綜　合　報　告

主成分分析法および因子分析法について

竹　内　　啓*
益　田　隆　司**

序　はしがき

因子分析法および主成分分析法は多変量解析の１分野に属するものである。一般に多変量解析は大きく分けて２つの分野に分けることができると考えられる。１つは，１或いは２変量の拡張として考えられるものであって，１変数の t 統計量を拡張したホテリングの一般化された T^2 統計量，或いは多変量分散分析等がこの典型的なものである。今一つは，多くの変量が存在する場合にそれらの変量をそのままの形では取り扱わないで，そこから何らかの意味で全体の変動を代表するような少数の変量を導き出し，その少数の変量により解析を進めて行く分野である。

１つの観測対象について多数の観測値が存在するものを多変量データと呼ぶが，現実問題としてそのような場合は非常に多い。したがって，多変量解析の手法は広く用いられるべきものであるが，統計の他の分野にくらべて，現在のところ，多変量解析の手法はあまり利用されていないように思われる。その原因はいろいろ考えられるが，まず第一の原因は，多変量解析には面倒な計算が必然的に伴うということであろう。また，他の原因は，ある手法によって多変量データを解析した場合，得られる結果の意味が１変量或いは２変量の場合にくらべて必ずしも明確ではないことが多いということである。特に上に述べた第二の分野ではそのように思われる。しかしがら，近年の電子計算機の発展は，この第一の困難性を完全に打破するものであり，多変量データの解析にしばしばあらわれる次元数の多い逆行列の計算或いは固有値問題等も，数分或いは数十分の order で解くことが可能になった。これに伴い，上記の第二の困難な点も次第に解決して行くものと思われる。したがって，多変量解析の分野には未知の問題も多いが，今後は種々の分野で広く利用されることになるものと考えられる。

この報告では，多変量解析法の第２の分野の中心的なテーマの一つである(主)成分分析法因子分析法について概説する。この二つは互いに密接な関連があるが，簡単に規定すれば，前者は記述的，後者は推測統計的な手法であるということができるであろう。そこで内容を部に分け，前篇(主)成分分析法を竹内が，後者因子分析法を益田が分担して執筆した。しかし原稿全体について二人が共同に責任を負うものであることはいうまでもない。なおこの稿の執筆にあたって，数年来共同で勉強をして来たＰＳＧ統計東京グループの人々に負うところが大きかったことを感謝する次第である。

東京大学*　日立製作所**　1966年4月5日受理　「経営科学」第9巻第3号

前　篇（主）成　分　分　析　法

1. 基　本　概　念

（主）成分分析 principal component analysis にはいろいろな解釈乃至説明法があるようであるが，ここでは，それを多数の観測値を少数の指標にまとめるような手法という観点から考えよう。

$$X_{ij} \quad i=1\cdots\cdots p \quad j=1\cdots\cdots n$$

を n 個の対象についての，p 種の観測値を表わす組のデータとする。このときこの p 個の観測値を代表するような指標を求めることと考えよう。

いま　　$T_j \quad j=1\cdots\cdots n$

を n 個の対象について与えられた指標の値であるとしよう。

X_{ij} が連続量であるとすれば，X_{ij} と T_j との関連の程度，或いは X_{ij} の変動が T_j によって代表される程度は，X_{ij} と T_j の相関係数の2乗

$$r^2(X_i, T)=\left\{\frac{\sum(X_{ij}-\bar{X}_i)(T_j-\bar{T})}{\sqrt{\sum(X_{ij}-\bar{X}_i)^2\sum(T_j-\bar{T})^2}}\right\}^2$$

$$\bar{X}_i=\frac{1}{n}\sum_j X_{ij} \qquad \bar{T}=\frac{1}{n}\sum_j T_j$$

によって表わすことができるであろう。

いま　$X_{ij}'=(X_{ij}-\bar{X}_i)\sqrt{\sum\ (X_{ij}-\bar{X}_i)^2}$

とおきかえれば，$\sum X_{ij}'=0$，$\sum X_{ij}'^2=1$ となり，

$$r^2(X_i, T)=r^2(\bar{X}_j', T)=\frac{\{X_{ij}'(T_j-\bar{T})\}^2}{\sum(T_i-\bar{T})^2}$$

と表わされる。

そこで，p 種の観測値について相関係数の2乗の和と

$$Q=\sum_{i=1}^{p}r^2(X_i, T)=\frac{\sum_i\{\sum_j X_{ij}'(T_j-\bar{T})\}^2}{\sum_j(T_j-\bar{T})^2}$$

$$=\sum_j\sum_k\{\sum_i X_{ij}'X_{ik}'\}(T_j-\bar{T})(T_k-\bar{T})\Big/\sum_j(T_j-\bar{T})^2$$

を最大にするように T_j を定めよう。

$$\sum X_{ij}'X_{i'k}'=l_{jk},\quad T_j-\bar{T}=T_j',\quad \sum T_j'=0$$

とおくと，

$$Q=\frac{\sum\sum l_{jk}T_j'T_k'}{\sum T_j'^2}$$

となる。これを最大にするには，明きらかに T_j' を，行列 $L=\{l_{jk}\}$ の最大固有根 λ_1 に対応す

る固有ベクトルの成分に比例するように定めればよい。ところで

$$\sum_k l_{jk}=0,\quad j=1\cdots\cdots n$$

だから L は固有根 0 を持ち，かつそれに対応する固有ベクトルは $(1,1\cdots\cdots1)'$ となる。従って 0 でない固有根に対応する固有ベクトルの成分については確かに

$$\sum_j T_j'=0$$

が成り立っている。

この T_j' を X_{ij} の第1成分（或いは第1主成分）と呼ぶ。またこのとき $\lambda_1=Q$ は，T_j' とと相関係数の乗の和を表わす。

ところで

$$\sum_k l_{jk}T_k'=\lambda_1 T_j',\quad j=1\cdots\cdots n$$

$$\sum_i\sum_k X_{ij}'X_{ij}'T_k'=\lambda_1 T_j'$$

より

$$\sum_j\sum_i\sum_k X_{hj}'X_{ij}'X_{ik}'T_k'=\lambda_1\sum X_{hj}'T_j',\quad h=1\cdots\cdots p$$

を得るから，$S_i=\sum_k X_{ik}'T_k',\ i=1\cdots\cdots p$

$$\sum_j X_{hj}'X_{ij}'=\sum(X_{hj}-X_h)(X_{ij}-X_i)\Big/\sqrt{\sum(X_{hj}-X_h)^2\sum(X_{ij}-X_i)^2}$$

$$=r_{hi}$$

とおくと

$$\sum_k r_{hi}S_i=\lambda_1 S_h,\quad h=1\cdots\cdots p$$

となる。すなわち λ_1 は X_{ij} の相関行列 $R=\{r_{hi}\}$ の固有根，S_i はそれに対応する固有ベクトルの成分になっている。

また逆に λ を R の任意の固有根として，

$$\sum r_{hi}S_i'=\lambda S_h'$$

として，$T_j''=\sum_i X_{ij}'S_j'$ とおくと，λ がやはり L の固有根になることがわかる。従って λ_1 は R の固有根の中でも最大のものになっていなければならない，すなわち S_i は R の最大固有根に対応する固有ベクトルの成分になる。

またこのとき

$$\sum_i X_{ij}'S_i=\sum\sum_k X_{ij}'X_{ik}'T_k'=\sum l_{jk}T_k'=\lambda_1 T_j'$$

となるから

$$T_j'=\lambda_1'\sum_i X_{ij}'S_j \qquad [1]$$

となる。すなわち T_j' は S_i を係数として，p 種の観測値の一次結合の形に表わされる。

ふつうには主成分は最初から〔1〕のような形に表わされることが多い。しかし上にのべたとこ

ろから明きらかなように，主成分は最初から観測値の一次式として表わされるものと決めておく必要はないのである。

ところで1個の指標のみで，X_{ij} の変動を十分よく代表することができないならば，2個以上の指標を計算しなければならないであろう。一般に r 個（$r<p$）の指標

$$T_{sj} \qquad s=1\cdots\cdots r,\ j=1\cdots\cdots n$$

を用いて X_{ij} の変動を表わすことと考えよう。個々の観測値について，その変動が T_{sj} によって代表される程度を，X_i と $T_1\cdots\cdots T_r$ との重相関係数の2乗によって表わし，その p 種のデータについての和を最大にするように T_{sj} を定めることを考えよう。

その場合，$T_{sj}'=T_{sj}-\bar{T}_s,\ \bar{T}_s=\dfrac{1}{n}\sum_j T_{sj}$

とするとき，

$$\sum_j T_{sj}'^2=1,\ \sum_j T_{sj}'T_{tj}'=0 \qquad 〔2〕$$

と仮定しても一般性を失わない。そうして最大にすべき量は

$$Q=\sum_i R^2(X_{ij}T_1\cdots\cdots T_r)=\sum_i\sum_s(\sum X_{ij}'T_{sj}')^2$$

と表わされる。

〔2〕の各式に対応するラグランシュ乗数と $\mu_{st}|s=1,\cdots\cdots r\ t=1\cdots\cdots r,\ \mu_{ts}=\mu_{st})$として，

$$\frac{\partial}{\partial T_{sj}'}\{Q-\sum_s\sum_t\sum_j\mu_{st}(T_{sj}'T_{tj}'-\delta_{st})\}=0$$

$$\delta_{st}=1,\ s=t,\ =0,\ s\neq t$$

とすると，

$$\sum_i\sum_k X_{ij}'X_{ikk}'T_{sk}'-\sum_t\mu_{st}T_{tj}'=0 \qquad s=1\cdots\cdots r$$

$$j=1\cdots\cdots n$$

$$\sum l_{jk}T_{sk}'=\sum\mu_{st}T_{tj}' \qquad 〔3〕$$

行列 $M=\{\mu_{st}\}$ は対称だから，適当な直交行列 $P=\{p_{st}\}$ を用いて

$$PMP'=\left\{\begin{matrix}\lambda_1 & & 0\\ & \ddots & \\ 0 & & \lambda_r\end{matrix}\right\}=\Lambda$$

と表わすことができる。T_{sk}' を要素とする行列を T とすると，〔3〕より

$$TL=MT$$

となるから，$PTL=(PMP')PT=\Lambda PT$ となる

したがって，$T_{sj}''=\sum_t p_{st}T_{tj}'$ とおけば，

$$\sum_k l_{jk}T_{sk}''=\lambda_s T_{sj}'' \qquad s=1\cdots\cdots r$$

$$j=1\cdots\cdots n$$

を得る。すなわち，T_{sj}'' はそれぞれ行列 L の固有ベクトルになっていなければならない。

またこのとき〔2〕より $TT'=I$ となるから

$$Q=\sum\sum\sum\sum X_{ij}'X_{ik}'T_{sj}'T_{sj}'$$

$$=hT\dot{L}T'=hMTT'=hM=\sum_s \lambda_s$$

従って，L の最大 r 個の固有根と $\lambda_1\cdots\cdots\lambda_r$ とし，それに対応する固有ベクトルの成分を T_{sj}'' とすれば，Q と最大にすることができる。またここで最初の T_{sj}' を任意に直交変換したものは，X_{ij} を説明するという点では同値であるから，T_{sj}'' を T_{sj}' としておいても一般性を失わない。すなわち T_{sj}' として L の固有ベクトルの成分をとればよい。このとき明きらかに

$$\sum_j T_{sj}'=0$$

$$\sum_j T_{sj}'T_{tj}'=0 \qquad s\neq t$$

がなり立つ。また $\sum T_{sj}'^2=1$ となるように定めることができることは明きらかである。

ここで再び先にのべたのと同様にして，$\lambda_1\cdots\cdots\lambda_r$ は，実は行列 $R=\{r_{hi}\}$ の最大 r 個の固有根に一致し，またそれらに対応する R の固有 r ベクトルの成分を S_{fi} とすると

$$T_{sj}'=\sum X_{hj}'S_{sh} \qquad s=1\cdots\cdots r,\ j=1\cdots\cdots n$$

となることが示される。

ここで，証明の途中で示したように，r 個の主成分 T_{sj}' と，それらを直交変換したものとは，全く同等であり，また実は条件〔2〕も便宜的に導入したものにすぎないことを考えれば，一般にそれらを正則一次変換したものは，すべて同等であると考えることができる。より抽象的にいえば，定められるのは個々の成分の値ではなく，p 次元空間内の一つの r 次元超平面であり，個の対称は，その上の個の点として表わされることになる。そうしてその超平面上の点と表示する座標が主成分の値となるが，その座標軸は任意にとってよいということになるのである。

従って逆に座標軸は全く任意に定めてよいことになる。すなわちいわゆる座標軸の変換は任意に行ってもよいということになる。

§2. 実用上の問題点

ところで前節にのべたような成分分析の方法を実際問題に適用するについては，いろいろな問題が生ずる。それは二つに分けて考えることができよう。第一は結果の解釈，至乃利用の問題であり，第二は，データに対する適用の可能性，或いは手法の妥当性の問題である。

結果の解釈という点については，得られた結果の具体的な意味が明確でないという苦情がしばしば聞かれる。この点については，二つの問題がある。一つは手法の意味に対する誤解が少くないということである。例えば主成分を計算するのに，本質的に異る単位を持つ数量を足したり引いたりすることは不合理であるというような疑問が唱えられることも少くない。しかしこのようなことをいう人も，例えば回帰式

$$Y=\beta_0+\beta_1x_1+\cdots\cdots+\beta_px_p+u$$

において説明変数 $x_1\cdots\cdots x_p$ が本質的に異る単位を持つものであっても，上の式の右辺のような和を作ることには異議を唱えないのは，不思議なことである。このようないわば単純な誤解は，手法の普及につれて，自づから解消して行くであろうが，しかし同時第二の問題点として主成分

というものは，本来何か実体的な量を表わすものではなく，あくまで多数のデータを少数のデータに縮約した代表値或いは指標にすぎないという観点を強調する必要があるように思われる。そうしてその点からすれば，重要なのは成分 T_j の値であって，係数 S_i ではないということになる。係数の大きさから成分の意味を考えることも勿論重要でないとはいえないが， T_j の動き自体の分析が重視されねばならないであろう。

