【数量化3類】ー対応分析(Correspondence Analysis)
とも ー

目的：主成分分析と同じ。

「多量の質的変数を、低い次元の
合成変数に変換し、データが有している
情報をより解釈しやすくする。」(ref. p132)

教科書の例では、

(表 10.1)　：　10 × 7 = 70
　　　　　　　サンプル数　変数の数

を

(表 10.8, 10.9)　：　7 × 2 + 10 × 2 = 34
　　　　　　　変数の数　合成変数数　サンプル数

に変換している。

さらに表 10.1, 10.2 でその解釈を与えている。

『「表10.6から求まる$x$と$y$の相関係数を最大にする$x$と$y$を与える」』

なぜ○を対角成分に集めることは$x$と$y$の相関係数を大きくすることなのか.

「並べかえ」を表現するために,

$$x_i \in \{1, 2, 3\} \quad \text{for } i \in \{1, 2, 3\}$$

ただし、$i \neq j$ で $x_i \neq x_j$ とする.

$y_i$ も同様に

$$y_i \in \{1, 2, 3, 4\} \quad \text{for } i \in \{1, 2, 3, 4\}$$

とする. 並べかえた後の表を作ると、次のようになる.

| $y_i$ \ $x_i$ | 1 | 2 | 3 |
|---|---|---|---|
| 1 | ○ | | |
| 2 | ○ | ○ | |
| 3 | | ○ | |
| 4 | | ○ | ○ |

$x_i$が小さいと $y_i$も小さい.
逆も同じ.

○を1つのサンプルとすると、$x$列、$y$列を表で書けて、

| $x$ | $y$ |
|---|---|
| $x_1$ | $y_4$ |
| $x_2$ | $y_1$ |
| $x_2$ | $y_2$ |
| $x_3$ | $y_3$ |
| $x_3$ | $y_4$ |

これらの相関係数を最大にすれば良い。

ここまでは「並べかえ」の説明のため
$x_i$ $y_i$ を $\{1, 2, 3\}$, $\{1, 2, 3, 4\}$ の値としたが、今後は実数値に拡張して
$x$, $y$ を求める。

P157 〜　　　の計算をやってみるが、読みとばしてOK.

(10.14) $u_1$ + (10.15) $u_2$ + (10.16) $u_3$

$$\Leftrightarrow \underbrace{\frac{w_4}{\sqrt{2}}u_1 + \frac{w_1}{2}u_2 + \frac{w_2}{\sqrt{2}}u_2 + \frac{w_1}{\sqrt{6}}u_3 + \frac{w_3}{\sqrt{3}}u_3 + \frac{w_4}{\sqrt{6}}u_3}_{= S_{xy}} = \lambda\underbrace{(u_1^2 + u_2^2 + u_3^2)}_{=1}$$

$$\Leftrightarrow \quad S_{xy} = \lambda$$

同様に (10.17) 〜 (10.20) からは.

$$S_{xy} = \eta$$

(10.17) 〜 (10.20) の $\eta$ を $\lambda$ で表し. $w_1 \sim w_4$ を (10.14) 〜 (10.16) に代入すると.

$$\frac{1}{\sqrt{2}}\frac{1}{\lambda}\left(\frac{u_1}{\sqrt{2}} + \frac{u_3}{\sqrt{6}}\right) = \lambda u_1$$

$$\Leftrightarrow \quad \frac{u_1}{2} + \frac{\sqrt{3}}{6}u_3 = \lambda^2 u_1$$

他も同様に…

「固有値を求める問題に帰着したが、一般形はどうなるのか？」

表10.6のx列、y列を、$\mathbf{x}$、$\mathbf{y}$とする。
$\mathbf{x}$は変数の数$p$として、$\{x_1, \dots, x_p\}$の要素で構成されるベクトルで、第$i$成分を$x_{\sigma_x(i)}$と書く。
$\mathbf{y}$はサンプルの数$n$として、$\{y_1, \dots, y_n\}$の要素で構成されるベクトルで、第$i$成分を$y_{\sigma_y(i)}$と書く。
反応の数を$N$とすると、どちらも長さ$N$である。

