"Factor Analysis" について

鍋谷清治

I. 序論

本論文では

Factor Analysis : a Synthesis of Factorial Methods

by Karl J. Holzinger and Harry H. Harman

に述べられてゐる統計的分析法を紹介する。

こゝに云はれてゐる factor analysis といふのは Fisher が実験計画法で取扱ってゐる要因分析法とは全然異るものであつて，最近米国に於て心理実験などに盛に用ひられてゐる。

その歴史的な事実は明らかでないが起源は Spearman の two factor theory あたりではないかと思はれる。

扨，factor analysis とはどんな事かを説明しよう。

例へば人体の発育状態を調べるにはその人の身長，腕の長さ，

脚の長さ，体重，胸囲，等色々なものがその尺度となるであらう．此の場合，一般的な成長を表はす factor F_0，身長の程度を表はす factor F_1，ずんぐりした程度を表はす factor F_2 といふ三つの假想的な factor を考へて，身長，体重等がこれ等三つの factor の一次結合として如何に表はされるかを考へる。勿論身長，腕の長さ，脚の長さに対しては F_0, F_1 が相当大なる正係数を以て，体重，胸囲に対しては F_0, F_2 が相当大なる正係数を以て現れる事は確かであらう。そこで F_0, F_1, F_2 が身長，体重 ------ 等に夫々如何程の寄與をなすかを調べて或る主張をなさうとする。これが上に述べた factor analysis の方法である。

或る一つの主張をなすには，假想的な factor F_0, F_1, F_2 ---- 等は勿論，豫め与へられたものでなく，又観測する変数（身長体重 ---- 等）といふのもその問題に適当なものを選ぶべきである。かう考へて来ると，或る一つの主張をなすにも factor analysis の解は一意的ではなくなる事になるが，それは他の諸科学統計の中で他の諸分野にも共通した事である。

即ち人の好みなども入って来て，直面した問題に適当な変数適当な factor を選んで以下に述べる factor analysis の方法を適用すべきである。

或る一つの問題に於て，実際に測定する量は n ヶあって，

それ等を Z_j ($j=1,2,\cdots n$) とする。$Z_1 \cdots Z_n$ は母集団に於ては正規分布をなし，その平均値は夫々 $\mu_1, \cdots \mu_n$，標準偏差は夫々 $\sigma_1, \cdots, \sigma_n$ であると假定すれば

$$z_j = \frac{Z_j - \mu_j}{\sigma_j} \qquad (j=1,2,\cdots n)$$

の各々は平均値 0，標準偏差 1 の所謂 standard form になる。以後常に変数はこのやうに標準形になつてゐると假定する。

かやうな母集団から N ヶの標本を抽出して i 番目の標本に対する j 番目の変数の値を z_{ji} ($j=1,2,\cdots n$; $i=1,2,\cdots,N$) として置く。こゝでは標本数 N を相当大きくとつて，z_j の標本に於ける分布も，必要とあれば，母集団のそれに近い事を假定する。

factors は假想的なものであるが，それはやはり正規分布をなすものとし，母集団に於ては平均値 0，標準偏差 1 の標準形になつてゐるとする。

factor は大きく二つに分類される。幾つかの変数に共通の factor，即ち common factor と，各変数に特有な factor 即ち，unique factor である。前者は更に，全部の変数に共通なもの general factor と，或る一部分の変数のみが共有する group factor の二種類に分類される。

Common factors は m ヶあつてそれ等を $F_1, F_2 \cdots F_m$

で表はす事にし unique factors は丁度変数の数 n ケだけあるのであるから，それ等を $U_1, U_2, \cdots U_n$ で表はす事にする。

- Factors
  - Common factors
    - general factors
    - group factors
  - unique factors

$\to F_1 \cdots F_m$

$\to U_1, \cdots U_n$

そして z_{ji} を次の様に表はす事を考へる。

(1) $$z_{ji} = a_{j1}F_{1i} + a_{j2}F_{2i} + \cdots + a_{jm}F_{mi} + a_jU_{ji} \qquad \begin{pmatrix} j=1,2,\cdots n \\ i=1,2,\cdots N \end{pmatrix}$$

茲に z_{ji} は上述の通りであるが F_{si} $(S=1,2,\cdots m;\ i=1,2,\cdots N)$ は i 番目の標本に対する factor F_s の値を表はし，U_{ji} は Z_j に対する unique factor U_j の値を表はす。

