日本における经营管理选集

コンピュータによって

10

管 理 工 学

下　册

《数 理 の 技 法》

日本における经営管理选集

コンピュータによって

X

《管理工学》

下　册

《数理の技法》

1981

目　　次

飽戸　弘

因子分析とはなにか

——その特徴と問題点——

I．因子分析の特徴

1．多変量解析における因子分析の位置

社会科学研究における多変量解析の適用は，近年にいたって急速に普及しつつある．政治学，経済学，経営学，社会学，社会心理学，心理学，法律学など，さまざまな分野で，多変量解析が活用されている．

これらの多変量解析の手法には，大きくわけて３つのクラスターがある．すなわち，a．**量の推定**，b．**質の推定**，c．**分類**の３つである　また用いることができる変数の種類によって，**量的変数**のみしか扱えないモデルと，**質的変数**が含まれても扱うことができるモデルとに２分することができる．これを組み合わせると，社会科学研究によく用いられている多変量解析の手法は，表１のように分類することができよう．

表１　おもな多変量解析の手法

外的基準		説明変数	おもな手法
あり	量的（量の推定）	量的変数のみ	重回帰分析，重相関分析
		質的変数も可	数量化理論（第Ⅰ類）
	質的（質の推定）	量的変数のみ	判別函数
		質的変数も可	数量化理論（第Ⅱ類）
なし（分類）		量的変数のみ	成分分析，因子分析，正準相関分析
		質的変数も可	数量化理論（第Ⅲ，Ⅳ類）

こうして，まず，成分分析，因子分析は，外的基準がない場合に，変数相互間の関連を利用して，変数を分類するためのモデルである点で，外的基準のあるモデルと根本的に異なるものであることがわかる．

成分分析，因子分析と，数量化理論（第Ⅲ，Ⅳ類）とのちがいは，扱う説明変数のタイプのちがいによることも，明らかになろう

こうして，成分分析，因子分析の特徴の第一は，外的基準のない場合で，かつすべての変数が量的変数であるような場合に，それらの変数間の遠近関係（または相関関係）を手がかりにして，変数の“似たもの集め”をすること．すなわち変数をいくつかのグループに分類することを目的とした多変量解析の一つの手法であるということがわかる．

2．成分分析と因子分析

では，成分分析と因子分析では，どこがちがうのであろうか．これはさまざまな角度から論ずることができるが，もっとも基本的なちがいは，それらのモデルの

philosophy のちがいであるといえよう．

すなわち，成分分析では，まず n 個の対象 O_1, O_2, …, O_n に関して，k 個の変数 X_1, X_2, …, X_k についてのデータが与えられる．すると，それらのデータをもっともよく表現できるような新しい変数 z を，たとえば，

$$z=\sum_{i=1}^{k}\beta_i X_i \tag{1}$$

というような一次結合で定義してやり，k 個の変数による情報をうまくこの z に吸収できれば，以下この k 個の変数のかわりに 1 個の z という合成変数をつかって，n 個の対象を分類することができるはずであると考える．

対象 O_1, O_2, …, O_n を変数 X_1, X_2, …, X_k を手がかりとして分類するために，X_1, …, X_k のすべてとの相関係数が最大になるような新しい z という 1 つの変数を合成し，その z を手がかりにして分類しようとするのである．それで不十分なら，${}^{(1)}z$, ${}^{(2)}z$ という 2 つの変数に合成し，2 次元の分類を試みるし，それでも不十分なら，${}^{(3)}z$, ${}^{(4)}z$, …, ${}^{(s)}z$（もちろん $k\gg s$）といった s 次元空間での分類を考えてもよい．いずれにしても，まずデータ (X) が与えられ，それからモデル (z) が求められる．

$$\begin{array}{c} \\ O_1 \\ O_2 \\ \vdots \\ O_n \end{array}\begin{array}{c} X_1,\ X_2,\ \cdots\cdots,\ X_k \\ \begin{bmatrix} x_{11}, & x_{12}, & \cdots\cdots, & x_{1k} \\ x_{21}, & x_{22}, & \cdots\cdots, & x_{2k} \\ \vdots & & & \vdots \\ x_{n1}, & x_{n2}, & \cdots\cdots, & x_{nk} \end{bmatrix} \end{array} \Rightarrow \begin{array}{c} z \\ \begin{bmatrix} z_1 \\ z_2 \\ \vdots \\ z_n \end{bmatrix} \end{array} \tag{2}$$

それに対して，因子分析の方は，はじめに，“精神能力”や“態度”や“イメージ”などについての，心理学的，または社会科学的モデルがまず設定される．スピアマンの“すべての精神能力には，知能というような一つの一般因子 general factor が存在する”というモデル，サーストンの，群因子 group factor についてのモデルなど，まずモデルがあり，得られたデータがそのモデルに適合するかどうかをチェックし，そのモデルにデータが適合する場合に限り，因子分析の計算を実行する．

こうして，どんなデータであろうと，データ X が与えられれば，それからモデル z を求めていける成分分析（正準相関分析も，同様であるが）と，まずモデルがあって，そのモデルにうまくデータがフィットするときにのみ計算が意味をもつ因子分析とは，発想がまったく異なるところに着目されたい．社会科学研究のなかで，問題がまずあって，それを解くために方法が開発されるという数少ないモデルの一つなのである．

3. 因子分析のモデル

こうして，因子分析においては，さまざまな研究主題，研究領域の心理学的理論，社会科学的仮説にあわせて，多くのモデルが開発されてきた．

スピアマン (Spearman) の「二因子説」
(two-factor theory)

トムソン (Thomson) の「見本説」
(sampling theory)

ケリー (Kelley) の「主軸法」
(method of principal axes)

ホテリング (Hotelling) の「主成分法」
(method of principal components)

バート (Burt) の「総和法」
(method of summation)

ホルツィンガー (Holzinger) の「Bi-因子法」
(bi-factor method)

サーストン (Thurstone) の「重心法」
(centroid method)

ガットマン (Guttman) の「ラデックス法」
(the Radex model)

などがその代表的なもので，因子分析というのはこれらの方法（モデル，技法）の総称であるという点でも上述の成分分析や正準相関分析などが単一のモデル，単一の手法に対応したものであるのと異なっている．

