愛媛経済論集第14巻　第1号（1994）

多変量統計解析法の理論と応用（上）

三　土　修　平

Ⅰ．序

現在，計量経済学の主流を占めている重回帰モデルおよび連立方程式モデルは，いずれも，説明変数と被説明変数との区別が最初からかなりはっきりしている現象の分析に適するものである。幸いにして経済学ではその種の現象が多い。これは，経済理論がすでに長年にわたる先人の努力の結果，何が何を規定するという因果連関を少なくとも仮説としてはかなり明確に打ち出せるまでに発達してきていることによる。事実，資源の賦存量，財・サービスの生産量，価格，利子率，投資量等の経済諸変数のあいだには，これらを関数的に結びつける数学的理論がすでに多数唱えられ，覇を競い合っている。

しかし，そのような狭い意味での経済学の領域から少し視野を広げ，経営学や社会学との競界領域に目を転じてみると，そこには，こうした因果的説明にはなじまない性質の変数がたいへん頻繁に見いだされる。たとえば，企業の経営指標と呼ばれるものがいくつもあるが，それらはおおむね，業績のよい企業では数値が高く，業績の悪い企業では数値が低いという連動関係にある。しかし，どれがどれを規定するという関係にはない。勤労者の勤労意欲，貯蓄意欲などを生活面での意識との関連でとらえようとして，多くの質問項目を設定してアンケート調査を行なったときの結果の分析なども，これと同様である。多くの質問項目への回答のあいだに連動性はあるが，因果的規定関係を想定することはむずかしい。

しかも，このような，変数間に顕著な連動関係のある現象に，あえてなんらかの因果的モデルを設定し，重回帰分析を実行した場合には，いわゆる多重線形性が障害となって，係数の推定値として安定した数値が得られないことが多い。

しかし，連動関係があるということは，ただちに，全変数が１次元的につながっているということではない。直接的には観測されないが，陰から働いている２つか３つぐらいの潜在的規定要因があって，それらの組み合わせで全変数の動きが決まっている結果，たとえばn変数の場合，その観測結果のちらばりぐあいがn次元空間全体にはゆきわたらず，２次元か３次元ぐらいの低次元の範囲内にほぼ収まっているという可能性がある。こういう場合，「多重線形性がある」と指摘しただけでは，ほとんど有意義な知見は得られないわけであり，多重線形的ではあるが１次元ではないこの独得の関係を，どうとらえ，どう表現するかが課題となる。本稿でとりあげる多変量統計解析法は，こうした要請にこたえる分析方法である。

なお，多変量統計解析法（multivariate statistical analysis）と総称されるものの中には，判別分析法，数量化理論なども含まれるが，本稿でとりあげるのは主成分分析と因子分析である。

Ⅱ．主成分分析と因子分析に共通する考え方

いま，３個の変数x_1, x_2, x_3があったとして，それらが２個の要因によって規定され，

$$\left.\begin{aligned} x_1 &= a_{11}f_1 + a_{12}f_2 \\ x_2 &= a_{21}f_1 + a_{22}f_2 \\ x_3 &= a_{31}f_1 + a_{32}f_2 \end{aligned}\right\} \quad (1)$$

という関係があったとする。f_1, f_2は変数であるが，a_{11}等は定数であるとする。

これをベクトルで書くと

$$\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = f_1 \begin{bmatrix} a_{11} \\ a_{21} \\ a_{31} \end{bmatrix} + f_2 \begin{bmatrix} a_{12} \\ a_{22} \\ a_{32} \end{bmatrix} \tag{2}$$

であるから，これら3変数の組は2次元ベクトル空間をなしており，3次元全体に広がってはいないことがわかる。実際，(1)の第1式と第2式から

$$\begin{bmatrix} f_1 \\ f_2 \end{bmatrix} = \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix}^{-1} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$$

$$= \frac{1}{\Delta} \begin{bmatrix} a_{22} & -a_{12} \\ -a_{21} & a_{11} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} \dfrac{a_{22}x_1 - a_{12}x_2}{\Delta} \\ \dfrac{-a_{21}x_1 + a_{11}x_2}{\Delta} \end{bmatrix}$$

$$(\Delta = a_{11}a_{22} - a_{12}a_{21}) \tag{3}$$

が得られ，これを第3式に代入すれば

$$x_3 = a_{31}\left(\frac{a_{22}x_1 - a_{12}x_2}{\Delta}\right) + a_{32}\left(\frac{-a_{21}x_1 + a_{11}x_2}{\Delta}\right)$$

$$= \frac{a_{31}a_{22} - a_{32}a_{12}}{\Delta} x_1 + \frac{-a_{31}a_{12} + a_{32}a_{11}}{\Delta} x_2 \tag{4}$$

となるから，これが平面の方程式になっていることは明らかである。規定要因たるf_1, f_2がいくら変化しても，x_1, x_2, x_3は3次元にはゆきわたらず，平面上のみを動くのである（図1）。このように，多くの変数があっても，規定要因がそれより少ない数であれば，変数の観測値の存在する空間は，規定要因数だけの次元しかもたない次元の低い空間になる。実際は，誤差があるから，観測値はこのような次元の低い空間の中には収まりきらず，変数の個数だけの次元を

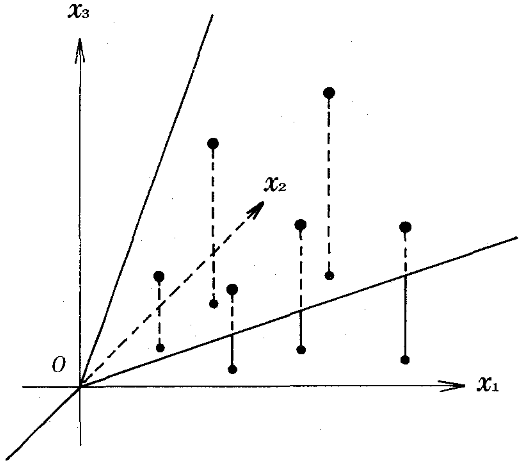

図１

もった空間の中に分布することになるだろうが，規定要因たる変数が変動することにともなって生じる低次元空間内でのちらばりの方が，誤差要因によってその空間からはみ出すはみ出し方のちらばりよりも大きいことになろう。そこで，全体の観測値のちらばりを何らかのやり方で分解することを通じて，当の低次元空間がどのあたりに存在するかの見当をつけることができるのではないかと考えられる。これが，主成分分析と因子分析とに共通する考え方である。

Ⅲ．数学的準備

主成分分析や因子分析に用いられる数学の中には，分散，共分散，相関係数，ラグランジュ乗数法，行列の固有値・固有ベクトルなど，経済学者によく知られている数学もあるが，写影作用素，スペクトル分解など，経済学者にあまり知られていない数学も含まれているので，それらを解説することから始めよう。

１．射影作用素

まず，２次元平面内での角 θ だけの回転を表わす行列 P は，「行列の第 i 列

は第i基本単位ベクトルの像である」という性質から，図2によって次のようになる。

$$P=\begin{bmatrix}\cos\theta & -\sin\theta\\ \sin\theta & \cos\theta\end{bmatrix} \tag{5}$$

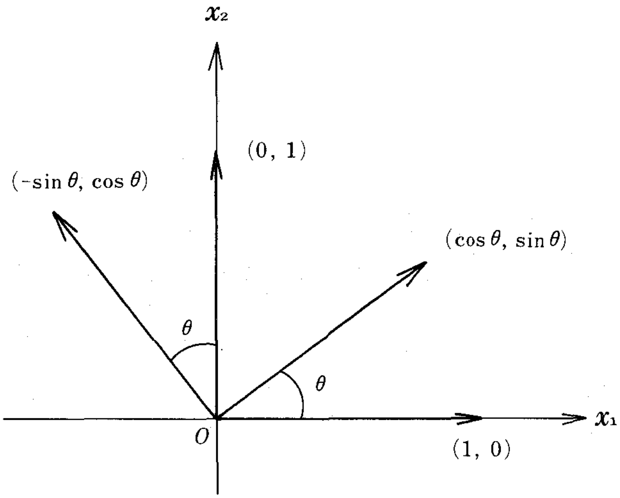

図2

次に，2次元平面内で原点を通って横軸となす角がθであるような直線ℓにかんして任意の点を対称移動する行列Qは，この変換が「角$-\theta$だけの回転」「横軸にかんする対称移動」「角θだけの回転」の合成変換であることを考慮して，次のように計算できる。

$$
\begin{aligned}
Q &= P\begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}P^{-1} \\
&= \begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix}\begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}\begin{bmatrix} \cos\theta & \sin\theta \\ -\sin\theta & \cos\theta \end{bmatrix} \\
&= \begin{bmatrix} \cos^2\theta - \sin^2\theta & \cos\theta\sin\theta + \sin\theta\cos\theta \\ \sin\theta\cos\theta + \cos\theta\sin\theta & \sin^2\theta - \cos^2\theta \end{bmatrix} \\
&= \begin{bmatrix} \cos 2\theta & \sin 2\theta \\ \sin 2\theta & -\cos 2\theta \end{bmatrix} \qquad (6)
\end{aligned}
$$

最後に，2次元平面内で任意の点を，原点を通って横軸となす角がθであるような直線ℓ上への正射影に写すような行列Rは，もとの点自身と対称移動した点との位置ベクトルを加えて2で割る操作に対応しているから（図3），

$$
\begin{aligned}
R &= \frac{1}{2}(I+Q) \\
&= \frac{1}{2}\left[\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} + \begin{bmatrix} \cos 2\theta & \sin 2\theta \\ \sin 2\theta & -\cos 2\theta \end{bmatrix}\right] \\
&= \begin{bmatrix} \dfrac{1+\cos 2\theta}{2} & \dfrac{\sin 2\theta}{2} \\ \dfrac{\sin 2\theta}{2} & \dfrac{1+\cos 2\theta}{2} \end{bmatrix} \\
&= \begin{bmatrix} \cos^2\theta & \sin\theta\cos\theta \\ \sin\theta\cos\theta & \sin^2\theta \end{bmatrix} \qquad (7)
\end{aligned}
$$

