

## 因子分析

データ数を  $n$ 、変量数を  $p$  とする。

$$\text{データを } [X] = \begin{bmatrix} x_{1,1} & \cdots & x_{1,p} \\ \vdots & \ddots & \vdots \\ x_{n,1} & \cdots & x_{n,p} \end{bmatrix} \quad \text{と表す。}$$

各変量ごとに標準化したデータを  $[Z]$  として

$$[Z] = \begin{bmatrix} z_{1,1} & \cdots & z_{1,p} \\ \vdots & \ddots & \vdots \\ z_{n,1} & \cdots & z_{n,p} \end{bmatrix} \quad \text{と表す。すなわち、}$$

上記の各列ベクトル  $\begin{Bmatrix} z_{1,j} \\ \vdots \\ z_{n,j} \end{Bmatrix} (j = 1 \sim p)$  の各成分  $z_{ij} (i = 1 \sim n, j = 1 \sim p)$  は

$\begin{Bmatrix} x_{1,j} \\ \vdots \\ x_{n,j} \end{Bmatrix}$  を標準化、つまり  $\mu_j$  を  $\begin{Bmatrix} x_{1,j} \\ \vdots \\ x_{n,j} \end{Bmatrix}$  の平均、 $\sigma_j$  を  $\begin{Bmatrix} x_{1,j} \\ \vdots \\ x_{n,j} \end{Bmatrix}$  の標準偏差

として  $z_{i,j} = \frac{x_{i,j} - \mu_j}{\sigma_j}$  としたもの。

因子分析モデルは 以下の式を前提とする。

ここで  $m$  は因子数を表す。

$$\begin{bmatrix} z_{1,1} & \cdots & z_{1,p} \\ \vdots & \ddots & \vdots \\ z_{n,1} & \cdots & z_{n,p} \end{bmatrix} = \begin{bmatrix} f_{1,1} & \cdots & f_{1,m} \\ \vdots & \ddots & \vdots \\ f_{n,1} & \cdots & f_{n,m} \end{bmatrix} \begin{bmatrix} a_{1,1} & \cdots & a_{p,1} \\ \vdots & \ddots & \vdots \\ a_{1,m} & \cdots & a_{p,m} \end{bmatrix} + \begin{bmatrix} e_{1,1} & \cdots & e_{1,p} \\ \vdots & \ddots & \vdots \\ e_{n,1} & \cdots & e_{n,p} \end{bmatrix}$$

改めて  $[F] = \begin{bmatrix} f_{1,1} & \cdots & f_{1,m} \\ \vdots & \ddots & \vdots \\ f_{n,1} & \cdots & f_{n,m} \end{bmatrix} = [\{f_1\} \cdots \{f_m\}]$  を因子スコア行列、

$[A] = \begin{bmatrix} a_{1,1} & \cdots & a_{1,m} \\ \vdots & \ddots & \vdots \\ a_{p,1} & \cdots & a_{p,m} \end{bmatrix}$  を因子負荷量行列 と呼ぶ。

また誤差については、

$$\begin{bmatrix} e_{1,1} & \cdots & e_{1,p} \\ \vdots & \ddots & \vdots \\ e_{n,1} & \cdots & e_{n,p} \end{bmatrix} \quad \text{であるが、} j = 1 \sim p, \quad \begin{Bmatrix} e_{1,j} \\ \vdots \\ e_{n,j} \end{Bmatrix} \quad \text{を平均} 0、\text{標準偏差} d_j \quad \text{として}$$

一般性を失うことなく、 $\begin{Bmatrix} e_{1,j} \\ \vdots \\ e_{n,j} \end{Bmatrix} = d_j \begin{Bmatrix} u_{1,j} \\ \vdots \\ u_{n,j} \end{Bmatrix}$  と表現でき、以下のようなになる。

$$[E] = \begin{bmatrix} e_{1,1} & \cdots & e_{1,p} \\ \vdots & \ddots & \vdots \\ e_{n,1} & \cdots & e_{n,p} \end{bmatrix} = [U][D] = \begin{bmatrix} u_{1,1} & \cdots & u_{1,p} \\ \vdots & \ddots & \vdots \\ u_{n,1} & \cdots & u_{n,p} \end{bmatrix} \begin{bmatrix} d_1 & 0 & \cdots & 0 \\ 0 & d_2 & \vdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & d_p \end{bmatrix}$$

