データ解析

http://coconut.sys.eng.shizuoka.ac.jp/data/

静岡大学工学部
安藤和敏

2006.01.25

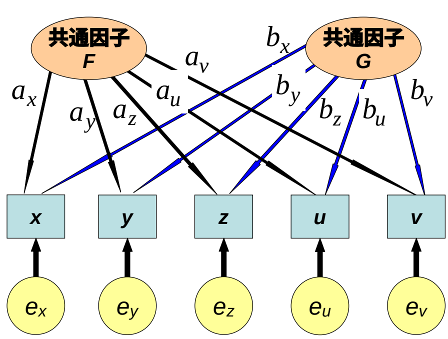
第4章 Excelで学ぶ因子分析

4-5　2因子モデルから学ぶ主因子法

因子分析のデータ (変数が5個の場合)

No	変数 x	変数 y	変数 z	変数 u	変数 v
1	x_1	y_1	z_1	u_1	v_1
2	x_2	y_2	z_2	u_2	v_2
⋮	⋮	⋮	⋮	⋮	⋮
i	x_i	y_i	z_i	u_i	v_i
⋮	⋮	⋮	⋮	⋮	⋮
n	x_n	y_n	z_n	u_n	v_n

因子分析法のパス図 (2因子5変数)



因子分析のモデル (5変数2因子)

$$x = a_x F + b_x G + e_x$$
$$y = a_y F + b_y G + e_y$$
$$z = a_z F + b_z G + e_z$$
$$u = a_u F + b_u G + e_u$$
$$v = a_v F + b_v G + e_v$$

仮定1

1因子ときの同様に,「共通因子と独自因子は,互いに無相関(⇒p.19)である」と仮定する.

$$s_{Fe_x}=s_{Fe_y}=s_{Fe_z}=s_{Fe_u}=s_{Fe_v}=s_{FG}=0,$$
$$s_{Ge_x}=s_{Ge_y}=s_{Ge_z}=s_{Ge_u}=s_{Ge_v}=0,$$
$$s_{e_xe_y}=s_{e_xe_z}=s_{e_xe_u}=s_{e_xe_v}=0,$$
$$s_{e_ye_z}=s_{e_ye_u}=s_{e_ye_v}=0,$$
$$s_{e_ze_u}=s_{e_ze_v}=0$$
$$s_{e_ue_v}=0\cdots\cdots(2)$$

仮定2

1因子のときと同様に,x, y, z, u, v, F, G は標準化されていると仮定する.

$$s_F^2=s_G^2=s_x^2=s_y^2=s_z^2=s_u^2=s_v^2=1,$$
$$s_{xy}=r_{xy},s_{xz}=r_{xz},s_{xu}=r_{xu},s_{xv}=r_{xv},$$
$$s_{yz}=r_{yz},s_{yu}=r_{yu},s_{yv}=r_{yv},$$
$$s_{zu}=r_{zu},s_{zv}=r_{zv},$$
$$s_{uv}=r_{uv}$$

基本方程式の導出(1)

1

0

$$s_x^2=a_x^2s_F^2+b_x^2s_G^2+s_{e_x}^2$$
$$+2a_xs_{Fe_x}+2b_xs_{Ge_x}+2a_xb_xs_{FG}$$

$$\therefore 1=a_x^2+b_x^2+s_{e_x}^2$$

基本方程式の導出(2)

1

0

$$r_{xy}$$
$$\parallel$$
$$s_{xy}=a_xa_ys_F^2+a_xb_ys_{FG}+a_xs_{Fe_y}$$
$$+b_xa_ys_{FG}+b_xb_ys_G^2+b_xs_{Ge_y}$$
$$+a_ys_{Fe_x}+b_ys_{Ge_x}+s_{e_xe_y}$$

$$\therefore r_{xy}=a_xa_y+b_xb_y$$

因子分析の基本方程式(5変数2因子の場合)

$$\begin{aligned}
1 &= a_x^2+b_x^2+s_{e_x}^2,\\
1 &= a_y^2+b_y^2+s_{e_y}^2,\\
1 &= a_z^2+b_z^2+s_{e_z}^2,\cdots\cdots(3)\\
1 &= a_u^2+b_u^2+s_{e_u}^2,\\
1 &= a_v^2+b_v^2+s_{e_v}^2,
\end{aligned}$$

