1. 回帰分析
1.1. ランダムモデルの回帰分析

Y :	r.v.	目的変数
$\boldsymbol{X} = (X_1, \cdots, X_p)'$:	$p \times 1$ r.vector	説明変数

$$E\begin{bmatrix} \boldsymbol{X} \\ Y \end{bmatrix} = \begin{bmatrix} \boldsymbol{\mu} \\ \nu \end{bmatrix}, \qquad \mathrm{Var}\begin{bmatrix} \boldsymbol{X} \\ Y \end{bmatrix} = \begin{bmatrix} \Sigma & \boldsymbol{\sigma} \\ \boldsymbol{\sigma}' & \sigma^2 \end{bmatrix}$$

$$E(\boldsymbol{X}) = \boldsymbol{\mu} \qquad \mathrm{Var}(\boldsymbol{X}) = \Sigma, \qquad \mathrm{Cov}(\boldsymbol{X}, Y) = \boldsymbol{\sigma}$$
$$E(Y) = \nu \qquad \mathrm{Var}(Y) = \sigma^2$$

仮定：$\begin{cases} \Sigma : \text{正値} \\ \sigma > 0 \end{cases}$

Prop.1.1.1. $Q_0(\alpha, \boldsymbol{\beta}) := E(Y - \alpha - \boldsymbol{\beta}'\boldsymbol{X})^2$
Y を α と $\boldsymbol{\beta}'\boldsymbol{X}$ で近似したときの誤差

問題 $\begin{cases} \min\limits_{\alpha, \boldsymbol{\beta}} \ Q_0(\alpha, \boldsymbol{\beta}) \\ \text{条件} \quad \alpha \in \boldsymbol{R}^1, \ \boldsymbol{\beta} : p \times 1 \end{cases}$

$\Longrightarrow$

解 $\begin{cases} \alpha = \alpha_0 := \nu - \boldsymbol{\beta}_0'\boldsymbol{\mu} \\ \boldsymbol{\beta} = \boldsymbol{\beta}_0 := \Sigma^{-1}\boldsymbol{\sigma} \end{cases}$

最小値 $= \sigma^2 - \boldsymbol{\sigma}'\Sigma^{-1}\boldsymbol{\sigma}$

証明：$Q_0(\alpha, \boldsymbol{\beta}) = \mathrm{Var}(Y - \alpha - \boldsymbol{\beta}'\boldsymbol{X}) + \{E(Y - \alpha - \boldsymbol{\beta}'\boldsymbol{X})\}^2$

第 1 項 $= \mathrm{Var}(Y) - 2\mathrm{Cov}(\boldsymbol{\beta}'\boldsymbol{X}, Y) + \mathrm{Var}(\boldsymbol{\beta}'\boldsymbol{X})$
$= \sigma^2 - 2\boldsymbol{\beta}'\boldsymbol{\sigma} + \boldsymbol{\beta}'\Sigma\boldsymbol{\beta}$

第 2 項 $= (\nu - \alpha - \boldsymbol{\beta}'\boldsymbol{\mu})^2$

第 1 項は、α を含まないから、Q_0 を最小にする α は第 2 項を最小にするという条件から決まる。

つまり、$\alpha = \nu - \boldsymbol{\beta}'\boldsymbol{\mu}$ とする。

このとき、$Q_0(\alpha, \boldsymbol{\beta}) = Q_0(\nu - \boldsymbol{\beta}'\boldsymbol{\mu}, \boldsymbol{\beta}) = \sigma^2 - 2\boldsymbol{\beta}'\boldsymbol{\sigma} + \boldsymbol{\beta}'\Sigma\boldsymbol{\beta}$

$$= (\sigma^2 - \boldsymbol{\sigma}'\Sigma^{-1}\boldsymbol{\sigma}) + \Big(\boldsymbol{\beta} - \Sigma^{-1}\boldsymbol{\sigma}\Big)'\Sigma\Big(\boldsymbol{\beta} - \Sigma^{-1}\boldsymbol{\sigma}\Big)$$

これを最小にする $\boldsymbol{\beta}$ は

$$\boldsymbol{\beta} = \Sigma^{-1}\boldsymbol{\sigma} = \boldsymbol{\beta}_0$$

このとき、$\alpha = \nu - \boldsymbol{\beta}_0'\boldsymbol{\mu} = \alpha_0$

Q_0 の最小値は、$\sigma^2 - \boldsymbol{\sigma}'\Sigma^{-1}\boldsymbol{\sigma}$ 証明終

Def.1.1.1. $y = \alpha_0 + \boldsymbol{\beta}_0'\boldsymbol{x}$ 母回帰関数（母回帰平面）
（$p = 1$ のとき、母回帰直線）