ここで  $[U] = \begin{bmatrix} u_{1,1} & \cdots & u_{1,p} \\ \vdots & \ddots & \vdots \\ u_{n,1} & \cdots & u_{n,p} \end{bmatrix} = [\{u_1\} \quad \cdots \quad \{u_p\}]$  を独自因子スコア行列、

$$[D] = \begin{bmatrix} d_1 & 0 & \cdots & 0 \\ 0 & d_2 & \vdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & d_p \end{bmatrix} \quad \text{を重み行列} \quad \text{と呼ぶ。}$$

因子分析モデルは以下のように書きなおせる。

$$[Z] = [F][A]^t + [U][D]$$

ここで  $[F]$  と  $[U]$  について以下を前提とする。

- (a) 因子スコア  $\{f_k\} (k = 1 \sim m)$  は標準化されている。  
すなわち 平均 0, 標準偏差 1。
- (b) 独自因子スコア  $\{u_j\} (j = 1 \sim p)$  は標準化されている。  
すなわち 平均 0, 標準偏差 1。
- (c) 因子スコア  $\{f_k\}$  同志は独立。つまり  
 $k_1, k_2 = 1 \sim m$  について  $\{f_{k_1}\}^t \{f_{k_2}\} = \delta_{k_1, k_2}$
- (d) 独自因子スコア  $\{u_j\} (j = 1 \sim p)$  同志は独立。つまり  
 $j_1, j_2 = 1 \sim p$  について  $\{u_{j_1}\}^t \{u_{j_2}\} = \delta_{j_1, j_2}$
- (e) 因子スコア  $\{f_k\} (k = 1 \sim m)$  と独自因子スコア  $\{u_j\} (j = 1 \sim p)$   
同志は独立。つまり  
 $k = 1 \sim m, j = 1 \sim p$  について  $\{f_k\}^t \{u_j\} = 0$

$[Z]$  は標準化されているので、 $[Z]^t [Z]$  は相関行列となる。

以下の手順で[A]、[F] を求める。

- 主因子法を用いて因子負荷量行列[A]を求める。
- バリマックス用いて因子負荷量行列を回転させる。
- [Z]、[A]を用いて、[F]を計算する。

#### (1) 因子負荷量行列[A]を求める

$$\begin{aligned}
 [Z]^t[Z](\equiv [\Sigma]) &= ([F][A]^t + [U][D])^t([F][A]^t + [U][D]) \\
 &= ([A][F]^t + [D]^t[U]^t)([F][A]^t + [U][D]) \\
 &= [A][F]^t[F][A]^t + && \text{(c)より}[F]^t[F] \text{ は } m \times m \text{ の単位行列} \\
 &\quad [A][F]^t[U][D] + && \text{(e)より}[F]^t[U] \text{ は } m \times p \text{ の } 0 \text{ 行列} \\
 &\quad [D]^t[U]^t[F][A]^t + && \text{(e)より}[U]^t[F] \text{ は } p \times m \text{ の } 0 \text{ 行列} \\
 &\quad [D]^t[U]^t[U][D] && \text{(d)より}[U]^t[U] \text{ は } p \times p \text{ の単位行列} \\
 &= [A][A]^t + [D]^t[D] \\
 &= [A][A]^t + \begin{bmatrix} d_1^2 & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & d_p^2 \end{bmatrix}
 \end{aligned}$$

既知は  $[Z]^t[Z]$ のみで、 $[D]$ は未知。

ここでは 収束計算により  $[A]$  を求める。

1) 最初[D]に適切な初期値 $[D^{(0)}]$ を設定する。

相関行列  $[R] = [Z]^t[Z]$  の逆行列  $[R]^{-1}$  とする。

$$[R]^{-1} = \begin{bmatrix} r^{1,1} & \cdots & r^{1,p} \\ \vdots & \ddots & \vdots \\ r^{p,1} & \cdots & r^{p,p} \end{bmatrix}$$

以下のように設定。

$$[D^{(0)}] = \begin{bmatrix} 1/r^{1,1} & 0 & \cdots & 0 \\ 0 & 1/r^{2,2} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 1/r^{p,p} \end{bmatrix}$$

