因子分析モデルにおけるいくつかの性質

大学入試センター 研究開発部　柳　井　晴　夫

因子分析モデル（母数モデル，変量モデル）における基本性質については，1940年代から今日までの過去半世紀にわたり多数の文献がある．丘本（1986），柳井 他（1990）においていくつかの基本性質が整理されているが，必ずしも十分でない面がある．本稿では筆者自身の関心といった点から因子分析の基本性質を5つの観点から整理し，さらに未解決の問題点の所在を明らかにしたい．なお，本稿では，n 個体，p 変数，m 因子モデルを考察し，次の2つの表記を前提とする．

(1)　$x=\Lambda f+\varepsilon$（x は p 個の変数に関する p 次元ベクトル）
(2)　$x_j=F\lambda_j+\varepsilon_j$（$x_j$ は変数 j の n 個の測定値を成分とする n 次元ベクトル）

1．母数モデルにおける性質

母数モデルの2つの性質を示す．

性質 1.1.（柳井 他（1990），性質 2.11）母数モデルによって推定される共通性の推定値は母数 Λ によって計算される共通性の不偏推定値ではない．

性質 1.2.（柳井 他（1990），性質 2.9）$F\lambda_j$ の最小二乗推定量を t_j，$e_j=x_j-t_j$ とおくと，

(i)　$(t_i,\ e_j)=0$，　$(x_i,\ e_j)=(e_i,\ e_j)=0$（$i\neq j$）
(ii)　$E(t_ie'_j)=O$，　$E(x_ie'_j)=E(e_ie'_j)=O$，　$E(e_je'_j)=\psi_jQ_F$（ψ_j は変数 j の独自性，$Q_F=I_n-P_F$）．

2．変量モデルにおける因子得点 f の推定（$\Psi>O$，$\Sigma>O$ を仮定しない場合）

母分散共分散行列 Σ が正則でない場合の因子得点 f の推定を行う．

性質 2.1.　変量モデルにおける f の最小二乗推定量（回帰推定量）は

$$\hat{f}=Px=E(fx')(E(xx'))^-x=\Lambda'\Sigma^-x$$

で表わされる．

性質 2.2.　$E(\hat{f})=(\Lambda'\Sigma^-\Lambda)f$ および $V(\hat{f})=\Lambda'\Sigma_r^-\Lambda$，ただし，$\Sigma_r^-$ は Σ の反射型一般逆行列である．

性質 2.3.　f の最小分散不偏推定量（バートレット推定量）は，Λ が column-full rank の場合

$$f_B=(\Lambda'T^{-1}\Lambda)^{-1}\Lambda'T^{-1}x.$$

ただし，$T=\Lambda U\Lambda'+\Psi$，$\mathrm{rank}(\Lambda,\ \Psi)=\mathrm{rank}(T)$ で与えられる．

3．識別可能性に関する Anderson Rubin の条件とその拡張

Σ が与えられた場合，識別可能な構造（Λ，Ψ）を求める問題に関して，Anderson and Rubin の条件はよく知られている．

性質 3.1. (Anderson and Rubin (1956)) $(p\times m)$ 因子負荷量 Λ 行列から任意の1行を取除いたとき2つの正則行列が存在すれば，構造 $(\Lambda,\ \Psi)$ は識別可能である．

この条件は Λ を

$$\Lambda=\begin{pmatrix}\Lambda_1\\ \Lambda_2\\ (\lambda_3)'\end{pmatrix}$$

と分解したとき，Λ_1，および Λ_2 が正則行列になる，すなわち，階数 m の正方行列となることを意味する．

ここで，上記の性質を拡張する．次の性質が成立する．

性質 3.2. (Kano (1989)) 上記の Λ_1, Λ_2 は階数 m であれば，正方行列である必要はない．すなわち，Λ_1, Λ_2 は column-full rank であればよい．

性質 3.3. Λ_1, Λ_2, および λ_3 が

$$\text{i)}\ \ \mathrm{rank}(\Lambda_1'\Lambda_2)=\mathrm{rank}(\Lambda_2), \qquad \text{ii)}\ \ \lambda_3\in S(\Lambda_2')$$

を満たす場合(ただし，$S(A)$ は A の列ベクトルによる部分空間)，λ_3 に対応する変数の独自性 ψ_3 が共分散行列 Σ の要素で表わされる．

証明. 補助定理 (Yanai (1990) の定理 2.1 の一部)

$$\mathrm{rank}(A'M)=\mathrm{rank}(A) \iff A(M'A)^-M'A=A.$$

上記の補助定理を用いて，以下のことが証明される．

$$\Lambda_2'(\Lambda_1'\Lambda_2)^-\Lambda_1\Lambda_2'=\Lambda_2' \Longrightarrow \Lambda_2'(\Lambda_1\Lambda_2')^-\Lambda_1\lambda_3=\lambda_3$$

$$\iff \Lambda_2'\Sigma_{12}^-\Sigma_{13}=\lambda_3 \Longrightarrow \Sigma_{32}\Sigma_{12}^-\Sigma_{13}=\lambda_3'\lambda_3\ \text{(左から}\ \lambda_3\ \text{をかける)}$$