因子分析の基本方程式(続き)(5変数2因子の場合)

$$\begin{aligned}
r_{xy} &= a_xa_y+b_xb_y, & r_{yu} &= a_ya_u+b_yb_u,\\
r_{xz} &= a_xa_z+b_xb_z, & r_{yv} &= a_ya_v+b_yb_v,\\
r_{xu} &= a_xa_u+b_xb_u, & r_{zu} &= a_za_u+b_zb_u,\\
r_{xv} &= a_xa_v+b_xb_v, & r_{zv} &= a_za_v+b_zb_v,\\
r_{yz} &= a_ya_z+b_yb_z, & r_{uv} &= a_ua_v+b_ub_v.
\end{aligned}$$

因子分析の基本方程式

未知変数の数が15であるのに対して，方程式の数は15個である．

1因子モデルのときと同様に行列で表現して解く．

因子分析の基本方程式

$$\begin{bmatrix} 1 & r_{xy} & r_{xz} & r_{xu} & r_{xv} \\ r_{xy} & 1 & r_{yz} & r_{yu} & r_{yv} \\ r_{xz} & r_{yz} & 1 & r_{zu} & r_{zv} \\ r_{xu} & r_{yu} & r_{zu} & 1 & r_{uv} \\ r_{xv} & r_{yv} & r_{zv} & r_{uv} & 1 \end{bmatrix} =$$

基本方程式(3)は行列を用いて，以下のように表すことができる．

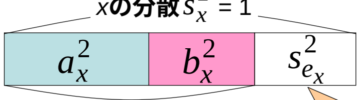
$$\begin{bmatrix} a_x^2+b_x^2+s_{e_x}^2 & a_xa_y+b_xb_y & a_xa_z+b_xb_z & a_xa_u+b_xb_u & a_xa_v+b_xb_v \\ a_xa_y+b_xb_y & a_y^2+b_y^2+s_{e_y}^2 & a_ya_z+b_yb_z & a_ya_u+b_yb_u & a_ya_v+b_yb_v \\ a_xa_z+b_xb_z & a_ya_z+b_yb_z & a_z^2+b_z^2+s_{e_z}^2 & a_za_u+b_zb_u & a_za_v+b_zb_v \\ a_xa_u+b_xb_u & a_ya_u+b_yb_u & a_za_u+b_zb_u & a_u^2+b_u^2+s_{e_u}^2 & a_ua_v+b_ub_v \\ a_xa_v+b_xb_v & a_ya_v+b_yb_v & a_za_v+b_zb_v & a_ua_v+b_ub_v & a_v^2+b_v^2+s_{e_v}^2 \end{bmatrix}$$

……(4)

共通性

$$a_x^2+b_x^2+s_{e_x}^2=1$$

xの分散$s_x^2=1$



h_x^2

独自因子の分散

xのもつ情報量のうちで共通因子が説明する情報の量
xの**共通性**と呼ばれれる．

$$h_x^2 = a_x^2+b_x^2 = 1-s_{e_x}^2$$

共通性

xの共通性 $h_x^2=a_x^2+b_x^2=1-s_{e_x}^2$

yの共通性 $h_y^2=a_y^2+b_y^2=1-s_{e_y}^2$

zの共通性 $h_z^2=a_z^2+b_z^2=1-s_{e_z}^2$

uの共通性 $h_u^2=a_u^2+b_u^2=1-s_{e_u}^2$

vの共通性 $h_v^2=a_v^2+b_v^2=1-s_{e_v}^2$

共通性の推定

$$\begin{bmatrix} 1-s_{e_x}^2 & r_{xy} & r_{xz} & r_{xu} & r_{xv} \\ r_{xy} & 1-s_{e_y}^2 & r_{yz} & r_{yu} & r_{yv} \\ r_{xz} & r_{yz} & 1-s_{e_z}^2 & r_{zu} & r_{zv} \\ r_{xu} & r_{yu} & r_{zu} & 1-s_{e_u}^2 & r_{uv} \\ r_{xv} & r_{yv} & r_{zv} & r_{uv} & 1-s_{e_v}^2 \end{bmatrix} =$$