4. サーストンのモデル

わが国でもっともよく用いられているサーストンのモデルについて，ごく簡単にその概略をまとめると，つぎのようになろう．

step 1：変数間の相関係数を総あたりで求め，相関行列 (correlation matrix) R を求める．

$$\begin{array}{c} \text{変数：} X_1,\ X_2,\ \cdots,\ X_n\ \text{変数} \\ R=\begin{bmatrix} r_{11} & r_{12} & \cdots & r_{1n} \\ r_{21} & r_{22} & \cdots & r_{2n} \\ \vdots & & & \vdots \\ r_{n1} & r_{n2} & \cdots & r_{nn} \end{bmatrix}\begin{array}{c} X_1 \\ X_2 \\ \vdots \\ X_n \end{array} \end{array} \tag{3}$$

仮説1：いまテスト j での個人 i の得点 S_{ji} は，r 個の共通因子（群因子）のウェイトづけられた和，

$$S_{ji}=f_{j1}x_{1i}+f_{j2}x_{2i}+\cdots+f_{jr}x_{ri} \tag{4}$$

と考え，

仮説2：したがって，テスト j の分散は，

$$\sigma_j^2=\underbrace{f_{j1}^2+f_{j2}^2+\cdots+f_{jr}^2}_{\substack{\text{共通性}(h_j^2) \\ \text{(communality)}}}+\underbrace{\sigma_{js}^2+\sigma_{je}^2}_{\substack{\text{固有性}(u_j^2) \\ \text{(uniqueness)}}} \tag{5}$$

（ただし，σ_{js}^2, σ_{je}^2 はそれぞれテスト j の特殊分散，誤差分散とする）と定義し，さらに，

仮説3：これらのウェイト F と相関係数 R との関係は，

$$r_{ij}=\sum_{k=1}^{r} f_{ik}f_{jk} \tag{6}$$

と仮定して，

step 2：求められた R を $F\times F'$ に分解して，ウェイト F を求める．この F のことを因子負荷量（factor-loading）行列という．すなわち，$F\times F'\rightarrow R$ が得られたと考え，R から逆に F を求める．これがサーストンの因子分析である．

II. 因子分析の問題点

このように，数学的には，3つの基本仮説にもとづいて，R を $F\times F'$ に分解することが，因子分析の特徴であるが，このプロセスでの因子分析の基本的な問題点について，二，三，整理しておこう．

1. 因子負荷量行列（F）と相関行列（R）の関係

R と F の関係は (6) 式したがって (7) 式のように想定されているが，この必然性については，批判が多い．一つのモデルであることを，認識しておく必要があろう．

$$
\begin{array}{r}
 \\ \\ \text{テスト}1\\ 2\\ 3\\ 4\\ 5
\end{array}
\begin{array}{c}
\begin{array}{cc}\text{第I} & \text{第II}\\ \text{因子} & \text{因子}\end{array}\\
\boxed{\begin{array}{cc} f_{11} & f_{12}\\ f_{21} & f_{22}\\ f_{31} & f_{32}\\ f_{41} & f_{42}\\ f_{51} & f_{52}\end{array}}\\
A
\end{array}
\times
\begin{array}{c}
\boxed{\begin{array}{ccccc} f_{11} & f_{21} & f_{31} & f_{41} & f_{51}\\ f_{12} & f_{22} & f_{32} & f_{42} & f_{52}\end{array}}\\
A'
\end{array}
$$

$$
=\begin{array}{c}
\begin{array}{ccccc}\text{テスト}1 & 2 & 3 & 4 & 5\end{array}\\
\boxed{\begin{array}{ccccc}
(h_1^2) & r_{12} & r_{13} & r_{14} & r_{15}\\
r_{21} & (h_2^2) & r_{23} & r_{24} & r_{25}\\
r_{31} & r_{32} & (h_3^2) & r_{34} & r_{35}\\
r_{41} & r_{42} & r_{43} & (h_4^2) & r_{45}\\
r_{51} & r_{52} & r_{53} & r_{54} & (h_5^2)
\end{array}}\\
R
\end{array}
\begin{array}{l}
 \\ 1\ \text{テスト}\\ 2\\ 3\\ 4\\ 5\\
\end{array}
\tag{7}
$$

2. 共通因子の個数と相関行列の階数

さて，スピアマンの「四価差」が0になるというスピアマンモデル適用可能の条件は，実は R の階数が1であるということにほかならないことにサーストンは気づいた．こうして一般に与えられた p 次のマトリックス R の階数が m のとき（$m\geqq p$），心理学的共通因子の数は m 個あることが証明された．

このときの R の対角線上には共通性 h^2 が入っている．すなわち，対角線上に，自己相関 $r_{ii}=1.00$ ではなく h^2 を入れるときに p 次の R の階数 m が，p より小さくなりうる．これが，対角線上に1.00を入れる成分分析と，根本的にちがう点の一つである．ここで R の次数 p と，階数 m との間には

$$2m+1-\sqrt{8m+1}\leqq 2p\leqq 2m+1+\sqrt{8m+1} \tag{8}$$

なる関係があるので，この条件をみたす p と m を求めてみると，次のようなものがある．

表 2　相関行列の次数と階数の関係

p（R の次数）	2	3	4	5	6	8	10	15
m（R の階数）	1	1	2	3	3	5	6	10

重要なのは，$m<p$ なる R の階数 m（心理学的に重要な共通因子の個数）を知ることなのである．しかし，スピアマンモデルのように，$m=1$ であったり，p の次数も階数も小さい場合には R の階数はかんじょうできるが，100行100列のマトリックスで階数が40とか50などというときには，R の階数など実際には計算できない．

3. 共通性 h^2 の推定

この共通因子の数を知るためには，対角線上に1.00ではなく h^2 を入れて因子分析を行なうことが必要なのであるが，この h^2 が，実は因子負荷量行列 F がすべて求められてはじめて算出可能なので，計算を始めるときにはわからない．そこで h^2 の近似値を何らかの方法で推定してやり，それを R の対角線上に入れてやらなければならない．この方法として，相関行列のなかで，