となる。この行列は射影作用素と呼ばれるが，この行列がたしかに任意の位置ベクトルをその正射影へと変える変換を表わしていることは，次のようにしても理解することができる。

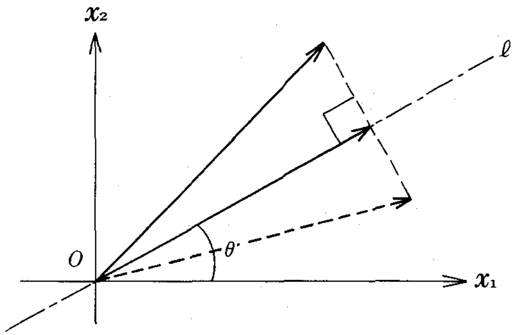

図3

図4で，直線ℓ方向に向いた長さ1の単位ベクトルを$\boldsymbol{a}$としよう。$\boldsymbol{a}=(\cos\theta,\ \sin\theta)'$である。

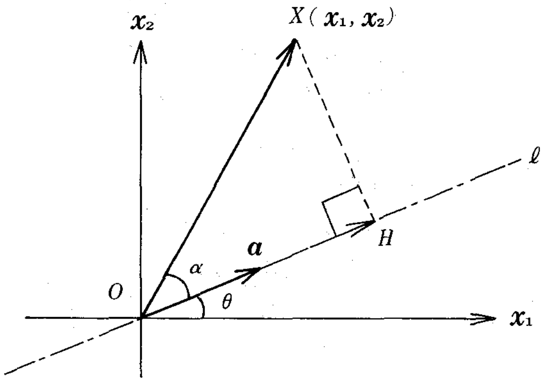

図4

任意の点Xの位置ベクトル$\boldsymbol{x}=(x_1, x_2)'$と$\boldsymbol{a}$とのなす角を$\alpha$として，両者の内積を作ると

$$\langle \boldsymbol{x}, \boldsymbol{a}\rangle = |\boldsymbol{x}||\boldsymbol{a}|\cos\alpha = |\boldsymbol{a}| \times |\boldsymbol{x}|\cos\alpha$$
$$= 1 \times \overline{\mathrm{OH}} \tag{8}$$

となり，この内積が正射影の長さになっていることがわかる。よって，ベクトル$\overrightarrow{\mathrm{OH}}$は$\boldsymbol{a}$にこの内積の値を掛けたものとなり，

$$\overrightarrow{\mathrm{OH}} = \langle \boldsymbol{x}, \boldsymbol{a} \rangle \boldsymbol{a} = \boldsymbol{a} \langle \boldsymbol{x}, \boldsymbol{a} \rangle \tag{9}$$

となる（ベクトルにスカラーを掛けるときは，前から掛けても後ろから掛けてもよい)。ところが，内積 $\langle \boldsymbol{x}, \boldsymbol{a} \rangle$ は行ベクトルと列ベクトルの積 $\boldsymbol{a}'\boldsymbol{x}$ に等しいので，

$$\begin{aligned}\overrightarrow{\mathrm{OH}} &= \boldsymbol{a}\boldsymbol{a}'\boldsymbol{x} \\ &= \begin{bmatrix} \cos\theta \\ \sin\theta \end{bmatrix} (\cos\theta, \sin\theta) \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} \\ &= \begin{bmatrix} \cos^2\theta & \sin\theta\cos\theta \\ \sin\theta\cos\theta & \sin^2\theta \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} \end{aligned} \tag{10}$$

となって，この写像を表わす行列が先の行列 R と同じであることがわかった。写影作用素とは，写影を落とす先の直線の方向ベクトルである単位ベクトルを，タテヨコ２つもってきて，掛け合わせた行列なのである。

2．対称行列の固有ベクトルの直交性

行列 S を対称行列とし，あい異なる固有値 λ_1, λ_2 に属する固有ベクトルをそれぞれ $\boldsymbol{a}_1$, $\boldsymbol{a}_2$ としよう。このとき，S が対称行列であることにより

$$S' = S \tag{11}$$

となり，また，固有値・固有ベクトルの性質により

$$S\boldsymbol{a}_1 = \lambda_1 \boldsymbol{a}_1 \tag{12}$$

$$S\boldsymbol{a}_2 = \lambda_2 \boldsymbol{a}_2 \tag{13}$$

の関係が成り立っている。いま，(12)式の両辺に左から $\boldsymbol{a}_2'$ を掛け，(13)式の両辺に左から $\boldsymbol{a}_1'$ を掛けると，次の２つの式になる。

$$\boldsymbol{a}_2'S\boldsymbol{a}_1 = \lambda_1\boldsymbol{a}_2'\boldsymbol{a}_1 \tag{14}$$

$$\boldsymbol{a}_1'S\boldsymbol{a}_2 = \lambda_2\boldsymbol{a}_1'\boldsymbol{a}_2 \tag{15}$$

これらの式の右辺の，固有値以外の部分は，掛ける順序が逆にはなっているが，どちらも $\boldsymbol{a}_1$ と $\boldsymbol{a}_2$ の内積になっている。左辺についてはやや複雑であるが，

$$\boldsymbol{b}_1 = S\boldsymbol{a}_1 \tag{16}$$

とおくと，

$$\boldsymbol{a}_2'S\boldsymbol{a}_1 = \boldsymbol{a}_2'\boldsymbol{b}_1 = \langle \boldsymbol{b}_1, \boldsymbol{a}_2 \rangle \tag{17}$$

となり，また，(11)式を考慮して

$$\begin{aligned}\boldsymbol{a}_1'S\boldsymbol{a}_2 &= \boldsymbol{a}_1'S'\boldsymbol{a}_2 \\ &= (S\boldsymbol{a}_1)'\boldsymbol{a}_2 = \boldsymbol{b}_1'\boldsymbol{a}_2 = \langle \boldsymbol{b}_1, \boldsymbol{a}_2 \rangle\end{aligned} \tag{18}$$

となる。よって，(14)式と(15)式は

$$\langle \boldsymbol{b}_1, \boldsymbol{a}_2 \rangle = \lambda_1 \langle \boldsymbol{a}_1, \boldsymbol{a}_2 \rangle \tag{19}$$

$$\langle \boldsymbol{b}_1, \boldsymbol{a}_2 \rangle = \lambda_2 \langle \boldsymbol{a}_1, \boldsymbol{a}_2 \rangle \tag{20}$$

となり，辺々差し引くことによって

$$0 = (\lambda_1 - \lambda_2)\langle \boldsymbol{a}_1, \boldsymbol{a}_2 \rangle \tag{21}$$

が得られる。ここで，λ_1，λ_2 があい異なる固有値であったことを思い出すと，

$$\langle \boldsymbol{a}_1, \boldsymbol{a}_2 \rangle = 0 \tag{22}$$

であるほかないことになって，$\boldsymbol{a}_1$ と $\boldsymbol{a}_2$ の直交性が証明できた。

3．対称行列のスペクトル分解

$p \times p$ の対称行列 S が，あい異なる固有値 λ_1，λ_2，……，λ_p とそれらに属する固有ベクトル $\boldsymbol{a}_1, \boldsymbol{a}_2, \ldots\ldots, \boldsymbol{a}_p$ をもつとしよう。固有ベクトルは長さ1に規準化

されているものとする。行列の固有値問題一般にかんして成り立つ定理により，

$$\Lambda = \begin{bmatrix} \lambda_1 & & 0 \\ & \ddots & \\ 0 & & \lambda_p \end{bmatrix} \tag{23}$$

$$P = [\boldsymbol{a}_1 \quad \boldsymbol{a}_2 \quad \cdots\cdots \quad \boldsymbol{a}_p] \tag{24}$$

として，

$$S = P\Lambda P^{-1} \tag{25}$$

が成立する。

ところで，各列が長さ1の互いに直交するベクトルからなっている「直交行列」においては，その転置行列がすなわち逆行列となる。なぜなら，

$$\begin{bmatrix} \boldsymbol{a}_1' \\ \boldsymbol{a}_2' \\ \vdots \\ \boldsymbol{a}_p' \end{bmatrix} [\boldsymbol{a}_1 \quad \boldsymbol{a}_2 \quad \cdots\cdots \quad \boldsymbol{a}_p] = \begin{bmatrix} 1 & 0 & \cdots\cdots & 0 \\ 0 & 1 & \cdots\cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots\cdots & 1 \end{bmatrix} \tag{26}$$

となるからである（(26)式の左辺の乗法を実行したとき，第 (i, j) 成分は $\boldsymbol{a}_i'\boldsymbol{a}_j = \langle \boldsymbol{a}_i, \boldsymbol{a}_j \rangle$ となり，この値はベクトルの直交性によって $i=j$ のときのみ1となり，$i \neq j$ のときは0となる）。