左辺の対角成分を，適当な数 $h_x^2,h_y^2,h_z^2,h_u^2,h_v^2$ で推定する．

$$\begin{bmatrix} a_x^2+b_x^2+s_{e_x}^2 & a_xa_y+b_xb_y & a_xa_z+b_xb_z & a_xa_u+b_xb_u & a_xa_v+b_xb_v \\ a_xa_y+b_xb_y & a_y^2+b_y^2+s_{e_y}^2 & a_ya_z+b_yb_z & a_ya_u+b_yb_u & a_ya_v+b_yb_v \\ a_xa_z+b_xb_z & a_ya_z+b_yb_z & a_z^2+b_z^2+s_{e_z}^2 & a_za_u+b_zb_u & a_za_v+b_zb_v \\ a_xa_u+b_xb_u & a_ya_u+b_yb_u & a_za_u+b_zb_u & a_u^2+b_u^2+s_{e_u}^2 & a_ua_v+b_ub_v \\ a_xa_v+b_xb_v & a_ya_v+b_yb_v & a_za_v+b_zb_v & a_ua_v+b_ub_v & a_v^2+b_v^2+s_{e_v}^2 \end{bmatrix}$$

共通性の推定

$$\begin{bmatrix} h_x^2 & r_{xy} & r_{xz} & r_{xu} & r_{xv} \\ r_{xy} & h_y^2 & r_{yz} & r_{yu} & r_{yv} \\ r_{xz} & r_{yz} & h_z^2 & r_{zu} & r_{zv} \\ r_{xu} & r_{yu} & r_{zu} & h_u^2 & r_{uv} \\ r_{xv} & r_{yv} & r_{zv} & r_{uv} & h_v^2 \end{bmatrix} =$$

左辺の対角成分を，適当な数 $h_x^2,h_y^2,h_z^2,h_u^2,h_v^2$ で推定する．

$$\begin{bmatrix} a_x^2+b_x^2 & a_xa_y+b_xb_y & a_xa_z+b_xb_z & a_xa_u+b_xb_u & a_xa_v+b_xb_v \\ a_xa_y+b_xb_y & a_y^2+b_y^2 & a_ya_z+b_yb_z & a_ya_u+b_yb_u & a_ya_v+b_yb_v \\ a_xa_z+b_xb_z & a_ya_z+b_yb_z & a_z^2+b_z^2 & a_za_u+b_zb_u & a_za_v+b_zb_v \\ a_xa_u+b_xb_u & a_ya_u+b_yb_u & a_za_u+b_zb_u & a_u^2+b_u^2 & a_ua_v+b_ub_v \\ a_xa_v+b_xb_v & a_ya_v+b_yb_v & a_za_v+b_zb_v & a_ua_v+b_ub_v & a_v^2+b_v^2 \end{bmatrix}$$

……(6)

因子決定行列

$$R_F = \begin{bmatrix} h_x^2 & r_{xy} & r_{xz} & r_{xu} & r_{xv} \\ r_{xy} & h_y^2 & r_{yz} & r_{yu} & r_{yv} \\ r_{xz} & r_{yz} & h_z^2 & r_{zu} & r_{zv} \\ r_{xu} & r_{yu} & r_{zu} & h_u^2 & r_{uv} \\ r_{xv} & r_{yv} & r_{zv} & r_{uv} & h_v^2 \end{bmatrix}$$