① その行（または列）のなかの r の最大の値を代用する

② その行（または列）のすべての r の平均値を代用する（アベロイド averoid 法という）

という2つの方法が経験的に用いられている．

理想的には，h^2 の第一近似値 $h^2_{(0)}$ を，①または②の方法で推定し，R の対角線上にその値を入れて因子分析を行ない，共通因子を $m_{(1)}$ 個抽出し（$m_{(1)}<p$），得られた因子負荷量行列 $F_{(1)}$ より $h^2_{(1)}$ を算出し（これが第二近似値）その値をもとの R の対角線上に入れなおし，もう一度因子分析を行ない，$m_{(2)}$ 個の共通因子を抽出し（$m_{(2)}\leqq m_{(1)}<p$），因子負荷量を算出する．以下この操作をくり返し，$m_{(k-1)}\fallingdotseq m_{(k)}$ でかつ $F_{(k)}$ が収斂したとき，この $m_{(k)}$，$F_{(k)}$ をもって，m，F と考え，解釈に供するというようなことを考えるべきであろう．しかし実

際には上述の①または②で $h^2_{(0)}$ を推定し，その結果で間に合わせることが多い（費用の関係上，だいたい経験的にこれで間に合うと考えられているので）．

このような共通性の推定の方法はいまだ未解決のままなのである．もちろんこのような方法でたいへんよく収斂したとしても そこで得られた h_{∞} は，真の共通性 h^2 ではなくて，あくまでもそれは h^2 の近似値なのであるが．

4. 因子負荷量 (factor loading) の算出

共通性が推定されると，つぎは R を $F \times F'$ に分解する計算に入るのであるが，ここでは主として，次の２通りの方法が用いられる．すなわち，

① 主成分法 (principal component analysis) （回帰直線ではなく主軸Ⅰ Ⅱに軸を通していく方法）（図1参照）

② 完全セントロイド法(complete centroid method)（重心に軸を通していく方法）（図2参照）

ケンドル (M.G Kendall) などは，よほどのことがないかぎり，②よりも①の方を用いることをすすめている．数学的には①の方がすっきりしており，厳密だが，解釈に難点があるので，心理学者は②の方を好んで用いる．このほかにもいろいろなモデルがあるが（1の1参照），このような数学的厳密さと，社会科学的解釈のしやすさとを両立させる新しいモデルが，求められているのである．

計算そのものだけに着目すれば，R の対角線に 100 を入れて，そのマトリックスに①の主成分法で因子分析を実行してやったものが，成分分析であるといえる．ただし，その内包は根本的に異なることは前述のとおりである．この点，つねに留意されたい．

5. 因子抽出を打ち切るとき

さて，一般のデータでは，上述のように一回の計算で因子負荷量を抽出してしまうので，R の階数は厳密には計算できない．階数 m 以下のところで打ち切るのがふつうであるが，この打ち切るべき時期については，明確な基準がない．よく用いられる基準は，

① タッカーのファイ (Tucker's Phi)

残余行列 (residual matrix) の絶対値の和の減少率をみる方法．すなわち，

$$\phi_{(s+1)/s}=\frac{\sum|\rho_{s+1}|}{\sum|\rho_s|}=\frac{\sum|\sum_{r,s+1}|+\sum h^2 \text{resid.}}{\sum|\sum_{r,s}|+\sum h^2_{\text{re-est}}} \tag{9}$$

$$\phi>(n-1)/(n+1) \tag{10}$$

となったら s 因子で打ち切ってよいというもの．

② ハンフリーの方法 (Humphrey's Rule)

因子行列表のある列のうちの２つの最高の因子負荷量の積$\leqq 2\times\frac{1}{\sqrt{N}}$

ならば，因子は有意でない．したがって，その前までで因子の抽出は打ち切ってよいというもの（この基準は非常に辛いので，あまり用いられない）．

③ クームの判定法 (Coomb's Criterion)