よって，(25)式はさらに次のように変形される。

$$S = [\boldsymbol{a}_1 \quad \boldsymbol{a}_2 \quad \cdots\cdots \quad \boldsymbol{a}_p] \begin{bmatrix} \lambda_1 & & 0 \\ & \ddots & \\ 0 & & \lambda_p \end{bmatrix} \begin{bmatrix} \boldsymbol{a}_1' \\ \boldsymbol{a}_2' \\ \vdots \\ \boldsymbol{a}_p' \end{bmatrix}$$

$$= [\lambda_1 \boldsymbol{a}_1 \quad \lambda_2 \boldsymbol{a}_2 \quad \cdots\cdots \quad \lambda_p \boldsymbol{a}_p] \begin{bmatrix} \boldsymbol{a}_1' \\ \boldsymbol{a}_2' \\ \vdots \\ \boldsymbol{a}_p' \end{bmatrix}$$

$$= \lambda_1 \boldsymbol{a}_1 \boldsymbol{a}_1' + \lambda_2 \boldsymbol{a}_2 \boldsymbol{a}_2' + \cdots\cdots + \lambda_p \boldsymbol{a}_p \boldsymbol{a}_p' \tag{27}$$

この最後の変形は直観的には理解しにくいが，次のように考えればよい。ベクトル $\boldsymbol{a}_1$, $\boldsymbol{a}_2$ 等を成分で

$$\boldsymbol{a}_1 = \begin{bmatrix} a_{11} \\ a_{21} \\ \vdots \\ a_{p1} \end{bmatrix}, \ \boldsymbol{a}_2 = \begin{bmatrix} a_{12} \\ a_{22} \\ \vdots \\ a_{p2} \end{bmatrix}, \ \cdots\cdots, \ \boldsymbol{a}_p = \begin{bmatrix} a_{1p} \\ a_{2p} \\ \vdots \\ a_{pp} \end{bmatrix}$$

と書くものと約束しておくと，行列 S の第（i, j）成分となるべきものは，行列

$$[\lambda_1 \boldsymbol{a}_1 \quad \lambda_2 \boldsymbol{a}_2 \quad \cdots\cdots \quad \lambda_p \boldsymbol{a}_p]$$

の第 i 行である

$$(\lambda_1 a_{i1}, \ \lambda_2 a_{i2}, \cdots\cdots, \ \lambda_p a_{ip})$$

と，行列

$$\begin{bmatrix} \boldsymbol{a}_1' \\ \boldsymbol{a}_2' \\ \vdots \\ \boldsymbol{a}_p' \end{bmatrix}$$

の第 j 列である

$$\begin{bmatrix} a_{j1} \\ a_{j2} \\ \vdots \\ a_{jp} \end{bmatrix}$$

との内積である。それを s_{ij} と名づけると，

$$s_{ij} = \lambda_1 a_{i1} a_{j1} + \lambda_2 a_{i2} a_{j2} + \cdots\cdots + \lambda_p a_{ip} a_{jp} \tag{28}$$

となる。この関係が 1 から p までのあらゆる i, j について成り立つ結果，(27)式の最後の変形が成立するのである。

こうして，行列 S は，固有ベクトルを射影軸とする射影作用素を固有値倍したものの和として表現できることがわかった。これを対称行列のスペクトル分解という。

スペクトル分解は，

$$\boldsymbol{b}_1 = \sqrt{\lambda_1}\boldsymbol{a}_1,\quad \boldsymbol{b}_2 = \sqrt{\lambda_2}\boldsymbol{a}_2,\quad \cdots\cdots,\quad \boldsymbol{b}_p = \sqrt{\lambda_p}\boldsymbol{a}_p$$

という置き換えをすると

$$S = \boldsymbol{b}_1\boldsymbol{b}_1' + \boldsymbol{b}_2\boldsymbol{b}_2' + \cdots\cdots + \boldsymbol{b}_p\boldsymbol{b}_p' \tag{29}$$

と表現することもできる。

射影作用素のそれぞれは，空間内のあらゆる点を 1 直線上に写像してしまう行列であるから，階数 1 の行列である。したがって，(29)式の右辺第 m 項までによって S を近似したとき，近似行列 $\tilde{S}$ は階数 m の行列となる。

Ⅳ．主成分分析

1．2変量の場合

主成分分析は，多数の変量についての観測結果が与えられたとき，観測値のちらばり方がどの方向において最大になっているかを見いだそうとするものである。2次元の場合でいえば，たとえば図5に示すように観測値の分布状況がほぼある方向に向いた楕円の形をしているとき，その楕円の長軸の方向を発見しようとするものである。

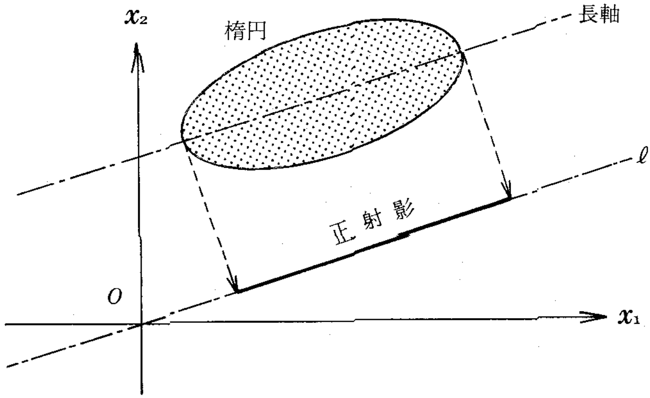

図5

その発見のためには，まず，あらゆる観測値（位置ベクトル）を原点を通る何らかの直線に射影したと考える。その正射影の長さは一般的に

$$z = a_1x_1 + a_2x_2 \tag{30}$$

ただし

$$a_1^2 + a_2^2 = 1 \tag{31}$$

と表現できる。(31)式の関係が満たされているかぎり，a_1 と a_2 は何らかの角度の余弦と正弦になる。その角を θ とする。ベクトル $(a_1, a_2)'$ は原点を通って横

軸となす角がθであるような直線ℓの方向ベクトルであり，かつ単位ベクトルである。したがって，(8)式により，a_1, a_2を係数とする合成変量zは，位置ベクトル$(x_1, x_2)'$を直線ℓ上に射影した正射影の長さになる。）

観測値をx_{1i}, x_{2i}（$i=1, 2, \cdots\cdots, n$）として，それらを(30)式にあてはめたとき得られる値をz_i（$i=1, 2, \cdots\cdots, n$）とする。そのz_i全体について分散を求め，それが最大になるようにa_1, a_2を決めれば，当の「長軸」の方向が求まる。

合成変量zの分散は

$$\mathrm{V}(z)=\frac{1}{n}\sum_{i=1}^{n}(z_i-\bar{z})^2$$

$$=\frac{1}{n}\sum_{i=1}^{n}\{a_1(x_{1i}-\bar{x}_1)+a_2(x_{2i}-\bar{x}_2)\}^2$$

$$=a_1^2\frac{\sum_{i=1}^{n}(x_{1i}-\bar{x}_1)^2}{n}+2a_1a_2\frac{\sum_{i=1}^{n}(x_{1i}-\bar{x}_1)(x_{2i}-\bar{x}_2)}{n}$$

$$+a_2^2\frac{\sum_{i=1}^{n}(x_{2i}-\bar{x}_2)^2}{n}$$

$$=a_1^2s_{11}+2a_1a_2s_{12}+a_2^2s_{22} \tag{32}$$

$$\left(\begin{array}{l}s_{11}: x_{1i}\text{ の分散，}\ s_{22}: x_{2i}\text{ の分散}\\ s_{12}: x_{1i}\text{ と }x_{2i}\text{ との共分散}\end{array}\right)$$

となる。V(z)の最大化のための1階の条件を求めるには，ラグランジュ乗数法を用いる。ラグランジュ関数は

$$L(a_1, a_2, \lambda)$$

$$=a_1^2s_{11}+2a_1a_2s_{12}+a_2^2s_{22}+\lambda(1-a_1^2-a_2^2) \tag{33}$$

である。これをa_1, a_2, λで偏微分して「＝0」とおき，

$$\frac{\partial L}{\partial a_1} \equiv 2\ (a_1 s_{11} + a_2 s_{12}) - 2\lambda a_1 = 0 \tag{34}$$

$$\frac{\partial L}{\partial a_2} \equiv 2\ (a_1 s_{12} + a_2 s_{22}) - 2\lambda a_2 = 0 \tag{35}$$

$$\frac{\partial L}{\partial \lambda} \equiv 1 - a_1^2 - a_2^2 = 0 \tag{36}$$

としたものが1階の条件である。

いま，分散共分散行列と係数ベクトルを

$$S = \begin{bmatrix} s_{11} & s_{12} \\ s_{21} & s_{22} \end{bmatrix}, \qquad \boldsymbol{a} = \begin{bmatrix} a_1 \\ a_2 \end{bmatrix}$$

と名づけると，(34)，(35)式は

$$S\boldsymbol{a} = \lambda \boldsymbol{a} \tag{37}$$

という固有値問題となる（s_{12} と s_{21} は同じものであるが，行列 S の成分を書く際には，形式を整える目的で左下の成分を s_{21} としてある)。

ところで，この問題の解として求められる固有値 λ は，合成変量の分散そのものに一致する。そのことは次のようにして確かめられる。

(31)式を成分で書いて

$$\begin{bmatrix} s_{11} & s_{12} \\ s_{21} & s_{22} \end{bmatrix} \begin{bmatrix} a_1 \\ a_2 \end{bmatrix} = \lambda \begin{bmatrix} a_1 \\ a_2 \end{bmatrix} \tag{38}$$

としよう。この式の両辺に左から行ベクトル（a_1，a_2）を掛けると

$$(a_1, a_2) \begin{bmatrix} s_{11} & s_{12} \\ s_{21} & s_{22} \end{bmatrix} \begin{bmatrix} a_1 \\ a_2 \end{bmatrix} = \lambda\ (a_1, a_2) \begin{bmatrix} a_1 \\ a_2 \end{bmatrix} \tag{39}$$