$$\sum_{1}^{N} x_{\sigma_x(i)} = 0, \quad \sum_{1}^{N} y_{\sigma_y(i)} = 0$$

とする。

$$S_{xy} = \sum^{N} xy - \underbrace{\frac{(\sum^{N} x)(\sum^{N} y)}{N}}_{=0}$$

$$= \mathbf{x}'\mathbf{y}$$

$$S_{xx} = \sum^{N} x^2 - \underbrace{\frac{(\sum^{N} x)^2}{N}}_{=0}$$

$$= \mathbf{x}'\mathbf{x}$$

$$S_{yy} = \mathbf{y}'\mathbf{y}$$

ここで、$S_{xx} = 1$ 、$S_{yy} = 1$ の元で

$S_{xy}$ の最大化を考える。

$$f(x_1, \cdots, x_p, y_1, \cdots, y_n, \lambda, \eta)$$

$$= x'y - \frac{\lambda}{2}(x'x - 1) - \frac{\eta}{2}(y'y - 1)$$

$x_1 \cdots x_p$ を並べたベクトルを $\dot{x} = \begin{pmatrix} x_1 \\ \vdots \\ x_p \end{pmatrix}$

$y_1 \cdots y_n$ を並べたベクトルを $\dot{y} = \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix}$

とすると、

$$\frac{\partial f}{\partial \dot{x}} = 0 \quad より、$$

$$\begin{cases} \dfrac{\partial x'}{\partial x_1} y - \dfrac{1}{2}\lambda \dfrac{\partial x}{\partial x_1} \cdot \dfrac{\partial (x'x)}{\partial x} = 0 \\ \quad \vdots \\ \dfrac{\partial x'}{\partial x_p} \cdot y - \dfrac{1}{2}\lambda \dfrac{\partial x}{\partial x_p} \cdot \underbrace{\dfrac{\partial (x'x)}{\partial x}}_{= 2x} = 0 \end{cases}$$

λかこむ

これをまとめて.

$$\underbrace{\frac{\partial x'}{\partial x}}_{p\times N}\underbrace{y}_{N\times 1} - \lambda \underbrace{\frac{\partial x'}{\partial x}}_{p\times N}\underbrace{x}_{N\times 1} = 0$$

$$\Leftrightarrow \frac{\partial x'}{\partial x}(y - \lambda x) = 0 \quad \text{—— ①}$$

同様に. $\frac{\partial f}{\partial y} = 0$ より.

$$\underset{n\times N}{\frac{\partial y'}{\partial y}}\underset{N\times 1}{(x - \eta y)} = 0 \quad \text{—— ②}$$

↳ ここまでできたが. これ以降の展開がわからず…

$\boxed{\text{別方針}}$

$$Z = \underbrace{\overbrace{\begin{pmatrix} 1 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 1 \\ & & \vdots & & \\ 0 & 1 & 0 & 0 & 1 \end{pmatrix}}^{p}}_{}\Bigg\} n \quad \overset{\text{和}}{\Longrightarrow} \quad Z_{n.} = \begin{pmatrix} 2 \\ 2 \\ \vdots \\ 2 \end{pmatrix}$$

$$\Downarrow \text{和}$$

$$Z_{.p}' = \begin{pmatrix} 3 & 4 & 3 & 1 & 2 \end{pmatrix}$$

$$x = \begin{pmatrix} x_1 \\ \vdots \\ x_p \end{pmatrix}, \quad y = \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix} \quad \text{とおく.}$$

(10.1) (10.2) より.

$$\begin{cases} Z_{.p}' \, x = 0 \\ Z_{n.}' \, y = 0 \end{cases}$$

$$S_{xy} = y' Z x$$

なぜなら

$$= \begin{pmatrix} y_1 & \cdots & y_n \end{pmatrix} \begin{pmatrix} 1 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 1 \\ & & \vdots & & \\ 0 & 1 & 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} x_1 \\ \vdots \\ x_p \end{pmatrix}$$

$$= \begin{pmatrix} y_1 & \cdots & y_n \end{pmatrix} \begin{pmatrix} x_1 + x_p \\ x_3 + x_5 \\ \vdots \\ x_2 + x_5 \end{pmatrix}$$

$$= x_1 y_1 + x_p y_n + x_3 y_2 + \cdots$$

$x$. $y$ を次のように変数変換する。 ref.(10.6) (10.7)

$$u = \begin{pmatrix} \sqrt{Z_{\cdot p}^{(1)}} & & & 0 \\ & \sqrt{Z_{\cdot p}^{(2)}} & & \\ & & \ddots & \\ 0 & & & \sqrt{Z_{\cdot p}^{(p)}} \end{pmatrix} x \equiv (\tilde{Z}_{\cdot p})^{\frac{1}{2}} x$$