これは R の値がすべてノン・ネガティブのとき（0にちかいマイナスは0と考える）用いられる．

符号反転後残余行列に残ったマイナスの r の個数を数える．

この数がクームが算出している臨界値と有意な差がな

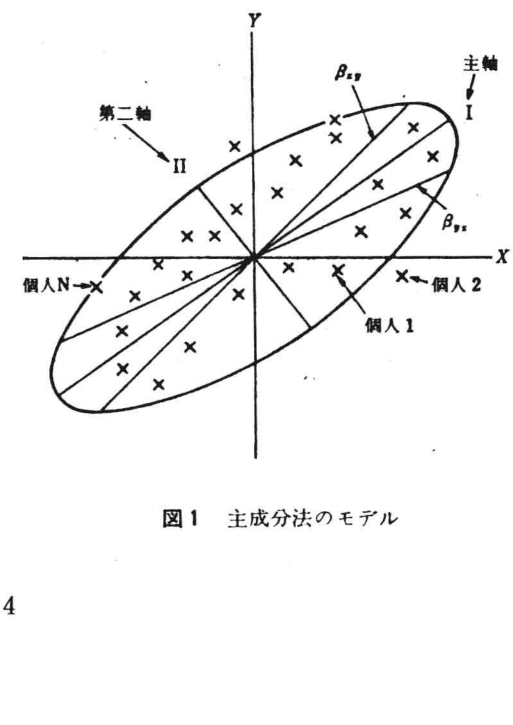

図１　主成分法のモデル

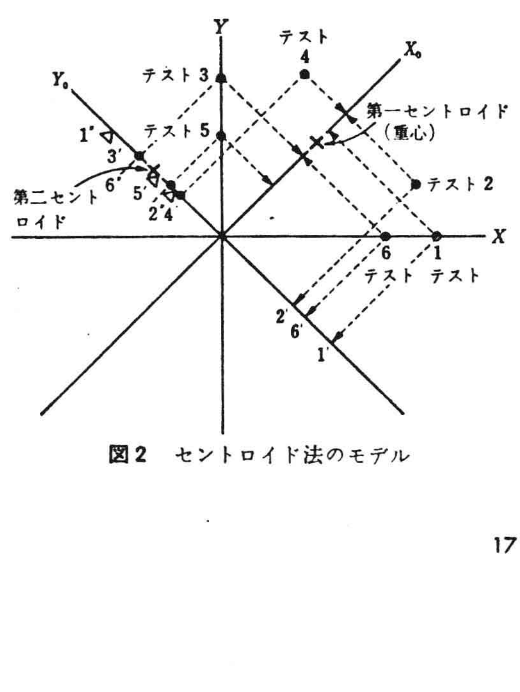

図２　セントロイド法のモデル

いときには，有意な因子はすでに抽出されつくしたと考える，というもの．

これらはいずれもまったく正しい基準というようなものではない．一つの立場による一つの基準である．このほか，いろいろな基準が考えられている（バーノン Vernon は 25 の基準を紹介しているほどである）．なかには，グラフや図で表現し，理解できるのは三次元空間までであるから，どんなデータでも三因子まで抽出したら，そこで因子抽出を打ち切るなどという社会科学者もたくさんいるというのが，笑えない現状なのである．そうかと思うと，いつもアイテの数だけ因子を抽出しつくしてしまうものもいる．そんなことならもとのローデータをそのまま考察すればよいし，どうしても多変量解析をしたければ成分分析を行なった方がよほどすっきりしている．なにも因子分析を行なう必要はない．R の元数より少ない因子におさめるところがミソなのだから．いずれにしても因子抽出の打切りは，このような基準を，いろいろ算出してみて，研究者がさいごにはそれらの結果をみながら，主観的に判断して因子抽出を打ち切るということしか，いまのところ方法はない．

6. 軸の回転

軸の回転は，数学的・統計的には何の意味もないが，心理学的にはきわめて重要である．バートらのように回転を否定する考えもあるが，一般に回転は心理学的にはきわめて重要である．特にセントロイド法では重要になる．というのは選ばれたテストや質問項目の性質によって，重心が変わってしまうので，回転によって不変性を獲得しなければならないからである．回転にあたって，もっとも重要な基準は，

① サーストンの単純構造 (Simple Structure)

であろう．これは，次のようなものである．

1) 回転因子行列の各行には少なくとも 1 個の 0 が含まれている．
2) 回転因子行列の各列には少なくとも r 個の 0 が含まれている．
3) 回転因子行列のすべての列の中から，2 列ずつ取って対にしたとき，一方が 0 なら他方が 0 でないようなテストまたは質問項目がいくつかある．
4) 上のようなすべての列の対に関して，その負荷量が両方の列とも 0 になるような対が，いくつか含まれている．
5) 同じくすべての列に関して，両方ともかなり大きい負荷量をもった対は，きわめて，少数でなければならない．

その他では

② 正のマニフォールド (positive manifold) の基準

では，できるだけすべての負荷量がプラスまたは 0 になるよう回転することを考えている．

③ キャッテル (Cattell, R. B.) の基準

さらに，キャッテルなどはずっと芸術的になって，

1) 心理学的発見と一致するように，
2) 過去の因子分析から得られた結果と一致するように，
3) 軸がプロットの中央をを貫通するように，
4) 連続的に決定された直交軸因子に一致するように，
5) 一般心理学的予想に適合するプロフィールがえられるように，
6) 同じような割合のプロフィールがえられるように，回転するべしとしている．

④ 主観的回転より分析的回転へ

回転に関して，ここ数年でのもっとも大きなできごとは，超大型電子計算器の開発・発展につれて，統計的基準によって，客観的に回転を行なうための理論と方法が開発され，従来の分度器とコンパスによる主観的な「グラフ法による回転」(graphic rotation) はすっかり影をひそめてしまったことであろう．このような客観的基準による分析的回転 (analytical rotation) の最初の試みは，得られた因子行列の行に関して，サーストンの単純構造を統計的に実現するように回転しようというものであった．すなわち，行ごとにできるだけ多くの因子で，因子負荷量が 0 または 0 に近い値をうるように，ということは，のこりの因子負荷量はできるだけ高い負荷量をうるように，回転を行なうという方法であった．キャロール (Carroll)，ニューハウス (Neuhaus)，リグレイ (Wrigley)，サウンダーズ (Saunders)，ファーグソン (Ferguson) らが，さまざまな方法を開発したが，いずれも代数的には同値であることが証明され，クォーティマックス (Quartimax) 法とよばれている．

それに対してカイザー (Kaiser) は，列に関しての分散が最大になるような回転を考えた．すなわち個々の因子ごとに，高い因子負荷量をもつ変数はごく少数になり，他の変数はすべて 0 に近くなるように回転するもので，クォーティマックスに対してバリマックス (Varimax) 法とよばれている．

これらのクォーティマックス，バリマックスは，直交回転だけでなく，斜交回転にも適用する工夫がこらされている．キャロールのクォーティミン (Quartimin) 法およびコヴァリミン (Covarimin) 法，その両者を組み合わせたオブリミン (Oblimin) 法など，さまざまなモデ

ルが開発されている．

こうした分析的回転によって，因子構造はつねに一定に保たれるという点では，心理学的には非常に重要であることは前述のとおりである．しかし，このような“近代化”の副産物として自分の仮説を，名人芸的グラフ回転で証明してみせるといったよろこびを社会科学者から奪ったということも，事実である．

7. 因子の解釈

因子分析においては，軸の回転が非常に重要なのである．数学的・統計的にはどうでもよいことであるが，心理学的には解釈がより容易になり，妥当になるように回転をするということが，非常に重要なのである．

このことをもっとも徹底して実現するためには，あらかじめ理論的にいくつかの軸を設定しておいて，そのような軸にもとづいてアイテムを設定し，因子分析し回転によってそのような軸を出していくという，アイゼンク (H. J. Eysenck) らの方法は，もっとも因子分析に適した活用方法といえよう．

すなわち，事実探索 (fact finding) 的用法（これはむしろ成分分析でよい）ではなく，仮説検証 (hypothesis testing) 的用法にこそ（それもできるだけ少ない軸におさまるような理論仮説が），因子分析には適しているのであり，そのためにはデータの事後解釈を無手勝流にやるのではなく，軸についての事前の設定とその吟味に，重点をおくべきであろう．