となり，左右両辺の乗法の計算を実行することによって

$$s_{11}a_1^2 + 2s_{12}a_1a_2 + s_{22}a_2^2 = \lambda(a_1^2 + a_2^2) \tag{40}$$

が導かれる。(40)式の左辺は(32)式で求めた$\mathrm{V}(z)$であり，右辺のカッコの中は(31)式（あるいは(36)式）によって１であるから，

$$\mathrm{V}(z) = \lambda \tag{41}$$

が得られる。

固有値が合成変量zの分散に等しい結果として，固有値は当然すべて非負である。

図５の「長軸」の方向を求めるには，分散を最大にするような射影軸を求めればよいわけだから，結論としては，固有方程式の根のうち大きい方をλ_1として，λ_1に属する規準化された固有ベクトル

$$\boldsymbol{a}_1 = \begin{bmatrix} a_{11} \\ a_{21} \end{bmatrix}, \qquad (a_{11}^2 + a_{21}^2 = 1)$$

の方向を射影軸とすればよい。この係数ベクトルによってウェイトづけされた合成変量

$$z_1 = a_{11}x_1 + a_{21}x_2 \tag{42}$$

のことを第１主成分と呼ぶ。個々の観測値をこの式にあてはめたときに得られるz_1の値

$$z_{1i} = a_{11}x_{1i} + a_{21}x_{2i} \tag{43}$$

のことを，その観測値の（第１主成分にかんする）主成分得点という。

小さい方の固有値をλ_2とすると，λ_2に属する固有ベクトル

$$\boldsymbol{a}_2 = \begin{bmatrix} a_{12} \\ a_{22} \end{bmatrix}, \qquad (a_{12}^2 + a_{22}^2 = 1)$$

を用いて第2主成分

$$z_2 = a_{12}x_1 + a_{22}x_2 \tag{44}$$

を求めることができる。ただし，2変量の場合，第2主成分は分散を最小化する射影軸の方向への正射影を表わすので，あまり分析上の意義はもたない。第2主成分以降の主成分が意義をもつのは，変量が3種類以上存在する場合である。

2．一般の場合

そこで，以下，そのような一般的な場合について考察することにしよう。

変量はp種類とする。合成変量は

$$z = a_1x_1 + a_2x_2 + \cdots\cdots + a_px_p \tag{45}$$

と定義される。ただし

$$a_1^2 + a_2^2 + \cdots\cdots + a_p^2 = 1 \tag{46}$$

という制約がついている。この(46)式の制約の下で

$$\begin{aligned} \mathrm{V}(z) &= \frac{1}{n}\sum_{i=1}^{n}(z_i - \bar{z})^2 \\ &= \frac{1}{n}\sum_{i=1}^{n}\{a_1(x_{1i} - \bar{x}_1) + \cdots\cdots + a_p(x_{pi} - \bar{x}_p)\}^2 \end{aligned} \tag{47}$$

を最大化するように$a_1, a_2, \cdots\cdots, a_p$を決めるのが課題である。

この場合の$\mathrm{V}(z)$はかなり複雑になるが，結論的には

$$\mathrm{V}(z) = (a_1, a_2, \cdots\cdots, a_p)\begin{bmatrix} s_{11} & s_{12} & \cdots\cdots & s_{1p} \\ s_{21} & s_{22} & \cdots\cdots & s_{2p} \\ \vdots & \vdots & & \vdots \\ s_{p1} & s_{p2} & \cdots\cdots & s_{pp} \end{bmatrix}\begin{bmatrix} a_1 \\ a_2 \\ \vdots \\ a_p \end{bmatrix} \tag{48}$$

となることが知られている。この問題にラグランジュ乗数法を適用すると，(38)式に準じた次の式が得られる。

$$\begin{bmatrix} s_{11} & s_{12} & \cdots\cdots & s_{1p} \\ s_{21} & s_{22} & \cdots\cdots & s_{2p} \\ \vdots & \vdots & & \vdots \\ s_{p1} & s_{p2} & \cdots\cdots & s_{pp} \end{bmatrix} \begin{bmatrix} a_1 \\ a_2 \\ \vdots \\ a_p \end{bmatrix} = \lambda \begin{bmatrix} a_1 \\ a_2 \\ \vdots \\ a_p \end{bmatrix} \tag{49}$$

ここに出てくる分散共分散行列を S，係数ベクトルを $\boldsymbol{a}$ とすると，(49)式の両辺に左から $\boldsymbol{a}'$ を掛けることにより

$$\boldsymbol{a}'S\boldsymbol{a} = \lambda\,\boldsymbol{a}'\boldsymbol{a} \tag{50}$$

が得られるが，(50)式の右辺の $\boldsymbol{a}'\boldsymbol{a}$ は内積 $\langle \boldsymbol{a}, \boldsymbol{a} \rangle$ に等しくてその値は 1，左辺は(48)式の右辺に等しいから，

$$\mathrm{V}(z) = \lambda \tag{51}$$

が得られる。なお，(50)式の左辺は対称行列 S によって表現される 2 次形式であり，これが(47)式のような 2 乗和でできた式に等しいということは，この 2 次形式が非負定符号（ベクトル $\boldsymbol{a}$ をどのようにとっても，つねに非負）であることを意味している。2 次形式が非負定符号であることは，その対称行列の固有値がすべて非負であることを意味している。このことは(50)式の関係から明らかである。

以上で，この固有値問題の解として得られる p 個の固有値 λ はすべて非負の実数であることがわかった。

さて，分散共分散行列 S は対称行列であるから，そのあい異なる固有値に属する固有ベクトルは必ず直交する。いま，最大の固有値を λ_1，それに属する固有ベクトルを $\boldsymbol{a}_1$，2 番目に大きい固有値を λ_2，それに属する固有ベクトルを $\boldsymbol{a}_2$ としよう。すると，第 1 主成分と第 2 主成分とのあいだに，次の関係が成立することがわかる。

$$\begin{aligned}
&\mathrm{Cov}(z_1, z_2)\\
&= \frac{1}{n}\sum_{i=1}^{n} \{a_{11}(x_{1i}-\bar{x}_1)+\cdots\cdots+a_{p1}(x_{pi}-\bar{x}_p)\}\\
&\qquad\qquad \{a_{12}(x_{1i}-\bar{x}_1)+\cdots\cdots+a_{p2}(x_{pi}-\bar{x}_p)\}\\
&= (a_{11},\cdots\cdots,a_{p1})\begin{bmatrix} s_{11} & \cdots\cdots & s_{1p}\\ \vdots & & \vdots \\ s_{p1} & \cdots\cdots & s_{pp}\end{bmatrix}\begin{bmatrix} a_{12}\\ \vdots \\ a_{p2}\end{bmatrix}\\
&= \boldsymbol{a}_1' S \boldsymbol{a}_2 = \boldsymbol{a}_1'(\lambda_2 \boldsymbol{a}_2) = \lambda_2 \boldsymbol{a}_1' \boldsymbol{a}_2 = \lambda_2 \langle \boldsymbol{a}_1, \boldsymbol{a}_2 \rangle = 0
\end{aligned} \tag{52}$$

つまり，両変量のあいだの共分散はゼロであり，両合成変量は無相関ということになる。このことは，第1主成分たる z_1 と無相関という限定つきで，分散が最大になるような合成変量を求めるならば，それがまさにベクトル $\boldsymbol{a}_2$ でウェイトづけされた第2主成分 z_2 だということを意味している。

(52)式と同様の「合成変量どうしが無相関」という関係は，もちろん，第1主成分と第2主成分とのあいだにだけ成り立つものではなく，任意の主成分どうしのあいだで成立する。

そこで，$z_2, z_3, \cdots\cdots, z_{p-1}$ という主成分はすべて，それ以前の番号の合成変量と無相関という限定つきで分散を最大化する合成変量を求めたときの答えということになる。最後の z_p だけは，分散を最小化する合成変量になっている。

こうして，p 次元空間内の観測値の分布が互いに直交する軸への正射影に分解される。このときたとえば，λ_1, λ_2 に比べて λ_3 以降の固有値がとても小さいならば，z_3 以降の軸の方向には観測値がほとんどちらばりをもたず，観測値は第1軸と第2軸で張られる2次元空間の中にほぼ収まっていることになる。

ところで，固有方程式にかんする理論の教えるところによれば，およそ行列の固有値の和はその行列の対角成分の和に等しい。$p \times p$ の分散共分散行列にこの法則を適用すれば，

$$\lambda_1 + \lambda_2 + \cdots\cdots + \lambda_p = s_{11} + s_{22} + \cdots\cdots + s_{pp} \tag{53}$$

となる。この式の右辺は，もとの全変量の分散の和である。固有値の方は合成変量の分散を表わしているから，(53)式は

（新たに構成された合成変量の分散の和）
　＝（もとの変量の分散の和）

という関係を述べていることになる。

ここで

$$\mu_i = \frac{\lambda_i}{\lambda_1 + \lambda_2 + \cdots\cdots + \lambda_p} = \frac{\lambda_i}{s_{11} + s_{22} + \cdots\cdots + s_{pp}} \tag{54}$$