第二に手法の妥当性に関連して，いくつの問題が考えられるが，その中で，データの比重の問題，線形性の問題，および時系列に適用する際の問題の３つについてのべよう。

§3. データの比重の問題

主成分を定めるとき，p 種のデータはすべて同じ比重で扱かわれていた。しかし現実にはこのことは疑問である。すなわち p 種のデータの間には，重要性の違いその他の理由によって，当然量比重を以って扱かわれるべきものが存在するかもしれない。そのような場合には，例えば一つの指標を求めるのに，単に相関係数の乗和でなく，適当な比重 w_i を用いて

$$Q_w=\sum_i w_i{}^2 r^2(X_i, T)$$

を最大にするように T_j を定めることが考えられる。そのためには T_j を行列 $L_w=\{\sum_i w_i{}^2 X_{ij}' X_{ik}'\}$ の最大固有根に対応する固有ベクトルに等しくすればよい。またこのとき

$$S_i=\sum_i w_i X_{ij}' T_j'$$

とおくと，S_i は行列 $R_w=\{w_i w_j r_{ij}\}$ の固有ベクトルの成分になる。そうしてまた

$$T_j' \propto \sum w_i X_{ij}' S_i$$

となることが示される。従って実際には，R_wの固有ベクトルから S_i を求め，次に T_j' を X_{ij}' の一次結合として求めればよい。

２つ以上の成分と求めようとする場合についても，全く同様である。

ところで問題は比重 w_i をどのようにして定めるかである。 例えば p 種のデータが本来同質のものであって，同一の単位例えば金額というようなもので表わされているならば，R_w として金額そのもののモーメント行列を用いればよいであろう。すなわち $\sum_j (X_{ij}-\bar{X}_i)(X_{hj}-\bar{X}_h)$ を要素とする行列の固有ベクトルを計算すればよい。

データが同種のものでなく，自然な共通の単位が存在しない場合には，比重を定める客観的な基準を求めることは困難である。その場合二つの観点が考えられる。すなわち一つはデータの重要性，すなわちある種のデータをよく説明することは重要だが，他のものについてはそれほど重要でないという場合である。例えば洋服の仕立ての場合，ある部分の長さは正確に合わさなければならないが，他の部分についてはそれほど正確である必要はないとであろう。このような場合人間の体型を表わすデータの中で，重要性には差異が生ずるから，人間の体型を表わす指標を構成するについても，データの重要性の違いを反映させねばならない。

第二の観点は，データの中での系統的な部分と偶然的な変動部分との割合を考慮することであ

る。もし指標は系統的な部分と反映すべきものであると考えるならば，偶然的な変動部分を多くふくむデータの比重は小さく偶然的な部分を少くふくむデータの比重は大きくすべきであろう。より精確にいえば，次のようになる。X_{ij} を

$$X_{ij}=Z_{ij}+V_{ij}$$

を表わし，V_{ij} は偶然的な変動部分，Z_{ij} は系統的な部分を表わすとし，$E(V_{ij})=0$, $E(V_{ij}^2)=\sigma_i^2$ とする。このとき T_j を求めるべき指標とするとき，最小２乗法的な考え方と用いれば,

$$Q_e'=\sum_i\frac{\sum_j(X_{ij}-\hat{X}_{ij})^2}{\sigma_i^2}$$

$$\hat{X}_{ij}=a_i+b_iT_j$$

を最小にするように T_j を定めるのが妥当であろう，すなわち,

$$Q'=\sum_i\frac{\sum(X_{ij}-\bar{X}_i)(T_j-\bar{T})^2}{\sigma_i^2\sum(T_j-\bar{T}^2)}$$

を最大にすればよい。いいかえれば比重 w_i を

$$w_i=\sum(X_{i},-\bar{X}_i)^2/\sigma_i^2 \qquad 〔4〕$$

に等しくとればよいことになる。２つ以上の指標を求める場合も全く同様である。

しかし一般には分散 σ_i^2 は未知であるから，何等かの形でそれを推定しなければならない。そこで厳密にモデルを設定して統計的推定を行うことは，後の章の主題である。ここでは二三の直観的な方法についてのべよう。まず σ_i^2 の推定量として，X_i の自分自身以外の変数に対する回帰式の残差分散を用いるという考え方がある。例えば X_1 について，$X_2\cdots\cdots X_p$ に対する回帰式の残差は，次のようにして求めることができる。まず X_1 と $X_2\cdots\cdots X_p$ との重相関係数の２乗は，相関行列 R の余因子を $|R_{hi}|$ というように表わすと

$$R^2(X_{1j}X_2\cdots\cdots X_p)=1-\frac{|R|}{|R_{11}|}$$

と表わされる。従って

$$\hat{\sigma}_1^2=\sum(X_{1j}-\bar{X}_1)^2(1-R^2)=\sum(X_{1j}-\bar{X}_1)^2\frac{|R|}{|R_{11}|}$$

すなわち〔　〕に代入すれば

$$w_1=\frac{\sum(X_{1j}-\bar{X}_1)^2}{\hat{\sigma}^2}=\frac{|R_{11}|}{|R|}$$

となる。これは R の逆行列の第(1.1)要素に一致する。同様にして一般に w_i は R^{-1} の第 i 対角要素に一致することになる。すなわち相関行列の逆行列の対角要素を比重とすればよいことになる。これが Jöreskog の方法である。

第二に，X_{ij} が時系列データ等の場合には，適当な階数の階差

$$\Delta X_{ij}=X_{ij+1}-X_{ij}$$

一般に　$$\Delta^kX_{ij}=\Delta^{k-1}X_{i1j+1}-\Delta^{k-1}X_{ij} \qquad k\geqq 2$$

を求め，その分散から σ^2 を計算することが考えられる。系統的な部分の変動はなめらかである

と考えることができるならば，k をある程度大きくすれば $\Delta^k X_{ij}$ はほとんど系統的な部分もふくまないものと考えることができるであろう。

また指標 T_j が与えられれば，逆に

$$\hat{\sigma}_i{}^2=\frac{1}{n}\sum(X_{ij}-\bar{X}_i)^2[1-r^2(X_i,T))]$$

という形で推定することができる。従って T_j と $\hat{\sigma}_i{}^2$ を連立させて求めることもできるであろう。実はこれは後にのべる Lawley の最尤法に一致する。しかしこの方法に関しては，解の一義性および，くり返えし計算の収足性に関して疑問があるので，必ずしも適当な方法であるとはいえない。原則としては，分散の程度はやはり外部的なデータから推定する方が妥当であると思う。

ただしここで，系統的な部分と偶然的な部分の区別は，実際問題においては絶対的なものではないということに注意しなければならない。それはいわば分析の目的によって定められる。経済時系列の分析においてトレンド，或いは比較的長期のサイクルを求めようとする場合には，一定周期以下の循環的な変動は偶然的なものと見なしてしまう方が適当である。

いづれにしても実際問題において比重を厳密に定めることは困難である。しかし他方，もし少数の主成分ですべてのデータの変動が十分よく表わされるならば，比重の相違は得られた指標にはほとんど影響を及ぼさない。極端な場合として，r 個の指標 $T_1\cdots\cdots T_r$ をとると，すべてのデータのこれらに対する重相関係数を1にすることができる。すなわち正確に

$$X_{ij}=a_i+b_{i1}T_{1j}+\cdots\cdots+b_{ir}T_{rj}\qquad \begin{array}{l} i=1\cdots\cdots p \\ j=1\cdots\cdots n \end{array}$$

と表わされるならば，このような $T_1\cdots\cdots T_r$ の値は X_{ij} の比重と無関係に定めることができることは明きらかである。

重相関が1にならなくても，もしすべてのデータの比重を1に等しくしたときの第1主成分を T とするとき，もし

$$Q=\sum_{i=1}^{p}r^2(X_i,T)$$

がほとんど p に近い，すなわち $r^2(X_i,T)$ がすべて1に近いならば，1と異る比重 w_i を用いたときの主成分 T_w は，もし比重の相対的な値が極端に違っていないならば，ほとんど T と同じ変動を示す，いいかえれば，

$$r(T_w,T)\fallingdotseq 1$$

となるであろう。

いくつかの主成分を用いると，重相関がほぼ1になるというような場合にも，事情は同じである。ただその場合，いくつかの固有ベクトルによって与えられる座標軸には若干の変動を生ずるかもしれない。しかし適当に座標変換を行えば，いくつかの指標によって定められる超平面自体はほとんど変化しないことが明きらかとなるであろう。

比重 w_i のうちの若干は0にしてもよいものに考えれば，この問題は，どれだけの観測値をデータの範囲の問題をもふくむものと見なすことができる。そうして p 種のデータの変動が比較的少数の指標によって十分よく表わされる場合，その中の一部 q 種（$a<p$）のデータから主成分を計算しても，得られる指標そのものはほとんど変らない場合も少くない。すなわちデータの範囲を狭くしても，得られる結果にはほとんど影響しない，或いは別の表現を用いれば，データの種類を増しても，ほとんど新たな情報は得られない場合も多いのである。実際もし R の最大固有根 λ_1 が十分 p に近いならば，

$$\sum_i \sum_h r_{ih}^2 = trR^2 = \sum \lambda_i^2 \geqq \lambda_1^2$$

だから，少くとも1つのについて

$$\sum_h r_{i0h}^2 \geqq \lambda_1/p$$

となる。すなわちただ1つのデータ X_{i0} のみを用いても，すべてのデータとの相関係数の乗の和は λ_1^2/p 以上になる。すなわちこれはかなり p に近くなっているであろう。

ところで偶然的な変動部分の大きさは，各データごとに異るのみならず，実は同じデータについても，対象によって異るかもしれない，すなわち

$$E(V_{ij}^2) = \sigma_{ij}^2$$

とするとき，この値は $j=1\cdots\cdots n$ のそれぞれについて異った値を示すかもしれない。実際企業についての経営データのようなものを扱うとき，企業規模によって観測値の大きさには著しい差異が見られるであろう。このような場合には偶然的な変動の大きさも，規模が大きくなれば大きくなると考えるのが自然であろう。

この問題は回帰分析における誤差分散の不均一性の問題と本質的には同じものである。従ってもし分散比がわかっていれば，X_{ij} をそれて割って基準化してから，すでにのべたような方法を適用すればよい。一般に分散比が未知ならば，適当な尺座を用いて例えば企業なら雇用者数，資本額等基準化しなければならない，実際分散比が正確に知られなくても結果に大きく影響することはないから，基準化に用いる尺度についてそれほど厳密な考慮を加える必要はないであろう。しかし全く基準化を行わないならば，例えば大企業の数字のみが強い影響を与えることになって結果に偏りを生ずる危険があると思われる。

§4.　非線形関係の処理

次に線形性の問題について考えよう。すでにのべたように，指標 T が観測値 X_i の一次結合の形で与えられるということは，数学的な展開の結果導びかれたものであって，最初から要求したものではなかった。しかし線形性の仮定は，Xと指標との関連を重相関係数によって表わすということの中にふくまれている。すなわち X を説明するのに T_j の線形式が用いられることが前提とされている。しかしデータの種類によっては，このような仮定があてはまらないと思われるような場合も存在するであろう。そのような場合には，データを適当に変換する，例えば対数をとるなどというような操作を行うことが望ましいかもしれない。好都合な場合には，このよ

うな変換によって前節で指摘した分散不均一性の問題を同時に解決することが可能であるかもしれない。

しかし場合によっては，適切な変換を見出すことが困難なこともあるであろう。そのとき，いわゆる順位相関を用いる方法が考えられる。すなわち X_i と X_h の順位相関係数，Kendall の τ_{ih}，或いは Spearman の ρ_{ih} を計算して，これらから主成分を求めることである。しかしこの場合直接 τ_{ih} や ρ_{ih} の行列の固有ベクトルを計算しても意味がない。この場合いくつかの方法が考えられる。

一つはデータ X_{ij} の背後に，何か多変量正規分布に従う量を仮定することである。すなわち何か未知の単調函数 f によって

$$X_{ij}=f_i(\xi_{ij})$$

と表わされ，かつ ξ_{ij}，$i=1\cdots\cdots p$，$j=1\cdots\cdots n$ は p 次元正規分布の n 個の独立な標本を与えると考える。そうすると X_i と X_h の順位相関係数は ξ_i と ξ_h の順位相関係数に一致するから，τ_{ih} 或いは ρ_{ih} から ξ_i と ξ_h の相関係数と次のような形で推定することができる。

$$\hat{r}_{ih}=\sin\frac{\pi}{2}\tau_{ih} \qquad \text{或いは} \qquad \hat{r}_{ih}'=2\sin\frac{\pi}{6}\rho_{ih}$$

そうして $\hat{r}_{ih}$ 或いは $\hat{r}_{ih}'$ の行列にすでにのべたような手法を適用すれば，ξ_{ij} のデータに主成分分析法と適用したのとほぼ同じ結果が得られるであろう。しかしこの場合，主成分を表わす指標 T の値を計算することはできない。

もっと後の処理に便利な方法は，X_{ij} を正規スコアに変換することである。すなわち各について $X_{i1}\cdots\cdots X_{in}$ を大きさの順に並べ，X_{ij} がこの中で R_{ij} 番目になっているならば，それを単位正規分布からの大きさの標本の R_{ij} 番目の順位統計量の期待値に変換する。その値を V_{ij} としよう。そうして V_{ij} について主成分分析を行う。今度は V_{ij} 自体を与えられたデータと見なせば，これまでにのべたことはすべて適用される。

しかしこのような方法の基礎にある多変量正規性の仮定は必らずしも現実に妥当するとは限らない。また正規スコアから作られる指標の具体的な意味も必ずしも明確でないかもしれない。従ってこのような手法が適用できるのは，人間の性向，或いは能力に関するテストの結果を分析するような場合に限られるであろう。

もっと別の考え方として，X_{ij} の変動を示す指標自体を順序として与えることが考えられる。すなわち $1\cdots\cdots n$ を適当な順序に並べて X_{ij} との順位相関となるべく大きくするようにする。X_{ij} の（$X_{i1}\cdots\cdots X_{in}$ の中での）順位を R_{ij} とし，また指標の順位を T_j とすると，X_i と T の Spearman 順位相関係数は

$$\sum_j R_{ij}T_j$$

の一次函数として表わされる。従って Spearman 順位相関係数の和を最大にするように T_j を定めるとすれば

$$\sum_j(\sum_i R_{ij})T_j$$

を最大にすればよい。すなわち $\sum R_{ij}$ すなわち p 個の順位の和の順に従って順位をつければよいことは明きらかである。ただしここで X_{ij} の中で全体として他のものと大きさの順が逆になっているものについては，順序を逆にしてもよい。すなわち適当に順序を逆転して上のような和を大きくするようにすればよい。

このような操作は Kendall τ を用いても行うことができるが，その計算手続きはより複雑になる。

しかしこのような方式は，2つ以上の指標を構成するのに直接拡張することはできない順位重相関係数というようなものは定義されてはいるが (Kendall "Rank Correlation Methods" 参照) その具体的な意味づけは難しい。また他の基準を用いるとしても，あまり適切な方式が見つからないようである。