こうして，因子分析に適した研究領域として，次のような領域があったし，今後もこのような領域において，因子分析は有効であろう．

① 知能および精神能力の研究

② パーソナリティの研究

表 3　ホルツィンガーとハーマン(1941)の地域研究の二因子仮説

	第Ⅰ因子	第Ⅱ因子
1. Lewis（民主党候補）への投票率	.69	−.28
2. Roosevelt（民主党候補）への投票率	.88	−.48
3. 投票（同じ政党に投票した率）	.87	−.17
4. 住居費（メディアン）	−.88	−.09
5. 自宅所有率	.28	.65
6. 失業率	.89	.01
7. モビリティ(現住所に1年以上居住率)	−.66	−.56
8. 教育度（在学10年以上の者の率）	−.96	−.15
分散比	62%	14%

(147 選挙区，アペロイド法，セントロイド法)

③ 社会的態度の研究

④ 意味，およびイメージの研究

これらはいずれももっともポピュラーな研究領域であり，いちいち紹介するまでもあるまい．この他，心理学以外の領域でも，比較的少数の軸のなかにうまくおさまるような理論仮説の検証には，因子分析は適してにる．次に紹介しているホルツィンガーとハーマン (Harman) の研究，バートとバンクス (Banks) の研究などは，いずれもそのような因子分析の特徴をよく生かした，優れた古典的研究といえよう．

表 4　バートとバンクス(1941)の体格の三因子説

	f_1	f_2	f_3
1. 身長	+	−	−
2. 座高	+	−	−
3. 腕の長さ	+	−	−
4. 脚の長さ	+	−	+
5. 腿の長さ	+	−	+
6. 腹まわり	+	+	+
7. 腰まわり	+	+	−
8. 肩幅	+	+	−
9. 体重	+	+	−
分散比	55%	13.5%	10.1%

(2,400 人，空軍兵，セントロイド法)

8. 因子分析の適用範囲

さいごに，因子分析そのものの問題ではないが，さまざまな変量系に因子分析を適用していくための理論と方法の開発の問題が残されている．キャッテルは，因子分析が適用できる変量として6つの系を考え，それぞれに，R-技法，Q-技法，P-技法などの名称を付している．

たとえば，

データはテスト（または質問項目）×被験者の平面で（すなわち，一時点内において），

R-技法：テストを変量と考えテスト間の相関よりその因子構造を明らかにする．

Q-技法：被験者を変量と考え，被験者間の相関行列より因子構造を求める．

データはテスト×時点の平面で（すなわち，一人の被験者に対して）

P-技法：テストの各時点や環境での変化を変量とし，分析するもの，などなどを意味している．

従来のR以外の技法を適用するためにはさまざまな統計的工夫が必要とされる．Q-技法におけるQ-分類法などはその一例である．このような方向での研究も，今後

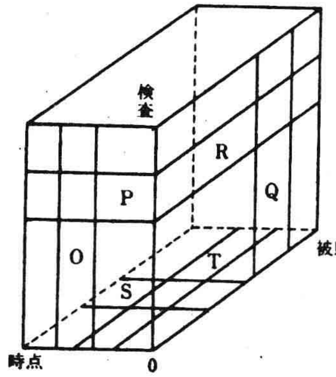

図3　キャテルの因子分析についての六技法

に残された重要な課題であろう*.

*.

本稿は，1970年2月12，13日に，統計数理研究所において行なわれたシンポジウム「データ解析の問題点」で，筆者が行なった報告"行動科学における因子分析的方法"の講演要旨に，若干手を加えてまとめなおしたものである．詳しくは飽戸(1970a)を参照されたい．

（あくと・ひろし，埼玉大助教授）

文献

飽戸弘　行動科学における因子分析的方法，『統計数理研究所シンポジウム記事』，2，1970a

飽戸弘　『イメージの心理学』潮出版社，1970b

飽戸弘　多変量解析法，『マネジメント・リサーチ・ハンドブック』，丸善出版，1967.

飽戸弘　数量化理論，『年報社会心理学』，5，勁草書房，1964.

飽戸弘　多次元解析法と数量化法(1)、(2)、『サンケイ・アド・マンスリー』，15，16，1963，1964.

飽戸弘　多次元解析の応用，『調査高等技術統計ゼミナールテキスト』，久保田宣伝研究所，1963.

Cooley. W W and Lohnes, P. R., *Multivariate Procedures for the Behavioral Sciences*, John Wiley. 1962

Fruchter, B., *Introduction to Factor Analysis*, D. Van Nostrand. 1954

Guilford, J. P., *Psychometric Methods*, McGraw-Hill, 1954

Kendall, M. G., *A Course in Multivariate Analysis*, Charles Griffin, 1965.

因子分析の解をたずねて

芝　祐順

因子分析は心理学における由緒ある方法として，ここ70年程の間に発達し，変遷をつづけてきた．いまでこそ多変量解析の一つとして，ひとり心理学の分野ばかりでなく，隣接する諸科学においても利用されるようになってきたが，それは一つにはコンピュータの使用が可能になったことによるもので，それまでの長い間，じつに忍耐強く，こつこつと生きのびてきたようにみえる．古くからの因子分析の発達をずっとみてみると，この間になされたかずかずの研究や，生まれた解の関係もよく理解できるが，いま，この時点でみると新旧とりまぜていくつもある解は必ずしも整然と並んでいるわけでなく，見通しもよくない．

人がものごとを考えるときに，なるべく原理的な方向へとつめていく場合と，逆に派生的な細かな差異を追跡していく場合とがある．前者の方向にそってみれば，因子分析は，データを成分とする行列を，より階数のひくい行列の積によって近似しようとする試みの一つにすぎない，ということになるだろう．このような見方からするといくつもある解に，もっともらしい名前をいちいちつけて区別するなどということは，いかにもわずらわしいことと思われる．しかし，因子分析の解が何を目的として分化してきたのか，いろいろな解の異同はどこにあるのかということは，因子分析をデータ処理の方法として利用する立場からは大変興味ぶかいことである．

そこで，それぞれの解がどんな特徴をもつものであるか簡単に整理してみよう．ふつう因子分析の解というと，いわゆる因子負荷をもとめる方法や，もとめられた因子構造をさす．ところで，たんに解の個々の名称をアルファベット順や年代順にならべていくのも能がない．解の差異を特徴づけている問題が何なのか，まず，これをあげ，その問題に関連する解を順次紹介していこう．その分類もあまり明快にはいかないので，いわば因子分析地方に散在する村々をはじから訪ねていく，というような感じである．