$L = 1, 2, \dots$  として以下を繰り返す。

2) 行列  $[\Sigma] - [D^{(L-1)}]$  の固有値と固有ベクトルを  $m \leq p$  個 計算する。

固有値を  $\{\lambda_1 \dots \lambda_m\}$ 、固有ベクトルを  $\{b_1\} \dots \{b_m\}$  とすると、

$$[\Sigma] - [D^{(L-1)}] \cong \sum_{k=1}^m \lambda_k \{b_k\} \{b_k\}^t$$

となり、 $[A^{(L)}] = [\sqrt{\lambda_1} \{b_1\} \dots \sqrt{\lambda_m} \{b_m\}]$  とすれば、

$$[\Sigma] - [D^{(L-1)}] \cong [A^{(L)}][A^{(L)}]^t \quad \text{となる。}$$

3) 行列  $[\Sigma] - [A^{(L)}][A^{(L)}]^t$  の対角要素を改めて  $[D^{(L)}]$  として、

$[D^{(L)}]$  が  $[D^{(L-1)}]$  に近ければ収束したとみなし、

$[A^{(L)}]$  を  $[A]$  の解とする。

そうでなければ、2) に戻る。

なお、収束した段階での、 $\begin{bmatrix} 1 & \dots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \dots & 1 \end{bmatrix} - [D]$  の対角成分を共通性と呼び、  
( $p \times p$  の単位行列)

$h_j$  ( $j = 1 \sim p$ ) として表示される。

## (2) バリマックス用いて因子負荷量行列を回転させる

因子負荷量が計算された後、各因子の解釈を容易にするため、因子軸の回転を行なう。ここでは 直交回転によるバリマックス法を用いる。

$$\text{回転前の因子負荷量行列を } [A] = \begin{bmatrix} a_{1,1} & \dots & a_{1,m} \\ \vdots & \ddots & \vdots \\ a_{p,1} & \dots & a_{p,m} \end{bmatrix}$$

$$\text{回転行列を } [T] = \begin{bmatrix} t_{1,1} & \dots & t_{1,m} \\ \vdots & \ddots & \vdots \\ t_{m,1} & \dots & t_{m,m} \end{bmatrix}$$

$$\text{回転後の因子負荷量行列を } [B] = \begin{bmatrix} b_{1,1} & \dots & b_{1,m} \\ \vdots & \ddots & \vdots \\ b_{p,1} & \dots & b_{p,m} \end{bmatrix} \quad \text{とすると、}$$

$$\begin{bmatrix} b_{1,1} & \cdots & b_{1,m} \\ \vdots & \ddots & \vdots \\ b_{p,1} & \cdots & b_{p,m} \end{bmatrix} = \begin{bmatrix} a_{1,1} & \cdots & a_{1,m} \\ \vdots & \ddots & \vdots \\ a_{p,1} & \cdots & a_{p,m} \end{bmatrix} \begin{bmatrix} t_{1,1} & \cdots & t_{1,m} \\ \vdots & \ddots & \vdots \\ t_{m,1} & \cdots & t_{m,m} \end{bmatrix} \text{ となる。}$$

回転後の  $[B]$  の第  $k$  列の  $b_{j,k}^2$  の分散は

$$\sigma_k^2 = \frac{1}{p} \left\{ \sum_{j=1}^p (b_{j,k}^2)^2 - \frac{1}{p} \left( \sum_{j=1}^p b_{j,k}^2 \right)^2 \right\} \text{ となる。}$$

これを  $m$  列全体について合計したものが最大化したものが、 $b_{j,k}^2$  を大きいものと小さいものに分極させることに等しい。が、実際には

$$V = \sum_{k=1}^m \left\{ p \sum_{j=1}^p \left( \frac{b_{j,k}}{h_j} \right)^4 - \left( \sum_{j=1}^p \frac{b_{j,k}^2}{h_j^2} \right)^2 \right\}$$

を最大となるよう回転を逐次行う。(  $h_j$  ( $j = 1 \sim p$ ) は前項で求めた共通性)  $m$  個の因子の中から 2 つの軸 ( $k_1, k_2$ ) を選び、上記  $V$  を最大にする回転角