を**因子決定行列**と呼ぶ．

対称行列のスペクトル分解（先週と同じ）

R_Fの固有値を$\lambda_1 > \lambda_2 > \lambda_3 > \lambda_4 > \lambda_5$とし，$w_1$, w_2, w_3, w_4, w_5をそれぞれの固有値に属する固有ベクトルで**長さが1のもの**とする．すると，R_F は以下のように書ける．（⇒付録Hを見よ．）

$$R_F = \lambda_1 w_1 {}^t w_1 + \lambda_2 w_2 {}^t w_2 + \lambda_3 w_3 {}^t w_3 + \lambda_4 w_4 {}^t w_4 + \lambda_5 w_5 {}^t w_5$$

（ここで，${}^t w_1$はw_1の転置を表す．）

因子決定行列の近似

$$R_F = \lambda_1 w_1 {}^t w_1 + \lambda_2 w_2 {}^t w_2 + \lambda_3 w_3 {}^t w_3 + \lambda_4 w_4 {}^t w_4 + \lambda_5 w_5 {}^t w_5$$

仮に，λ_1とλ_2が他の固有値に比べて十分大きいとすれば，上式の第3項以降を無視して，

$$R_F \quad \lambda_1 w_1 {}^t w_1 + \lambda_2 w_2 {}^t w_2$$

と書ける．

因子決定行列の近似

$$w_1 = \begin{bmatrix} w_{1x} \\ w_{1y} \\ w_{1z} \\ w_{1u} \\ w_{1v} \end{bmatrix}, w_2 = \begin{bmatrix} w_{2x} \\ w_{2y} \\ w_{2z} \\ w_{2u} \\ w_{2v} \end{bmatrix},$$

のとき，直前の近似式は以下の(7)のように書ける．

$$R_F \quad \lambda_1 \begin{bmatrix} w_{1x} \\ w_{1y} \\ w_{1z} \\ w_{1u} \\ w_{1v} \end{bmatrix} [w_{1x} \ w_{1y} \ w_{1z} \ w_{1u} \ w_{1v}] + \lambda_2 \begin{bmatrix} w_{2x} \\ w_{2y} \\ w_{2z} \\ w_{2u} \\ w_{2v} \end{bmatrix} [w_{2x} \ w_{2y} \ w_{2z} \ w_{2u} \ w_{2v}]$$

$$\cdots\cdots(7)$$

(6)式の右辺

(6)式の右辺は，以下のよう書けることに注意しよう．

$$\begin{bmatrix} a_x^2 + b_x^2 & a_x a_y + b_x b_y & a_x a_z + b_x b_z & a_x a_u + b_x b_u & a_x a_v + b_x b_v \\ a_x a_y + b_x b_y & a_y^2 + b_y^2 & a_y a_z + b_y b_z & a_y a_u + b_y b_u & a_y a_v + b_y b_v \\ a_x a_z + b_x b_z & a_y a_z + b_y b_z & a_z^2 + b_z^2 & a_z a_u + b_z b_u & a_z a_v + b_z b_v \\ a_x a_u + b_x b_u & a_y a_u + b_y b_u & a_z a_u + b_z b_u & a_u^2 + b_u^2 & a_u a_v + b_u b_v \\ a_x a_v + b_x b_v & a_y a_v + b_y b_v & a_z a_v + b_z b_v & a_u a_v + b_u b_v & a_v^2 + b_v^2 \end{bmatrix}$$

$$= \begin{bmatrix} a_x \\ a_y \\ a_z \\ a_u \\ a_v \end{bmatrix} [a_x \quad a_y \quad a_z \quad a_u \quad a_v] + \begin{bmatrix} b_x \\ b_y \\ b_z \\ b_u \\ b_v \end{bmatrix} [b_x \quad b_y \quad b_z \quad b_u \quad b_v] \cdots\cdots(8)$$

主因子法

$$\begin{bmatrix} a_x \\ a_y \\ a_z \\ a_u \\ a_v \end{bmatrix} = \sqrt{\lambda_1} \begin{bmatrix} w_{1x} \\ w_{1y} \\ w_{1z} \\ w_{1u} \\ w_{1v} \end{bmatrix}, \quad \begin{bmatrix} b_x \\ b_y \\ b_z \\ b_u \\ b_v \end{bmatrix} = \sqrt{\lambda_2} \begin{bmatrix} w_{2x} \\ w_{2y} \\ w_{2z} \\ w_{2u} \\ w_{2v} \end{bmatrix}$$