§5. 時系列データの場合

次に時系列データへの応用に際しての問題を考えよう。すなわち各について

$$X_{ij},\ \ j=1\cdots\cdots n$$

が時間的な経過を表わしている場合である。

このようなデータに成分分析法を適用する場合，二つの問題がある。一つはトレンド，季節変動等の処理の問題である。例えば経済時系列から，景気循環を表わす指標を求めようとする場合，得られた指標は，なるべく単調なトレンド，或いは季節変動をふくまないものとして与えられることが望ましい。このとき主成分分析を行う前にデータを加工してトレンド，或いは季節変動を除く方法と，生のデータを用いて成分分析を行い，得られた結果からトレンド或いは季節変動を除く方法の二つの接近法が考えられる。勿論ここで分析目的がトレンド或いは季節変動をふくまない指標を構成するという点にあることが明確であり，しかもトレンド或いは季節変動の形が明確に知られているならば，第一の方法が正しいであろう。しかし現実にはトレンドや季節変動の形自体をデータから求めなければならない場合が多い。そのような場合成分分析法を適用する前に，事前の操作をあまり行うと，得られた結果の意味が明確でなくなってしまう恐れがある。このような場合には第二の接近法の方が望ましい。実際ある場合には成分分析の結果，明確な季節変動の形が現われて来るという場合もある。またとくにトレンドと循環変動とは，多くの場合2つの成分（或いはそれ以上）の中に混って現われて来る。そのような場合には適当な座標変換によって，トレンドを現わす成分と循環を表わす成分を分離することができるであろう。またそれだけでは不十分であれば，更にそこに適当な手法を加えて，トレンドを除けばよい，このような事後の操作を用いれば，その操作の影響の程度が明確にされるので，結果の評価には好都合である。

時系列データにともなう第二の問題は，主成分分析法の手法そのものは，データの時系列的性格を全く考慮に入れていないところから生ずる。勿論この方法によって多くのデータの時間的な

変動のパターンを表わすような指標が得られることは事実である。しかしこの手法はデータの時間的な順序関係を全く無視しているために，得られた指標の時間的変化があまりなめらかでない場合が多い。特に第2第3成分にはそのような傾向が見られる。そのために移動平均法その他の手法によって指標の時系列を平滑化する必要がある場合も少くない。またこのことは指標を何等かの意味で“予測”のために用いたり，或いは傾向の変化を観測したりする場合には一つの難点となる。このような点からすれば，多変量時系列データの分析法としては成分分析法は必らずしも十分満足すべきものにはいえないかもしれない。そのためのよりよい手法の開発は今後の課題であるように思われる。

なお時系列データについて

$$X_t,\ X_{t+1}\cdots\cdots,\ X_{t+p+1},\ t=1,\ 2\cdots\cdots$$

を一組の多変量データとみなし成分分析法と適用したらどうなるであろうかを検討してみよう。

まづこのようなデータの相関行列は，(両端の影響を無視すれば)，i 行 j 列に，自己相関係数 $\rho_{|j-i|}$ をならべたものになる。ただし

$$\rho_k=\frac{\sum_{t=1}^{n-k}({}_t-X)(X_{t+k}-X')/(n-{}_k+1)}{\sqrt{\sum_{t=1}^{n-k})X_t-X)^2\sum^{n-k}(X_{t+h}-X')^2/(n-k+1)^2}}$$

である。

ここで更に時系列が循環的に定義されている，すなわち $X_{t+cn}\equiv X_t$, $c=1,2\cdots\cdots$ であるとすれば，自己相関係数は次のように定義することができる。

$$\rho_k'=\frac{\sum_{t==}^{n}(X_t-X)(X_{t+h}-X)}{\sum(X_t-X)^2}$$

そうすると，$\rho'_{n-k}=\rho_h'$ となるから，

$$X_t,\ X_{t+1}\cdots\cdots,\ X_{t+n-1}$$

の相関行列は

$$\begin{pmatrix} 1 & \rho_1 & \rho_2\cdots\cdots\rho_{n-1} \\ \rho_1 & 1 & \rho_1\cdots\cdots\rho_{n-2} \\ \rho_2 & \rho_1 & 1 \\ \vdots & \vdots & \end{pmatrix} \qquad \text{ただし}\quad \begin{aligned}\rho_{n-1}&=\rho_1\\ \rho_{n-2}&=\rho_2\end{aligned}$$

という形になる。これは巡回行列であるから，その固有根は，

$$\omega+\rho_1\omega^2+\cdots\cdots+\rho_{n-1}\omega^{n-1}$$

$$\omega^n=1$$

という形で与えられる。すなわち

$$\lambda_m=\sum_{k=0}^{n-1}e^{ik\theta_m}\rho_k,\quad \theta_m=2m\pi/n,\quad \omega=e^{i\theta}$$

となり，固有ベクトルは $(1.\ \omega,\omega^2\cdots\cdots\omega^{n-1})'$ で与えられる。λ_m は丁度 $\theta=2m\pi/n$ に対応する

"スペクトル"を表わしていこるとになる。そうして最大のスペクトルに対応する成分が，第一主成分を表わすということになる。

ここで時系列が定常であることを仮定した上で，$n\to\infty$ とすれば，スペクトルは一般に

$$\lambda_\theta=\sum_{k=0}^{\infty}e^{ik\theta}\rho_k$$

という形に表わされることになる。

すなわち成分分析とスペクトル分析には相通ずるものがあることが示されたといってよいであろう。そこで多変量データについても，

$$X_{1t}\cdots\cdots,\ X_{1t+k-1},\ X_{2t}\cdots\cdots,\ X_{2t+h-1}\cdots\cdots,\ X_{pt}\cdots\cdots,\ X_{pt-k+1}$$

という次元データに成分分析を行うことによって，p 種のデータに共通な波動の大きさと同期を定めることができるであろう。ただしそのような分析をどのような形で利用することができるかについては，よくわかってはいない。

§6. 若干の応用

次に成分分析法の直接的でない利用法，およびその変形について考えよう。

まづデータ X_{ij}，$i=1\cdots\cdots p$，$j=1\cdots\cdots n$ は p 次元観測値の組の値と考えたが，すでにのべたところから明きらかなように，成分分析において，データの添字は全く対称的に扱かわれている。したがってこれを n 次元データの p 組に見ることもできるし，また $p\times n$ の行列データと見なすことにもできる。行列データとみれば，結局 X_{ij}' について，

$$X_{ij}'=S_{i1}T_{j1}'+\cdots\cdots+S_{ir}T_{jr}'+V_{ij}'$$

として，$\sum\sum V_{ij}^2=\min$ となるように S_i，T_j' を定める問題と考えることができるし勿論ここで X_{ij} を基準化して X_{ij}' を求めるときには $j=1\cdots\cdots n$ がくり返えしを表わすということを用いてはいるが，もとのデータ X_{ij} については

$$X_{ij}=\bar{X}_i+S_{i1}T_{j1}+\cdots\cdots+S_{ir}T_{jr}+V_{ij}$$

と表わすことができる。すなわち $X_{ij}-\bar{X}_i$ を要素とする行列がほとんど階数 r の行列に等しいという仮定の下に，そのような階数 r の行列を求める問題と考えることができる。このような考え方はある種の元分類データを分析するのに用いることができるであろう。

次に成分分析法は，回帰分析，或いは分散分析と結びつけて用いることもできる。例えば X_{ij} について，t_{ij} を説明変数とする適当な線形モデル

$$X_{ij}+\beta_{i0}+\beta_{i1}t_{ij}+V_{ij}$$

が仮定されているとき，この式から計算される残差 $\hat{V}_{ij}=X_{ij}-\hat{\beta}_{i0}-\hat{\beta}_{i1}t_{ij}$ に成分分析法を適用することが考えられる。同じ様な考え方によって，くり返えしのない2元配置の場合残差の行列に主成分分析法を適用することによって交互作用の構造をしらべることができるであろう。

また，p 次元データ自体がいろいろな構造を持っている場合には，成分分析をいろいろな形で適用することができる。例えば2元配置データ

$$X_{ik.j}=\mu_j+\alpha_{i.j}+\beta_{k.j}$$

の場合 $\hat{\alpha}_{i.j}$ の行列，$\hat{\beta}_{k.j}$ の行列，$\hat{u}_{i.k.j}$ の行列のそれぞれについて成分分析を適用することができ，それによって各要因効果，および残差の主成分を別個に求めることができる。このような方法は要因効果がどのような方向に現われるかを分析する場合に興味深い結果を与えるであろう。

また回帰分析において，$X_1\cdots\cdots X_p$ を説明変数 Y を従属変数として，

$$Y_j=\beta_0+\beta_1X_{ij}+\cdots\cdots+\beta_pX_{pj}+U_j,\quad j=1\cdots\cdots n$$

とする。このとき p が大きいと，最小2乗法によって β を推定するとき，いわゆる multicollinearity を生じて係数の推定量の誤差が非常に大きくなってしまう場合が多い。このような場合には，$X_1\cdots\cdots X_p$ からいくつかの主成分 $T_1\cdots\cdots T_r$ を求め，この主成分について Y の回帰式を求めるという方法が提案されている。

いま $X_1\cdots\cdots X_p$ の Y に対する影響の大きさがほぼ同じ程度であるとすれば，$X_1\cdots\cdots X_p$ を基準化して，

$$Y_j-\bar{Y}=\beta_1'X_{1j}'+\cdots\cdots+\beta_p'X_{pj}'+u_j$$

とするとき，係数 $\beta_1'\cdots\cdots\beta_p'$ はほぼ同じ程度の大きさとなるであろう。そうして $X_1\cdots\cdots X_p$ から計算される p 個の主成分をすべて求めて，対応する R の固有値の大きさの順に $T^{1\prime}\cdots\cdots T_p'$ とし，かつ

$$T_{ij}'=a_{i1}X_{ij}'+\cdots\cdots+a_{ip}X_{pj}'$$

とするとき $\sum_{h=1}^{p}b_{ih}^2=1$ となるようにすれば，

$$\sum_{j=1}^{n}T_{ij}'^2=\lambda_i\ (=R\ の\ i\ 番目の固有根)$$

となる。従って

$$Y_j-Y=r_1T_{1j}'+\cdots\cdots+r_pT_{pj}'+u_j$$

とすると，$r_1\cdots\cdots r_p$ はほぼ同じ程度の大きさとなるであろうが，その推定量の分散は

$$V(\hat{r}_j)=\sigma/\lambda_i$$

となるから，小さい λ_i に対応する T_i の係数の推定量の分散は，母数 r_i の値にくらべて非常に大きくなるであろう。このような場合には Y_j の期待値を予測するとき，適当な r をとって

$$\hat{E}(Y_0)'=\hat{r}_1T_1+\cdots\cdots+\hat{r}_rT_r \qquad \text{〔a〕}$$

とした方が，すべての成分を用いて

$$\hat{E}(Y)=\hat{r}_1T_1+\cdots\cdots+\hat{r}_pT_p \qquad \text{〔b〕}$$

とするより誤差が小さくなるであろう，〔a〕は偏りを持ち，〔b〕は不偏であるが，〔b〕は誤差分散が大きくなるから，平均平方誤差

$$E\{\hat{E}(Y)-E(Y)\}^2$$

は〔b〕の方が大きくなるであろう。

§7. 一つの拡張

最後に主成分分析法の一つの拡張を紹介しよう。

これまで指標 $T_1\cdots\cdots T_r$ はすべてデータ $T_1\cdots\cdots T_p$ をすべて用いて計算されるものとなっていた。しかし後はものべるように，或る場合には，指標を計算するために用いるデータと，その指標によって説明されるデータとが，必らずしも一致しなくてもよいとした方が便利である。いま指標を，データ $Y_1\cdots\cdots Y_q$ の一次式として

$$T_j=a_1Y_{1j}+\cdots\cdots+a_qY_{qj},\quad j=1\cdots\cdots n$$

とし，これと $X_1\cdots\cdots X_p$ の相関係数の2乗和

$$Q=\sum_k r^2(X_k,T)$$

を最大にするように T の係数 $a_1\cdots\cdots a_q$ を定めることを考えよう。 いま $Y_1\cdots\cdots Y_q$ は基準化されていて，$\sum_j Y_{ij}=0$, $\sum Y_{ij}^2=1$ となっているものとし，$r_{ih}=\sum_j Y_{ij}Y_{hj}$, $r_{ik}{}^*=\sum_j Y_{ij}X_{kj}{}'$ とすれば，

$$r^2(X_kT)=(\sum_i r_{ik}{}^*a_i)^2/\sum\sum r_{ih}a_ia_h$$

と表わされるから

$$Q=\frac{\sum_i\sum_n(\sum_k r_{ik}{}^*r_{hk}{}^*)a^2a_h}{\sum_i\sum_h r_{ih}a_ia_h}$$

となる。従って r_{ih} を要素とする行列を R_{xy}, $r_{ih}{}^*$ を要素とする行列を R_{yx}, その転置行列を R_{xy} とすれば，a_i を成分とするベクトルを a_1 とするとき，

$$Q=\frac{a_1{}'R_{yx}R_{xy}a}{a_1{}'R_{yy}a_1}$$

に表わされる。従ってこれを最大にするには a_1 を方程式 $|R_{yx}R_{xy}-\lambda R_{yy}|=0$ の最大根 λ_1 に対応して，$R_{yx}R_{xy}a_1=\lambda_1R_{yy}a_1$ を満足するように定めればよい。

ここで $Y_1\cdots\cdots Y_q$ のうちの全部或いは一部は $X_1\cdots\cdots X_p$ の中のどれかに一致してもよい。もし X の組と Y の組とが完全に一致するならば

$$R_{xx}R_{xx}a_1=\lambda_1R_{xx}a_1 \text{ より}$$

$$R_{xx}a_1=\lambda_1a_1$$

を得るから，これはすでにのべた場合に一致する。

ここでもし λ_1 が十分 p に近くないならば $|R_{yx}R_{xy}-\lambda R_{yy}|=0$ の根をいくつかとり，それに対応するいくつかのベクトルを係数とする指標をそれぞれ計算し，いくつかの指標を用いて $X_1\cdots\cdots X_p$ を説明することにすればよい。

すでにのべたように，$X_1\cdots\cdots X_p$ が比較的少数の成分によって十分よく説明される場合，実は $X_1\cdots\cdots X_p$ の中の少数のののから指標と計算しても全体をよく説明できるであろうと考えられることが多い。このような場合には $X_1\cdots\cdots X_p$ の一部を $Y_1\cdots\cdots Y_q$ にみなして上記の手法を適用

することができる。もしそれぞれによって十分よく $X_1\cdots\cdots X_p$ を説明することができるならば，新しい対象について指標を求める場合の観測と計算の手間を省くことができるであろう。

