数理解析研究所講究録
60巻 1968年 1-12

因子分析における数値実験
(収束の一意性について)

芝浦工大　福富和夫

§1. 序

因子分析において因子抽出の方法は種々考え出されて来たがいずれも iteration を用いるものである(Rao [9], Hemmerle[4], Lawley[6] の C.F.A. , Jöreskog [5] の方法, P.F.A. など)。ところで手法として iteration の収束の一意性は当然保証されねばならないものであるがどうであろうか。以下この点について検討を加える為試みた数値実験の結果を述べることにする。先ず実際のデータについて考察する前に解の存在とその構造がはっきりしている model について考えていく

因子分析の model として次のものを考える。

(1) $\underset{\sim}{x} = \underset{\sim}{\mu} + \Lambda \underset{\sim}{f} + \underset{\sim}{\varepsilon}$

こゝで $\underset{\sim}{f}$ は m 変量ベクトルで共通因子、F は $p\times m$ $(p>m)$ の因子行列、$\underset{\sim}{\varepsilon}$ は p 変量ベクトルで specific と error を

合せたもので $\underset{\sim}{f}$ とは無相関である。

$\underset{\sim}{x}$ の分散共分散行列 Σ は

(2) $\qquad \Sigma = \Lambda \Phi \Lambda' + \Psi$

ただし $\Phi = E(\underset{\sim}{f}\underset{\sim}{f}')$ 、$\Psi = E(\underset{\sim}{e}\underset{\sim}{e}')$.

こゝで因子分析では、先ず因子数 m を与え Σ の推定値 $\hat{\Sigma}$ より一定の criterion の下で $\hat{\Psi}$ を求めて因子空間を定め、オ二段階として実際的な意味を持つように $\hat{\Lambda}, \hat{\Phi}$ を定めるのである。こゝではオ一段階についてのみ取扱うので $\Phi = I$ と仮定しておく。

§2. 解とその一意性

Σ が対角行列 Ψ と rank m の行列との和に分解されるとき、Ψ を Σ の m に対する解と呼ぶことにする。§1 で与えた model では Ψ は positive definite $\Sigma - \Psi$ は positive semidefinite でなければならないが、これを満足しないとき improper な解と呼ばれる。（$\Phi = I$ の仮定から Ψ が決れば Λ は右からの直交変換を除いて決るので、その意味で Λ を解と呼ぶこともある。）

次に共通因子についてであるが、§1 で述べたようにオ一段階は因子空間である。従ってある座標軸の取り方によって specific factor になるものは共通因子から除かれるべきである。即ち共通因子行列 Λ はその右からの全ての直交変換に対

しどの列も少くとも2ヶ以上の nonzero 要素を含むも とする。

次に解の一意性については Anderson-Rubin [1] の与えた次の十分条件がある。

> Σ の $m=r$ における解 Λ, F について $FF'=\Lambda\Lambda'$ が成立つための十分条件は、Λ のどの一行を除いても互に disjoint な rank r の行列が2つ含まれることである。（よって $p \geqq 2r+1$ ）。

次に Reiersøl [10] は Σ が $m=r$ で解を持つとき $m=r+1$ で解を持つことを述べているがこれはその解の作り方から Λ に specific を追加したものに過ぎない。$m>r$ での共通因子行列の一意性については次の定理が得られる。

> [定理] Σ が $m=r$ で解を持ち、全ての r 次 off diagonal 行列が正則ならば、Σ は $p-r-1 \geqq m$ では異なる解を持たない。

この定理の証明は次の lemma から明らかである。いま Σ は $m=r$ で解 Λ を持ち、また $m=r'$ $(r<r')$ で $\text{off diag}\,\Sigma = \text{off diag}\,FF'$ でかつ $\text{rank}\,F=r'$ を満たす $F\,(p\times r')$ が存在するとする。

rank $F = r'$ より行を適当に置換えて $F = \begin{bmatrix} F_1 \\ F_2 \end{bmatrix}$, $|F_2| \neq 0$ にできる。こゝで次の lemma が得られる。

