No. 1

Date

統計学3 中間レポート 6321120 横溝尚也

[1]

$$\lambda = \frac{\{\underset{\sim}{w}(\bar{\underset{\sim}{x}}_1 - \bar{\underset{\sim}{x}}_2)\}^2}{\underset{\sim}{w}^T S \underset{\sim}{w}}$$

ただし $\underset{\sim}{w} = \begin{pmatrix} w_1 \\ w_2 \end{pmatrix}$, $\bar{\underset{\sim}{x}}_1 = \begin{pmatrix} \bar{x}_1^{(1)} \\ \bar{x}_2^{(1)} \end{pmatrix}$, $\bar{\underset{\sim}{x}}_2 = \begin{pmatrix} \bar{x}_1^{(2)} \\ \bar{x}_2^{(2)} \end{pmatrix}$

$S = \frac{1}{n_1 + n_2 - 2}\{(n_1 - 1)S_1 + (n_2 - 1)S_2\}$,

$S_1 = \begin{pmatrix} s_{11}^{(1)} & s_{12}^{(1)} \\ s_{21}^{(1)} & s_{22}^{(1)} \end{pmatrix}$, $S_2 = \begin{pmatrix} s_{11}^{(2)} & s_{12}^{(2)} \\ s_{21}^{(2)} & s_{22}^{(2)} \end{pmatrix}$

$$\lambda = \frac{\{\underset{\sim}{w}(\bar{\underset{\sim}{x}}_1 - \bar{\underset{\sim}{x}}_2)\}^2}{\underset{\sim}{w}^T S \underset{\sim}{w}}$$

$$= \frac{\underset{\sim}{w}^T(\bar{\underset{\sim}{x}}_1 - \bar{\underset{\sim}{x}}_2)(\bar{\underset{\sim}{x}}_1 - \bar{\underset{\sim}{x}}_2)^T \underset{\sim}{w}}{\underset{\sim}{w}^T S \underset{\sim}{w}}$$

$$\frac{\partial \lambda}{\partial \underset{\sim}{w}} = \frac{1}{(\underset{\sim}{w}^T S \underset{\sim}{w})^2}\left\{2(\bar{\underset{\sim}{x}}_1 - \bar{\underset{\sim}{x}}_2)(\bar{\underset{\sim}{x}}_1 - \bar{\underset{\sim}{x}}_2)^T \underset{\sim}{w}(\underset{\sim}{w}^T S \underset{\sim}{w}) - 2(\underset{\sim}{w}^T(\bar{\underset{\sim}{x}}_1 - \bar{\underset{\sim}{x}}_2)(\bar{\underset{\sim}{x}}_1 - \bar{\underset{\sim}{x}}_2)^T \underset{\sim}{w}) S \underset{\sim}{w}\right\}$$

$$= 0$$

$$\Leftrightarrow (\bar{\underset{\sim}{x}}_1 - \bar{\underset{\sim}{x}}_2)(\bar{\underset{\sim}{x}}_1 - \bar{\underset{\sim}{x}}_2)^T \underset{\sim}{w}\, \underset{\sim}{w}^T S \underset{\sim}{w} = \underset{\sim}{w}^T(\bar{\underset{\sim}{x}}_1 - \bar{\underset{\sim}{x}}_2)(\bar{\underset{\sim}{x}}_1 - \bar{\underset{\sim}{x}}_2)^T \underset{\sim}{w} S \underset{\sim}{w}$$

$$\Leftrightarrow \underset{\sim}{w}^T(\bar{\underset{\sim}{x}}_1 - \bar{\underset{\sim}{x}}_2)(\bar{\underset{\sim}{x}}_1 - \bar{\underset{\sim}{x}}_2)^T \underset{\sim}{w} S \underset{\sim}{w} = (\underset{\sim}{w}^T S \underset{\sim}{w})(\bar{\underset{\sim}{x}}_1 - \bar{\underset{\sim}{x}}_2)(\bar{\underset{\sim}{x}}_1 - \bar{\underset{\sim}{x}}_2)^T \underset{\sim}{w}$$