そして $a_{j1}, a_{j2}, \cdots, a_{jm}$ 並に a_j $(j=1,2,\cdots n)$ は，母集団に於ける各要素には無関係に，Z_j 並に F_s, U_j の定義のみから定まる係数である。随って上の式は母集団に於ける要素には無関係に

(2) $$Z_j = a_{j1}F_1 + a_{j2}F_2 + \cdots + a_{jm}F_m + a_jU_j \qquad (j=1,2,\cdots n)$$

と書いてもよい訳である。

かう書いた時今後常に $U_1, U_2, ---- U_n$ は互に独立に，しかも $F_1, F_2, --- F_m$ とも独立に分布してゐると假定する。

$F_1, F_2 --- F_m$ の間では独立であつてもなくてもよい。

（2）を $j = 1, 2, --- n$ について書並べて

$$
\begin{cases}
Z_1 = a_{11}F_1 + a_{12}F_2 + \cdots\cdots + a_{1m}F_m + a_1 U_1 \\
Z_2 = a_{21}F_1 + a_{22}F_2 + \cdots + a_{2m}F_m \qquad + a_2 U_2 \\
\cdots\cdots\cdots\cdots\cdots\cdots \\
Z_n = a_{n1}F_1 + a_{n2}F_2 + \cdots + a_{nm}F_m \qquad + a_n U_n
\end{cases}
$$

と書くとき，これと（factor）pattern といふ。

Z_j と F_s の間の相関係数を $r_{Z_j F_s}$（$S = 1, 2, ---- m$），Z_j と U_j との間の相関係数を $r_{Z_j U_j}$，F_s と F_t との間の相関係数を $r_{F_s F_t}$（$s \neq t,\ s, t = 1, 2, --- m$）と書けば，（2）の両辺と $F_1, F_2, ---- F_m, U_j$ との間の相関係数を考へる事によつて次の式を得る。

$$
\begin{cases}
r_{Z_j F_1} = a_{j1} + a_{j2} r_{F_1 F_2} + \cdots\cdots + a_{jm} r_{F_1 F_m} \\
r_{Z_j F_2} = a_{j1} r_{F_2 F_1} + a_{j2} \quad + \cdots\cdots + a_{jm} r_{F_2 F_m} \\
\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots \qquad (j = 1, 2, --, n) \\
r_{Z_j F_m} = a_{j1} r_{F_m F_1} + a_{j2} r_{F_m F_2} + \cdots + a_{jm}
\end{cases}
$$

$$r_{Z_j U_j} = a_j$$

これ等の式を(factor) structureといふ。

一般に変数 Z_j ($j=1,2,\cdots n$) と因子 F_s ($S=1,2,\cdots m$) との関係を調べるには pattern と structure の両方が必要であるが，特に $F_1, F_2, \cdots F_m$ が互に無相関のときは structure は単に

$$\begin{cases} r_{z_jF_1}=a_{j1}, \quad r_{z_jF_2}=a_{j2}, \cdots\cdots, r_{z_jF_m}=a_{jm}, \\ r_{z_jU_j}=a_j \end{cases} \quad (j=1,2,\cdots,n)$$

となつて事情は簡単である。

以後当分の間 Common factors $F_1, F_2, \cdots, F_m$ が互に無相関の場合を取扱ふが，この場合には(2)の両辺の分散を計算すれば

$$1=a_{j1}^2+a_{j2}^2+\cdots\cdots+a_{jm}^2+a_j^2$$

所で U_j といふ factor は Z_j のみに現れるのであって，この中には更に他の変数を幾つか附加へれば Common factors となつて現れる部分もあれば，純然たる測定誤差による部分もあると考へられる。上の a_jU_j の中，前者を b_jS_j，後者を c_jT_j とすれば (S_j, T_j は共に平均値0，標準偏差1で互に無相関)

$$a_jU_j=b_jS_j+c_jT_j$$
$$a_j^2=b_j^2+c_j^2$$

そこで次の定義を用いる。（既に説明したものも纏めて置く）

$h_j^2 = \sum_{s=1}^{m} a_{js}^2 = \text{communality}$　　$F_s = \text{Common factor}$

$b_j^2 = \text{specificity}$　　$S_j = \text{specific factor}$

$c_j^2 = \text{unreliability}$　　$T_j = \text{unreliable factor}$

$a_j^2 = b_j^2 + c_j^2 = 1 - h_j^2 = \text{uniqueness}$　$U_j = \text{unique factor}$