という指標を定義すると，μ_i（$i = 1, 2, \cdots\cdots, p$）の総和は1であって，$\mu_i$は全分散の何％が第$i$主成分の分散で説明できるかを示す指標だと解釈できる。これを第i主成分の寄与率という。さらに

$$\nu_i = \sum_{k=1}^{i} \mu_k \tag{55}$$

を，第i番目までの主成分の累積寄与率といい，分析結果の記述にあたって第何番目までの主成分を採用すべきかを考えるときの判断の基準になる。

たとえば$i = 3$までで累積寄与率が95％に達したという場合，3つの軸の方向への分散だけで分散全体の95％が説明できていることになるので，4番目以降の主成分はほとんど意味を持たないとして，記述から落とすことが許されよう。このことは言い換えれば，分散共分散行列のスペクトル分解

$$S = \lambda_1 \boldsymbol{a}_1 \boldsymbol{a}_1' + \lambda_2 \boldsymbol{a}_2 \boldsymbol{a}_2' + \cdots\cdots + \lambda_p \boldsymbol{a}_p \boldsymbol{a}_p' \tag{56}$$

の第4項以下がひじょうに小さくて，

$$\tilde{S} = \lambda_1 \boldsymbol{a}_1 \boldsymbol{a}_1' + \lambda_2 \boldsymbol{a}_2 \boldsymbol{a}_2' + \lambda_3 \boldsymbol{a}_3 \boldsymbol{a}_3' \tag{57}$$

によるSの近似がかなり有効性をもつということにほかならない。

3．相関行列による主成分分析

いま，最初の変量を x_1, x_2 として，これに標準化をほどこすことを考える。標準化とは，もとの変量の平均値からの偏差をとり，それを標準偏差で割ったものを新たな変量とすることである。変量 x_1 と x_2 の分散をそれぞれ $V(x_1)$, $V(x_2)$ として，新変量は

$$y_{1i} = \frac{x_{1i} - \bar{x}_1}{\sqrt{V(x_1)}} \tag{58}$$

$$y_{2i} = \frac{x_{2i} - \bar{x}_2}{\sqrt{V(x_2)}} \tag{59}$$

によって定義される。これらの変量の平均である $\bar{y}_1$ と $\bar{y}_2$ がともに0となることは，説明を要しない自明のことであろう。分散については，次のようにして，その値が1になることが証明できる。

$$V(y_1) = \frac{\sum_{i=1}^{n} (y_{1i} - \bar{y}_1)^2}{n} = \frac{\sum_{i=1}^{n} y_{1i}^2}{n}$$

$$= \frac{1}{n} \cdot \frac{\sum_{i=1}^{n} (x_{1i} - \bar{x}_1)^2}{V(x_1)} = \frac{1}{n} \cdot \frac{\sum_{i=1}^{n} (x_{1i} - \bar{x}_1)^2}{\frac{\sum_{i=1}^{n} (x_{1i} - \bar{x}_1)^2}{n}} = 1 \tag{60}$$

$V(y_2)$ についても同様である。

y_1 と y_2 のあいだの共分散は，次のようになる。

$$\mathrm{Cov}(y_1, y_2) = \frac{\sum_{i=1}^{n}(y_{1i}-\overline{y_1})(y_{2i}-\overline{y_2})}{n} = \frac{\sum_{i=1}^{n} y_{1i}y_{2i}}{n}$$

$$= \frac{1}{n}\cdot\sum_{i=1}^{n}\frac{(x_{1i}-\overline{x_1})(x_{2i}-\overline{x_2})}{\sqrt{\mathrm{V}(x_1)}\sqrt{\mathrm{V}(x_2)}}$$

$$= \frac{\dfrac{\sum_{i=1}^{n}(x_{1i}-\overline{x_1})(x_{2i}-\overline{x_2})}{n}}{\sqrt{\mathrm{V}(x_1)}\sqrt{\mathrm{V}(x_2)}}$$

$$= \frac{\mathrm{Cov}(x_1, x_2)}{\sqrt{\mathrm{V}(x_1)}\sqrt{\mathrm{V}(x_2)}} = r(x_1, x_2) \tag{61}$$

つまり，新変量間の共分散は，もとの変量のあいだの相関係数なのである。ところが，新変量のあいだの相関係数も，実はこれと同じになる。なぜなら，

$$r(y_1, y_2) = \frac{\mathrm{Cov}(y_1, y_2)}{\sqrt{\mathrm{V}(y_1)}\sqrt{\mathrm{V}(y_2)}} = \mathrm{Cov}(y_1, y_2) = r(x_1, x_2) \tag{62}$$

だからである。

結論として，以上のように標準化された変量にかんする分散共分散行列は

$$R = \begin{bmatrix} r_{11} & r_{12} \\ r_{21} & r_{22} \end{bmatrix} \tag{63}$$

で与えられる。この行列の成分となっている相関係数は，もとの変量についての相関係数と考えてもよいし，標準化された変量についての相関係数と考えてもよい。この行列 R を相関行列という。

主成分分析は，変量が標準化されていない場合には，特有の不都合が起こることがあるので，最近では，相関行列を用いた主成分分析がしばしば推奨されている。変量を標準化しない場合，分析上どういう問題が生じるかは，次の例からわかる。

たとえば，変量 x_1 が cm 単位，変量 x_2 が g 単位で測定されていて，その分散共分散行列が

$$S = \begin{bmatrix} 1 & 1 \\ 1 & 5 \end{bmatrix}$$

で与えられていたとする。固有方程式は

$$\lambda^2 - 6\lambda + 4 = 0$$

であるから，固有値および寄与率は

$$\lambda = 3 \pm \sqrt{5} = \begin{cases} 5.236 & (87.3\%) \\ 0.764 & (12.7\%) \end{cases}$$

となる。

次に，同じデータについて，変量 x_1 の測定単位を cm から mm に変えたとしたら，s_{11} は100倍，s_{12} と s_{21} は10倍されるため，

$$S = \begin{bmatrix} 100 & 10 \\ 10 & 5 \end{bmatrix}$$

となる。固有方程式は

$$\lambda^2 - 105\lambda + 400 = 0$$

となるので，固有値及び寄与率は

$$\lambda = 105 \pm \frac{\sqrt{9425}}{2} = \begin{cases} 101.04 & (96.2\%) \\ 3.96 & (\ 3.8\%) \end{cases}$$

となる。このように，単位のとり方で寄与率は変わるのである。主成分のウェイトそのものも変わる。もちろん，単位が違うことによるウェイトの違いはあって当然だが，それだけではない。新単位で測った変量につくウェイトを計算してから，変量を旧単位に戻して，それに見合った倍率をウェイトの方に掛けて

やっても，式はもとに戻らないのである。たとえば上の問題で，第１主成分は，

cm 単位のとき：　$z_1 = 0.230x_1 + 0.973x_2$

mm 単位のとき：　$z_1 = 0.995x_1^* + 0.104x_2$

$= 9.95x_1 + 0.104x_2$

（x_1^* は mm 単位，x_1 は cm 単位）

となって，分散を計算する際の測定単位が異なると，後に単位を戻しても主成分の式は同じにはならないのである。

このように，変量によって重さと長さのように測定される単位が異なる場合，しかも，一方の変量をどういう単位で測るかについて必然的な根拠がない場合は，たまたま選んだ単位によって結果が異なるのは好ましくないので，分散共分散行列による主成分分析は適切でないといえよう。こうした場合，変量を標準化した相関行列による主成分分析を用いるべきである。

V．因　子　分　析

１．基本概念

いま，n 人について，国語，社会，数学，理科の４科目のテストの結果が得られたとしよう。それらの得点数を加工して，平均０，標準偏差１の変量に直したものを z_1, z_2, z_3, z_4 としよう。個々の受験者についての観測値は $z_{1i}, z_{2i}, z_{3i}, z_{4i}$ （$i = 1, 2, \cdots\cdots, n$）となる。

観測結果をざっと概観したとき，次のような傾向が観察されたと仮定しよう。(a)全体として，ある科目のよくできる人は他の科目もよくできるという正の相関関係が観察される。(b)得点のあいだの正の相関は国語と社会のあいだで強く，また，数学と理科のあいだで強い。両グループのあいだではあまり強くない。

このような観測結果を説明するのに，次のような仮説を立てるのは，ごく自然であろう。

(i)　得点に影響する主要な因子として「総合学力因子」と「文系理系適性因

子」とがあり，数的に標準化できる。

(ii) 各個体は自己固有の「総合学力因子」と「文系理系適性因子」のスコアをもつ。

(iii) 個々人の各科目の得点は両因子の影響を受けて決まる。どの因子のスコアにどのくらい影響されるかの感応度は数量化することができる。この感応度は科目に依存して決まるもので，あらゆる人を通じて共通である。

(iv) もちろん，人の実力が一定であっても，個々のテストの成績は受験するときのコンディションなどの影響を受けるから，個々の観測値は上の仮説だけでは説明できない誤差を含む。

以上を数式化すると，次のようなモデルとなる。

$$\left.\begin{aligned} z_{1i} &= a_{11}f_{1i} + a_{12}f_{2i} + e_{1i} \\ z_{2i} &= a_{21}f_{1i} + a_{22}f_{2i} + e_{2i} \\ z_{3i} &= a_{31}f_{1i} + a_{32}f_{2i} + e_{3i} \\ z_{4i} &= a_{41}f_{1i} + a_{42}f_{2i} + e_{4i} \end{aligned}\right\} \quad (i = 1, 2, \cdots\cdots, n) \tag{64}$$

f_{1i}, f_{2i} は個人ごとに異なる値をとるが，科目によっては変わらない共通因子と呼ばれるものである。共通因子の原点と単位は自由にとることができるから，一般性を失うことなく，平均0，標準偏差1に標準化されているものと仮定してよい。$a_{11},\cdots\cdots, a_{41}, a_{12},\cdots\cdots, a_{42}$ の8つの係数は因子負荷量と呼ばれ，観測される個々人が誰であるかによらない定数である。たとえば a_{32} は，第3科目の成績に第2因子がどれだけ反映するかの影響度を表わしている。$e_{1i}, e_{2i}, e_{3i}, e_{4i}$ は，個々人の個々の科目の受験結果固有の変動を表わす独自因子と呼ばれる項で，平均0，標準偏差 d_1, d_2, d_3, d_4 であるとする。

e_1, e_2, e_3, e_4 は互いに無相関，それらと f_1, f_2 とのあいだも無相関と仮定する。共通因子相互間の相関については2通りの考え方があり，共通因子が互いに無相関という仮定のもとで得られる解を直交解，そのような仮定をおかずに得られる解を斜交解という。