$\widehat{Y} := \alpha_0 + \boldsymbol{\beta}_0'\boldsymbol{X}$	回帰
$e := Y - \widehat{Y}$	残差
$R_0 := \sqrt{\boldsymbol{\sigma}'\Sigma^{-1}\boldsymbol{\sigma}}/\sigma$	母重相関係数
R_0^2	寄与率（決定係数）

$$0 \leqq R_0 \leqq 1$$
$$\left(\text{Min } Q_0 = \sigma^2(1 - R_0^2) \geqq 0\right)$$

$p = 1$ のとき、

$$E\begin{bmatrix} X \\ Y \end{bmatrix} = \begin{bmatrix} \mu \\ \nu \end{bmatrix}, \quad \text{Var}\begin{bmatrix} X \\ Y \end{bmatrix} = \begin{bmatrix} \sigma_{xx} & \sigma_{xy} \\ & \sigma_{yy} \end{bmatrix} = \begin{bmatrix} \sigma_x^2 & \rho\sigma_x\sigma_y \\ & \sigma_y^2 \end{bmatrix}$$

として、

$$\beta_0 = \frac{\sigma_{xy}}{\sigma_{xx}} = \rho\frac{\sigma_y}{\sigma_x}, \quad R_0 := \frac{\sqrt{\boldsymbol{\sigma}'\Sigma^{-1}\boldsymbol{\sigma}}}{\sigma} = \frac{\sqrt{\rho\sigma_x\sigma_y\sigma_x^{-2}\rho\sigma_x\sigma_y}}{\sigma_y} = |\rho|$$

Prop.1.1.2. (i) $E(\widehat{Y}) = \nu, \quad \text{Var}(\widehat{Y}) = \sigma^2 R_0^2$
(ii) $E(e) = 0, \quad \text{Var}(e) = \sigma^2(1 - R_0^2)$
(iii) $\text{Cov}(\widehat{Y}, e) = 0, \quad \text{Cov}(\boldsymbol{X}, e) = \boldsymbol{0}$

証明：(i) $\widehat{Y} = \alpha_0 + \boldsymbol{\beta}_0'\boldsymbol{X} = \nu + \boldsymbol{\beta}_0'(\boldsymbol{X} - \boldsymbol{\mu})$ (1)
ゆえに、$E(\widehat{Y}) = \nu$
$\text{Var}(\widehat{Y}) = \boldsymbol{\beta}_0'\text{Var}(\boldsymbol{X})\boldsymbol{\beta}_0 = \boldsymbol{\sigma}'\Sigma^{-1}\boldsymbol{\sigma} = \sigma^2 R_0^2$

(ii) $e = Y - \widehat{Y} = (Y - \nu) - \boldsymbol{\beta}_0'(\boldsymbol{X} - \boldsymbol{\mu})$ (2)
ゆえに、$E(e) = 0$
$\text{Var}(e) = Q_0(\alpha_0, \boldsymbol{\beta}_0) = \sigma^2(1 - R_0^2)$

(iii) まず (2) 式より、$\text{Cov}(\boldsymbol{X}, e) = \text{Cov}(\boldsymbol{X}, Y - \boldsymbol{\beta}_0'\boldsymbol{X}) = \boldsymbol{\sigma} - \Sigma\boldsymbol{\beta}_0 = \boldsymbol{0}$
ゆえに (1) 式より、$\text{Cov}(\widehat{Y}, e) = \text{Cov}(\boldsymbol{\beta}_0'\boldsymbol{X}, e) = \boldsymbol{\beta}_0'\text{Cov}(\boldsymbol{X}, e) = 0$ 証明終

[注意] 要因の分解

変量	$Y =$	$\widehat{Y}$	$+$	e
平均	$\nu =$	ν	$+$	0
分散	$\sigma^2 =$	$\sigma^2 R_0^2$	$+$	$\sigma^2(1 - R_0^2)$
		回帰		残差

Prop.1.1.3. 問題 $\left\{ \begin{array}{l} \max\limits_{\boldsymbol{\beta}} \ \mathrm{cor}(\boldsymbol{\beta}'\boldsymbol{X}, Y) \\ 条件 \quad \boldsymbol{\beta} : p \times 1 \end{array} \right.$

$\Longrightarrow$

解：$\boldsymbol{\beta} \propto \boldsymbol{\beta}_0$
最大値 $= R_0$

証明：$\mathrm{cor}(\boldsymbol{\beta}'\boldsymbol{X}, Y) = \dfrac{\mathrm{Cov}(\boldsymbol{\beta}'\boldsymbol{X}, Y)}{\sqrt{\mathrm{Var}(\boldsymbol{\beta}'\boldsymbol{X})}\sqrt{\mathrm{Var}(Y)}} = \dfrac{\boldsymbol{\beta}'\boldsymbol{\sigma}}{\sqrt{\boldsymbol{\beta}'\Sigma\boldsymbol{\beta}}\ \sigma}$