また時系列データの場合には $Y_1\cdots\cdots Y_q$ を $X_1\cdots\cdots X_p$ の何期か前のデータとすることができる。それによって何期か後のデータの動きに最もよく説明するような，すなわち予測向きの指標を作ることができるであろう。

勿論 Y データの一部或いは全部は X データとは全く別種のものであってもよい。全く別のデータならば，それはいわゆる正準相関分析 cannonical correlation analysis に非常に近いものとなるであろう。

いづれにしてもこのような手法の応用範囲はかなり広いように思われる。

後 篇 因 子 分 析 法

§1. は し が き

因子分析法にも，前篇で述べられた主成分分析法と同様に非常に面倒な計算が必要になるが，これまで主として心理学の分野で，重心法等の近似的な解法を用いて利用されて来た。因子分析法，主成分分析法は互いに非常に深い関係にある。一般的に言って，主成分分析法の考え方はかなり直観的であって，最大の変動をあらわす変量（主成分）をいくつか決めて行くものであるのに対し，因子分析法は，それを統計的に定式化したものであり，相関に最も大きく寄与している変量（因子）をいくつか見つけ出そうという考え方に基づいているものということができる。したがって，この両者ははっきりとは区別できるものではないのであって，ある条件のもとでは両者から得られる結果は全く一致することが示される。さらに，この両者の中間的な方法もいくつか存在する。主因容分析法，正準因容分析法等がそれである。

以下，主として因子分析法に関するこれまでの理論の大要を述べる。

§2. 定 式 化

互いにある程度の相関を有する p 個の変量 $x_1, x_2, \cdots x_p$ が存在する場合に，その各々の変量が新しい m 個 ($m<p$) の変量 $f_1, f_2, \cdots f_m$ を用いて，次式のような形にあらわすことができるのではないかということが考えられる。以下，ベクトルは $\boldsymbol{f}$ であらわすことにする。

$$\boldsymbol{x}=\boldsymbol{\mu}+\Lambda\boldsymbol{f}+\boldsymbol{e} \tag{2.1}$$

ここに，

$\boldsymbol{x}=\{x_1, x_2, \cdots x_p\}$: p 次元の列ベクトル

$\boldsymbol{\mu}=\{\mu_1, \mu_2, \cdots, \mu_p\}$: 〃

$\boldsymbol{e}=\{e_1, e_2, \cdots, e_p\}$: 〃

$\boldsymbol{f}=\{f_1, f_2, \cdots, f_m\}$: 〃

$$\Lambda=(\lambda_{ij}):pxm \text{ の行列 } (i=1,2,\cdots,p,\ \ j=1,2,\cdots,m)$$

であるとする。μ は一般平均，e は誤差項をあらわす。この (2.1) が通常の因子分析のモデルである。

モデル (2.1) を分かりやすく説明するために，心理学に於ける知能テストの問題を例にとる。r 種類のテストが存在するとした場合。一般にそれら p 個のテストの点数は互いに独立ではなくて，ある程度の相関を有するのが普通である。そこで各テストの点数は，テスト間に共通に含まれている $m(<p)$ 個の変量（人間のある側面の基本的な能力をあらわす変量）の1次結合としてあらわされるのではないかということが考えられる。各テスト毎に，それら m 個の共通に含まれる変量の重みが異なるのは当然である。モデル (2.1) でこの p 種類のテストの点数をあらわす変量が x，m 個の共通に含まれる変量が f，各テスト毎に f にかかる重みが Λ である。また，μ は x の一般平均，e は誤差項である。厳密には，e は m 個の共通な変量によってはあらわされない各テスト毎に特有な変量と測定誤差の和であるが，実験にくりかえしがない場合には，この両者を区別することはできない。(2.1) のモデルでは e は単に測定誤差をあらわすものとする。通常，f の各成分を共通因子(特に，f が1次元の場合には一般因子)，Λ の各要素を因子負荷量と呼んでいる。

以上より分がるように，(2.1) は共通因子の数が m 個であるという仮説のもとに於けるモデルに於いて，x，f，e の分散共分散行列をそれぞれ Ψ，M，$\sum$ であらわし，モデルを以下のような仮定のもとで考えていくことにする。

1° $\varepsilon e=0$

2° e の各成分と f の各成分が独立

3° $\sum(p\times p)$：正値対角行列

4° $M(m\times m)$：正値行列

5° $\mathrm{rank}\,\Lambda=m$

これらの仮定はいずれも当然の仮定である。たとえば，2°，3°については，テスト間の相関はすべて m 個の共通因子によってあらわされることになるので，共通因子と誤差項が互いに独立になることあるいは誤差項の間に相関がないことは当然認めてもよい仮定である。

モデル (2.1) に対して2通りの場合が考えられる。第1は f の各成分を確率変数と考える場合であり，この場合には容易に分かるように $\varepsilon f=0$ として一般性を失わない。また，この場合には，f，e は正規分布にしたがうと考えるのが普通である。第2は f の各成分が確率変数ではなくて，各個人に特有なパラメータ（α 番目の個人に対しては $f=f_\alpha$）であると考える場合である。このときには，第1の場合に対応して

$$\frac{1}{n}\sum_{\alpha=1}^{n} f_\alpha=0,\quad \frac{1}{n}\sum_{\alpha=1}^{n} f_\alpha f_\alpha'=M \tag{2.2}$$

と仮定する。

さて，以上の議論より分かるように，推定すべき本質的なパラメータは，Λ, M, Σ であるがこれらはモデル (2.1) のもとに正しく推定することができるであろうか。それを調べるために，まずモデル (2.1) の構造について考えて見る。

前に述べたように，$\varepsilon \boldsymbol{e}=0$, $\varepsilon \boldsymbol{ee}'=\Sigma$ と仮定すると，(2.1) およびこれまでの仮定により，

$$\varepsilon \boldsymbol{x}=\boldsymbol{\mu} \tag{2.3}$$

$$\begin{aligned}\varepsilon(\boldsymbol{x}-\boldsymbol{\mu})(\boldsymbol{x}-\boldsymbol{\mu})' &= \varepsilon(\Lambda\boldsymbol{f}+\boldsymbol{e})(\Lambda\boldsymbol{f}+\boldsymbol{e})' \\ &= \varepsilon(\Lambda\boldsymbol{ff}'\Lambda'+\boldsymbol{ef}'\Lambda'+\Lambda\boldsymbol{fe}'+\boldsymbol{ee}') \\ &= \Lambda(\varepsilon\boldsymbol{ff}')\Lambda'+\varepsilon\boldsymbol{ee}' \\ &= \Lambda M\Lambda'+\Sigma \\ &= \Psi\end{aligned} \tag{2.4}$$

であることが分かる。すなわち，$\boldsymbol{x}$ の分散共分散行列 Ψ が2つの部分に分離されることになる。第1の部分 $\Lambda M\Lambda'$ は階数 $\boldsymbol{m}$ の半正値行列であって，共通因子によりあらわされる部分の $\boldsymbol{x}$ の各成分の分散共分散行列をあらわしている。この行列の非対角要素は $\boldsymbol{x}$ の分散共分散行列の非対角要素に等しい。また一般にこの行列の対角要素を共通度と呼んでいる（通常，共通度という言葉は，$\boldsymbol{x}$ の各成分の分散が1に規準化されているときに用いられる場合が多い)。第2の部分 Σは，誤差の分散共分散行列である。

そこで，パラメータ Λ, M, Σ を推定する問題を考える前にまず，$\boldsymbol{x}$ の分散共分散行列を与えた場合に，(2.4) およびこれまでの仮定を満足する Λ, M, Σ の値が得られるか，また解が得られた場合に，その解は一意的なものであるか否かという問題を考察しておかなければならない。解の存在の有無は無論 $\boldsymbol{m}$ の値に依存している。$\boldsymbol{m}$ が大きければ解が存在する可能性は強くなるであろう。特に $\boldsymbol{m}=p-1$ の場合には，解は必ず存在することが容易に示される。また，もしある $\boldsymbol{m}$ の値に対して Λ, M, Σ という解が得られた場合，この解は一意的なものであろうか。いま，任意の $\boldsymbol{m}\times\boldsymbol{m}$ の正列行列を $\boldsymbol{H}$ とすると，次式より定まる Λ^*, M^*, Σ もやはり同じ $\boldsymbol{m}$ の値に対して，前とは異なった解になっていることが分かる。

$$\Lambda^*=\Lambda H^{-1} \tag{2.5}$$

$$M^*=HMH' \tag{2.6}$$

さらにまた，

$$\boldsymbol{f}^*=H\boldsymbol{f} \tag{2.7}$$

とすると，

$$\Lambda^*\boldsymbol{f}^*=\Lambda\boldsymbol{f} \tag{2.8}$$

であり，したがって (2.1) は

$$\boldsymbol{x}=\boldsymbol{\mu}+\Lambda^*\boldsymbol{f}^*+\boldsymbol{e} \tag{2.9}$$

となり，$\boldsymbol{x}$ を観測することにより，(2.1) と (2.9) を区別することができなくなる。これは今まで考えてきたモデルについての基本的な不確定要素であって，このよとなことの一部を無くする

には，$M=I$ ($I: m\times m$ の単位行列) としてやればよい。この場合，因子は直交しているという。また，M が対角行列ではない場合には，因子は斜交しているという。$M=I$ の場合には(2.6) より分るように，$HH'=I$，すなわち，H は直交行列になり，$\Lambda\theta$ (θ：任意の直交行列) の集合だけが Λ と同等の角になる。そこで実際に問題を解く場合には，$\{\Lambda\theta\}$ からある特定の Λ を選びだすために，Λ に関しさらに何らかの制約をつけ加えなければならない。

また，一般に行列 $A(p\times p)$ を与えた場合に，それが BB' ($B: p\times m$) とあらわされるための必要十分条件は，A が階数 m の半正値行列であるという性質を用いて，今の場合の Λ，Σ についての解の存在性に関し，次の定理が得られる。

＜定理＞ $m: \Psi$ を与えた場合に，Λ，Σ についての解が存在するための必要十分条件は，$\Psi-\Sigma$ (Σ：正の対角要素を有する対角行列) が階数 m の半正値行列になるようなΣが存在することである。

この定理を満足するΣが存在した場合に，そのΣが一意的なものであるかどうかは分からない。これに対してこれまで種々の試みがなされて来たが，いくつかの特殊な場合を除いてはまだ一般には解かれていない。

§3. 主因子分析法，正準因子分析法

§2で通常の因子分析のモデルについて述べたが，同じ問題をもう少し異なった立場からとらえて見る。(2.1) で，

$$z=\Lambda f \tag{3.1}$$

すなわち，

$$z_i=\lambda_{i1}f_1+\lambda_{i2}f_2+\cdots\cdots+\lambda_{im}f_m\ (i=1,2,\cdots,p) \tag{3.2}$$

(ここに，z_i は p 次元ベクトル z の成 i 分) とおくと，

$$x=\mu+z+e \tag{3.3}$$

となる。§2と同様，x，e の分散共分散行列を Ψ，Σ であらわす。また，z の分散共分散行列は $\Lambda M\Lambda'$ であるが，これを $\Gamma=(r_{ij})$ ($i,j=1,2,\cdots,p$) であらわす。Γ の階数は m である。すると，(2.4) に対応して

$$\Psi=\Gamma+\Sigma \tag{3.4}$$

$$\text{nondiag } \Psi=\text{nondiag } \Gamma \tag{3.5}$$

が成立する。

我々の目的は Ψ の推定値から Γ の推定値を求めることであるが，(3.4)，(3.5) および前に述べた種々の仮定を満足する任意の Γ を求めることではなく，条件を満足するΓのうち，ある意味で最も簡単な，いいかえればxの間の関係を簡単に表現する行列 Γ を求めることである。

Γ の階数が m であるから，$f_1, f_2, \cdots, f_m$ はいずれも，$z_1, z_2, \cdots, z_p$ の何らかの1次結合であらわすことができる。因子の数は少いほど良いわけであるから，我々の目的は条件を満足するΓのうち，階数が最小になるようなものを見つけ出すことにある。Γ の階数は一定であっても，

$f_1, f_2, \cdots, f_m$ の定め方は一意ではない。そこで，$f_1, f_2, \cdots, f_m$ をどのように定めるかということを定式化する必要がある。それに対応して以下に2通りの考え方を述べる。いずれも主成分分析の考え方にかなり近い考え方である。

（I） 主因子分析

上に述べたように，$f_1, f_2, \cdots, f_m$ は $\boldsymbol{z}$ の各成分 $z_1, z_2, \cdots, z_p$ の1次結合としてあらわすことができる。まず，f_1 を求めることを考えよう。

$$f_1 = \boldsymbol{l}'\boldsymbol{z} = l_1 z_1 + l_2 z_2 + \cdots + l_p z_p \tag{3.6}$$

とあらわすことにする。$\boldsymbol{x}$ の i 番目の成分 x_i の変動のうち，この $\boldsymbol{l}'\boldsymbol{z}$ により説明される変動部分は，

$$\frac{\cos^2(x_i, \boldsymbol{l}'\boldsymbol{z})}{V(\boldsymbol{l}'\boldsymbol{z})} = \frac{(l_1 r_{i1} + l_2 r_{i2} + \cdots + l_p r_{ip})^2}{\boldsymbol{l}'\Gamma\boldsymbol{l}} \tag{3.7}$$

で与えられる。そこで，$x_1, x_2, \cdots, x_p$ すべての変量について，$\boldsymbol{l}'\boldsymbol{z}$ により説明される変動は，

$$\frac{\sum_{i=1}^{p}(l_1 r_{i1} + l_2 r_{i2} + \cdots + l_p r_{ip})^2}{\boldsymbol{l}'\Gamma\boldsymbol{l}} = \frac{\boldsymbol{l}'\Gamma\Gamma\boldsymbol{l}}{\boldsymbol{l}'\Gamma\boldsymbol{l}} \tag{3.8}$$

であらわされることになる。そこで，(3.8) を最大にするようなベクトル $\boldsymbol{l}$ を求めることを求めることを考える。$\boldsymbol{l}'\Gamma\boldsymbol{l}=1$ としても一般性を失わないから，$\boldsymbol{l}'\Gamma\boldsymbol{l}=1$ の条件のもとで，$\boldsymbol{l}'\Gamma\Gamma\boldsymbol{l}$ を最大ならしめる $\boldsymbol{l}$ を求めればよいことになる（$\boldsymbol{l}'\Gamma\boldsymbol{l}=1$ は共通因子の分散が1であるということで§2の場合と対応している)。ラグランジュ乗数を d として，

$$\phi = \boldsymbol{l}'\Gamma\Gamma\boldsymbol{l} - d(\boldsymbol{l}'\Gamma\boldsymbol{l} - 1) \tag{3.9}$$