$r_j = 1 - c_j^2 = h_j^2 + b_j^2 = \text{reliability}$

これ等は factor 並にその係数に対する名称である。Z_j の分散 1 の中 common factors による部分が h_j^2，測定誤差に関する分散が c_j^2（これは誤差に属するのだから信用出來ぬといふ意味であらう。）---- といつた類である。

そして窮極に於ては common factors に分解さるべき分散の中 $Z_1, Z_2, ---, Z_n$ といふ n ヶの変数を用ひることによつて分解された部分は，実際に何 % であつたかを示すものとして

$$H_j = \text{index of completeness of factrization}$$

$$= \frac{100 h_j^2}{h_j^2 + b_j^2} = 100 \frac{\text{communality}}{\text{reliability}}$$

が擧げられる。

以上の様に factor analysis に於ては変数も factors も標

準形になつてゐるので，その解を求めるには，観測より求めた変数相互間の相関係数のみを用ひる訳である。

Z_j と Z_k（$j \neq k$；$j,k = 1,2,\cdots,n$）の間の相関係数の観測値を r_{jk} とし，一度解を求めてから再生して出来た相関係数を $r'_{j,k}$ とする。即ち

$$r'_{j,k} = a_{j1}a_{k1} + a_{j2}a_{k2} + \cdots + a_{jm}a_{km} \quad (j \neq k;\ j,k=1,2,\cdots n)$$

標本数を相当大きくとれば

$$\overline{r}_{j,k} = r_{j,k} - r'_{j,k}$$

は 0 に近いものであつて，これを residual といふ。その標準偏差は大体 $\frac{1}{\sqrt{N}}$ となる。

次に $\overline{r}_{j,k} = 0$ と仮定すれば，即ち sampling の誤差を無視すれば

$$r_{j,k} = a_{j,1}a_{k,1} + a_{j,2}a_{k,2} + \cdots + a_{j,m}a_{k,m} \quad (j,k=1,2,\cdots n)$$

（但し　$r_{jj} = h_j^2 =$ communality　とする。）

そこで

$$R = \begin{pmatrix} r_{1,1} & \cdots & r_{1,n} \\ r_{2,1} & \cdots & r_{2,n} \\ \cdots & \cdots & \cdots \\ r_{n,1} & \cdots & r_{n,n} \end{pmatrix}, \quad A = \begin{pmatrix} a_{1,1} & \cdots & a_{1,m} \\ a_{2,1} & \cdots & a_{2,m} \\ \cdots & \cdots & \cdots \\ a_{n,1} & \cdots & a_{n,m} \end{pmatrix}$$

但し $r_{jj} = h_j^2$

と置けば

$$R = AA'$$

であり，Aは $n \times m$ 行列なので R の階段は高々 $m(\leq n)$ といふ事がわかる。故に，$\boxed{\text{定理}}$ $z_1, z_2, ---, z_n$ の correlation matrix（但し対角線には Communalities を置く）の階段が m であるならば，これ等 n ヶの変数を幾つかの factors の一次結合として表はすとき，その為に必要な common factors の最小数は m である。

次に如何なる factor solution が望ましいものであるかそれを述べて置かう。

勿論これは人の好みや問題の性質などの関係上，一概には云へない事は既に述べたのであるが，一応次の様な事が考へられる。

1. 各変数は common, specific, unreliable といふ三種の factors の一次結合となってゐる。勿論 specific, unreliable といふ二種の factors は，実際に解き求めるときには分解出来ないが，この構成を裏書きするやうなものである事が第一に望ましい。
2. 自然現象は一定の函数関係に従って解釈する場合には，その公式は簡潔で考へ易いものであることが望ましい。これと同様な事が factor analysis に於ても云はれる。

即ち次の二点である。

(a) common factors の数 m は，変数の数 n より遥かに小であること。

(b) 一つ一つの変数に於ても，それに 0 でない係数を以って含まれてゐる common factors の数は少い程よい。

3. Common factors の間は互に無相関の方がよい。かやうな解を orthogonal solution といふ。然らざるときの解は oblique solution といふ。

4. factor F_s $(s=1,2,\cdots,m)$ の，すべての変数に対する contribution を $d_s=\sum_{j=1}^{n} a_{js}^2$ で表はすことが出来るが，この場合，問題の性質によって次の三つの何れかが適当である。

(a) $d_1, d_2, \cdots, d_m$ は順次小さくなる.