Prop.B2 より、解 $\boldsymbol{\beta} \propto \Sigma^{-1}\boldsymbol{\sigma} = \boldsymbol{\beta}_0$
最大値 $= \sqrt{\boldsymbol{\sigma}'\Sigma^{-1}\boldsymbol{\sigma}}/\sigma = R_0$ 証明終

Prop.1.1.4 Σ を正則と仮定しない場合

(i) Σ^- を Σ の一般逆行列とすると
$\boldsymbol{\sigma}'\Sigma^-\boldsymbol{\sigma}$, $\boldsymbol{\sigma}'\Sigma^-(\boldsymbol{X}-\boldsymbol{\mu})$ は、Σ^- の選択に依存しない。

(ii) $Q_0(\alpha, \boldsymbol{\beta})$ を最小にする解 $\widehat{Y} = \alpha_0 + \boldsymbol{\beta}_0'\boldsymbol{X}$ は、

$$\widehat{Y} = \nu + \boldsymbol{\sigma}'\Sigma^-(\boldsymbol{X}-\boldsymbol{\mu})$$

最小値：$\sigma^2 - \boldsymbol{\sigma}'\Sigma^-\boldsymbol{\sigma}$

証明：$\Sigma = V\Lambda V' = \begin{bmatrix} V_1 & V_2 \end{bmatrix} \begin{bmatrix} \Lambda_1 & O \\ O & O \end{bmatrix} \begin{bmatrix} V_1' \\ V_2' \end{bmatrix}$：スペクトル分解

$\boldsymbol{Z} = \begin{bmatrix} \boldsymbol{Z}_1 \\ \boldsymbol{Z}_2 \end{bmatrix} := V'(\boldsymbol{X}-\boldsymbol{\mu})$, $\boldsymbol{c} = \begin{bmatrix} \boldsymbol{c}_1 \\ \boldsymbol{c}_2 \end{bmatrix} := V'\boldsymbol{\sigma}$ とおく。

このとき、$\boldsymbol{Z}$ の期待値と分散行列を計算すると、以下のようになる。

$$E(\boldsymbol{Z}) = V'E(\boldsymbol{X}-\boldsymbol{\mu}) = \boldsymbol{0}$$

$$\mathrm{Var}(\boldsymbol{Z}) = V'\mathrm{Var}(\boldsymbol{X}-\boldsymbol{\mu})V = V'\Sigma V = V'V\Lambda V'V = \Lambda = \begin{bmatrix} \Lambda_1 & O \\ O & O \end{bmatrix}$$

ここで、$E(\boldsymbol{Z}_2) = \boldsymbol{0}$, $\mathrm{Var}(\boldsymbol{Z}_2) = O$ であるから、

$$\boldsymbol{Z}_2 = \boldsymbol{0}(w.p.1) \qquad \cdots \quad (1)$$

がわかる。また、$\mathrm{Cov}(\boldsymbol{Z}, Y) = V'\mathrm{Cov}(\boldsymbol{X}, Y) = V'\boldsymbol{\sigma} = \boldsymbol{c}$ であるから、

$$\boldsymbol{c} = \begin{bmatrix} \boldsymbol{c}_1 \\ \boldsymbol{c}_2 \end{bmatrix} = \begin{bmatrix} V_1'\boldsymbol{\sigma} \\ V_2'\boldsymbol{\sigma} \end{bmatrix} = \begin{bmatrix} \mathrm{Cov}(\boldsymbol{Z}_1, Y) \\ \mathrm{Cov}(\boldsymbol{Z}_2, Y) \end{bmatrix}$$

と、(1) 式を併せて考えると、

$$\boldsymbol{c}_2 = \mathrm{Cov}(\boldsymbol{Z}_2, Y) = \boldsymbol{0} \qquad \cdots \quad (2)$$

がわかる。

ここで、$\Sigma = V\Lambda V' = V_1\Lambda_1 V_1'$ より、$P_1 = Q_1 = V_1$ と考えられるから、どんな一般逆行列 Σ^- に対しても、$P = Q = V$ の場合を考えよう。
すると、Prop.A.16 より、$C = V'\Sigma^- V$ とおくと、$C_{11} = \Lambda_1^{-1} \quad \cdots (3)$ である。

このとき、$\boldsymbol{\sigma}'\Sigma^-\boldsymbol{\sigma} = \boldsymbol{\sigma}'V(V'\Sigma^-V)V'\boldsymbol{\sigma} = \boldsymbol{c}'(V'\Sigma^-V)\boldsymbol{c} = \begin{bmatrix} \boldsymbol{c}_1' & \boldsymbol{c}_2' \end{bmatrix} \begin{bmatrix} C_{11} & C_{12} \\ C_{21} & C_{22} \end{bmatrix} \begin{bmatrix} \boldsymbol{c}_1 \\ \boldsymbol{c}_2 \end{bmatrix}$.