$\varphi_{k_1, k_2}$  を以下で求める。

$$x_j = \frac{a_{j,k_1}}{h_j}, \quad y_j = \frac{a_{j,k_2}}{h_j} \quad (j = 1 \sim p)$$

$$A = \sum_{j=1}^p (x_j^2 - y_j^2)$$

$$B = 2 \sum_{j=1}^p x_j y_j$$

$$C = \sum_{j=1}^p \left\{ (x_j^2 - y_j^2)^2 - 4x_j^2 y_j^2 \right\}$$

$$D = 4 \sum_{j=1}^p x_j y_j (x_j^2 - y_j^2)$$

$$\tan(4\varphi_{k_1, k_2}) = \frac{D - 2AB/p}{C - (A^2 - B^2)/p}$$

$$(D - 2AB/p) \sin(4\varphi_{k_1, k_2}) > 0$$

回転行列  $T$  は 以下のように単位行列のうち、 $k_1$  行、 $k_2$  行、 $k_1$  列、 $k_2$  列、について回転成分を施した行列となる。

$$\begin{bmatrix} t_{1,1} & \cdots & t_{1,m} \\ \vdots & \ddots & \vdots \\ t_{m,1} & \cdots & t_{m,m} \end{bmatrix} = \begin{matrix} & & k_1 \text{ 列} & & k_2 \text{ 列} & & \\ \begin{bmatrix} 1 & 0 & \cdots & \cdots & 0 & \cdots & 0 \\ \vdots & \ddots & \cdots & \cdots & \cdots & \cdots & \vdots \\ \vdots & \vdots & \cos(\varphi_{k_1, k_2}) & \cdots & -\sin(\varphi_{k_1, k_2}) & \cdots & \vdots \\ \vdots & \vdots & \vdots & \cdots & \vdots & \cdots & \vdots \\ \vdots & \vdots & \sin(\varphi_{k_1, k_2}) & \cdots & \cos(\varphi_{k_1, k_2}) & \cdots & \vdots \\ \vdots & \vdots & \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \cdots & \cdots & \cdots & 1 \end{bmatrix} \end{matrix}$$

上記の変換を  $m$  個の因子の組み合わせの数  $m(m-1)/2$  だけ逐次行い、 $V$  が収束した時点で終了する。

(3)  $[Z]$ 、 $[A]$ を用いて、因子スコア $[F]$ を計算する

$$[Z] = [F][A]^t + [U][D]$$

から $[F]$ を求めることになる。 $[A]$ と $[D]$ とは(1)で求まっているが、 $[U]$ は求められないので、仕方なくここでは、

$$[U][D] = \begin{bmatrix} 0 & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & 0 \end{bmatrix} \text{ とする。}$$

ここで  $[A]$  はm個までの因子を求めたものであるが、

p 個まで求めた因子行列を  $[A_{(p)}]$ 、対応する因子スコアを $[F_{(p)}]$ とする。

すると

$$[Z] = [F_{(p)}][A_{(p)}]^t \quad \text{なので} \quad [F_{(p)}] = [Z] \left( [A_{(p)}]^t \right)^{-1}$$

$$\text{一方} \quad [R] = [Z]^t [Z] = [A_{(p)}][A_{(p)}]^t$$

$$[R]^{-1} = \left( [A_{(p)}][A_{(p)}]^t \right)^{-1} = \left( [A_{(p)}]^t \right)^{-1} [A_{(p)}]^{-1}$$

$$\left( [A_{(p)}]^t \right)^{-1} = [R]^{-1} [A_{(p)}] \quad \text{となるから、}$$

$$[F_{(p)}] = [Z][R]^{-1}[A_{(p)}]$$

実際には因子はm個までを求めているので、

$$[F] = [\{f_1\} \quad \cdots \quad \{f_m\}], \quad [A] = [\{a_1\} \quad \cdots \quad \{a_m\}] \quad \text{として}$$

$$[F] = [Z][R]^{-1}[A] \quad \text{となる。}$$

$[Z]$  は標準化された値であり、 $[F]$ 、 $[R]$ 、 $[A]$ もそれに準ずるので、

実際のデータに戻すには、最初の

元のデータ  $\begin{Bmatrix} x_{1,j} \\ \vdots \\ x_{n,j} \end{Bmatrix}$  の平均  $\mu_j$ 、標準偏差  $\sigma_j$  を用いて変換する必要がある。