$\Leftrightarrow \underset{\sim}{w} = S^{-1}(\bar{\underset{\sim}{x}}_1 - \bar{\underset{\sim}{x}}_2)$ ∵ $(\bar{\underset{\sim}{x}}_1 - \bar{\underset{\sim}{x}}_2)^T \underset{\sim}{w}$ はスカラーより、
$\bar{\underset{\sim}{x}}_1 - \bar{\underset{\sim}{x}}_2$ に平行なベクトル。

$\therefore \underset{\sim}{w} = S^{-1}(\bar{\underset{\sim}{x}}_1 - \bar{\underset{\sim}{x}}_2)$ //

2

回帰モデル $y_i = \sum_{j=1}^{m} w_j b_j(\underline{x}_i) + \varepsilon_i \ (i=1,\cdots,n)$

$\varepsilon_i \sim N(0, \sigma^2)$ を仮定。

このとき $y_i \sim N(\sum_{j=1}^{m} w_j b_j, \sigma^2)$ となる。

ε_i は互いに無相関より y_i と y_j は無相関 $(i \neq j)$。

よって、尤度関数 $L(\underline{y} | X ; \underline{w}, \sigma^2)$ は

$$L(\underline{y} | X ; \underline{w}, \sigma^2) = \frac{1}{(2\pi\sigma^2)^{\frac{n}{2}}} \exp\left[-\frac{1}{2\sigma^2}(\underline{y} - B\underline{w})^T(\underline{y} - B\underline{w})\right],$$

$$\left(\because B = \begin{bmatrix} b_1(\underline{x}_1) & \cdots & b_m(\underline{x}_1) \\ b_1(\underline{x}_2) & & \\ \vdots & \ddots & \vdots \\ b_1(\underline{x}_n) & & b_m(\underline{x}_n) \end{bmatrix} \right)$$

対数尤度関数 l は

$$l(\underline{w}, \sigma^2) = \log L$$

$$= -\frac{n}{2}\log 2\pi - \frac{n}{2}\log\sigma^2 - \frac{1}{2\sigma^2}(\underline{y} - B\underline{w})^T(\underline{y} - B\underline{w})。$$

ここで正則化項を $-\frac{\lambda}{2}\underline{w}^T K \underline{w} \ (\lambda > 0, K \geq 0)$ として課した

正則化対数尤度関数 $l_\lambda(\underline{w}, \sigma^2)$ は

$$l_\lambda(\underline{w}, \sigma^2) = -\frac{n}{2}\log 2\pi - \frac{n}{2}\log\sigma^2 - \frac{1}{2\sigma^2}(\underline{y} - B\underline{w})^T(\underline{y} - B\underline{w}) - \frac{\lambda}{2}\underline{w}^T K \underline{w}$$

と表せる。パラメータ $\underline{w}$ と σ^2 で偏微分し $(=0)$ を解き最尤推定量を求める。

$$\frac{\partial l_\lambda}{\partial \underline{w}} = \frac{1}{\sigma^2}(B^T\underline{y} - B^T B\underline{w}) - \lambda K\underline{w} \ (=0)$$

$$\left(\because \frac{\partial}{\partial \underline{w}}\underline{w}^T B^T \underline{y} = \frac{\partial}{\partial \underline{w}}\underline{y}^T B\underline{w} = B^T\underline{y}, \ \frac{\partial}{\partial \underline{w}}\underline{w}^T B^T B\underline{w} = 2B^T B\underline{w}\right)$$

$$\Leftrightarrow \hat{\underline{w}} = (B^T B + \lambda\sigma^2 K)^{-1} B^T \underline{y} \quad \cdots ①$$

$$\frac{\partial l_\lambda}{\partial \sigma^2} = -\frac{n}{2\sigma^2} + \frac{1}{2\sigma^4}(\underline{y} - B\underline{w})^T(\underline{y} - B\underline{w}) \ (=0)$$

$$\Leftrightarrow \hat{\sigma}^2 = \frac{1}{n}(\underline{y} - B\underline{w})^T(\underline{y} - B\underline{w}) \cdots ②$$

①②より $\underline{w}$ と σ^2 の正則化最尤推定量 $\hat{\underline{w}}, \hat{\sigma}^2$ は

$$\begin{cases} \hat{\underline{w}} = (B^T B + \lambda\hat{\sigma}^2 K)^{-1} B^T \underline{y} \\ \hat{\sigma}^2 = \frac{1}{n}(\underline{y} - B\underline{w})^T(\underline{y} - B\underline{w}) \end{cases} //$$