を $\boldsymbol{l}$ について偏微分して0とおくと，

$$\frac{\partial\phi}{\partial\boldsymbol{l}} = 2\Gamma\Gamma\boldsymbol{l} - 2d\Gamma\boldsymbol{l} = 0$$

$$\Gamma\Gamma\boldsymbol{l} - d\Gamma\boldsymbol{l} = 0 \tag{3.10}$$

あるいは，

$$(\Gamma - dI)\Gamma\boldsymbol{l} = 0 \tag{3.11}$$

すなわち d は Γ の固有根，$\Gamma\boldsymbol{l}$ が対応した固有ベクトルである。また，このとき (3.8) の値は d になっている。容易に確かめられるように，$\boldsymbol{l}$ としては Γ の固有ベクトルそのものを取ればよい。結局，(3.8) の最大値は Γ の最大固有根 d_1 で与えられ，$\boldsymbol{l}$ としては d_1 に対応した固有ベクトルを取ればよいことになる。このようにして第1因子を求めることができる。

次に第2因子は，第1因子に直交して，第1因子で説明されない残りの変動をできるだけ多く説明するように定める。すなわち，

$$f_2 = \boldsymbol{h}'\boldsymbol{z} = h_1 z_1 + h_2 z_2 + \cdots + h_p z_p \tag{3.12}$$

とあらわすことにすれば，

$$\boldsymbol{h}'\boldsymbol{l}=0,\ \boldsymbol{h}'\Gamma\boldsymbol{h}=1 \tag{3.13}$$

の条件のもとで,

$$\frac{\boldsymbol{h}'\Gamma\Gamma\boldsymbol{h}}{\boldsymbol{h}'\Gamma\boldsymbol{h}} \tag{3.14}$$

を最大ならしめるように $\boldsymbol{h}$ を定めればよいことになる。上で $\boldsymbol{l}$ を求めたのと同様にして,ラグランジュ乗数法を用いると,容易に $\boldsymbol{h}$ としては Γ の第2番目の固有根 d_2 に対応した固有ベクトルを取ればよいことが分かる。また,このとき (3.14) の値は d_2 で与えられる。以下同様にして,結局行列 Γ の階数に等しい数だけの因子が得られることになる。

また,第1因子の因子負荷量は次のようにして求めることができる。(3.2) より,

$$\begin{aligned}\cos(z_i, f^1)&=\cos(\lambda_{i1}f_1+\lambda_{i2}f_2+\cdots+\lambda_{im}f_m, f_1)\\&=\lambda_{i1}V(f_1)\\&=\lambda_{i1}\end{aligned}$$

また,(3.6) を用いると,

$$\begin{aligned}\cos(z_i, f_1)&=\cos(z_i, \boldsymbol{l}_1 z_1+\boldsymbol{l}_2 z_2+\cdots+\boldsymbol{l}_p z_p)\\&=\boldsymbol{l}_1 r_{i1}+\boldsymbol{l}_2 r_{i2}+\cdots+\boldsymbol{l}_p r_{ip}\\&=d_1 \boldsymbol{l}_i \end{aligned} \tag{3.16}$$

(3.15), (3.16) より,

$$\lambda_{i1}=d_1 l_i\ \ (i=1, 2, \cdots, p) \tag{3.17}$$

他の因子に対する因子負荷量も全く同様にして求めることができる。また,このとき誤差分散 σ_{ii} は以下の (3.19) を満足している。

$$x_i=\mu_i+\lambda_{i1}f_1+\lambda_{i2}f_2+\cdots+\lambda_{im}f_m+e_i \tag{3.18}$$

より,各共通因子の直交性を用いると,

$$\psi_{ii}=\lambda_{i1}^2+\lambda_{i2}^2+\cdots+\lambda_{im}^2+\sigma_{ii} \tag{3.19}$$

$$=d_1^2 l_i^2+d_2^2 h_i^2+\cdots+d_m^2 k_i^2+\sigma_{ii} \tag{3.20}$$

ここに,h_i, k_i はそれぞれ,d_2, d_m に対応した固有ベクトルの第 i 成分をあらわす。

(II) 正準因子分析

主因子分析の考え方は,x の各成分の変動のうち,できるだけ多くの部分を説明するような因子を求めることであった。同じモデルについて,次のような意味で,$x_1, x_2 \cdots, x_p$ に最も深く結びついている因子を求めることを考える。

$\boldsymbol{l}'\boldsymbol{z}$, $\boldsymbol{q}'\boldsymbol{x}$ をそれぞれ $z_1, z_2, \cdots, z_p$ および $x_1, x_2, \cdots x_p$ の1次結合とするとき,この2つの1次結合の相関係数の2乗は次式で与えられる。

$$\frac{(\boldsymbol{l}'\Gamma\boldsymbol{q})^2}{(\boldsymbol{l}'\Gamma\boldsymbol{l})(\boldsymbol{q}'\Psi\boldsymbol{q})} \tag{3.21}$$

そこで，これを最大にするようなベクトル $\boldsymbol{l}, \boldsymbol{q}$ を求めることを考える。$\boldsymbol{l}'\Gamma\boldsymbol{l}=\boldsymbol{q}'\Psi\boldsymbol{q}=1$ としても一般性を失わないから，この2つの制限条件のもとで，$(\boldsymbol{l}'\Gamma\boldsymbol{q})^2$ を最大にするベクトル $\boldsymbol{l}, \boldsymbol{q}$ を求めればよい。η, θ をラグランジュ乗数として，

$$\phi=(\boldsymbol{l}'\Gamma\boldsymbol{q})^2-\eta(\boldsymbol{l}'\Gamma\boldsymbol{l}-1)-\theta(\boldsymbol{q}'\Psi\boldsymbol{q}-1) \tag{3.22}$$

を $\boldsymbol{l}, \boldsymbol{q}$ について偏微分して0とおくと，

$$\frac{\partial\phi}{\partial\boldsymbol{l}}=2(\boldsymbol{l}'\Gamma\boldsymbol{q})\Gamma\boldsymbol{q}-2\eta\Gamma\boldsymbol{l}=0$$

$$\frac{\partial\phi}{\partial\boldsymbol{q}}=2(\boldsymbol{l}'\Gamma\boldsymbol{q})\Gamma\boldsymbol{l}-2\theta\Psi\boldsymbol{q}=0$$

すなわち，

$$(\boldsymbol{l}'\Gamma\boldsymbol{q})\Gamma\boldsymbol{q}-\eta\Gamma\boldsymbol{l}=0 \tag{3.23}$$

$$(\boldsymbol{l}'\Gamma\boldsymbol{q})\Gamma\boldsymbol{l}-\theta\Psi\boldsymbol{q}=0 \tag{3.24}$$

(3.23)，(3.24) および $\boldsymbol{l}'\Gamma\boldsymbol{l}=\boldsymbol{q}'\Psi\boldsymbol{q}=1$ より，$\eta=\theta=(\boldsymbol{l}'\Gamma\boldsymbol{q})^2$ であり，(3.21) の値が η になっていることが分かる（これより，$0<\eta<1$，$\eta=1$ の場合には，Σ の対角要素のあるものが0になってしまう。(3.23)，(3.24)をすぐ上に述べたことを用いて，次式のようにかきかえることができる。

$$(\Gamma-\eta\Psi)\boldsymbol{l}=0 \tag{3.25}$$

$$(\Gamma-\eta\Psi)\boldsymbol{q}=0 \tag{3.26}$$

あるいは，$\delta=1/(1-\eta)$ とおくと，

$$(\Psi-\delta\Sigma)\boldsymbol{l}=0 \tag{3.27}$$

$$(\Psi-\delta\Sigma)\boldsymbol{q}=0 \tag{3.28}$$

すなわち，$\boldsymbol{l}$ と $\boldsymbol{q}$ は互いに比例関係にある。これより分かるように，第1因子は，(3.27) において最大の固有根 δ_1 に対応した固有ベクトル $\boldsymbol{l}$ から求められる。また，第1因子と直交する第2因子は (3.27) の第2番目に大きい固有根 δ_2 に対応した固有ベクトルから求められる。以下同様にして，Γ の階数に等しい数（あるいは (3.27) において1よりも大きい固有根の数）だけの因子を求めることができる。

次に第1因子について，その因子負荷量を求めることを考える。(3.25)，(3.27)を用いると，

$$\Gamma\boldsymbol{l}=(\delta-1)\Sigma\boldsymbol{l} \tag{3.29}$$

であることが分かるから，

$$\begin{aligned}\cos(x_i, f_1)&=\cos(x_i, \boldsymbol{l}_1\boldsymbol{z}_1+\boldsymbol{l}_2\boldsymbol{z}_2+\cdots+\boldsymbol{l}_p\boldsymbol{z}_p)\\&=\boldsymbol{l}_1\boldsymbol{r}_{i1}+\boldsymbol{l}_2\boldsymbol{r}_{i2}+\cdots+\boldsymbol{l}_p\boldsymbol{r}_{ip}\\&=(\delta_1-1)\boldsymbol{l}_i\sigma_{ii}\end{aligned} \tag{3.30}$$

また，(3.15) と同様，$\cos(x_i, f_1)=\lambda_{i1}$ であるから，

$$\lambda_{i1}=(\delta_1-1)\boldsymbol{l}_i\sigma_{ii}\ \ (\boldsymbol{i}=1, 2, \cdots, \boldsymbol{p}) \tag{3.31}$$

他の因子に対する因子負荷量も全く同様にして求めることができる。

このとき，誤差分散 σ_{ii} は (3.19) を求めたのと同様にして，次式を満足していることが分かる。

$$\Psi_{ii}=\lambda_i^2+\lambda_{i2}^2+\cdots+\lambda_{im}^2+\sigma_{ii} \tag{3.23}$$
$$=(\delta_1-1)^2\boldsymbol{l}_i^2\sigma_{ii}^2+(\delta_2-1)^2\boldsymbol{h}_i^2\sigma_{ii}^2+\cdots$$
$$+(\delta_m-1)^2\boldsymbol{k}_i^2\sigma_{ii}^2+\sigma_{ii} \tag{3.33}$$

ここに，$\boldsymbol{h}_i, \boldsymbol{k}_i$ は (3.27) に於いて，δ_2 および δ_m に対応した固有ベクトルの第 $\boldsymbol{i}$ 成分をあらわす。

§4. パラメータの推定

§2，§3で因子分析の種々のモデルについて述べたが，本節ではその各々の場合について，因子負荷量，誤差分散等のパラメータを推定する問題を考察する。

（I） §2に対応したパラメータの推定

前に述べたように，共通因子の分散共分散行列 M については，$M=\boldsymbol{I}$ として議論を進める。

まずはじめに，因子ベクトル $\boldsymbol{f}$ が確率変数で $N(\Theta, \boldsymbol{I})$ にしたがうものと仮定する場合について述べる。誤差項 $\boldsymbol{e}$ は $N(\Theta, \Sigma)$ にしたがうものとする（Σ：対角行列)。すなわち，$\boldsymbol{x}=\boldsymbol{\mu}+\Lambda \boldsymbol{f}+\boldsymbol{e}$ は $N(\boldsymbol{\mu}, \Psi)$ にしたがうここで，$\Psi=\Lambda\Lambda'+\Sigma$ である)。このような母集団から $\boldsymbol{n}$ 個の独立な標本 $\boldsymbol{x}_1, \boldsymbol{x}_2, \cdots, \boldsymbol{x}_n$ をとりだした場合，それに基づいて，$\boldsymbol{\mu}, \Lambda, \Sigma$ を推定しようというわけである。

$\boldsymbol{\mu}, \Lambda, \Sigma$ の最尤推定量を求めることを考える。尤度函数を $L(\boldsymbol{\mu}, \Lambda, \Sigma)$ であらわすと，

$$\log L(\boldsymbol{\mu}, \Lambda, \Sigma)=-\frac{1}{2}pn\log 2\pi-\frac{1}{2}n\log|\Sigma+\Lambda\Lambda'|$$
$$-\frac{1}{2}\sum_{i=1}^{n}(\boldsymbol{\lambda}_i-\boldsymbol{\mu})'(\Sigma+\Lambda\Lambda')^{-1}(\boldsymbol{x}_i-\boldsymbol{\mu})$$
$$=-\frac{1}{2}pn\log 2\pi-\frac{1}{2}n\log|\Sigma+\Lambda\Lambda'|$$
$$-\frac{1}{2}n\,\mathrm{tr}[A(\Sigma+\Lambda\Lambda')^{-1}]$$
$$-\frac{1}{2}n(\bar{\boldsymbol{x}}-\mu)_{\prime}(\Sigma+\Lambda\Lambda')^{-1}(\bar{\boldsymbol{x}}-\boldsymbol{\mu}) \tag{4.1}$$

ここに，

$$\bar{\boldsymbol{x}}=\frac{1}{n}\sum_{i=1}^{n}\boldsymbol{x}_i\text{：平均ベクトル} \tag{4.2}$$

$$A=\frac{1}{n}\sum_{i=1}^{n}(\boldsymbol{x}_i-x)(\boldsymbol{x}_i-\bar{\boldsymbol{x}})'\text{：標本分散共分散行列} \tag{4.3}$$

対数尤度函数 (4.1) を最大ならしめるように μ, Λ, Σ を定めればよい。

(4.1) を μ について偏微分して0とおくと直ちに $\boldsymbol{\mu}$ の最尤推定量として，$\hat{\boldsymbol{\mu}}=\bar{x}$ を得る。次にこれを (4.1) の右辺に代入すると第4項は0になり，また第1項は定数項であるから，Λ, Σ

の推定に必要な項は第２項，第３項だけである。この部分を改めて L とおくと，

$$L=-\frac{1}{2}n\log|\Sigma+\Lambda\Lambda'|-\frac{1}{2}n\,\mathrm{tr}\,A(\Sigma+\Lambda\Lambda')^{-1} \tag{4.4}$$