(b) $d_1, d_2, \cdots, d_m$ は大凡等しい大さである。

(c) d_1 のみが特に大きく，$d_2, \cdots, d_m$ は小さくて大体等しい大きさである。

5. (a) 互に相関の高い一群の変数は唯一つの common factors によって表はされてゐると考へることが出来る。m 次元空間内で $P_1(a_{1,1}\cdots a_{1,m})$，$P_2(a_{2,1}\cdots a_{2,m})$ $\cdots$ $P_n(a_{n1}\cdots a_{n,m})$ といふ n ヶの点が考へられるが，相関が高い一群の変数があれば，それ等に対応する

vectors OP_j は一つの軸の周りに互に小なる角度をなしてゐるわけである。この様な時には，その中心軸に相当する factor を選ぶとよい。

(b) 一つの軸の周りでなく，原点Oを通る一つの平面の周りに幾つかの OP_j が集ってゐる時には，その平面内に二つの factors をとる事が適当である。

(C) 一般に $(m-1)$ 次元，或はそれより小なる或る次元の linear space の周りに横たはってゐる場合にも同様の事がいへる。

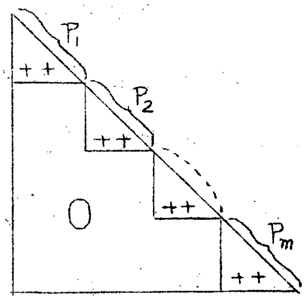
6. $(z_{1i}, z_{2i}, ---, z_{ni})$ $(i=1,2,---N)$ といふ n 次元空間内の N ヶの点を考へれば，それ等は原点を含む或る m 次元空間の近くにある訳である。そしてその中で密度の等しい点は，この m 次元空間の楕円面になってゐるわけである。この楕円の主軸の方向に当る factors をとる事が望ましい。

これ等の基準を満足する解として如何なるものが挙げられてゐるかを次に示さう。

1. Uni-factor Pattern

これは各変数が唯一つの factor を含む場合で，その係数は正である。これは最も簡単な解であって factor solution としては理想的なものである。unique factor を除けば次の様に書ける。

$$
\begin{aligned}
Z_1 &= a_{1,1}F_1 \\
Z_2 &= a_{2,1}F_1 \\
&\vdots \\
Z_{p_1} &= a_{p_1,1}F_1 \\
Z_{p_1+1} &= \qquad a_{p_1+1,2}F_2 \\
&\vdots \\
Z_{p_1+p_2} &= \qquad a_{p_1+p_2,2}F_2 \\
&\vdots \\
Z_n &= \qquad\qquad a_{nm}F_m
\end{aligned}
$$



又，かやうな pattern に於ける correlation matrix は上の如く示される。ここで ++ は相当大きな正の量を表はし，O は 0 に近い値を表はす。

factor analysis の起りと思はれる Spearman の two-factor pattern といふのは uni-factor pattern の特別な場合であつて，common factor が唯一つの時である。

即ち Z_j ($j=1,2,\cdots n$) は general factor F_0 と unique factor U_j との一次結合になつてゐる場合である。

$$
\begin{aligned}
Z_1 &= a_{10}F_0 + a_1U_1 \\
Z_2 &= a_{20}F_0 \qquad + a_2U_2 \\
&\cdots\cdots\cdots \\
Z_n &= a_{n0}F_0 \qquad\qquad + a_nU_n
\end{aligned}
$$

#71

2. Bi-factor Pattern

これは各変数が一つの general factor と一つの group factor を含む場合で，前の様に factors の係数はすべて正なる事を必要とする。 Unique factors を除いた pattern 及び correlation matrix の形を書けば次の様になる。

$$Z_1 = a_{10}F_0 + a_{11}F_1$$

$$Z_2 = a_{20}F_0 + a_{21}F_1$$

$$\vdots$$

$$Z_{p_1} = a_{p_1 0}F_0 + a_{p_1 1}F_1$$

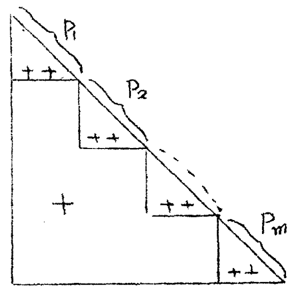
$$Z_{p_1+1} = a_{p_1+1\,0}F_0 \qquad\qquad + a_{p_1+1,2}F_2$$

$$\vdots$$

$$Z_{p_1+p_2} = a_{p_1+p_2\,0}F_0 \qquad\qquad + a_{p_1+p_2,2}F_2$$

$$\vdots$$

$$Z_n = a_{n,0}F_0 \qquad\qquad\qquad\qquad + a_{nm}F_m$$



ここで

$$r'_{j.k}\begin{cases} = a_{j,0}a_{k,0} & (Z_j と Z_k が異る group factor を有する場合) \\ = a_{j,0}a_{k,0} + a_{j,s}a_{k,s} & (Z_j と Z_k が同じ group factor F_s を有する場合) \end{cases}$$