はじめの村は，因子分析の中でも，もっとも古い解のあつまりである．その生い立ちからもわかるように，知能理論と密接な関係をもっている．すなわち，一つの顕著な共通因子をもって各変量の大部分を説明し，のこりの変動は各変量に固有の独自因子として因子構造を考え

特集　因子分析

る二因子解や，複数の共通因子で因子構造を説明しようとする重因子解は，この村での代表的存在である．このほか，知能の二因子説を修正した形の階層因子説，階層群因子説では，まず，一般因子からはじめて順次に小さくなる群による因子の階層的構造がもとめられるが，そのもっとも簡単な構造の因子解として双因子解がある．この解では一つの優勢な一般因子とあといくつかの群因子とによって共通因子が構成される．そして各変量は一般因子と一つの群因子に高い負荷を示すが二つ以上の群因子に同時に高い負荷をもたないことをめざす．知能理論と密着した議論では，それが知能の構造に関するものなのか，方法としての因子分析に関するものなのか，はっきりわけにくくさえなってしまう．こんなこともいくらか尾をひいてきたが，大すじとしては数量的方法としての因子分析は，知能理論からはなれていった．

因子分析の発生地ともいえるこの村をはなれ，以下では因子分析法が遭遇してきた問題別に，新しい村々を訪ねてみよう．因子構造は相関行列（ただし，共通性をその対角成分に入れたもの）をつぎのように分解することによってえられる．

$$R-D^2=ALA'$$

ここで，R は相関行列を，D^2 は独自性を対角成分とする対角行列を，A は因子負荷の行列を，そして L は因子間の相関の行列をあらわす．これを満足する解 A は一つに定まらず，任意の他の解は変換 T によって

$$B=AT$$

とあらわされる．この基本的な過程のさまざまな段階において，さまざまな異なった解が生まれるわけである．

まず，はじめに，この相関行列 R を用意するとき，どのような変量のモデルを考えていくかによって，いくとおりかの因子分析を区別することがある．これらは技法(technique)とよばれ，O．P．Q．R．S．T などの名称がつけられている．心理学の場合には，いくつかの変量について被験者を観測のくりかえしにとって，変量間の相関をもとめるという例がもっとも多いが，これとは対照的にいくつかの変量についての観測をもとに，被験者と被験者との相関をもとめるということもある．データを規定する要因としては，このほか，時間を異にして観測するという場合もある．何を観測のくりかえしとし，何と何との間の相関をとるかという組合せの6通りを区別したのがこの技法である．さきにあげた2つの場合のうち，はじめの組合せであるもっとも普通の組合せによる方法を R 技法とよび，もう一方の技法を Q 技法とよぶ．残りの4技法もあわせて，心理学的にはその区別は重要だし，いろいろ詳しい議論もある．しかし，これは因子分析本来の問題ではなく，いわばずっとむこうのはずれの村になるので，その名にふれる程度にして，次の村へいこう．

順序からいくと，いきなりもっとも重要な問題に入りこむことになる．上の式で D^2 であらわされた独自性の推定の問題，ふつう，これを相関行列の対角成分である1からひいて共通性，

$$H^2=R-D^2$$

の問題とよんでいるが，この共通性推定の方法の差異に関連して，いくつかの方法があげられる．そのもっとも簡単なものは相関行列の非対角成分のうち各行ごとに絶対値最大のものを共通性の推定値としてセントロイド法（後に少し詳しくふれる）を行なうものだが，その一つに完全セントロイド法がある．これを出発点として因子負荷をもとめたら，その値から共通性を逆算して新しい共通性の推定値をもとめる．それを相関行列に入れてふたたび因子負荷をもとめる．この計算過程をくりかえして収束をまつのが，いわゆる逐次近似（セントロイド法ともいう）である．これに対し非対角成分の各行の絶対値の平均値をもって共通性の推定値とする方法はアヴェロイド法とよばれている．共通性を直接推定するかわりに共通性の値を指定せず，相関行列の非対角成分だけについて，1因子ごとに残差が最小となるように解をもとめていく方法がミンレス法である．このほか最尤推定解も共通性の推定と密接な関係をもつ解であるが，これはのちにのべるとして，もう一つ，共通性に関係深いものとしてイメージ因子分析をあげることができる．イメージ因子分析はやはり直接に共通性 H^2 を推定するのではないが，偏回帰的方法によって各変量の共通成分による変動（これをイメージと呼んでいる）をもとめ，この共通成分の間の共分散行列を因子分析することによって，共通因子の因子負荷をもとめようとするものである．

以上のような共通性に関係の深いいくつかの解は，たがいにまとまりがなく，村をなしているとはいいがたい．それぞれ別のいくつかの村に共通性の問題をかかえて入

りこんでいるといった格好である．

おそらく因子分析地方の，もっとも大きな村の一つは，つぎにのべる直接解の村である．直接解とは相関行列（共通性の推定が終わったとして）から直接因子負荷をもとめる解をいう．ここでは，そのとき因子を直交するようにとるか，斜交も許すかによって解はわかれる．また，実際上さきの式のように相関行列を分析しても，これを完全に行なうのではなく，相関行列の次数にくらべて，比較的少数の因子によってよい近似をえようとする．したがって，その因子の数を決める仕方によっても解は特長づけられている．

まず，直交解の中で，かつてもっともよくもちいられたのがセントロイド法である．しかし，計算機の利用後はこれにかわって，もっぱら主因子解（主軸法ともいう）がもちいられるようになった．主因子解はつぎのようにしてあたえられる　相関行列$(R-D^2)$の固有値を成分とする対角行列をλとし，これに対する固有ベクトルを各列にもつ行列をQとすると，因子負荷Aは

$$A=Q\lambda^{\frac{1}{2}}$$

となる．この解はまた，因子負荷の平方和を因子ごとに最大とする解でもある．すなわち，各因子をPで表わし，変量を$j=1$からnまでとすると，因子負荷a_{pj}について

$$\sum_{j=1}^{n}a^2_{pj}$$

をとり，これを第1因子より順次最大とする解である．

セントロイド法は，その結果も主因子解によく近似することが多いが，このように因子負荷の平方和ではなく，絶対値を順次最大とする解になっている．これらを，因子負荷について，各因子ごとに順次そのモーメント（原点からの）を最大とする解であると考えると，さらに高次のモーメントによる解が予想される．4次のモーメントを最大とする解は，直接コーティマックス解とでもよぶべきもので，のちにのべるコーティマックス解の変形となる．モーメントの次数ばかりでなく，モーメントの中心のとり方によっても異なった解が生まれる．まず第1因子の因子負荷について，その平方の平均を考え，これを中心にして，この平方の2次のモーメント

$$\sum_{j=1}^{n}(a^2_{j1}-\frac{1}{n}\sum_{k=1}^{n}a^2_{k1})^2$$

を最大とするような解をもとめる．このようにして順次の直交因子をもとめていくのが直接バリマックス解である．これは，のちにのべる間接解としてのバリマックス解の変形である．因子負荷の各種モーメントを最大とする解については，国鉄労研の柏木氏らが包括的に検討中であるという．