〔lemma〕 Σ が $m = r$ で解 Λ を持ち、上の $p \times r'$ 行列 F について $F_1 F_2'$ のどの一行を除いても r 次の正則行列が残るならば、次の直交行列 T が存在する。

$$F T = (\Lambda, S) \quad \text{但し} \quad \text{offdiag}\, S S' = 0$$

(証明)

$F_1 F_2' = \Lambda_1 \Lambda_2'$ より rank $F_1 F_2' =$ rank $\Lambda_1 \Lambda_2' \leqq r$

$|F_2| \neq 0$ より rank $F_1 \leqq r$

F_1 の任意の行を $\underset{\sim}{f}_\alpha'$ とし、これを除いた残りから r 行、また F_2 から r 行を選び $F_{(1)}, F_{(2)}$ で表わすとき $F_{(1)} F_{(2)}'$ で r 次の正則なものがある。

rank $\begin{bmatrix} F_{(1)} \\ \underset{\sim}{f}_\alpha' \end{bmatrix} [\underset{\sim}{f}_\alpha, F_{(2)}'] \leqq r$ より

$$\begin{vmatrix} F_{(1)} \underset{\sim}{f}_\alpha & F_{(1)} F_{(2)}' \\ \underset{\sim}{f}_\alpha' \underset{\sim}{f}_\alpha & \underset{\sim}{f}_\alpha' F_{(2)}' \end{vmatrix} = 0 \quad \therefore \begin{vmatrix} \Lambda_{(1)} \underset{\sim}{\lambda}_\alpha & \Lambda_{(1)} \Lambda_{(2)}' \\ \underset{\sim}{f}_\alpha' \underset{\sim}{f}_\alpha & \underset{\sim}{\lambda}_\alpha' \Lambda_{(2)}' \end{vmatrix} = \begin{vmatrix} \Lambda_{(1)} \underset{\sim}{\lambda}_\alpha & \Lambda_{(1)} \Lambda_{(2)}' \\ \underset{\sim}{\lambda}_\alpha' \underset{\sim}{\lambda}_\alpha & \underset{\sim}{\lambda}_\alpha' \Lambda_{(2)}' \end{vmatrix} = 0$$

$$\therefore \underset{\sim}{f}_\alpha' \underset{\sim}{f}_\alpha |\Lambda_{(1)} \Lambda_{(2)}'| + g(\Lambda) = \underset{\sim}{\lambda}_\alpha' \underset{\sim}{\lambda}_\alpha |\Lambda_{(1)} \Lambda_{(2)}'| + g(\Lambda)$$

$|\Lambda_{(1)} \Lambda_{(2)}'| = |F_{(1)} F_{(2)}'| \neq 0$ より $\underset{\sim}{f}_\alpha' \underset{\sim}{f}_\alpha = \underset{\sim}{\lambda}_\alpha' \underset{\sim}{\lambda}_\alpha$

$\therefore F_1 F_1' = \Lambda_1 \Lambda_1'$

よって $F_1 T = (\Lambda_1, 0)$ なる直交行列 T が存在するのでその T により

$$FT = \begin{pmatrix} \Lambda_1 & 0 \\ A & B \end{pmatrix}$$ とおくと　$F_1 F_2' = \Lambda_1 A' = \Lambda_1 \Lambda_2$

Λ_1 は r 次正則行列を含むので　$A = \Lambda_2$

off diag $F_2 F_2' =$ off diag $(\Lambda_2 \Lambda_2' + BB') =$ off diag $\Lambda_2 \Lambda_2'$　より

off diag $BB' = 0$.　　$S = \begin{bmatrix} 0 \\ B \end{bmatrix}$ とおけば lemma の結果をうる。

上の定理は improper な解についても成立つので $m = r$ で improper な解を持ち、定理の条件を満たすときは $p-r-1 < m$ でなければ proper な解は持たないことになる。