だから上の様な correlation matrix の形が得られる訳である。即ち同じ group factor を有する変数の間では相関係数が大きな正の値で，異る group factor を有するものの間では小さな正数である。

3. Multiple-factor Pattern

これは general factor がなくて pattern の各行各列に少くも一つ 0 があり，係数は皆正なるものである。

その correlation matrix は bi-factor pattern の場合と同様だから省略して，multiple-factor pattern の一例を示して置かう。

$$
\begin{aligned}
Z_1 &= a_{1,1}F_1 + a_{1,2}F_2 + a_{1,3}F_3 \\
Z_2 &= a_{2,1}F_1 + a_{2,2}F_2 \qquad\qquad + a_{2,4}F_4 \\
Z_3 &= a_{3,1}F_1 + a_{3,2}F_2 \\
Z_4 &= \qquad\qquad a_{42}F_2 + a_{43}F_3 + a_{44}F_4 \\
Z_5 &= a_{5,1}F_1 \qquad\qquad + a_{53}F_3 + a_{54}F_4 \\
Z_6 &= \qquad\qquad a_{62}F_2 \qquad\qquad + a_{64}F_4
\end{aligned}
$$

4. Principal Factor Pattern

これは少くも一つの general factor を含み，他の factors はその contribution d_s が順次小さくなるものである。

1，2，3，の場合は何れも正係数しか許されなかつたが今度の場合はさうとは限らない。

以上四つの patterns の有する性質を次に要約して置かう。

Factor Pattern	Assumptions	Complexity
uni-	1. 2a. 2b, 3, 4b, 5a	1
Bi-	1. 2a, 2b. 3, 4C, 5b	2
multiple-	1, 2a, 2b, 3, 4b, 5C.	< m
Principal	1 2a 3.4a.6	m

4.73

これだけ準備して置いて factor analysis の解法を次に説明して行く。

factor solution を得る為に先づ最初になすべき事は common factors の数 m を決定することである。

これは上掲の定理によれば，correlation matrix の階数より決めらるべきものである。併しその correlation matrix は対角線に communalities を置かなくてはならないが，communalities は観測の結果からは得られない。

そこで，逆の方法をとって，先づ correlation matrix の階数即ち common factors の数を別の面から計算し，その階数になるやうに communalities を定めるのである。

Common factors の数とか，どの変数がどの factors を含むとかいふ事は主として過去の研究から解ることなのであるが，それがない場合には次に述べる B－係数の方法を用ひる。B－係数を定義する為に次の記号を用ひる。

- n ＝ 全体の変数の数
- u ： 変数全体の或る部分集合であって，これが B－係数の定義域をなす。
- p ＝ 部分集合 u の中に入ってゐる変数の数
- C.u ＝ u に入らない変数の集合

$$S = \sum_{\substack{j \lessgtr k \\ j,k \in u}} r_{j,k} \qquad T = \sum_{\substack{j \in u \\ \alpha \in cu}} r_{j,\alpha}$$

この時，u といふ集合の B－係数 B（u）を次の式で定義する．

$$B(u) = 100 \frac{\dfrac{S}{\binom{P}{2}}}{\dfrac{T}{P(n-P)}} = \frac{200(n-P)S}{(P-1)T}$$

ここで S は u に属する変数間の相関係数の和であつて $\dfrac{S}{\binom{P}{2}}$ はその平均，T は u に属する変数と cu に属する変数との間の相関係数の和であつて $\dfrac{T}{P(n-P)}$ はその平均となる。随つて B（u）はこれ等二つの平均の比の 100 倍といふ事になる。

例を挙げれば

$n = 7$，$u = (1, 3, 5)$，$P = 3$，$cu = (2, 4, 6, 7)$，$n-P=4$

として計算すると

$$B(u) = 400 \frac{r_{1.3} + r_{1.5} + r_{3.5}}{r_{1.2} + r_{1.4} + r_{1.6} + r_{1.7} + r_{3.2} + r_{3.6} + r_{3.7} + r_{5.2} + r_{5.4} + r_{5.6} + r_{5.7}}$$