Λ,Σ の最尤推定量は，(4.4) を Λ および Σ の各要素について偏微分して０とおくことにより得られる。すなわち，

$$\frac{\partial L}{\partial\lambda_{ir}}=0\ \ (i=1,2,\cdots,p,\ \ r=1,2,\cdots,m) \tag{4.5}$$

$$\frac{\partial L}{\partial\sigma_{ii}}=0\ \ (i=1,2,\cdots,p) \tag{4.6}$$

途中の計算はかなり面倒であるが，最終的には (4.5) に対応して，

$$\Lambda'\Psi^{-1}-\Lambda'\Psi^{-1}A\Psi^{-1} \tag{4.7}$$

の (i,r) 要素が０，(4.6) に対応して，

$$\Psi^{-1}-\Psi^{-1}A\Psi^{-1} \tag{4.8}$$

の i 番目の対角要素が０であるという結果を得ることができる。

ところで，§２で述べたように，Λ と $\Lambda\Theta$ (Θ：任意の直交行列) はすべて同等の解になっているので上に得られた結果からだけでは，一意的な解を得ることはできない。$\Lambda\Theta$ からある特定の Λ を選びだすために，Λ に関しさらに何らかの制約を付け加えなければならない。種々の制約条件が考えられるが，ここでは，

$$J=\Lambda'\Sigma^{-1}\Lambda \tag{4.9}$$

が対角行列であって，その対角要素が大きさの順に並んでいるという条件を付け加えることにする。以上をまとめると，結局，パラメータ Λ, Σ の推定量 $\hat{\Lambda}$, $\hat{\Sigma}$ を求めるための式として，

$$\hat{\Lambda}'\hat{\Psi}^{-1}-\hat{\Lambda}'\hat{\Psi}^{-1}AA^{-1}=0 \tag{4.10}$$

$$\mathrm{diag}(\hat{\Psi}^{-1}-\hat{\Psi}^{-1}AA^{-1})=0 \tag{4.12}$$

$$\hat{J}=\hat{\Lambda}'\hat{\Sigma}^{-1}\hat{\Lambda}：\text{対角行列} \tag{4.12}$$

$$\hat{\Psi}=\hat{\Lambda}\hat{\Lambda}'+\hat{\Sigma} \tag{4.13}$$

が得られたことになる。(4.10)～(4.13)をより取り扱いやすい形にするために若干変形する。以下便宜上，推定量のを省略する。

(4.10) の後からを乗ずると，

$$\hat{\Lambda}'-\hat{\Lambda}'\hat{\Psi}^{-1}A=0 \tag{4.14}$$

(4.11)の前後から，$\Sigma=\hat{\Psi}-\hat{\Lambda}\hat{\Lambda}'$ を乗じさらに(4.14)を用いると，

$$\mathrm{diag}(\hat{\Psi}-A)=0$$

また，(4.14)は恒等式

$$\hat{\Psi}^{-1}=\hat{\Sigma}^{-1}-\hat{\Sigma}^{-1}\hat{\Lambda}(I+J)^{-1}\hat{\Lambda}'\hat{\Sigma}^{-1}$$

を用いると，

$$\hat{\Lambda}'=\hat{J}^{-1}\hat{\Lambda}'\hat{\Sigma}^{-1}(\hat{A}-\hat{\Sigma}) \tag{4.17}$$

あるいは，

$$\hat{\Lambda}'=\hat{J}^{-1}\hat{\Lambda}'\Sigma^{-1}(A-\hat{\Sigma}) \tag{4.17}$$

という形に変形できる。さらに，(4.17)より，

$$G=\hat{\Lambda}'\hat{\Sigma}^{-1}(A-\hat{\Sigma})\hat{\Sigma}^{-1}\hat{\Lambda} \tag{4.18}$$

は $\hat{J}^2$ に等しいことが容易に確かめられる。

結局，(4.10)，(4.11) が (4.15)，(4.17) の形にかきあらためられたことになる。(4.15)，(4.17)，(4.12(，(4.13) が，Λ, Σ の推定量を求める際に用いられる式である。これら 4 つの式から直接に Λ，Σ の推定量は求めることができないので，実際には Λ，Σ に適当な初期値を与え，繰り返し法により解を求めるわけである。具体的な手順については後述の附録に述べる。

さて次に因子ベクトル f の各成分が確率変数ではなくて，各個人に特有なパラメータ (α 番目の個人に対しては $f=f_\alpha$) であると考えた場合の推定の問題を考えよう。この場合，モデル (2.1) をもう少し正確な形であらわすと，

$$x_\alpha=\mu+\Lambda f_\alpha+e \quad (\alpha=1,2,\cdots,n) \tag{4.19}$$

となる。この場合にも，e が正規分布にしたがうという仮定のもとで，ある形のパラメータの最大推定量を求めることができるが，得られる結果は非常に複雑であって実用には適さない。そこで，ここでは $\varepsilon e=0$，$\Sigma=\sigma^2 I$ (これは本質的な仮定である) と仮定して，最小二乗法を用いてパラメータを推定することを考える。§ 2 でも述べたように，

$$\frac{1}{n}\sum_{\alpha=1}^{n} f_\alpha=0,\quad \frac{1}{n}\sum_{\alpha=1}^{n} f_\alpha f_\alpha'=I \tag{4.20}$$

と仮定する。この条件のもとで

$$\sum_{\alpha=1}^{n}(x_\alpha-\mu-\Lambda f_\alpha)'(x_\alpha-\mu-\Lambda f_\alpha) \tag{4.21}$$

を最小にするように，μ, Λ, f_α ($\alpha=1,2,\cdots,n$) を定めればよいことになる。途中の計算は省略して，最終的に以下のような結果が得られる。

$p\times m$ の行列 Y を次のように定義しておく。

$$Y=(x_{ij}-\bar{x}_i) \tag{4.22}$$
$$(i=1,2,\cdots,p,\quad j=1,2,\cdots,n)$$

1° Λ の k 番目の列 $\lambda_{(k)}$ の推定量 $\hat{\lambda}_{(k)}$ は，YY'($p\times p$ の行列) の k 番目の固有値 d_k に対応した固有ベクトルとして与えられる。$\hat{\lambda}_{(k)}$ は $\hat{\lambda}_{(k)}'\hat{\lambda}_{(k)}=d_k/n$ によって規準化され，$\hat{\lambda}_{(k)}$，$\hat{\lambda}_{(l)}$ ($k\neq l$) は互いに直交している。

2° n 人についての因子ベクトル $f_\alpha(\alpha=1,2,\cdots,n)$ の k 番目の成分の推定量は，$Y'Y$ ($n\times n$ の行列) の k 番目の固有値 d_k (YY' の k 番目の固有値に等しい) に対応した固有ベクトルとして与えられ，(4.20) により規準化されている。

標本分散共分散行列$A=\frac{1}{n}YY'$ であり，上の 1° で得られた結果は通常の主成分分析により得られる結果と全く一致していることが分かる。

（II） 主因子分析の場合

標本分散共分散行列 A を用いて，§3―(I) で述べた Γ および Σ を推定することができる。§3―(I) で $\boldsymbol{l}$ に対しては次式が得られた。

$$(\Gamma-d\boldsymbol{I})\boldsymbol{l}=0 \tag{4.23}$$

$$\boldsymbol{l}'\Gamma\boldsymbol{l}=1 \text{ あるいは } \boldsymbol{l}'\boldsymbol{l}=1/\alpha \tag{4.24}$$

因子負荷量と $\boldsymbol{l}$ の間には，(3.17) の関係があることを用いると，上式は次のようにかきかえることができる。

$$(\Gamma-d\boldsymbol{I})\boldsymbol{\lambda}=0 \tag{4.25}$$

$$\boldsymbol{\lambda}'\boldsymbol{\lambda}=\alpha \tag{4.26}$$

ここに，$\boldsymbol{\lambda}$ は Λ のある列をあらわす。ここで，$\Gamma=\Psi-\Sigma$ とおきかえ，Ψ に A を代入すると次式を得る。

$$(A-\Sigma-d\boldsymbol{I})\boldsymbol{\lambda}=0 \tag{4.27}$$

$$\boldsymbol{\lambda}'\boldsymbol{\lambda}=d \tag{4.28}$$

また，(3.19) に対応して

$$a_{ii}=\lambda_{i1}^2+\lambda_{i2}^2+\cdots+\lambda_{im}^2+\sigma_{ii} \tag{4.29}$$

すなわち，

$$\sigma_{ii}=a_{ii}-\lambda_{i1}^2-\lambda_{i2}^2-\cdots-\lambda_{im}^2 \tag{4.30}$$

$$(i=1,2,\cdots,p)$$

そこで実際に計算を行う場合にはまず (4.27) で Σ に適当な初期値を与え，大きさの順に m 個の固有値 $d_1,d_2,\cdots,d_m$ を求め，(4.28) により規準化された対応した固有ベクトル $\boldsymbol{\lambda}_{(1)},\boldsymbol{\lambda}_{(2)},\cdots,\boldsymbol{\lambda}_{(m)}$ を求める。次にこうして求めた $\boldsymbol{\lambda}_{(1)},\boldsymbol{\lambda}_{(2)},\cdots\boldsymbol{\lambda}_{(m)}$ を (4.30) に代入して新しい Σ の推定値を得る。これを再び (4.27) に代入し新しい $A-\Sigma$ の固有値，固有ベクトルを計算し，収束値が得られるまでこのループをくりかえせばよい。

（III） 正準因子分析の場合

この場合には前出の (3.27)，(3.32) において，Ψ に標本分散共分散行列 A を代入した式を用いればよい。

まず，(3.27) に対応して，

$$(A-\delta\Sigma)\boldsymbol{l}=0 \tag{4.31}$$

$$\boldsymbol{l}'\Gamma\boldsymbol{l}=1 \tag{4.32}$$

(4.32) は前に述べたように，共通因子の分散が1に規準化されていることをあらわす。$\boldsymbol{l}'\Gamma\boldsymbol{l}$ は (3.25)，(3.27) を用いれば容易に次のようにかきかえることができる。

$$\boldsymbol{l}'\Gamma\boldsymbol{l}=(\delta-1)\boldsymbol{l}'\Sigma\boldsymbol{l} \tag{4.33}$$

すなわち，(4.32) の規準化の条件は，

$$\boldsymbol{l}'\Sigma\boldsymbol{l}=1/(\delta-1) \tag{4.34}$$

とあらわすことができる。実際に推定したいのは因子負荷量であるから，(4.31)，(4.34)を前に述べた (3.31) を用いて，因子負荷量ベクトル $\boldsymbol{\lambda}$ (Λ のある列) についての式にかきなおすことを考える。(3.31)を一般的な形にあらわすと，

$$\boldsymbol{\lambda}=(\delta-1)\Sigma\boldsymbol{l} \tag{4.35}$$

すなわち，

$$\boldsymbol{l}=\frac{1}{\delta-1}\Sigma^{-1}\boldsymbol{\lambda} \tag{4.36}$$

これを (4.31)，(4.34) に代入すると容易に次式を得る。

$$(A\Sigma^{-1}-\delta\boldsymbol{I})\boldsymbol{\lambda}=0 \tag{4.37}$$

$$\boldsymbol{\lambda}'\Sigma^{-1}\boldsymbol{\lambda}=\delta-1 \tag{4.38}$$

また，(3.32) に対応して，

$$a_{ii}=\lambda_{i1}^2+\lambda_{i2}^2+\cdots+\lambda_{im}^2+\sigma_{ii}$$

すなわち，

$$\sigma_{ii}=a_{ii}-\lambda_{i1}^2-\lambda_{i2}^2-\cdots-\lambda_{im}^2 \tag{4.40}$$

$$(i=1,2,\cdots,p)$$

このようにして得られた (4.37)，(4.38)，(4.40) を用いて，因子負荷量等のパラメータを推定することができる。

実際に計算を行う場合には，まず (4.37) で Σ に適当な初期値を与え，$A\Sigma^{-1}$ の大きい方の m 個の固有根，および (4.38) により規準化された固有ベクトルを求める。 次に求めた固有ベクトルを (4.40) に代入して，新しい Σ の推定値を得る。これを再び (4.37) に代入して，収束値が得られるまでこのループをくり返すわけである。

最後に正準因子分析により得られた因子負荷量の推定値とこの節のはじめに述べた最尤法による因子負荷量の推定値が完全に一致することを証明しておく。

最尤法による因子負荷量の推定値は，前に述べた (4.15)，(4.17，(および，$J=\Lambda'\Sigma^{-1}\Lambda$ が対角行列であって，その対角要素が大きさの順に並んでいるという条件) をもとにして得られる。(4.17) をかきなおすと，

$$(A-\hat{\Sigma})\hat{\Sigma}^{-1}\hat{\Lambda}=\hat{\Lambda}\hat{J} \tag{4.41}$$

あるいは，

$$A\hat{\Sigma}^{-1}\hat{\Lambda}=(\hat{J}+\boldsymbol{I}) \tag{4.42}$$

となる。また，(4.15) より直ちに，

$$\sigma_{ii}=a_{ii}-\lambda_{i1}^2-\lambda_{i2}^2-\cdots-\lambda_{im}^2 \tag{4.43}$$

$$(i=1,2,\cdots,p)$$

これらを今得られた正準因子分析についての式 (4.37)，(4.38)，(4.40) とくらべると，明らかに両者より得られる結果は一致していることが分かる。また，対角行列 J の各対角要素と (4.