以下では，直交解（そのときの求められた因子を直交因子という）の場合の

みをとりあげる。

2．因子負荷量の推定

直交解の場合，たとえば変量 z_1 の分散は

$$\mathrm{V}(z_1) = \mathrm{V}(a_{11}f_1 + a_{12}f_2 + e_1)$$

$$= a_{11}^2\mathrm{V}(f_1) + a_{12}^2\mathrm{V}(f_2) + \mathrm{V}(e_1)$$

$$= a_{11}^2 + a_{12}^2 + d_1^2 \tag{65}$$

であり，変量 z_1 と z_2 との共分散は

$$\begin{aligned}\mathrm{Cov}(z_1, z_2) &= \mathrm{Cov}(a_{11}f_1 + a_{12}f_2 + e_1,\ \ a_{21}f_1 + a_{22}f_2 + e_2)\\ &= a_{11}a_{21}\mathrm{V}(f_1) + a_{12}a_{22}\mathrm{V}(f_2)\\ &= a_{11}a_{21} + a_{12}a_{22}\end{aligned} \tag{66}$$

となる。このようにして，分散共分散行列（変量が標準化されているため相関行列に等しい）は，

$$S = R = \begin{bmatrix} r_{11} & r_{12} & r_{13} & r_{14} \\ r_{21} & r_{22} & r_{23} & r_{24} \\ r_{31} & r_{32} & r_{33} & r_{34} \\ r_{41} & r_{42} & r_{43} & r_{44} \end{bmatrix}$$

$$= \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \\ a_{31} & a_{32} \\ a_{41} & a_{42} \end{bmatrix} \begin{bmatrix} a_{11} & a_{21} & a_{31} & a_{41} \\ a_{12} & a_{22} & a_{32} & a_{42} \end{bmatrix} + \begin{bmatrix} d_1^2 & & & 0 \\ & d_2^2 & & \\ & & d_3^2 & \\ 0 & & & d_4^2 \end{bmatrix} \tag{67}$$

のようになる。つまり，変量の個数を行数とし，想定する因子数を列数とする

ような縦長の矩形行列 A というものを考え，行列 S を

$$S = AA' + D \tag{68}$$

と分解するような A と非負対角行列 D とを求めることに，この問題は帰着するのである。

ここまでくると，因子分析と主成分分析との類似性が明らかになる。

主成分分析は要するに，分散共分散行列 S をスペクトル分解して，適当な項までで近似する作業であった。たとえば，4つの変量にかんする4行4例の分散共分散行列がスペクトル分解の第2項までで近似されるというとき，

$$\tilde{S} = \lambda_1 \boldsymbol{a}_1 \boldsymbol{a}_1' + \lambda_2 \boldsymbol{a}_2 \boldsymbol{a}_2' = \boldsymbol{b}_1 \boldsymbol{b}_1' + \boldsymbol{b}_2 \boldsymbol{b}_2'$$

$$= [\boldsymbol{b}_1 \quad \boldsymbol{b}_2] \begin{bmatrix} \boldsymbol{b}_1' \\ \boldsymbol{b}_2' \end{bmatrix} = [\sqrt{\lambda_1}\boldsymbol{a}_1 \quad \sqrt{\lambda_2}\boldsymbol{a}_2] \begin{bmatrix} \sqrt{\lambda_1}\boldsymbol{a}_1' \\ \sqrt{\lambda_2}\boldsymbol{a}_2' \end{bmatrix}$$

$$= \begin{bmatrix} \sqrt{\lambda_1}a_{11} & \sqrt{\lambda_2}a_{12} \\ \sqrt{\lambda_1}a_{21} & \sqrt{\lambda_2}a_{22} \\ \sqrt{\lambda_1}a_{31} & \sqrt{\lambda_2}a_{32} \\ \sqrt{\lambda_1}a_{41} & \sqrt{\lambda_2}a_{42} \end{bmatrix} \begin{bmatrix} \sqrt{\lambda_1}a_{11} & \sqrt{\lambda_1}a_{21} & \sqrt{\lambda_1}a_{31} & \sqrt{\lambda_1}a_{41} \\ \sqrt{\lambda_2}a_{12} & \sqrt{\lambda_2}a_{22} & \sqrt{\lambda_2}a_{32} & \sqrt{\lambda_2}a_{42} \end{bmatrix} \tag{69}$$

という階数2の行列で S を近似しているのである。このときに現れる，固有ベクトルに固有値の平方根を掛けて並べた行列を $\tilde{A}$ と名づけるならば，主成分分析は

$$\underset{p\times p}{\tilde{S}} = \underset{p\times m}{\tilde{A}}\,\underset{m\times p}{\tilde{A}'} \tag{70}$$

というものを計算していることになる（ただし p は変量の個数，m は第何主成分までを採用するかの個数）。

これに対して因子分析の方は，

$$\underset{p\times p}{S} - \underset{p\times p}{D} = \underset{p\times m}{A}\,\underset{m\times p}{A'} \tag{71}$$

であって，S から適当な対角行列を引くことによって，右辺が近似ではなくちょうどぴったり $p \times m$ 行列とその転置行列との積になるように，D と A とを同時決定しようということなのである。

実際の計算にあたっては，まず D について恣意的な初期値を設定し，$S - D$ を主成分分析法で分解し，第 m 項までを採用する。すると，一般に $S - D$ がぴったり $p \times m$ 行列とその転置行列との積には分解されない結果，残余が出るであろう。その残余がだんだん収束して小さくなるように，D と A とを交互に修正してゆくのである。具体的な手順をステップで示せば，次のようになる。

ステップ 1．独自因子の分散の初期値 $D^{(0)}$ の設定。

ステップ 2．k 回目の反復での推定値 $D^{(k)}$ を用いて，$S - D^{(k)}$ の固有値 $\lambda_1 \geqq \lambda_2 \geqq \cdots\cdots \geqq \lambda_m$ と，対応する固有ベクトル $\boldsymbol{c}_1, \boldsymbol{c}_2, \cdots\cdots, \boldsymbol{c}_m$ を計算し，

$$A^{(k)} = [\sqrt{\lambda_1}\boldsymbol{c}_1 \quad \sqrt{\lambda_2}\boldsymbol{c}_2 \quad \cdots\cdots \quad \sqrt{\lambda_m}\boldsymbol{c}_m]$$

とおく。

ステップ 3．$S - A^{(k)}A^{(k)\prime}$ の対角成分を $D^{(k+1)}$ とおく。

ステップ 4．もし $D^{(k)}$ と $D^{(k+1)}$ の各成分が十分近ければ終了。そうでなければ，$D^{(k+1)}$ をあらたな $D^{(k)}$ として，ステップ 2 に戻る。

このようにして，因子負荷量が求められる。

なお，直交因子の場合，想定する因子の個数を m とすると，

$$\mathrm{V}(z_j) = a_{j1}^2 + a_{j2}^2 + \cdots\cdots + a_{jm}^2 + d_j^2 \tag{72}$$

となり，この値が行列 S の第 j 番目の対角成分と等しくなることから，

$$a_{j1}^2 + a_{j2}^2 + \cdots\cdots + a_{jm}^2 = s_{jj} - d_j^2 \tag{73}$$

となるが，変量が標準化され S が相関行列 R に一致しているときは $s_{jj} = r_{jj} = 1$ であるから，上記の値は $1 - d_j^2$ となる。これは，変量 z_j の全分散のうち共通因子で説明される割合を示しており，共通性と呼ばれる。（逆に d_j^2 のことを独自性という。）

3．回転の不定性

いま，⒁式で示された4変数2因子のモデルについて，$a_{11},\cdots\cdots,a_{41},a_{12},\cdots\cdots,a_{42}$の8つの因子負荷量と，各人の因子得点$f_{1i},f_{2i}$（$i=1,2,\cdots\cdots,n$）が得られたと仮定しよう。因子得点は，変数$f_1$，$f_2$を両軸にとった座標平面内で各個体を表現する点がもつ座標と考えることができる。その座標軸を角θだけ回転すると，同じ点が別の座標で読まれることになる。もとの座標を（f_{1i},f_{2i}），新しい座標を（f^*_{1i},f^*_{2i}）とすると，それらのあいだの関係は，点を角$-\theta$だけ回転した場合と同じなので，

$$\begin{bmatrix} f^*_{1i} \\ f^*_{2i} \end{bmatrix} = \begin{bmatrix} \cos\theta & \sin\theta \\ -\sin\theta & \cos\theta \end{bmatrix} \begin{bmatrix} f_{1i} \\ f_{2i} \end{bmatrix} \tag{74}$$

となる。逆に書けば

$$\begin{bmatrix} f_{1i} \\ f_{2i} \end{bmatrix} = \begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix} \begin{bmatrix} f^*_{1i} \\ f^*_{2i} \end{bmatrix} \tag{75}$$

である。この⒂式を，⒁式の行列表示である

$$\begin{bmatrix} z_{1i} \\ z_{2i} \\ z_{3i} \\ z_{4i} \end{bmatrix} = \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \\ a_{31} & a_{32} \\ a_{41} & a_{42} \end{bmatrix} \begin{bmatrix} f_{1i} \\ f_{2i} \end{bmatrix} + \begin{bmatrix} e_{1i} \\ e_{2i} \\ e_{3i} \\ e_{4i} \end{bmatrix} \tag{76}$$