ただし

$$B = \begin{bmatrix} \exp\left[-\frac{\|\underline{x}_1 - \bar{\underline{x}}_1\|^2}{2s_1^2}\right] & \cdots & \exp\left[-\frac{\|\underline{x}_1 - \bar{\underline{x}}_m\|^2}{2s_m^2}\right] \\ \exp\left[-\frac{\|\underline{x}_2 - \bar{\underline{x}}_1\|^2}{2s_1^2}\right] & \ddots & \vdots \\ \vdots & & \\ \exp\left[-\frac{\|\underline{x}_n - \bar{\underline{x}}_1\|^2}{2s_1^2}\right] & & \exp\left[-\frac{\|\underline{x}_n - \bar{\underline{x}}_m\|^2}{2s_m^2}\right] \end{bmatrix}$$

[3] $\lambda_1, \dots, \lambda_m$ が A の固有値のとき $A > O \iff \lambda_i > 0$ for all i の証明

(i) $A > O \Rightarrow \lambda_i > 0$ for all i について
A は正定値行列より $\forall \underset{\sim}{x}$ ($m \times 1$ ベクトル), $\underset{\sim}{x}^t A \underset{\sim}{x} > 0$ となる。
$\underset{\sim}{x}$ を A の固有ベクトル, λ を $\underset{\sim}{x}$ に対応する固有値とすると
$A\underset{\sim}{x} = \lambda \underset{\sim}{x}$ より $\underset{\sim}{x}^t A \underset{\sim}{x} = \underset{\sim}{x}^t \lambda \underset{\sim}{x} = \lambda \|\underset{\sim}{x}\|^2 > 0$.
$\therefore \lambda > 0$ ($\because \|\underset{\sim}{x}\|^2 > 0$)

$\lambda_i > 0$ for all $i \Rightarrow A > O$ について
A が対称行列より $U^{-1} A U = \begin{pmatrix} \lambda_1 & & O \\ & \ddots & \\ O & & \lambda_m \end{pmatrix} = \Lambda$ と対角化できる。
($\because U$ はユニタリ行列)
$U^{-1} = U^t$ より $A = U \Lambda U^t$。
$\underset{\sim}{x}^t A \underset{\sim}{x} = \underset{\sim}{x}^t U \Lambda U^t \underset{\sim}{x}$
$= \underset{\sim}{k} \Lambda \underset{\sim}{k}^t$ ($\because \underset{\sim}{k} = U^t \underset{\sim}{x}$)
$= \sum_{i=1}^{m} \lambda_i k_i^2 \cdots ①$ と表せる。
$\lambda_1, \dots, \lambda_m > 0$ より $① > 0$.
$\therefore \underset{\sim}{x}^t A \underset{\sim}{x} > 0$ となり $A > O$

よって題意は示された。//

[4]
(1) $\underset{\sim}{x}$ を $n \times 1$ ベクトルとし $A\underset{\sim}{x} = 0$ のとき $A^t A \underset{\sim}{x} = 0$,
$AA\underset{\sim}{x} = 0$ のとき. $\langle A^t A \underset{\sim}{x}, \underset{\sim}{x} \rangle = \langle A\underset{\sim}{x}, A\underset{\sim}{x} \rangle = \|A\underset{\sim}{x}\|^2 = 0$
より $A\underset{\sim}{x} = 0$
$\therefore \mathrm{Ker}(A^t A) = \mathrm{Ker}(A)$

$\mathrm{rank}(A^t A) = n - \dim(\mathrm{Ker}\, A^t A)$
$= n - \dim(\mathrm{Ker}\, A)$
$= r$
$\therefore \mathrm{rank}(A^t A) = r$
また. $(A^t A)^t = A^t A$ より $A^t A$ は対称行列となり対角化可能。
よって $A^t A$ の正の個有値は r。//

(2) $r = n \Rightarrow \mathrm{rank}(A^t A) = n$, $A^t A$ は $n \times n$ 行列より
$A^t A$ の全ての固有値が正。
[3] より $A^t A > O$ //