直接解をもとめる数学的方法として，相関行列を三角行列の形の因子負荷行列に分解する平方根法は，対角法あるいは階層法などともよばれてきたが，今日では実際に因子解として利用されることはほとんどない．

このほか，群因子解では互いに相関の高い変量をあらかじめ群にまとめ，この各群のセントロイド解（ただし群ごとに第1因子のみをとる）となるように群の数だけの因子を定める．これはグループセントロイド法とよばれそのままでは一般には斜交解となるが，これを直交化することもできる．群ごとのセントロイド解をもとめるかわりに，群ごとの主因子解をとった場合にはグループ主因子解となる．

つぎに，これもまた有名な，単純構造という名の村がある．さきにあげた因子構造の変換ATは，何かの目標にむかっておこなわれる．その目標の中でももっとも重要なものは，因子と変量との関係を単純化するというものである．この単純化をごく大まかにいえば，一つの変量が二つ以上の因子において大きな負荷を示さないような因子構造にすることをさす　あたえられた条件の中で，このような意味での単純構造へ変換をするため，因子構造をまず任意の2因子の組合せでとりあげ，これを幾何学的に平面上に図にかきあらわす．これの視察によって新しい座標へ変換する．これを何とか主観的な判断に頼らず，全く客観的方法によってできないものかといろいろ工夫された．このため提案された諸方法の中で，頂点にあるものは，いわゆるバリマックス解である．これは，すでにのべた直接バリマックス解同様，因子負荷の平方平均（因子ごとにとったもの）から偏差の平方和を最大とする解である．直接バリマックスとちがうところは，直接バリマックス解はこのような最大値を因子ごとに順次もとめるという解であるという点で，ここにあげるバ

リマックス解では，初期解としてあたえられた因子数すべてについて同時にこの平方和を最大とするように変換がおこなわれる．これに対し因子負荷の平方の平均をとるときに，バリマックス解のように因子ごとに平均をとるのではなく，変量ごとに平均をとるという考え方も，やはり因子構造を単純化するときの基準として使えるが，この場合には因子負荷の変量ごとの平方の和をとるとこれが共通性となって変換に際しては定数となるので，結果的に因子負荷の4次のモーメント（原点からの）を最大とする解をもとめるという基準となる．これによる解が**コーティマックス解**である．このようにバリマックス解のための基準と，コーティマックス解の基準とを比べると，因子負荷の平方の2次のモーメントをもとめるときの中心が異なっている．これら二つの基準をあらわす式中，バリマックス基準のみにあらわれる項に全く恣意的に$^1/_2$の重み係数をつけて，両者の中間的基準とした場合えられる解を**エカマックス解**という．これをさらに一般化し重みを任意に選ぶことによってえられる一群の解を総称して**オーソマックス解**とよんでいる．その基準の意味するところはかならずしも明らかではないが，経験的には重みが0の場合（コーティマックス解にあたる）や，1の場合（バリマックス解にあたる）よりもよい結果がえられることがある．なおバリマックス解では，変換に際して共通因子の高い変量が因子構造を定めるときに強い影響をもちすぎることをおさえるため，変換をおこなうまえに，すべての因子負荷をそれぞれの変量の共通性の平方根で規準化しておくことが多い．このことを幾何学的にいえば，共通因子空間において長さ（ノルム）の異なった各変量のベクトルを，すべて同じ長さ1に方向をかえずになおし，そののちにバリマックス基準を満たすように座標軸の回転をおこなうことにあたる．この方法をとくに区別するときには，**ノーマル・バリマックス解**とよび，規準化しない方法を**ロー・バリマックス解**とよぶ．この区別を一般化するなら**ノーマル・オーソマックス解**と**ロー・オーソマックス解**の2群が区別される．

主観的な判断にしたがって幾何学的に単純構造をもとめるのは，今日ほとんどはやらないが，この主観的な判断の過程をそのまま客観化するという発想で提案された柏木氏の**ベクター法**（**ジェオマックス解**ともよぶ）は注目に値する．たんに，計算が簡単で計算所要時間が短いというばかりでなく，これによってえられる解が，バリマックス解などに比べ，むしろ従来の主観解によく一致するため，実用上高く評価されてよいものである．ベクター法の原理は，任意の2因子からなる平面の上で，セントロイド法による回転をおこない，第1因子軸より上側（第1象限）にくる変量群と，下側（第4象限）にくる変量群とに全体の変量をわける．それぞれの群ごとに，変量ベクトルの合成をおこなって，できた二つのベクトルが，それぞれなるべく大きな射影をもつように，二つの因子軸を回転する．この間，いくつかの注意が必要であるが，だいたいこの原理で全因子の組合せについて回転をくりかえしていけばよい．この方法を少し修正して，ちょうど一つのセントロイド（重心）を，二つの変量群の重心に分割し，これに2軸が近くなるような解をもとめてみるとこれもやはりよい結果をしめした．これを**バイセントロイド法**とよぶことにしよう．

単純構造付には，このようなたくさんの解に対応して，さらに，因子間の直交性を必ずしも仮定しない一群の解がある．直交解の場合には，もともと出発点となる発想を異にした解が，直交であるという条件のため，結局同じ結果となるという事情があって，むしろ解の数は少なくなっている．これに対し斜交解では，単純構造の基準のあたえ方によって，さらに多くの解が生まれた．ここではそれをいちいち説明する余裕もないので，名称だけをあげておこう．**オブリマックス，オブリミン，コバリミン，コーティミン，バイノーマミン，斜交バリマックス**などである．