§3. 収束の一意性に関する数値実験

実験は次の三つの場合に分けられる。

〔1〕構造が既知である相関行列 R に関する実験――ここでは Λ を与え、Ψ は $\mathrm{diag}(I - \Lambda\Lambda')$ とした。

〔2〕与えられた Λ, Ψ に正規乱数 f_α, e_α を発生させ $x_\alpha = \Lambda f_\alpha + \Psi^{\frac{1}{2}} e_\alpha$ より作った $\hat{R}$ についての実験。

〔3〕実験のデータについての実験。

実験の結果をまとめて述べると

〔1〕の場合　Λ の構造により更に次のように分けられる。

(A) 構造が一意であるとき収束も一意になることが経験的に確められた。即ち

イ) Λ が Anderson-Rubin の条件を満たし真の因子数 r が与えられたとき、収束は初期値によらず一意で真の Λ が得られる。

ロ) Λ B で m が Tumura-Fukutomi の条件を満たすときは，真の因子行列 Λ と $m-r_\gamma$ の specific（r は真の因子数）が得られ、specific の入る変数及びその loading は初期値により決まる。

(B)構造が一意でないときは当然のことながら収束も一意でない。特に $r \geq \frac{p}{2}$ のとき（この場合一般に構造は一意にならない），$m < \frac{p}{2}$ として推定を行うとしばしば improper な解が得られ、それも初期値により種々の値に収束している。

[2]の場合　このときは[1]とほゞ平行した結果が得られている。即ち R の構造が一意であるか否かにより $\hat{R}$ の場合の収束も大体決定されているがこの場合については更に実験を重ねる必要がある。

[3]の場合　これについては Mattsson, Olsson, Rosén [7] が多くの例で improper な解が得られたことを報告している。ここでは更に improper が得られたケースは全て収束が一意にならないことが確められた。

〔1〕の実験

(A)の D)のケース　　＊印は specific が入った変数の Communality

因子行列			真の COMM.	推定された comm. m=4	m=5	m=5
.8	.		.64	.640	.640	.640
.7	.6		.85	.850	.850	.850
.2	.1	.1	.06	.060	*.573	*.300
.6	.4	.2	.56	.560	*.591	.560
.3	.5	-.4	.50	.500	.500	.500
.8	.5	.2	.93	*.988	.930	.930
.4	.3	-.6	.61	.610	.610	.610
.5	-.2	.7	.78	.780	.780	.780
.6	-.3	-.3	.54	.540	.540	*.799

(B)の例1. 構造が一意でないので収束も一意でない。

因子行列				真の COMM.	m=4	m=4
.7				.49	.490	.490
.4	.6			.52	.520	.520
.3	.7			.58	.580	.580
.4	.7			.65	.650	.650
.6	.3			.45	.450	.450
.3	.3	.4		.34	.331	.297
.4	.2	.5	.4	.61	.623	.544
.2	.4	.6	.3	.65	.647	.736
.3	.4	.5	.2	.54	.542	.647

(B)の例2.　この因子行列はAnderson-Rubinの条件は満しているがTumura-Fukutomiの条件（m>3）は満していない。m=3では収束が一意なったがm=4では一意にならなかった。

因子行列			真のcomm.	m=4	m=4
.7			.49	.490	.490
.4	.7		.65	.650	.650
.6	.6		.72	.720	.720
.4	.5		.41	.410	.410
.7	.2		.53	.530	.530
.5	.4		.41	.410	.410
.2	.2	.8	.72	.606	.518
		.7	.49	.604	.854
		.4	.16	.650	.352

(B)の例3.　この因子行列は $r \geq \frac{p}{2}$ で一意ではない。m=3のとき初期値により三種のimproper解が得られたもの。

因子行列					真のcomm.	m=3	m=3	m=3
.8					.64	*.995	.838	.547
.7	.5				.74	.656	.631	.798
.6	.3	.2			.49	.689	.390	.561
.5	.5	.4	.3		.75	.680	.733	.663
.4	.4	.3	.4	.2	.61	.818	.768	*.995
.3	.5	.2	.2	.5	.63	.563	*.995	.539
.2	.6	.4	.3	.2	.69	.665	.678	.717