$$w = \begin{pmatrix} \sqrt{Z_{n\cdot}^{(1)}} & & & 0 \\ & \sqrt{Z_{n\cdot}^{(2)}} & & \\ & & \ddots & \\ 0 & & & \sqrt{Z_{n\cdot}^{(n)}} \end{pmatrix} y \equiv (\tilde{Z}_{n\cdot})^{\frac{1}{2}} y$$

$$\therefore \begin{cases} x = (Z_{\cdot p}')^{-\frac{1}{2}} u \\ y = (Z_{n\cdot}')^{-\frac{1}{2}} w \end{cases}$$

これを使って相関係数を $u$, $w$ で表す。

$$\begin{cases} S_{xy} = w' \underbrace{(\tilde{Z}_{n\cdot})^{-\frac{1}{2}} Z (\tilde{Z}_{\cdot p}')^{-\frac{1}{2}}}_{= K \text{ とおく.}} u \\ S_{xx} = u'u \\ S_{yy} = w'w \end{cases}$$

よって. $S_{xx} = 1$ . $S_{yy} = 1$ の元で. ラグランジュの式は.

$$f = \underset{1\times n}{w'} \underset{n\times p}{K} \underset{p\times 1}{u} - \frac{\lambda}{2}(u'u - 1) - \frac{\eta}{2}(w'w - 1)$$

$\dfrac{\partial f}{\partial u} = 0$ , $\dfrac{\partial f}{\partial w} = 0$ より.

$$\begin{cases} K' w - \lambda u = 0 & \text{—— ①} \\ K u - \eta w = 0 & \text{—— ②} \end{cases}$$

$u' \cdot$ ① と $w' \cdot$ ② より.

$$S_{xy} = u' K' w = \lambda = \eta$$

これと ② より.

$$w = \frac{K}{\lambda} u$$

① に代入し.

$$K' K u = \lambda^2 u$$

よって. $K'K$ の固有値問題に帰着する.

$$K'K = (\tilde{Z}_{\cdot p})^{-\frac{1}{2}} Z' (\tilde{Z}_{n\cdot})^{-1} Z (\tilde{Z}_{\cdot p})^{-\frac{1}{2}}$$

$p \times p$ $\quad$ $p \times n$ $\quad$ $n \times n$ $\quad$ $n \times p$ $\quad$ $p \times p$

次に $w$ の方を求める。①より、

$$u = \frac{1}{\lambda} K' w$$

これを②に代入し、

$$K K' w = \lambda^2 w$$

よって $KK'$ の固有値問題に帰着する。

P159 数量化3類では、つねに固有値1で、…

わからん

成分は{(pとnの小さい方)−1}コまで求めることができる

KK' と K'K の固有値は nコ と pコ で、

第1固有値は常に使えないので、

{(pとnの小さい方)−1}コ

が最大個数になる。

## 【問題 (10.2)】

(1) (10.28) に、(10.6) を代入すると。

$$
\begin{cases}
\frac{1}{2}x_1 + \frac{1}{2}x_3 = \lambda^2 x_1 \\
\frac{3\sqrt{2}}{4}x_2 + \frac{\sqrt{2}}{4}x_3 = \lambda^2\sqrt{2}\,x_2 \\
\frac{\sqrt{3}}{6}x_1 + \frac{\sqrt{3}}{6}x_2 + \frac{2\sqrt{3}}{3}x_3 = \lambda^2\sqrt{3}\,x_3
\end{cases}
$$

$$
\Leftrightarrow \begin{pmatrix} \frac{1}{2} & 0 & \frac{1}{2} \\ 0 & \frac{3}{4} & \frac{1}{4} \\ \frac{1}{6} & \frac{1}{6} & \frac{2}{3} \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \lambda^2 \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix}
$$

(2) $\lambda_1^2 = 1$ $\quad x_1' = (0.58,\ 0.58,\ 0.58)$

$\lambda_2^2 = 0.67$ $\quad x_2' = (0.69,\ -0.69,\ 0.23)$

$\lambda_3^2 = 0.25$ $\quad x_3' = (0.87,\ 0.22,\ -0.44)$

$x_2' \cdot x_3 = 0.35 \neq 0$ (直交しない)