37) の第 i 番目の固有根 δ_i の間には，

$$j_i=\delta_i-1 \tag{4.41}$$

という関係があることが分かる。ここに，j_i は対角行列 J の i 番目の対角要素をあらわす。

§5. 因子の数に関する検定

パラメータの推定については前節で述べたが，その場合には共通因子の数はあらかじめ与えられているものとして議論を進めた。しかしながら，一般には共通因子の数は前以っては定まっておらず，共通因子の数をいくつにすればよいかという問題を考察しておかなければならない。そこでこの節では，共通因子の数 m に関する検定の問題について述べる。p 次元ベクトル x は平均 μ，分散共分散行列 Ψ の多変量正規分布にしたがっているものとする。

前節の最後に，正準因子分析で得られる解は最尤法から得られる解と全く一致することを示したので，この両者についての検定の問題は同じ形で取り扱うことができる。そこでまず前節のはじめに述べた最尤法によりパラメータを推定した場合に対応した検定の問題から考察する。

標本の数 n がかなり大きい場合には，次のような形の尤度比検定を考えることができる。

帰無仮説は $\Psi=\Sigma+\Lambda\Lambda'(\mathrm{rank}\,\Lambda=m)$ であり，対立仮説は Ψ が任意の正値行列であることである。尤度函数を $L(A,\Psi,\mu)$ であらわすことにすれば，

$$L(A,\Psi,\mu)=(2\pi)^{-pn/2}|\Psi|^{-n/2}\exp\left\{-\frac{1}{2}\left[n\,\mathrm{tr}\,A\Psi^{-1}\right.\right.$$
$$\left.\left.+n(\bar{x}-\mu)'\Psi^{-1}(\bar{x}-\mu)\right]\right\} \tag{5.1}$$

対立仮説のもとでの μ, Ψ の最尤推定量は明らかに $\hat{\mu}=\bar{x}$, $\hat{\Psi}=A$ であり，帰無仮説のもとでの μ, Ψ の最尤推定量は，$\bar{\mu}=\bar{x}$, $\hat{\Psi}=\hat{\Lambda}\hat{\Lambda}'+\hat{\Sigma}$ である。すると，尤度比は次式で与えられることになる。

$$\frac{L(A,\hat{\Sigma}+\hat{\Lambda}\hat{\Lambda}',\hat{\mu})}{L(A,A,\hat{\mu})}=\frac{|A|^{n/2}e^{np/2}}{|\hat{\Sigma}+\hat{\Lambda}\hat{\Lambda}'|^{n/2}e^{n\,\mathrm{tr}\,A(\hat{\Sigma}+\hat{\Lambda}\hat{\Lambda}')^{-1}/2}} \tag{5.2}$$

ここで，(4.11)，(4.14) を用いれば容易に，

$$\mathrm{tr}\,A(\hat{\Sigma}+\hat{\Lambda}\hat{\Lambda}')^{-1}=p \tag{5.3}$$

であることが確かめられるから，尤度比 (5.2) を K とすると，

$$K=\frac{|A|^{n/2}}{|\hat{\Sigma}+\hat{\Lambda}\hat{\Lambda}'|^{n/2}}=|A(\hat{\Sigma}+\hat{\Lambda}\hat{\Lambda}')^{-1}|^{n/2} \tag{5.4}$$

となり，$-2\log K$ が漸近的に χ^2 分布にしたがうことを用いて検定することができる。

(5.4) を実際に利用する場合の形に変形する。

$$\begin{aligned}|A(\hat{\Sigma}+\hat{\Lambda}\hat{\Lambda}')^{-1}|&=|A\hat{\Sigma}^{-1}|\cdot|I+\hat{\Lambda}\hat{\Lambda}'\hat{\Sigma}^{-1}|^{-1}\\&=|A\hat{\Sigma}^{-1}|\cdot|I+\hat{\Lambda}'\hat{\Sigma}^{-1}\hat{\Lambda}|^{-1}\\&=|A\hat{\Sigma}^{-1}|\cdot|I+\hat{J}|^{-1}\end{aligned} \tag{5.5}$$

ここで，(4.17) をかきなおした (4.42) を用いると，

$$|\Lambda\hat{\Sigma}^{-1}|=\prod_{i=1}^{p}(1+\hat{j}_i) \tag{5.6}$$

ここに $\hat{j}_i$ は $\hat{J}$ の第 i 対角要素である。これを用いると，

$$|\Lambda(\hat{\Sigma}+\hat{\Lambda}\hat{\Lambda}')^{-1}|=\sum_{i=1}^{p}(1+\hat{j}_i)/\prod_{i=1}^{m}(1+\hat{j}_i)$$

$$=\sum_{i=1}^{p}(1+\hat{j}_i) \tag{5.7}$$

すなわち，

$$-2\log K=-2\log\left[\prod_{i=m+1}^{p}(1+\hat{j}_i)\right]^{n/1}$$

$$=-n\sum_{i=m+1}^{p}\log(1+\hat{j}_i) \tag{5.8}$$

検定統計量 (5.8) が漸近的に χ^2 分布にしたがうことになり，それを用いて因子の数 m についての検定を行うことができるわけである。また，このときの自由度は，仮説によりパラメータにかせられた制約の数に等しいから次のようになる。Ψ に制約をつけ加えないときのパラメータの数は $p(p+1)/2$, また，仮説のもとでの実質的なパラメータの数は次のように考えられる。パラメータの数は Λ について pm 個，Σ について p 個であるが，その他に $\Lambda\Theta$ (Θ：任意の直交行列）からある特定の Λ を選びだすために，$J=\Lambda'\Sigma^{-1}\Lambda$ が対角行列であるという条件を付け加えたので，この条件の分だけ仮説のもとでの実質的なパラメータの数は減ずることになる。$J=\Lambda'\Sigma^{-1}\Lambda$ は $m\times m$ の行列であるから，これにより加えられる制約条件の数は $m(m-1)/2$ である。以上をまとめると結局 χ^2 の自由度は，

$$\frac{p(p+1)}{2}-\left\{pm+p-\frac{m(m-1)}{2}\right\}$$

$$=\frac{1}{2}\left\{(p-m)^2-(p+m)\right\} \tag{5.9}$$

で与えられることになる。今の場合，(5.8) で n を若干かえて，

$$n-\frac{1}{6}(2p+1)-\frac{2}{3}m$$

とした方が，χ^2 に対する近似が若干よくなることが示されている。

また，正準因子分析の場合は最尤法と全く一致するわけであるから，(4.44)を用いると検定統計量 (5.8) は

$$-n\sum_{i=m+1}^{p}\log\delta_i \tag{5.10}$$

となる。

次に，因子ベクトル f が確率変数ではないと考えた場合には，$\Sigma=\sigma^2I$ の仮定のもとで，それは主成分分析に全く一致することを示したので，それに関する検定の問題はここでは省略する。

§6. 因子スコアの推定

§2で示した通常の因子分析のモデルに於いて，我々は因子ベクトル f を確率変数と考え，

パラメータ $\boldsymbol{\mu}$, Λ, Σ を推定することを考えてきた。しかしながら，実際問題に於いては $\boldsymbol{n}$ 個の標本 $\boldsymbol{x}_1, \boldsymbol{x}_2, \cdots, \boldsymbol{x}_n$ を観測した場合，$\boldsymbol{\mu}$, Λ, Σ だけでなく，各個人についての因子スコアベクトル $\boldsymbol{f}_\alpha$ $(\alpha=1,2,\cdots,\boldsymbol{n})$ の値がどの位になっているかを推定したい場合が非常に多い。§4で述べた最尤法により，$\boldsymbol{\mu}$, Λ, Σ だけでなく，$\boldsymbol{f}_1, \boldsymbol{f}_2, \cdots, \boldsymbol{f}_n$ を同時に推定することはできないので，何らかの方法により $\boldsymbol{f}', \boldsymbol{f}_2, \cdots, \boldsymbol{f}_n$ を推定することを考えなければならない。これに対して通常用いられている方法は，パラメータ $\boldsymbol{\mu}$, Λ, Σ の値が $\hat{\boldsymbol{\mu}}$, $\hat{\Lambda}$, $\hat{\Sigma}$ に等しいと考えそれらを既知の定数として扱い，然る後に $\boldsymbol{f}_1, \boldsymbol{f}_2, \cdots, \boldsymbol{f}_n$ を推定する方法である。以下これにしたがう2, 3の方法を簡単に述べる。

i）正規性の仮定のもとに $\boldsymbol{x}_1, \boldsymbol{x}_2, \cdots, \boldsymbol{x}_n$ の尤度を考え，それを $\boldsymbol{f}_1, \boldsymbol{f}_2, \cdots, \boldsymbol{f}_n$ に関し最大にする。これは，

$$\sum_{i=1}^{p} \frac{(x_{i\alpha}-\mu_i-\sum_{v=1}^{m}\lambda_{iv}f_{v\alpha})^2}{\sigma_{ii}} (\alpha=1,2,\cdots,\boldsymbol{n}) \tag{6.1}$$

を最小にすることに等しく，重みずけをした最小二乗法であり，$\boldsymbol{f}_\alpha$ の推定値として，

$$\hat{\boldsymbol{f}}_\alpha=(\Lambda'\Sigma^{-1}\Lambda)^{-1}\Lambda'\Sigma^{-1}(\boldsymbol{x}_\alpha-\boldsymbol{\mu}) \tag{6.2}$$

を得る。

ii）$\varepsilon \boldsymbol{f}\boldsymbol{f}'=\boldsymbol{I}$ と仮定すると，$\boldsymbol{x}$ と $\boldsymbol{f}$ の分散共分散行列は，

$$\varepsilon\begin{pmatrix}\boldsymbol{x}-\boldsymbol{\mu}\\ \boldsymbol{f}\end{pmatrix}\begin{pmatrix}\boldsymbol{x}-\boldsymbol{\mu}\\ \boldsymbol{f}\end{pmatrix}'=\begin{pmatrix}\Sigma+\Lambda\Lambda' & \Lambda\\ \Lambda' & \boldsymbol{I}\end{pmatrix} \tag{6.3}$$

となるので，$\boldsymbol{x}$ の上への $\boldsymbol{f}$ の回帰を考えることにより，

$$\begin{aligned}\hat{\boldsymbol{f}}_\alpha&=\Lambda'(\Sigma+\Lambda\Lambda')^{-1}(\boldsymbol{x}_\alpha-\bar{\boldsymbol{x}})\\ &=(\boldsymbol{I}+\boldsymbol{J})^{-1}\Lambda'\Sigma^{-1}(\boldsymbol{x}_\alpha-\bar{\boldsymbol{x}})\end{aligned} \tag{6.4}$$

を得る。

iii）$\dfrac{1}{\boldsymbol{n}}\sum_{\alpha=1}^{n}\boldsymbol{f}_\alpha\boldsymbol{f}_\alpha'=\boldsymbol{I}$ の条件のもとで，

$$\sum_{\alpha=1}^{n}(\boldsymbol{x}_\alpha-\bar{\boldsymbol{x}}-\Lambda f_\alpha)'\Sigma^{-1}(x_\alpha-\bar{\boldsymbol{x}}-\Lambda f_\alpha) \tag{6.5}$$

を最小にする。ラグランジュ乗数行列を Θ（対称行列）として，

$$\sum_{\alpha=1}^{n}(\boldsymbol{x}_\alpha-\bar{\boldsymbol{x}}-\Lambda\boldsymbol{f}_\alpha)'\Sigma^{-1}(\boldsymbol{x}_\alpha-\bar{\boldsymbol{x}}-\Lambda\boldsymbol{f}_\alpha)$$

$$+\mathrm{tr}\,\Theta(\sum_{i=1}^{n}\boldsymbol{f}_\alpha\boldsymbol{f}_\alpha'-\boldsymbol{n}\boldsymbol{I}) \tag{6.6}$$

を $\boldsymbol{f}_\alpha$ に関し偏微分して0とおくと，$\boldsymbol{f}_\alpha$ の推定値として，

$$\boldsymbol{f}_\alpha=(\Lambda'\Sigma^{-1}\Lambda+\Theta)^{-1}\Lambda'\Sigma^{-1}(\boldsymbol{x}_\alpha-\bar{\boldsymbol{x}}) \tag{6.7}$$

を得る。$\sum_{\alpha=1}^{n}\boldsymbol{f}_\alpha\boldsymbol{f}_\alpha'=n\boldsymbol{I}$ より，

$$(\Lambda'\Sigma^{-1}\Lambda+\Theta)^2=\Lambda'\Sigma^{-1}\Lambda\Sigma^{-1}\Lambda \tag{6.8}$$

であり，さらに，

$$\Lambda'\Sigma^{-1}\Lambda\Sigma^{-1}\Lambda=J(I+J) \tag{6.9}$$

であるから，結局 (6.7) は

$$\hat{\boldsymbol{f}}_\alpha=[J(I+J)]^{-1/2}\Lambda'\Sigma^{-1}(\boldsymbol{x}_\alpha-\bar{\boldsymbol{x}})$$

となる。

§7. 単位の問題

この節では，通常の因子分析のモデルで $\boldsymbol{p}$ 次元ベクトル $\boldsymbol{x}$ の各成分のスケールを変換した場合に，因子負荷量，誤差分散等にどのような影響があらわれるかを考察しておく。

そのために，

$$\boldsymbol{z}=D\boldsymbol{x} \quad (D：対角行列) \tag{7.1}$$

というスケール変換を考える。すると，

$$\varepsilon\,\boldsymbol{z}=D\boldsymbol{\mu}=\boldsymbol{\mu}^* \tag{7.2}$$

$$\begin{aligned}\varepsilon\,(\boldsymbol{z}-D\boldsymbol{\mu})(\boldsymbol{z}-D\boldsymbol{\mu})'&=D\Psi D\\&=D\Lambda(D\Lambda)'+D\Sigma D\\&=\Psi^*\end{aligned} \tag{7.3}$$

となる。Ψ^* を，

$$\Psi^*=\Lambda^*\Lambda^{*\prime}+\Sigma^* \tag{7.4}$$

とあらわすことにすると，あきらかに Σ^* としては $D\Sigma D$, $\Lambda^*\Lambda^{*\prime}$ としては $D\Lambda(D\Lambda)'$ をとることができる。すなわち，$\Lambda^*=D\Lambda\theta$ (θ：直交行列) とすることができる。また，前に解の不確定性を取り除くために，$J=\Lambda'\Sigma^{-1}\Lambda$ が対角行列であるという制約をつけ加えた。それに対応して今の場合 $J^*=\Lambda^{*\prime}\Sigma^{*\prime}\Lambda^*$ とおくと，

$$\Lambda^{*\prime}\Sigma^{*-1}\Lambda^*=(D\Lambda\theta)'(D\Sigma D)^{-1}(D\Lambda\theta)=\theta'\Lambda'\Sigma^{-1}\Lambda\theta \tag{7.5}$$

となり，$\theta=I$ と選ぶことにより，

$$\Lambda^{*\prime}\Sigma^{*-1}\Lambda^*=\Lambda'\Sigma^{-1}\Lambda \tag{7.6}$$

となり，

$$\Lambda^*=D\Lambda,\quad \Sigma^*=D\Sigma D \tag{7.7}$$

とすることができる。

また，推定の場合について，スケールの変換の影響を考えて見ると，$\boldsymbol{z}_i=D\boldsymbol{x}\boldsymbol{z}$ ($\boldsymbol{x}_i$：i 番目のデータについての $\boldsymbol{p}$ 次元ベクトル) とすると，スケール変換後の標本分散共分散行列 A^* は

$$A^*=DAD \tag{7.8}$$

であることが分かるから，この場合の最尤推定量を求めるための式は，(4.10) ～ (4.13) におい

て，A, $\hat{\Psi}$, $\hat{\Lambda}$, $\hat{\Sigma}$ をそれぞれ A^*, $\hat{\Psi}^*$, $\hat{\Lambda}^*$, $\hat{\Sigma}^*$ におきかえることにより得られる。それらの式より直ちに，$\hat{\Lambda}$, $\hat{\Sigma}$ が (4.10)～(4.13) の解になっている場合には，$\hat{\Lambda}^*=D\hat{\Lambda}$, $\hat{\Sigma}^*=D\hat{\Sigma}D$ が上の解になっていることが確かめられる。以上より分かるように，$J=\Lambda'\Sigma^{-1}\Lambda$ が対角行列であるという制約をつけ加えた場合には，$\boldsymbol{x}$ のスケール変換は，本質的には解に影響を与えていないことになる。