注意 1. しかし，i), ii) の条件を Λ に仮定しない場合，構造 $(\Lambda,\ \Psi)$ は必ずしも識別可能ではない．Ihara and Kano (1986) に従って，$\Sigma=\Lambda\Lambda'+\Psi=\Gamma\Gamma'+V$，という分解を仮定する．$\Gamma$ も Λ と同様に $\Gamma'=(\Gamma_1',\ \Gamma_2',\ \gamma_3)$ と分解する．Ψ，V は対角行列．従って，

$$\Sigma_{12}=\Lambda_1\Lambda_2'=\Gamma_1\Gamma_2', \qquad \Sigma_{13}=\Lambda_1\lambda_3=\Gamma_1\gamma_3, \qquad \Sigma_{23}=\Lambda_2\lambda_3=\Gamma_2\gamma_3.$$

従って，

$$\lambda_3'\lambda_3=\Sigma_{32}\Sigma_{12}^-\Sigma_{13}=\gamma_3'\Gamma_2'(\Gamma_1\Gamma_2')^-\Gamma_1\gamma_3.$$

Γ_1 と Γ_2 が column-full rank であれば $\Gamma_2'(\Gamma_1\Gamma_2')^-\Gamma_1$ は単位行列となり，$\lambda_3'\lambda_3=\gamma_3'\gamma_3$ となり，識別可能となる (性質 3.2 の別証になっている)．また，Γ_1, Γ_2 と λ_3 が i), ii) を満たせば，同様に識別可能となる．性質 3.3 を考慮して，性質 3.2 の拡張が可能に思われるが，今後の検討課題である．

4. 共通性の下限に関する性質

性質 4.1. (Yanai and Ichikawa (1990)) p 次の相関係数行列 Σ ($\Sigma>O$ と仮定) の j 番目に大きい固有値を $\lambda_j(\Sigma)$ とするとき，$(1-\lambda_p(\Sigma))$ は少なくともひとつの変数に対し，SMC の値を上回る共通性の下限を与える．

注意 2. Σ が与えられた場合，SMC を上回る共通性の下限が一般的にいくつあるかについては今後の研究課題である．

性質 4.2. 因子分析モデル（母数モデル）$X=F\Lambda'+E$ において，$X=(X_1, X_2)$, $\Lambda'=(\Lambda_1', \Lambda_2')$ と分割する．このとき，$n^*=n-1$ とおくと

i) $$\hat{\Lambda}\hat{\Lambda}'\geqq(X_2)'X_1(X_1'X_1)^-X_1'(X_2),\ \text{ただし，}\ \hat{\Lambda}=(1/n^*)F'X_2$$

が成立する．$X_2=(x_p)$ のときは，上記の結果は $h_p^2\geqq \mathrm{SMC}_{(p)}$ を意味する．さらに，

ii) $$\Lambda_2\Lambda_2'\geqq\Sigma_{21}\Sigma_{11}^-\Sigma_{12}$$

が成立する．

5. Rank-reducibility の条件

Σ を p 次の相関係数行列として，$\Sigma=\Lambda\Lambda'+\Psi$ を満たす構造 (Λ, Ψ) を探す必要がある．Λ, Ψ に条件をいれない場合には，無数の構造 (Λ, Ψ) が得られる．そこで，一般に

i) $(\Sigma-\Psi)$ が非負定符合行列，または，$\Sigma\geqq\Psi$
ii) $m=\mathrm{rank}(\Sigma-\Psi)$ が最小
iii) $\mathrm{tr}(\Sigma-\Psi)$ が最小
iv) $\Psi\geqq O$

になるという条件で構造 (Λ, Ψ) を求める研究が行われている．iii) の条件で構造 (Λ, Ψ) を求める方法は MTFA (minimum trace factor analysis (Bentler (1972))) と呼ばれる．MTFA は improper solution を生じやすいので，iii) に iv) の条件を含めたものを constrained minimum trace factor analysis (CMTFA) と呼んでいる (Woodhouse and Jackson (1977))．さらに，i) および iv) の条件を考慮せずに，ii) の条件で構造 (Λ, Ψ) を求めると，そのとき得られる共通因子数の下限は Ledermann の境界に一致すると推測されている．

性質 5.1. (Shapiro (1982)) A reduced rank of the $p\times p$ covariance matrix is greater than or equal to Ledermann's bound almost surely.