単純構造への変換の解の付は，このように大変にぎやかであるが，変換の目標は単純構造ばかりとはかぎらない．特定の因子負荷を高くしたり低くしたりするように，仮説的構造を設け，これになるべく近くなるような解へ変換をするということもおこなわれる．さきにのべた群因子解も，その一種とみなすことができる．このやり方の中で，もっとも典型的な解は，あらかじめ与えられた仮設的因子負荷にもっとも近い因子構造をあたえる**プロクラステス解**であろう．この方法は目指す構造が単純構造に限らず，どんなものでもよい．もちろん，いくら仮説的にあたえられた因子構造に近づけるからといってももともとあたえられた変量の間の関係が許す範囲においてのことである．

なお，さきにあげた双因子解では第1因子の抽出には，セントロイド法や主因子解をもちいることができるが，第2以下の因子，すなわち群因子をもとめるときには，どの変量をまとめて一群にするかについて，主観的な判断が入ってくる．それをさけるために，変量間の相関関係の大小にもとづいて，群化をおこなうための基準などが提案されている．しかし，いままでにのべた諸方法を組みあわせることによって，より客観的に双因子解をもとめることができる．まず，相関行列から主因子解によって必要なだけの因子を抽出しておく，そのあと，第1因子はそのままのこし，第2因子以下についてだけオーソマックスのような直交変換を行なって単純構造をもとめる．このようにすると，第2因子以下では，一つの変量が2つ以上の因子にまたがって高い負荷量を示すことが少なくなるから，この結果と第1因子とをあわせると，双因子解の意図する構造がえられるのである．このように特定の因子空間についてだけ単純構造をもとめるというやり方を偏オーソマックスとよんではどうだろうか．はじめに一時，除いておくのは第1因子だけとはかぎらない．必要とあらば2因子以上を一時除いて変換をするということもできる．この偏オーソマックス解では，さきのノーマルバリマックス解のように規準化することは，むしろ好ましくないようで，ロー・バリマックス解，もっと一般的にはロー・オーソマックス解を利用するのがよい．

因子軸の変換に関係した分類はこのくらいにするが，以上の村々はどちらかというと，実用上の興味が高く，その評価も実際役に立つかどうかという点からなされがちである．これに対し，理論的な興味からとりあげられる村にはカノニカル因子分析やアルファ因子分析などがある．前者は各変量について，それに適宜重みをつけてえられる合成変量と，共通因子空間での変動とが最大の相関（形としてはいわゆるカノニカル相関）を示すようにまず第1因子を定め，以下，残差の変動をもちいて順次の因子を同様に定めていくものである．

アルファ因子分析では各変量の共通因子空間での変動を合成して一つの合成変量を作るとき，その合成変量のアルファ係数（テスト理論で合成変量の信頼性をあらわす係数）が最大となるようにし，これをまず第1因子とする．以下，残された共通因子空間で，やはりアルファ係数を最大とするよう因子を定めていく．アルファ係数が負にならない因子のみを心理学的に有意味であるとし，因子数をこの条件によってきめるとしている．

以上の議論にはほとんど統計的推測の問題がくみ入れられていなかった．あたえられたデータが想定された母集団からの標本であるとすれば，母集団における因子数や因子負荷は統計的な手続きによって推定されなければならない．このような考え方で最尤推定解が生まれた．このような考慮を払うことは因子分析を実用的目的から使うものにも必要なことだが，現時点ではまだこの解も理論村に籍をおいているようである．

以上の因子分析地方の散策も，その内側だけをさっとまわってきたもので，ことさらその周辺地区をさけてきた．周辺地区の外側には，古くからの村々，新興の村々ととりまぜて，どうも境界のはっきりしないものが多い．人がどこから因子分析地方ですか，とたずねてもはっきり答えにくい．このことはむしろ因子分析にとってのぞましいことで，今後領域の内外の関係はますます密接になっていくだろう．

（しば・すけより，成蹊大学助教授・文学部）

因子分析の諸法とそれらの特性

浅野長一郎 *

はじめに

生物学から心理学への橋わたしを試みた H. Spencer の時代，精神能力というものは，情動やその他の知的でない因子の影響を除けば，知能という一因子説にもとづいて，賢い人は何をやっても同様にいつも賢く，愚かな人は何をやってもいつも愚かであると一因子説が考えられていた．しかし，1904 年に C. Spearman は g と s の因子の存在をモデル化して，所謂二因子説を提唱したわけで，因子分析法の発端は，じつにいまをさる約70年にさかのぼると考えられている．

爾来，因子分析の諸法は，多種の観測特性の内で抽象的かつ本質的な原因系を探究する巧妙な統計的多変量解析法として，実質科学と数理の両面から種々の模型や立場のもとに研究開発されてきている．

とくに，近年の電子計算機の普及にともなって，情報化システムや情報処理への関心がすすむにつれ，この範疇にも入る因子分析法の適用と研究が，今日ふたたび盛んになってきている．

しかし，この因子分析法は適用の仕方や考え方のうえで，かなりの難解な点を含んでいることも事実である．この方法には，歴史的に年輪を経てきた種々の思想的背景があり，また，数理的にも少なからぬ困難を残してきている．人間の発想や意図などの思考は深遠で，現象自体もまた広範で具体的かつ個別的に複雑で，さらに数理面でもきれいごとではすまぬ吟味すべき難問を多くかかえている．このために，これらの適用や改良には，かなりの前提や取組み方に関する理解が必要になってくる．

1. 諸解法の立場と分類

因子分析の諸法には，種々の意図や立場にもとづき，異ったアプローチによる解法が展開されている．

まず，因子抽出の基準 (criterion) のたて方という，もっとも基本的な観点から，**対内的因子分析法** (internal factor analysis) と**対外的因子分析法** (external factor analysis) に分類される．前者は，多種の計測特性による一群の観測資料だけにもとづき，その資料変動の内から因子抽出をはかる．後者は，観測特性群または対象の群の数を複数にしておいて，これらの間の相対的関係を基調にした外的基準により因子を抽出しようとする．また，このような外的基準は，のちに，Kuder-Richardson

特集　因子分析