すると、(2),(3) 式より、$\boldsymbol{\sigma}'\Sigma^-\boldsymbol{\sigma} = \begin{bmatrix} \boldsymbol{c}_1' & \boldsymbol{0}' \end{bmatrix} \begin{bmatrix} C_{11} & C_{12} \\ C_{21} & C_{22} \end{bmatrix} \begin{bmatrix} \boldsymbol{c}_1 \\ \boldsymbol{0} \end{bmatrix} = \boldsymbol{c}_1'\Lambda_1^{-1}\boldsymbol{c}_1$ とわかる。

次に、$\boldsymbol{\sigma}'\Sigma^-(\boldsymbol{X}-\boldsymbol{\mu}) = \boldsymbol{c}'(V'\Sigma^-V)\boldsymbol{Z} = \begin{bmatrix} \boldsymbol{c}_1' & \boldsymbol{c}_2' \end{bmatrix} \begin{bmatrix} C_{11} & C_{12} \\ C_{21} & C_{22} \end{bmatrix} \begin{bmatrix} \boldsymbol{Z}_1 \\ \boldsymbol{Z}_2 \end{bmatrix}$.

また、(1),(3) 式より、$\boldsymbol{\sigma}'\Sigma^-(\boldsymbol{X}-\boldsymbol{\mu}) = \boldsymbol{c}_1'\Lambda_1^{-1}\boldsymbol{Z}_1$ とわかる。
（ここまでで、2 つの量は、Σ^- によらない定まっている量のみで表現できた。）

これらのことを用いて、$Q_0(\alpha, \boldsymbol{\beta})$ を最小化しよう。$\boldsymbol{\gamma} := V'\boldsymbol{\beta}$ とおくと、

$$
\begin{aligned}
Q_0(\alpha, \boldsymbol{\beta}) &= \mathrm{Var}(Y - \boldsymbol{\beta}'\boldsymbol{X}) + \{E(Y - \alpha - \boldsymbol{\beta}'\boldsymbol{X})\}^2 \\
&= \sigma^2 - 2\boldsymbol{\beta}'\boldsymbol{\sigma} + \boldsymbol{\beta}'\Sigma\boldsymbol{\beta} = \sigma^2 - 2\boldsymbol{\beta}'VV'\boldsymbol{\sigma} + \boldsymbol{\beta}'(V\Lambda V')\boldsymbol{\beta} = \sigma^2 - 2\boldsymbol{\gamma}'\boldsymbol{c} + \boldsymbol{\gamma}'\Lambda\boldsymbol{\gamma} \\
&= \sigma^2 - 2\begin{bmatrix} \boldsymbol{\gamma}_1' & \boldsymbol{\gamma}_2' \end{bmatrix} \begin{bmatrix} \boldsymbol{c}_1 \\ \boldsymbol{0} \end{bmatrix}. + \begin{bmatrix} \boldsymbol{\gamma}_1' & \boldsymbol{\gamma}_2' \end{bmatrix} \begin{bmatrix} \Lambda_1 & O \\ O & O \end{bmatrix} \begin{bmatrix} \boldsymbol{\gamma}_1 \\ \boldsymbol{\gamma}_2 \end{bmatrix} \\
&= \sigma^2 - 2\boldsymbol{\gamma}_1'\boldsymbol{c}_1 + \boldsymbol{\gamma}_1'\Lambda_1\boldsymbol{\gamma}_1
\end{aligned}
$$

これを最小化する $\boldsymbol{\gamma}_1 = \Lambda_1^{-1}\boldsymbol{c}_1$ であるから、このときの最小値は

$$\text{最小値} = \sigma^2 - \boldsymbol{c}_1'\Lambda_1^{-1}\boldsymbol{c}_1 = \sigma^2 - \boldsymbol{\sigma}'\Sigma^-\boldsymbol{\sigma}$$

であり、その時の予測値 $\widehat{Y} = \nu + \boldsymbol{c}_1'\Lambda_1^{-1}\boldsymbol{Z}_1 = \nu + \boldsymbol{\sigma}'\Sigma^-(\boldsymbol{X}-\boldsymbol{\mu})$ である。 証明終