B－係数の方法が用ひられるのは，すべての相関係数が正なる場合であるが，これを用ひて次の方法で変数を分類する。

先づ相関係数の最大のものを求めて，それが例へば $r_{1.3}$ であれば $u = (1, 3)$ として $B((1, 3))$ を計算する。

次に Z_1, Z_3 の両方と，大きな相関係数を有する変数をとつて，それが例へば Z_5 であれば，今度は $\mathfrak{U} = (1, 3, 5)$ として B((1,3,5)) を計算する。 B((1,3,5)) が B((1,3)) に比して相当小さくなれば Z_5 を除いて (Z_1, Z_3) を一つの群とする。

B((1,3,5)) が B((1,3)) に比してあまり小さくなつてゐない時は更に他の変数例へば Z_4 をとつて B((1,3,4,5)) を計算する。 そしてこれが B((1,3,5)) に比して相当小さくなつていれば Z_4 を除いて (Z_1, Z_3, Z_5) を一つの群とする。

B((1,3,5)) と B((1,3,4,5)) との差がそれ程大きくない時は更に同様の方法を続けて行く。

一つの群が定まれば，今度はその余集合の間で最大の相関係数を有する二つの変数から始めて，又同じ方法で他の群を定めて行く。------- と云つた工合である。但し今度は前に定めた群に属する変数は，$C\mathfrak{U}$ の中から除いて計算を進めるのである。

著者は更に，或る一つの集合 $\mathfrak{U}$ に属する変数が群を作る為には $B(\mathfrak{U}) \geqq 130$ を要請している。尚，上で B(1,3,5) が B(1,3) に比して相当小さくなると云ふ言葉を使つたが，ここは厳密には有意検定を行ふべき所であらう。併しこの方面の研究は進んでゐない。

このやうにして n ヶの変数が幾つかの群に分れるが，その

各群の変数の間では相関係数が大なので，それらは同一の factors から成つておると考へられる。そこで

$$\text{common factors の数} = \text{群の数} = m$$

と置く事が一応考へられる。次にはかう假定して計算を進める。

扨

$$R = \begin{pmatrix} h_1^2, & r_{12}, & \cdots\cdots & r_{1,n} \\ r_{2,1}, & h_2^2, & \cdots\cdots & r_{2,n} \\ \cdots & \cdots & \cdots & \cdots \\ r_{n,1} & r_{n,2} & \cdots\cdots & h_n^2 \end{pmatrix}$$

に於て R の階数が m であるといふ事は

1. R の或る m 次の主座行列式 R' が 0 でない。

2. R' を含む $m+1$ 次並に $m+2$ 次の主座行列式は全部 0 である

ここで條件 2. は $V_m = (n-m) + \binom{n-m}{2} = \frac{(n-m)(n-m+1)}{2}$ だけの式が 0 になる事を意味する。R に於て r_{ik} ($i \neq k$) は観測の結果より与へられるものであつて $h_1^2, h_2^2 - h_n^2$ を未知数と考へれば、未知数は n ケあつて R が階数 m になる為の條件には V_m ケの式を必要とする。之等の式が独立な事は嚴密な證明を要するが，其の場合には

$$\phi(m) = V_m - n = \frac{m^2 - (2n+1)m + n(n-1)}{2} \leqq 0$$

即ち

$$\frac{2n+1+\sqrt{8n+1}}{2} \geqq m \geqq \frac{2n+1-\sqrt{8n+1}}{2}$$

なるとき 解 $h_1^2, h_2^2, \cdots\cdots h_n^2$ が存在する。即ち $h_1^2, h_2^2, \cdots\cdots, h_n^2$ を適当にとれば R の階数を m 以下にする事が出來るのである。次に $\phi(m)$ の表を掲げて置かう。

