{\sigma}^2$ を (11.7) で定義したものとし, マローズ (Mallows) の Cp 統計量と呼ばれる次の統計量

$$C(p) \equiv \frac{RSS_p}{\hat{\sigma}^2} + 2p - n$$

を最小にする p によって, 母数の次元を決定する方法がよく使われる.

11.3 ノンパラメトリック法とセミパラメトリック法

(11.1) の線形回帰モデルの回帰係数ベクトルの順位推定量と M 推定量がそれぞれプリ • セン (Puri and Sen (1985)) とフーバー (Huber (1981)) に論述され, (11.9) の線形仮説の検定法と線形仮説の下での回帰係数ベクトルの推定法のノンパラメトリック論とセミパラメトリック論がサレー • 白石 (Saleh and Shiraishi (1989) (1993)) に載せられている. これらの手法は計算機アルゴリズムが大変で明解なアルゴリズムが構築されるとよいかと思われる.

演習問題 11

11.1. 11.1 節の設定で, (11.2) のように $\boldsymbol{X} = (\boldsymbol{1}_n, \boldsymbol{x}_2, \cdots, \boldsymbol{x}_k)$ とおき, $\boldsymbol{P} \equiv \boldsymbol{X}({}^t\boldsymbol{X}\boldsymbol{X})^{-1}{}^t\boldsymbol{X}$ とおくと, $\{\lambda_1 \boldsymbol{1}_n + \sum_{i=2}^{k} \lambda_i \boldsymbol{x}_i : \lambda_1, \cdots, \lambda_k \in R\} = \{\boldsymbol{X} \cdot \boldsymbol{\lambda} : \boldsymbol{\lambda} \in R^k\}$ 上の全ての元 $\boldsymbol{z}$ に対して $\boldsymbol{P}\boldsymbol{z} = \boldsymbol{z}$ であることを示せ.

11.2. $\boldsymbol{X} = (\boldsymbol{1}_n, {}^t(x_1, \cdots, x_n))$, $\boldsymbol{\theta} = {}^t(\theta_1, \theta_2)$, $GMN(\boldsymbol{Y}; \boldsymbol{X}\boldsymbol{\theta}, \sigma^2 \boldsymbol{I}_n)$ において, $\theta_1 + \theta_2 x$ の $100(1-\alpha)$ パーセント信頼区間は, $\bar{x}. \equiv \dfrac{1}{n} \sum_{i=1}^{n} x_i$ として

$$(\hat{\theta}_1 + \hat{\theta}_2 x - t_{n-2,\frac{\alpha}{2}} \hat{\sigma} h_n,\ \hat{\theta}_1 + \hat{\theta}_2 x + t_{n-2,\frac{\alpha}{2}} \hat{\sigma} h_n)$$

で与えられることを示せ,

ただし, $h_n \equiv \sqrt{\dfrac{1}{n} + \dfrac{(x - \bar{x}.)^2}{\sum_{i=1}^{n} (x_i - \bar{x}.)^2}}$, $\hat{\sigma}^2 = \dfrac{\sum_{i=1}^{n} (Y_i - \hat{\theta}_1 - \hat{\theta}_2 x_i)^2}{n-2}$,

$$\hat{\theta}_1 = \bar{Y}. - \hat{\theta}_2 \cdot \bar{x}., \ \hat{\theta}_2 = \frac{\sum_{i=1}^{n}(x_i - \bar{x}.)(Y_i - \bar{Y})}{\sum_{i=1}^{n}(x_i - \bar{x}.)^2} \text{ とする.}$$

11.3. $GMN(\boldsymbol{Y}; \boldsymbol{X\theta}, \sigma^2 \boldsymbol{I}_n)$ において, 対立仮説 H_1 の下で T_S は自由度 $(r, n-k)$ 非心度 $\delta^2 \equiv \dfrac{1}{\sigma^2}{}^t\boldsymbol{\theta}^t\boldsymbol{X}(\boldsymbol{P}-\boldsymbol{Q})\boldsymbol{X\theta}$ の非心 F 分布に従うことを示せ, ただし $H_1, T_S, \boldsymbol{P}, \boldsymbol{Q}$ は 11.1 節で定義したものとする.

11.4. 11.1 節のモデル設定で, $\boldsymbol{c}_0$ を r 次元定数列ベクトルとする. また $\boldsymbol{\theta}_0$ を $\boldsymbol{A\theta}_0 = \boldsymbol{c}_0$ を満たすベクトルとし, $\boldsymbol{Z} \equiv \boldsymbol{Y} - \boldsymbol{X\theta}_0$ とする. このとき, 線形制約 $\boldsymbol{\Theta}_0 = \left\{\boldsymbol{\theta} \in R^k : \boldsymbol{A\theta} = \boldsymbol{c}_0\right\}$ の下での $\boldsymbol{\theta}$ の制限最小二乗推定量は,

$$\hat{\boldsymbol{\theta}}^{\#} = \boldsymbol{\theta}_0 + \boldsymbol{C}_1{}^t\boldsymbol{XZ}$$

で与えられることを示せ.

11.5. 11.4 の設定で, 正規線形モデル $GMN(\boldsymbol{Y}; \boldsymbol{X\theta}, \sigma^2 \boldsymbol{I}_n)$ において, 回帰係数が線形制約を満たすかどうかの 帰無仮説 $H_0' : \boldsymbol{A\theta} = \boldsymbol{c}_0$ と 対立仮説 $H_1' : \boldsymbol{A\theta} \neq \boldsymbol{c}_0$ を検定するための統計量は

$$T_S' = \frac{(||\boldsymbol{Y} - \boldsymbol{X}\hat{\boldsymbol{\theta}}^{\#}||^2 - ||\boldsymbol{Y} - \boldsymbol{X}\hat{\boldsymbol{\theta}}||^2)/r}{||\boldsymbol{Y} - \boldsymbol{X}\hat{\boldsymbol{\theta}}||^2/(n-k)}$$

である. このとき T_S' は帰無仮説 H_0' の下で自由度 $(r, n-k)$ の F 分布に従うことを示せ.

11.6. 11.5 の設定で, T_S' は対立仮説 H_1' の下で自由度 $(r, n-k)$ 非心度 $\delta^2 \equiv \dfrac{1}{\sigma^2}{}^t(\boldsymbol{\theta} - \boldsymbol{\theta}_0)^t\boldsymbol{X}(\boldsymbol{P}-\boldsymbol{Q})\boldsymbol{X}(\boldsymbol{\theta} - \boldsymbol{\theta}_0)$ の非心 F 分布に従うことを示せ.