性質 5.2. (Bekker and De Leeuw (1987))

(a) $\rho_{ik}\rho_{jl}=\rho_{il}\rho_{jk}\ (i\neq j,\ k,\ l,\quad j\neq k,\ l,\quad k\neq l)$,　　(b) $\rho_{jk}-\rho_{ik}\rho_{ij}>0$

は，$\Sigma=\lambda\lambda'+\Psi$ が成立するための必要十分条件である．

性質 5.3. (Bekker and De Leeuw (1987)) $(\Sigma-\Psi)$ の最小階数が $m=p-1$ になる場合の必要十分条件は，変数の符号をいれかえるという条件で，Σ^{-1} のすべての成分が正 (strictly positive) になることである．

今後，識別可能性の条件，共通性の推定法についての相互の関連を踏まえながら，もっと整理した議論をする必要があろう．

参 考 文 献

Anderson, T.W. and Rubin, H. (1956). Statistical inference in factor analysis, *Proc. Third Berkeley Symp. on Math. Statist. Prob.*, Vol. 5, 111-150, Univ. of California Press, Berkeley.

Bekker, P.A. and De Leeuw, J. (1987). The rank of reduced dispersion matrices, *Psychometrika*, **52**, 125-135.

Bentler, P.M. (1972). A lower-bound method for the dimension-free measurement of internal consistency, *Social Science Research*, **1**, 343-357.

Ihara, M. and Kano, Y. (1986). A new estimator of the uniqueness in factor analysis, *Psychometrika*, **51**, 563-566.

Kano, Y. (1989). A new estimation procedure using g-inverse matrix in factor analysis, *Math. Japon.*, **34**, 43-52.

丘本　正 (1986).『因子分析の基礎』, 日科技連, 東京.

Shapiro, A. (1982). Rank-reducibility of a symmetric matrix and sampling theory of minimum trace factor analysis, *Psychometrika*, **47**, 187-199.

Woodhouse, B. and Jackson, P.M. (1977). Lower bounds for the reliability of the total score on a test composed of nonhomogeneous items, *Psychometrika*, **42**, 579-591.

Yanai, H. (1990). Some generalized forms of least squares g-inverse, minimum norm g-inverse, and Moore-Penrose inverse matrices, *Comput. Statist. Data Anal.*, **10**, 251-260.

Yanai, H. and Ichikawa, M. (1990). New lower and upper bounds for communality in factor analysis, *Psychometrika*, **55**, 405-409.

柳井晴夫, 繁桝算男, 前川眞一, 市川雅教 (1990).『因子分析 —— その理論と方法 ——』, 朝倉書店, 東京.

潜在変数または顕在変数に関する尺度不変因子分析モデルのある拡張

鉄道総合技術研究所　小笠原　春　彦

1. 問題

独立な集団間での因子構造や因子パターンの不変性または相違の問題は, 心理学においては factorial invariancy の問題として古くから論じられてきた. 因子負荷行列を異なる集団間で不変とするモデルのうち, 因子の分散共分散行列の構造化モデルのひとつは, Harshman の PARAFAC2 モデルである. このモデルは, 因子回転の不定性から自由であるというユニークな性質を持っている. ところで, 観測個体が属する集団が複数個の集団に離散的に分かれるのではなく, 外的変数 (例えば年齢や所得) とともに連続的に変化する事態において, 分散共分散行列が変化するモデルが Ogasawara (1989, 1990) により提案されている. また, その特殊ケースとして PARAFAC2 モデルの連続的に変化する多母集団への拡張が小笠原 (1989) により行われた. ここでは, PARAFAC2 モデルを潜在変数である共通因子のレベルでの尺度不変モデルとしてとらえるとともに, これに対応する顕在変数に関する尺度不変モデルを提案し, 適用例を示す.

2. モデルとモデルにおけるパラメータの推定

小笠原 (1989) は PARAFAC2 モデルのある拡張として次のモデルを提案した. p 個の観察された変量で構成されるベクトルを $\boldsymbol{s}$ であらわし, i 番目のサンプルの値を $\boldsymbol{s}_i$ とすると, $\boldsymbol{s}_i$ は次のように記述される.

$$\boldsymbol{s}_i=\boldsymbol{\mu}_i+\boldsymbol{\Lambda}\boldsymbol{f}_i+\boldsymbol{\varepsilon}_i$$

ここで, $\boldsymbol{\Lambda}$ は $p\times k$ の集団間で不変な因子負荷行列, $\boldsymbol{f}_i$ は長さ k の共通因子のベクトル, $\boldsymbol{\varepsilon}_i$ は長さ p の独自因子のベクトルであり, $E(\boldsymbol{s}_i)=\boldsymbol{\mu}_i$, $E(\boldsymbol{f}_i)=\boldsymbol{0}$, $E(\boldsymbol{\varepsilon}_i)=\boldsymbol{0}$ である. $\boldsymbol{f}_i$ と $\boldsymbol{\varepsilon}_i$ は, $i=1,\ldots,N$ のサンプルについて, 互いに独立に次の多変量正規分布に従うとする.