§8. 因子分析法と主成分分析法の関係

以上，因子分析法についていろいろ述べたが，最後に前篇で述べられた主成分分析法とこれまで述べてきた因子分析法の相違について若干ふれておく。いずれの方法も目的は，p 個の特性値を有する多変量データが存在した場合に，実質的には p よりも少い m 個の新しい変量を用いて，その p 個の特性値を代表させることができるのではないかということである。

その場合に，m 個の新しい変量をどのように定めるかという定め方において，主成分分析では，最初の p 個の変量に対して，分布型には無関係に通常の最小二乗法的な考え方（あてはめた場合の誤差を最小にするということは，いいかえれば，主成分によりあらわされる変動の割合を最大にするということになる）を用いているのに対し，因子分析法では，最初に (2.1) のようなモデルを立て，各変量が正規分布にしたがうという仮定のもとで，パラーメタの最尤推定量を求めることを考えたわけである。

因子分析のモデルに於いて，誤差の分散共分散行列の対角要素がすべて等しい場合，すなわち，$\Sigma=\sigma^2 I$ の場合には，因子分析から得られる結果と主成分分析から得られる結果は全く一致することが分かる。以下にこれを示す。

§4―(I)の因子負荷量および誤差分散の最尤推定量を得るためのいくつかの方程式に対し，$\Sigma=\sigma^2 I$ とすると容易に次のような方程式が得られる。

$$A\hat{\Lambda}=\hat{\Lambda}[\sigma^2(J+I)] \tag{8.1}$$

$$p\hat{\sigma}^2=\mathrm{tr}(A-\hat{\Lambda}\hat{\Lambda}') \tag{8.2}$$

$$\hat{J}=\frac{1}{\hat{\sigma}^2}\hat{\Lambda}'\hat{\Lambda}：対角行列$$

今，A の固有値を大きさの順に $d_1, d_2, \cdots, d_p$ とし，対応した規準化された固有ベクトルを $\boldsymbol{l}_1, \boldsymbol{l}_2, \cdots, \boldsymbol{l}_p$ ($\boldsymbol{l}_i'\boldsymbol{l}'=1$) とする。大きい方の m 個の固有根 $d_1, d_2, \cdots, d_m$ を対角要素とする $m\times m$ の対角行列を D とし，それに対応した固有ベクトル行列を $L(p\times m)$ とすると，

$$AL=LD \tag{8.4}$$

$$L'L=I \tag{8.5}$$

である。すると，D, L は一意的に定まり，(8.1) と対応付けるとまず，

$$D=\hat{\sigma}^2(J+I) \tag{8.6}$$

であることが分かる。すると，(8.2), (8.3) を用いて，

$$\begin{aligned} p\hat{\sigma}^2&=\mathrm{tr}(A-\hat{\Lambda}\hat{\Lambda}') \\ &=\mathrm{tr}\,A-\mathrm{tr}\,\hat{\Lambda}'\hat{\Lambda} \end{aligned}$$

$$=\mathrm{tr}\,A-\hat{\sigma}^2\,\mathrm{tr}\,\hat{J}$$
$$=\mathrm{tr}\,A-\hat{\sigma}^2\,\mathrm{tr}\left(\frac{1}{\hat{\sigma}^2}D-I\right)$$
$$=\sum_{i=1}^{p}d_i-\sum_{i=1}^{m}d_i+m\hat{\sigma}^2$$

すなわち，

$$\hat{\sigma}^2=\frac{1}{p-m}\sum_{i=m+1}^{p}d_i \tag{8.7}$$

であることが分かる。

また，$\hat{A}$ の各列は，$d_1, d_2, \cdots, d_m$ に対応した固有ベクトルになっているが，規準化は次のようになされている。

$$\hat{A}'\hat{A}=\hat{\sigma}^2J=\hat{\sigma}^2\left(\frac{1}{\hat{\sigma}^2}D-J\right)$$
$$=D-\hat{\sigma}^2I \tag{8.8}$$

以上で，$\Sigma=\sigma^2I$ のときには両者から得られる結果が一致することが示されたことになる。また，$\Sigma=\sigma^2I$ のときには前に述べた主因子分析から得られる結果もこれらに一致することが容易に示される。

<**附録**>～因子負荷量，誤差分散の最尤推定量を求めるための計算手順

因子負荷量および誤差分散の最尤推定量を求める場合の計算の手続きは，前述の (4.15)，(4.17)，(4.18) および $J=A'\Sigma^{-1}A$ が対角行列であるということに基いている。

これらの方程式から直接に A，Σ の最尤推定量を求めることはできないので，A，Σ に適当な初期値を与えて繰返し法により問題を解かなければならない。そのためにこれらの方程式を繰返し法に適するように変形する。

A の各列を $\boldsymbol{\lambda}_{(1)}, \boldsymbol{\lambda}_{(1)}, \cdots, \boldsymbol{\lambda}_{(m)}$ であらわし，また，J の対角要素を $j_1, j_2, \cdots j_m$ であらわすことにする。

まず，第1因子負荷量 $\lambda_{(1)}$ について考える。(4.17) より，

$$\boldsymbol{\lambda}_{(1)}=\frac{1}{j_1}(A-\Sigma)\Sigma^{-1}\boldsymbol{\lambda}_{(1)} \tag{1}$$

また，(4.18) より，

$$j_1{}^2=\boldsymbol{\lambda}_{(1)}\Sigma^{-1}(A-\Sigma)\Sigma^{-1}\boldsymbol{\lambda}_{(1)} \tag{2}$$

そこで(1)の右辺に $\boldsymbol{\lambda}_{(1)}$，$\Sigma$ の初期値を代入し，新しい $\boldsymbol{\lambda}_{(1)}$ を求める。ただし，このとき j_1 は(2)より計算する。

次に，第2因子負荷量 $\lambda_{(2)}$ については，(4.17) より，

$$(\boldsymbol{\lambda}_{(1)}, \boldsymbol{\lambda}_{(2)})\begin{pmatrix} j_1 & \boldsymbol{\lambda}_{(1)}'\Sigma^{-1}\boldsymbol{\lambda}_{(2)} \\ \boldsymbol{\lambda}_{(2)}\Sigma^{-1}\boldsymbol{\lambda}_{(1)} & j_2 \end{pmatrix}=(A-\Sigma)\Sigma^{-1}(\lambda_1, \lambda_2)$$

であるから，

$$\boldsymbol{\lambda}_{(1)}\boldsymbol{\lambda}_{(1)}'\Sigma^{-1}\boldsymbol{\lambda}_{(2)}+j_2\boldsymbol{\lambda}_{(2)}=(A-\Sigma)\Sigma^{-1}\boldsymbol{\lambda}_{(2)}$$

すなわち，

$$\boldsymbol{\lambda}_{(2)}=\frac{1}{j_2}(A-\Sigma-\boldsymbol{\lambda}_{(1)}\boldsymbol{\lambda}_{(1)}')\Sigma^{-1}\boldsymbol{\lambda}_{(2)} \tag{3}$$

が得られる。また，このとき，

$$j_2{}^2=\boldsymbol{\lambda}_{(2)}'\Sigma^{-1}(A-\Sigma-\boldsymbol{\lambda}_{(1)}\boldsymbol{\lambda}_{(1)}')\Sigma^{-1}\boldsymbol{\lambda}_{(2)} \tag{4}$$

であることが容易に確かめられる。そこで，(3)の右辺に $\boldsymbol{\lambda}_{(2)}$，$\Sigma$ の初期値および上で得た $\boldsymbol{\lambda}_{(1)}$ を代入し，新しい $\boldsymbol{\lambda}_{(2)}$ を求める。ただし，このとき j_2 は(4)から計算される。

以下全く同様にして，一般に第 i 因子負荷量の新しい推定値を得るには次式を利用すればよい。

$$\boldsymbol{\lambda}_{(i)}=\frac{1}{j_i}(A-\Sigma-\sum_{k=1}^{i-1}\boldsymbol{\lambda}_{(k)}\boldsymbol{\lambda}_{(k)}')\Sigma^{-1}\boldsymbol{\lambda}_{(i)} \tag{5}$$

$$j_i{}^2=\boldsymbol{\lambda}_{(i)}'\Sigma^{-1}(A-\Sigma-\sum_{k=1}^{i-1}\boldsymbol{\lambda}_{(k)}\boldsymbol{\lambda}_{(k)}')\Sigma^{-1}\boldsymbol{\lambda}_{(i)} \tag{6}$$

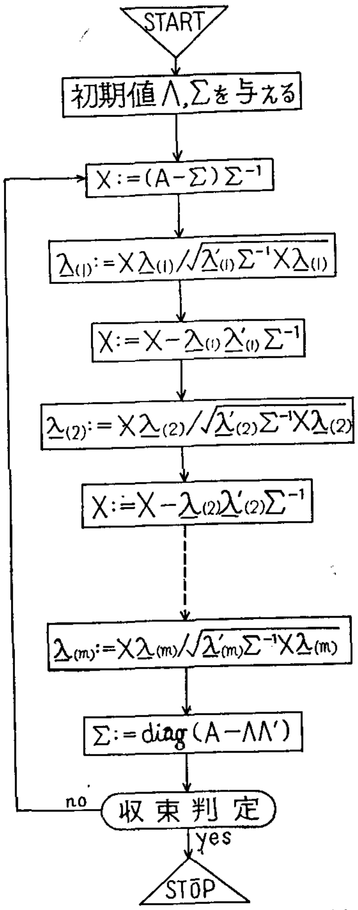
このようにして，$\boldsymbol{\lambda}_{(1)}, \boldsymbol{\lambda}_{(2)}, \cdots, \boldsymbol{\lambda}_{(m)}$ の新しい推定値を求めることができる。

次に (4.15) を変形した。

$$\Sigma=\mathrm{diag}(A-\Lambda\Lambda') \tag{7}$$

に今求めた $\boldsymbol{\lambda}_{(1)}, \boldsymbol{\lambda}_{(2)}, \cdots, \boldsymbol{\lambda}_{(m)}$ を代入し新しい Σ の推定値を求めることができる。こうして求めた Σ をもとにして，再び $\boldsymbol{\lambda}_{(1)}, \boldsymbol{\lambda}_{(2)}, \cdots \boldsymbol{\lambda}_{(m)}$ の新しい推定値を求めるわけである。収束した結果が得らさるまで，このループを繰返す。しかしながら，一般に収束は非常に遅い。特に初期値の与え方がわるい場合には遅い。

以上を流れ図にあらわすと次のようになる。



(注) 上記のフロー・チャートは東京大学吉沢正氏による。

<参 考 文 献>

〔1〕 B.A. Albert, "The matrices of factor analysis," Proc. Nat. Acad. Sci., Vol. 30 (1944), pp. 90—95.

〔2〕 A.A. Albert, "The minimum rank of a correlation matrix", Proc. Nat. Acad. Sci., Vol. 30 (1944), pp. 144—146.

〔3〕 T.W Anderson, An Introduction to Multivariate Statistical Analysis, New York, Wiley, 1958.

〔4〕 T.W. Anderson, "Asymptotic theory for principal component analysis", Annals of Math. Stat., Vol. 34 (1963), pp. 122—148.

〔5〕 T.W. Anderson and H. Rubin, "Statistical inference in factor analysis", Proc. Third Berkeley Symp. on Math. Statist. and Prob., Vol. 5 (1956), pp. 111—150.

〔6〕 M.S. Bartlett, "The statistical conception of mental facters", Brit. Jour. Psych., Vol. 28 (1937), pp. 97—104.

〔7〕 M.S. Bantlett, "Tests of significance in factor analysis", Brit. Jour. Psych. (Stat. Sec.), Vol. 3 (1950), pp. 77—85.

〔8〕 M.S. Bartlett, "A further note on tests of significance in factor analysis", Bsit. Jour. Psych. (Stat.), Vol. 4 (1951), pp. 1—2.

〔9〕 M.S. Bartlett, "The effect of standerdization on a χ^2 approximation in factor analysis", Biometrica, Vol. 38 (1951), pp. 337—344.

〔10〕 H. Hotelling, "Analysis of a complex of statistical variables into principal components", Jour. Educational Psych., Vol. 24 (1933), pp. 417—441, 498—520.

〔11〕 H.H. Harman, Modern Factor Analysis, Chicago, University of Chicago Press, 1960.

〔12〕 K.G. Jöreskog, Statistical Estimation in Factor Analysis, Uppsala, Almqvist and Wiksell, 1963.

〔13〕 M.G. Kendall, A Course in Multivariate Analysis, London, Charles Griffin, 19 57.

〔14〕 D.N. Lawley, "The estimation of factor loading by the method of maximum likelihood", Proc. Roy. Soc. Edin., Vol. 60 (1940), pp. 64—82.

〔15〕 D.N. Lawley, "Further investigations in factor estimation", Proc. Roy. Soc. Edin., Vol. 61 (1942), pp. 176—185.

〔16〕 D.N. Lawley, "A modified method of estimation in factor analysis and some large sample results", Uppsala Symposium on Psychological Factor Analysis, 17—19 March 1953, Uppsala, Almqvist and Wiksell, 1953, pp. 35—42.

〔17〕 D.N. Lawley, "A statistical examination of the centroid method", Proc. Roy. Soc. Edin., Sec. A, Vol. 64 (1955), pp 175—189.

〔18〕 D.N. Lawley, "Estimation in factor analysis under various initial assumptions", Brit. Jour. Stat. Psych., Vol. 11 (1958), pp. 1—12.

〔19〕 D.N. Lawley and A.E. Maxwell, Factor Analysis as a Statistical Method, London, Butterworths, 1963.

〔20〕 C.R.Rao, Advanced Statistical Methods in Biometric Research, New York, Wiley, 1952.

〔21〕 C.R.Rao, "Estimation and tests of significance in factor analysis", Psychometrica, Vol. 20 (1955), pp. 93—111.

〔22〕 C.R. Rao, "The use and interpretation of principal component analysis applied research", Stanford University Technical Report, No. 9 (1964), pp. 1—50.

〔23〕 L.L. Thurstone, Multiple-factor Analysis, Chicago, University of Chicago Press, 1947.

〔24〕 P. Whittle, "On principal components and least square methods of factor analysis", Skand. Aktuar., Vol. 35 (1952), pp. 223—239.

〔25〕 G. Tintner, "Econometrics" John Wiley 1959.

〔26〕 第8回日科技連数学計画シンポジウム"因子分析について" (1965年11月箱根小涌谷) 報告書 (ガリ版)

〔27〕 竹内啓 "多変量解析法の問題点—経済—データへの応用の観点から" 経済学論集, 1965年3月