$m \backslash n$	2	3	4	5	6	7	8	9	10	11	12	n
1	-1	0	2	5	9	14	20	27	35	44	54	$\binom{n-1}{2}-1=\frac{n(n-3)}{2}$
2	-2	-2	-1	1	4	8	13	19	26	34	43	$\binom{n-2}{2}-2=\frac{n(n-5)}{2}+1$
3		-3	-3	-2	0	3	7	12	18	25	33	$\binom{n-3}{2}-3=\frac{n(n-7)}{2}+3$
4			-4	-4	-3	-1	2	6	11	17	24	$\binom{n-4}{2}-4=\frac{n(n-9)}{2}+6$
5				-5	-5	-4	-2	1	5	10	16	$\binom{n-5}{2}-5=\frac{n(n-11)}{2}+10$
6					-6	-6	-5	3	0	4	9	$\binom{n-6}{2}-6=\frac{n(n-13)}{2}+15$
m												$\binom{n-m}{2}-m$

$\phi(m) \leqq 0$ ならば，観測の結果與へられた r_{jk} $(j \neq k)$ の如何に係らず，R の階数が m 以下になるやうに h_1^2, h_2^2, --- h_n^2 を定める事が出来るといふ事は，上述の通りであるが，$\phi(m) > 0$ ならば R の階数を m 以下にする為には，r_{jk} $(j \neq k)$ の間に $\phi(m)$ ヶ の條件が成立することを必要とする。

勿論ここにいふ條件といふのは *sampling error* の範囲内で成立つ等式であればよいのである。

これ等の條件といふのは実際にどういふ式で表はされるであらうか。

Rの階数を１にする為には互に相異る e, j, k に対して

$$\begin{vmatrix} h_e^2 & r_{ej} \\ r_{ek} & r_{jk} \end{vmatrix} = 0$$

これより

$$h_e^2 = \frac{r_{ej}\, r_{ek}}{r_{jk}}$$

を得る。この式は e を止めて置くとき，すべての j, k の組について成立すべきであるから，それ等の組については右辺は皆等しくなければならない。例へば $e=1$ のとき，

$$\frac{r_{12}r_{13}}{r_{23}} = \frac{r_{12}r_{14}}{r_{24}} = \cdots\cdots = \frac{r_{12}r_{1n}}{r_{2n}} = \frac{r_{13}r_{14}}{r_{34}} = \cdots = \frac{r_{13}r_{1n}}{r_{3n}} = \cdots = \frac{r_{1,n-1}r_{1n}}{r_{n-1,n}}$$

これが成立てば，同様な式が $e = 2, 3, \cdots, n$ に対しても成立つ事は明らかである。これは丁度 $\phi(1)$ ヶの式であって，Rの階数を1にする為の條件である。

但し，上の表から明らかなやうに，$n=2$ 又は $n=3$ であれば $\phi(1) \leqq 0$ となるから，その場合にはこの條件は不要である。

次に R の階数が1 になるといふのはどういふ事かを考へよう。これは $Z_1, Z_2, \cdots Z_n$ を表はす common factor が唯一つである事を意味する。即ち一つの general factor F_0 をとって

$$\begin{aligned}
Z_1 &= a_{10}F_0 + a_1 U_1 \\
Z_2 &= a_{20}F_0 \qquad\quad + a_2 U_2 \\
&\cdots\cdots\cdots\cdots\cdots \\
Z_n &= a_{n0}F_0 \qquad\qquad\qquad + a_n U_n
\end{aligned}$$

と表はされる場合である。これは前に述べた Spearman の two-factor pattern である。

この場合 $\overline{r}_{j,k}=0$ 即ち $r_{j,k}=r'_{j,k}$ $(j \neq k)$ と假定すれば

$$r_{j,k} = r'_{j,k} = a_{j,0}\, a_{k,0}$$

両辺に a_{e0}^2 を掛ければ

$$a_{e0}^2\, r_{jk} = (a_{e0}\, a_{j0})(a_{e0}\, a_{k0}) = r_{ej}\, r_{ek}$$

ここで e, j, k は皆互に相異るものとして

$$(3) \qquad a_{e,0}^2 = h_e^2 = \frac{1}{\binom{n-1}{2}} \sum_{\substack{j<k \\ j,k \neq e}} \frac{r_{ej}\, r_{ek}}{r_{jk}}$$

又は

$$a_{e,0}^2 = h_e^2 = \frac{\sum_{\substack{j<k \\ j,k \neq e}} r_{ej}\, r_{ek}}{\sum_{\substack{j<k \\ j,k \neq e}} r_{jk}}$$

によって a_e が求まる。

1. Uni-factor Pattern.

これを求めるには先づ B-係数によって変数の群の数 m を求め，その群を G_1, G_2, ---- G_m で表はす事にする。そして之れ等は夫々 P_1, P_2 --- P_m ヶの変数を含むものとして夫々の群の中で Spearman の two-factor solution を求めれば，これによって uni-factor solution が得られる。