とすれば，

主因子法

$$\begin{bmatrix}a_x\\a_y\\a_z\\a_u\\a_v\end{bmatrix}\begin{bmatrix}a_x & a_y & a_z & a_u & a_v\end{bmatrix}+\begin{bmatrix}b_x\\b_y\\b_z\\b_u\\b_v\end{bmatrix}\begin{bmatrix}b_x & b_y & b_z & b_u & b_v\end{bmatrix}$$

$$=\lambda_1\begin{bmatrix}w_{1x}\\w_{1y}\\w_{1z}\\w_{1u}\\w_{1v}\end{bmatrix}\begin{bmatrix}w_{1x} & w_{1y} & w_{1z} & w_{1u} & w_{1v}\end{bmatrix}+\lambda_2\begin{bmatrix}w_{2x}\\w_{2y}\\w_{2z}\\w_{2u}\\w_{2v}\end{bmatrix}\begin{bmatrix}w_{2x} & w_{2y} & w_{2z} & w_{2u} & w_{2v}\end{bmatrix}$$

R_F

を得る．以上が**主因子法**と呼ばれる手法である．

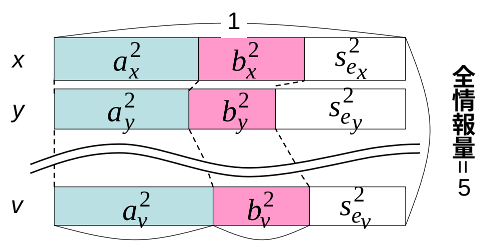
因子負荷行列

$$A=\begin{bmatrix}a_x & b_x\\a_y & b_y\\a_z & b_z\\a_u & b_u\\a_v & b_v\end{bmatrix}$$ を**因子負荷行列**と呼ぶ．

因子負荷行列を使うと(6)の右辺は以下のように書ける．

$$(6)\text{の右辺}=A^tA$$

寄与率と総共通性



因子Fの情報量 $= a_x^2+a_y^2+a_z^2+a_u^2+a_v^2$

因子Gの情報量 $= b_x^2+b_y^2+b_z^2+b_u^2+b_v^2$

寄与率と総共通性

因子Fの情報量 $= a_x^2+a_y^2+a_z^2+a_u^2+a_v^2$

因子Gの情報量 $= b_x^2+b_y^2+b_z^2+b_u^2+b_v^2$

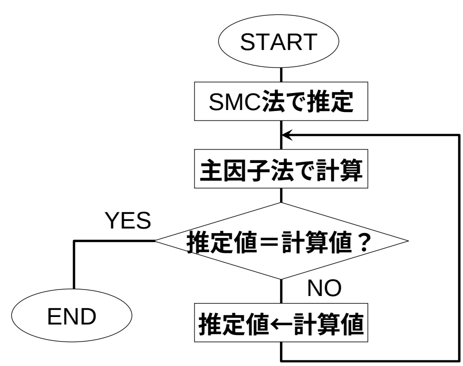
総共通性 $h^2 = (a_x^2+a_y^2+a_z^2+a_u^2+a_v^2)$
$+(b_x^2+b_y^2+b_z^2+b_u^2+b_v^2)$

因子Fの寄与率C_F $= \dfrac{a_x^2+a_y^2+a_z^2+a_u^2+a_v^2}{5}$

因子全体の寄与率 $= \dfrac{\text{総共通性}}{5}$

4-6　2因子モデルから学ぶ反復主因子法

反復主因子法



Excelで学ぼう

ファイル：第4章/4_5, 4_6

本日のまとめ

- 因子分析のモデル（2因子の場合）を理解した．
- 因子分析の基本方程式の導出（2因子の場合）を理解した．
- 寄与率と総共通性の意味を理解した．
- 反復主因子法によって，因子負荷量を求める方法を理解した．
- 反復主因子法による因子負荷量を，Excelを用いて，計算する方法を理解した．