に代入すると，

$$\begin{bmatrix} z_{1i} \\ z_{2i} \\ z_{3i} \\ z_{4i} \end{bmatrix} = \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \\ a_{31} & a_{32} \\ a_{41} & a_{42} \end{bmatrix} \begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix} \begin{bmatrix} f^*_{1i} \\ f^*_{2i} \end{bmatrix} + \begin{bmatrix} e_{1i} \\ e_{2i} \\ e_{3i} \\ e_{4i} \end{bmatrix}$$

$$
= \begin{bmatrix} a_{11}\cos\theta + a_{12}\sin\theta & -a_{11}\sin\theta + a_{12}\cos\theta \\ a_{21}\cos\theta + a_{22}\sin\theta & -a_{21}\sin\theta + a_{22}\cos\theta \\ a_{31}\cos\theta + a_{32}\sin\theta & -a_{31}\sin\theta + a_{32}\cos\theta \\ a_{41}\cos\theta + a_{42}\sin\theta & -a_{41}\sin\theta + a_{42}\cos\theta \end{bmatrix} \begin{bmatrix} f^*_{1i} \\ f^*_{2i} \end{bmatrix} + \begin{bmatrix} e_{1i} \\ e_{2i} \\ e_{3i} \\ e_{4i} \end{bmatrix} \tag{77}
$$

となる。ここにあらわれた行列をあらためて

$$
A^* = \begin{bmatrix} a^*_{11} & a^*_{12} \\ a^*_{21} & a^*_{22} \\ a^*_{31} & a^*_{32} \\ a^*_{41} & a^*_{42} \end{bmatrix} \tag{78}
$$

と名づけるならば，A^* の成分であるような因子負荷量と，回転された因子得点（f^*_{1i}, f^*_{2i}）もまた，元のモデルを満たす解であることになる。新しい因子負荷量ともとの因子負荷量との関係は

$$
\begin{bmatrix} a^*_{j1} \\ a^*_{j2} \end{bmatrix} = \begin{bmatrix} \cos\theta & \sin\theta \\ -\sin\theta & \cos\theta \end{bmatrix} \begin{bmatrix} a_{j1} \\ a_{j2} \end{bmatrix} \quad (j = 1, \cdots, 4) \tag{79}
$$

であり，因子得点にほどこした回転と同じ回転を因子負荷量のほうにもほどこした関係になっている。

このような解の不定性を回転の不定性という。

したがって，先に示した因子負荷量の計算手続きで 1 種類の解が得られたとき，実はそれはありうる解のうちのひとつであるにすぎず，それを元にして因子軸の回転をほどこして得られる解はすべて解となるのである。

実際の分析においては，適当な方法で 1 組の推定値を得た後，実質科学的に適切に因子の意味づけができるように，因子軸の回転を行なう。

4．因子数の決め方

主成分分析では，採用する主成分の数をあらかじめ決めておかなくても計算

はできる。計算の結果として出てきた固有値から累積寄与率を計算し，その値がたとえば60％とか80％とかいうあらかじめ設定した基準値を超えるのは何番目の固有値においてであるかを調べ，そこまでの固有値を採用し，対応する固有ベクトルによって主成分を構成すればよいのである。

ところが因子分析の場合は，あらかじめいくつまでの因子を想定するかを決めておかねば計算が開始できない。ここで，因子分析が本質的には主成分分析と同じく分散共分散行列の分解であり，両者の相違はただ，独自性の項を考慮するか否かの違いにすぎないことを想起しよう。あつかう分散共分散行列は同じものであるから，それの固有値として大きい値をもつものがいくつあるかを数え，その個数だけの因子を想定することで，ほぼ妥当な結果が得られるはずである。

したがって，その判定基準としては，主成分分析の場合と同じ累積寄与率を用いてもよい。しかし，因子分析の場合，最初から分散共分散行列のかわりに相関行列を用いることを考慮すると，その固有値の総和は変量の個数pと必ず一致している（なぜなら，相関行列の対角成分r_{jj}がつねに1であることにより，対角成分の和はpになるから）。そのことから，1以上の値をもつ固有値の個数を数えてそれを想定する因子の個数とするという考え方もある。因子とは，いくつかの変量を統一的に説明できる総合的指標であると考えれば，因子はもとの変量1個分を超える情報を持っているべきであると考えられるからである。

5．因子軸の回転

いま，7変量に対して因子2個を想定する因子分析を行なって，変量別の因子負荷量の組合せが1種類求められたとして，それをプロットしたものが図6のようになっていたとしよう。このままでは因子の解釈が困難であるが，もし座標軸を角θだけ回転してⅠ′Ⅱ′のように直したとすると，変量②，④，⑤，⑦は第Ⅰ′因子の負荷量が大きく第Ⅱ′因子の負荷量は小，変量①，③，⑥は第Ⅱ′因子の負荷量が大きく第Ⅰ′因子の負荷量は小，という関係が明瞭になる。

ここでかりに，この調査が食品に対する好みの強度を尋ねたもので，②，④，⑤，⑦は「ステーキ」「トンカツ」「シチュー」「ピッツァ」，①，③，⑥は「すし」「かばやき」「てんぷら」であったとすると，Ⅰ′軸が洋食嗜好の軸，Ⅱ′軸が和食嗜好の軸と解釈できるであろう。このようにして，ある特定因子の負荷量はある一群の変量で大きく，他の変量では小さいという明確な分離が生じたとき，これを単純構造という。因子軸の回転を通じてなるべく単純構造を導き出すのが,実質科学の観点から分析を有効ならしめる秘訣であると考えられる。

このような単純構造を得るための作業は，因子数が多くなると図解ができないためになかなか面倒になるが，計算によって最適な回転を見つける方法がいろいろと提案されている。

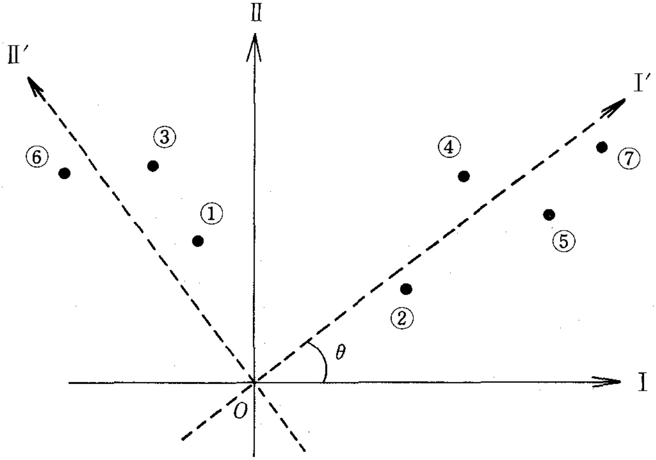

図 6

その代表的なものはバリマックス法である。

先の例の場合，回転後の因子負荷量を表わす行列ともとの因子負荷量を表わす行列とのあいだの関係は

$$\begin{bmatrix} a^*_{11} & a^*_{12} \\ a^*_{21} & a^*_{22} \\ a^*_{31} & a^*_{32} \\ a^*_{41} & a^*_{42} \end{bmatrix} = \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \\ a_{31} & a_{32} \\ a_{41} & a_{42} \end{bmatrix} \begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix} \tag{80}$$

となったが，これを一般化すれば

$$\underset{\substack{p \times \mathrm{m} \\ (\text{新因子負荷量行列})}}{\begin{bmatrix} b_{11} & \cdots\cdots & b_{1m} \\ \vdots & & \vdots \\ b_{p1} & \cdots\cdots & b_{pm} \end{bmatrix}} = \underset{\substack{p \times \mathrm{m} \\ (\text{旧因子負荷量行列})}}{\begin{bmatrix} a_{11} & \cdots\cdots & a_{1m} \\ \vdots & & \vdots \\ a_{p1} & \cdots\cdots & a_{pm} \end{bmatrix}} \underset{\substack{m \times m \\ (\text{回転行列})}}{\begin{bmatrix} t_{11} & \cdots\cdots & t_{1m} \\ \vdots & & \vdots \\ t_{m1} & \cdots\cdots & t_{mm} \end{bmatrix}} \tag{81}$$

となる。回転後の因子負荷量行列 B の第 k 列における因子負荷量の2乗 b^2_{jk} の分散は

$$\sigma^2_k = \frac{1}{p} \left\{ \sum_{j=1}^{p} (b^2_{jk})^2 - \frac{1}{p} \left(\sum_{j=1}^{p} b^2_{jk} \right)^2 \right\} \tag{82}$$

となるが，これを大きくすることが，個々の成分の2乗 b^2_{jk} を小さいものと大きいものの2つに分極させることに対応すると考えられる。これを m 個の列全部について合計した

$$\sigma^2 = \sum_{k=1}^{m} \sigma^2_k = \frac{1}{p^2} \sum_{k=1}^{m} \left\{ p \sum_{j=1}^{p} (b^2_{jk})^2 - \left(\sum_{j=1}^{p} b^2_{jk} \right)^2 \right\} \tag{83}$$

を最大化するという基準にもとづく回転法を粗バリマックス法（raw varimax method）と呼ぶ。しかし，この基準では，共通性の大きい変量では平均的に各因子負荷量も大きく，したがって回転におよぼす影響も大きいので，各変量の因子負荷量を共通性 $h^2_j = 1 - d^2_j$ で修正して，$b^2_{jk} \rightarrow b^2_{jk} / h^2_j$ として

$$V = \sum_{k=1}^{m} \left\{ p \sum_{j=1}^{p} \left(\frac{b_{jk}}{h_j}\right)^4 - \left(\sum_{j=1}^{p} \frac{b_{jk}^2}{h_j^2} \right)^2 \right\} \tag{84}$$