〔3〕の実験　数字はimproper解が得られたときのcommu ty の値を示す。(1)はMattsson et. の結果、(2)は別の初期値で得られた値。

Maxwell's example ($p=10$, m = 4)

(1)	.615	.377	.699	.362	.653	.222	.714	*.995	.310	.400
(2)	.564	.333	.725	.349	.879	*.995	.682	.339	.323	.379

Rao's example ($p=9$, m= 2)

(1)	*.995	.656	.227	.185	.467	.813	.704	.441	.701
(2)	.673	.665	.248	.974	.455	.803	.701	*.995	.699

Harman's example 1 ($p=13$, m= 5)

(1)	.541	.217	.410	.384	*.995	.754	.763	.526	.701	.670
	.831	.628	.592							
(2)	.578	.648	.322	.388	.682	.705	.728	.539	.723	.614
	*.995	.643	.574							

Harman's example 2 ($p=8$, m= 3)

(1)	.873	*.995	.807	.843	.910	.641	.589	.510
(2)	.830	.893	.836	.800	.841	.692	*.995	.476

Bechtoldt's example (sample 1)($p=17$, m = 6)

(1)	.326	*.995	.830	.865	.772	.605	.684	.433	.635	.763
	.695	.538	.860	.516	.731	.641	.535			
(2)	*.995	.254	.840	.869	.770	.620	.670	.439	.630	.770
	.681	.559	.803	.510	.729	.641	.539			

こゝで用いた推定の方法はJöreskog-Lawleyの方法である。iterationの過程で$\hat{\Psi}$の要素に.005に達したものが出たとき解をimproperと呼び，収束の一意性を見るときはその変数を除去して条件付分散共分散行列を作ってさらに収束させた。(下の例)

iterationの収束は尤度函数のgradientの全ての要素が.00005以下のとき打切るようにしている。使用した計算機はIBM 1620である。

〔1〕の実験　この因子行列では第三変数のcommunalityが1を越している所謂Heywood caseになっている。しかしこの場合も一意性の条件を満しているので収束は一意になった。

因子行列			真のcomm.	m=3
.9			.81	.829
.8	.2		.68	.674
.8	.3	.6	1.09	1.000
.6	.4	.4	.68	.715
.6	.5	.3	.70	.701
.5	.6	.3	.70	.701
.4	.7	.4	.81	.810
.3	.7	.4	.74	.741
.2	.8	.5	.93	.928

References

[1] Anderson, T.W. & Rubin, H. (1956): "Statistical inference in factor analysis," Proceedings of 3rd Berkeley Symposium, V, 111-150.

[2] Bechtoldt, H.P. (1961): "An empirical study of the factor analysis stability hypothesis," Psychometrika, 26, 405-432.

[3] Harman, H.H. (1960): Modern Factor Analysis, Univ. of Chichago Press.

[4] Hemmerle, W. J. (1965): "Obtaining maximum likelihood estimates of factor loadings and communalities using an easily implemented iterative procedure," Psychometrika, 30, 291~302

[5] Jöreskog, K. G. (1966): "Some contributions to maximum likelihood factor analysis," Psychometrika, 32, 443-482.

[6] Lawley, D. N. & Maxwell, A.E. (1963): Factor Analysis as a Statistical Method, London, Butterworths.

[7] Mattsson, A., Olsson, U. & Rosén, M. (1966): "The maximum likelihood method in factor analysis with special consideration to the problem of improper solutions," Research Report, Univ. of Uppsala.

[8] Maxwell, E.A. (1961): "Recent trend in factor analysis," J.R.S.S. Series A. 124.

[9] Rao, C. R. (1955): "Estimation and tests of significance in factor analysis," Psychometrika, 20, 93-111.

[10] Reiersøl, C. (1950): "On the identifiability of parameters in Thurstone's multiple factor analysis," Psychometrika, 15, 121-149.