のような量を最大化する基準が提案された。この基準にもとづく回転法を基準バリマックス法（normal varimax method）と呼ぶ。通常たんにバリマックス法といえば，この基準バリマックス法をさしている。

このバリマックス法の具体的な計算手順や，そのために必要な数式の導出過程はかなり繁雑であるが，現代ではしかるべきコンピュータ・ソフトにそうした手順は組み込まれているので，応用の立場からは，いちいち数式を記憶している必要はない。

6．因子得点の推定

因子分析が主成分分析と異なる点のひとつは，因子負荷量が求められても個体のもつ因子得点はまだ定まらないという点である。これは，因子分析が現象の背後に特定のモデルを想定した分析であって，実際の観測値を，モデルで記述される法則が偶然的なぶれを含みながら発現した結果とみなしているからである。その意味で，因子分析は回帰分析などに近い性格を備えている。このため，因子分析の最終段階である因子得点の推定にあたっては，最小自乗法によってモデルの理論値と実際の観測値とのあいだの誤差を最小にする手続きが必要になる。この段階で活用されるのは重回帰分析である。

重回帰分析は，たとえば独立変数 x_1, x_2 と従属変数 y について n 個の観測値が与えられたとき，

$$y = a_0 + a_1x_1 + a_2x_2. \tag{85}$$

というモデルを立て，x_1, x_2 に実際の観測値 x_{1i}, x_{2i} （$i = 1, 2, \cdots\cdots, n$）をあてはめたとき得られる y の理論値と，実際の y の観測値 y_i （$i = 1, 2, \cdots\cdots, n$）とのあいだの誤差の2乗和を最小にしようとするものであるが，よく知られているように，その解である係数 a_0, a_1, a_2 は正規方程式

$$\begin{bmatrix} n & \sum x_{1i} & \sum x_{2i} \\ \sum x_{1i} & \sum x_{1i}^2 & \sum x_{1i}x_{2i} \\ \sum x_{2i} & \sum x_{2i}x_{1i} & \sum x_{2i}^2 \end{bmatrix} \begin{bmatrix} a_0 \\ a_1 \\ a_2 \end{bmatrix} = \begin{bmatrix} \sum y_{1i} \\ \sum y_i x_{1i} \\ \sum y_i x_{2i} \end{bmatrix} \tag{86}$$

を解くことで得られる。ところでこの正規方程式は，現象と理論モデルをつなぐ式である。

$$\begin{array}{cccccccc} \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{bmatrix} & = & \begin{bmatrix} 1 & x_{11} & x_{21} \\ 1 & x_{12} & x_{22} \\ \vdots & \vdots & \vdots \\ 1 & x_{1n} & x_{2n} \end{bmatrix} & & \begin{bmatrix} a_0 \\ a_1 \\ a_2 \end{bmatrix} & + & \begin{bmatrix} u_1 \\ u_2 \\ \vdots \\ u_n \end{bmatrix} \\ \equiv & & \equiv & & \equiv & & \uparrow \\ \boldsymbol{y} & & X & & \boldsymbol{b} & & (\text{誤差項}) \end{array} \tag{87}$$

にあらわれるベクトルや行列を用いて

$$X'X\boldsymbol{b} = X'\boldsymbol{y}$$

$$\therefore \boldsymbol{b} = (X'X)^{-1}X'\boldsymbol{y} \tag{88}$$

と表現することができる。

因子分析の場合の因子得点の推定には，これと同じ手法を用いる。いま，先の $p = 4$，$m = 2$ のモデルを例にとって，とりあえず第1因子得点だけについて考察しよう。

回帰式として

$$f_1 = b_{11}z_1 + b_{12}z_2 + b_{13}z_3 + b_{14}z_4 \tag{89}$$

という式を考える。$i = 1, 2, \cdots\cdots, n$ なるすべての観測値に対してこの式をあてはめ，最も当てはまりのよい係数 $b_{11}, \cdots\cdots, b_{14}$ を求めようというのである。

現象と理論モデルをつなぐ式は

$$\begin{bmatrix} f_{11} \\ f_{12} \\ \vdots \\ f_{1n} \end{bmatrix} = \begin{bmatrix} z_{11} & z_{21} & z_{31} & z_{41} \\ z_{12} & z_{22} & z_{32} & z_{42} \\ \vdots & \vdots & \vdots & \vdots \\ z_{1n} & z_{2n} & z_{3n} & z_{4n} \end{bmatrix} \begin{bmatrix} b_{11} \\ b_{12} \\ b_{13} \\ b_{14} \end{bmatrix} + \begin{bmatrix} u_1 \\ u_2 \\ \vdots \\ u_n \end{bmatrix} \tag{90}$$

$$\begin{array}{cccc} \shortparallel & \shortparallel & \shortparallel & \uparrow \\ \boldsymbol{f}_1 & Z & \boldsymbol{b}_1 & (\text{誤差項}) \end{array}$$

となる。(88)式の関係により，最適解として求められる $\boldsymbol{b}_1$ は

$$\boldsymbol{b}_1 = (Z'Z)^{-1}Z'\boldsymbol{f}_1 \tag{91}$$

となるはずである。

ところが

$$Z'Z = \begin{bmatrix} z_{11} & \cdots\cdots\cdots & z_{1n} \\ z_{21} & \cdots\cdots\cdots & z_{2n} \\ z_{31} & \cdots\cdots\cdots & z_{3n} \\ z_{41} & \cdots\cdots\cdots & z_{4n} \end{bmatrix} \begin{bmatrix} z_{11} & z_{21} & z_{31} & z_{41} \\ \vdots & \vdots & \vdots & \vdots \\ z_{1n} & z_{2n} & z_{3n} & z_{4n} \end{bmatrix} \tag{92}$$

となるので，変量 z_j（$j = 1, \cdots\cdots, 4$）が平均 0 に標準化されていたことを思い出すと，この行列の各成分は z_j（$j = 1, \cdots\cdots, 4$）の分散・共分散の n 倍になっており，この場合，分散共分散行列が相関行列に等しいことをも考慮すると

$$Z'Z = nR \tag{93}$$

となる。

次に，$Z'\boldsymbol{f}_1$ については，(64)式を考慮するとともに，共通因子相互間および共通因子と独自因子とのあいだは無相関という仮定，さらに f_1 の平均は 0 で分

散は1という仮定を思い出しながら変形すると，

$$
Z'\boldsymbol{f}_1 = \begin{bmatrix} z_{11} & \cdots\cdots\cdots & z_{1n} \\ z_{21} & \cdots\cdots\cdots & z_{2n} \\ z_{31} & \cdots\cdots\cdots & z_{3n} \\ z_{41} & \cdots\cdots\cdots & z_{4n} \end{bmatrix} \begin{bmatrix} f_{11} \\ \vdots \\ f_{1n} \end{bmatrix}
$$

$$
= \begin{bmatrix} \sum_{i=1}^{n} z_{1i} f_{1i} \\ \vdots \\ \sum_{i=1}^{n} z_{1i} f_{1i} \end{bmatrix} = \begin{bmatrix} \sum_{i=1}^{n} (a_{11} f_{1i} + a_{12} f_{2i} + e_{1i}) f_{1i} \\ \vdots \\ \sum_{i=1}^{n} (a_{11} f_{1i} + a_{12} f_{2i} + e_{1i}) f_{1i} \end{bmatrix}
$$

$$
= \begin{bmatrix} a_{11} \sum_{i=1}^{n} f_{1i}^2 \\ \vdots \\ a_{11} \sum_{i=1}^{n} f_{1i}^2 \end{bmatrix} = n \begin{bmatrix} a_{11} \\ \vdots \\ a_{41} \end{bmatrix} = n\boldsymbol{a}_1 \tag{94}
$$

となる。つまり，因子負荷量行列の第1列のn倍になるのである。

こうして，

$$
\boldsymbol{b}_1 = (nR)^{-1} n\boldsymbol{a}_1 = R^{-1}\boldsymbol{a}_1 \tag{95}
$$

によって第1因子得点の係数が求められることがわかった。同様にして，第2因子得点の係数は

$$
\boldsymbol{b}_2 = R^{-1}\boldsymbol{a}_2 \tag{96}
$$

によって求められる。

〔(下) に続く〕

参　考　文　献

[1]　竹内啓・柳井晴夫『多変量解析の基礎』東洋経済新報社，1972.

[2]　瀧好英『経済分析のための因子分析法』国元書房，1978.

[3]　芝祐順『行動科学における相関分析法』(第2版) 東京大学出版会，1982.

[4]　田中豊・脇本和昌『多変量統計解析法』現代数学社，1983.

[5]　大村平『多変量解析のはなし』日科技連，1985.

[6]　柳井晴夫・高木廣文編著『多変量解析ハンドブック』現代数学社，1986.

[7]　有馬哲・石村貞夫『多変量解析のはなし』東京図書，1987.

[8]　圓川隆夫『多変量のデータ解析』朝倉書店，1988.

[9]　塩谷實『多変量解析概論』朝倉書店，1990.

[10]　R.A. ジョンソン・D.W. ウィッチャン著，西田俊夫訳『多変量解析の徹底研究』現代数学社，1992.

[11]　B.F.J. マンリー著，村上正康・田栗正章訳『多変量解析の基礎』培風館，1992.

[12]　菅民郎『初心者がらくらく読める多変量解析の実践』上下，現代数学社，1993.