2. Bi-factor Pattern

これを求めるにも，やはり先づ変数の群 $G_1, G_2, \cdots G_m$ を定め，それ等は夫々 $P_1, P_2, \cdots, P_m$ ヶの変数を含むものと假定して置く。そして bi-factor pattern を

$$
\begin{aligned}
Z_1 &= a_{10}F_0 + a_{11}F_1 && + a_1 U_1 \\
&\vdots \\
Z_{p_1} &= a_{p_1 0}F_0 + a_{p_1 1}F_1 && + a_{p_1} U_{p_1} \\
Z_{p_1+1} &= a_{p_1+1\,0}F_0 + a_{p_1+1,2}F && + a_{p_1+1} U_{p_1+1} \\
&\vdots \\
Z_n &= a_{n0}F + a_{nm}F_m && + a_n U_n
\end{aligned}
$$

とする。但し

$$G_1 = (1, 2, \cdots P_1),\ G_2 = (P_1+1, \cdots P_1+P_2), \cdots\cdots,\ G_m = (n-P_m+1, \cdots n)$$

となるやうに変数の番号を附け変へたものとして置く。

このとき $e \in G_u,\ j \in G_s,\ k \in G_t$ として u, s, t が互に相異る場合には，前と同様に $r_{jk} = r'_{jk}$ $(j \neq k)$ として次の式が成立する。

$$r_{j,k} = a_{j,0}\, a_{k,0}$$

両辺に a_{e0}^2 を掛ければ

$$a_{e0}^2 r_{jk} = (a_{e0} a_{j0})(a_{e0} a_{k0}) = r_{ej} r_{ek}$$

これより

$$a_{e0}^2 = \frac{\sum_{j,k} r_{ej} r_{ek}}{\sum_{j,k} r_{jk}} \quad \left(又、\ a_{e0}^2 = \frac{1}{\nu} \sum_{j,k} \frac{r_{ej} r_{ek}}{r_{jk}} \right)$$

但し j, k は夫々 $j \in G_s$, $k \in G_t$ う $s < t$, $s, t \neq u$ となるやうなすべての j, k に対する和を作るやうに動き，ν はそれに用ひる和の数である。

例へば　$G_1 = (1, 2, 3)$, $G_2 = (4, 5, 6, 7)$, $G_3 = (8, 9, 10, 11)$, $G_4 = (12, 13, 14)$ とするとき

$$a_{10}^2 = \frac{N}{D}$$

但し

$$N = \left(\sum_{j=4}^{7} r_{1,j} \right) \left(\sum_{k=8}^{14} r_{1,k} \right) + \left(\sum_{j=8}^{11} r_{1,j} \right) \left(\sum_{k=12}^{14} r_{1,k} \right)$$

$$D = \sum_{j=4}^{7} \sum_{k=8}^{14} r_{jk} + \sum_{j=8}^{11} \sum_{k=12}^{14} r_{j,k}$$

かうして general factor の係数 a_{j0} ($j = 1, 2, \cdots n$) が求まった後は，$r_{j,k}$ の剰余即ち general factor residual を

$$\dot{r}_{jk} = r_{jk} - a_{j0} a_{k0} \qquad (j \neq k)$$

によって定義する。若しも bi-factor pattern の假定が正しく，変数の群別が正当に行はれてゐるならば，$\dot{r}_{j,k}$ は j と

kが異る群に属するときは sampling error を除いて 0 となり，同一の群の中では $r_{j,k}$ のなす行列は階数 1 となる。

j, k は G_s の中を動くとすれば

$$\dot{r}_{j,k} = a_{j,s}\, a_{k,s} \qquad (j \neq k,\ j,k \in G_s)$$

故に前と同様に両辺に $a_{e,s}^2$ $(e \in G_s)$ を掛ければ

$$a_{e,s}^2\, \dot{r}_{jk} = (a_{e,s}\, a_{j,s})(a_{e,s}\, a_{k,s}) = \dot{r}_{ej}\, \dot{r}_{ek}$$

随って

$$a_{es}^2 = \frac{1}{\binom{P_s - 1}{2}} \sum_{j,k} \frac{\dot{r}_{ej}\, r_{ek}}{\dot{r}_{jk}}$$

又は

$$a_{es}^2 = \frac{\sum_{j,k} \dot{r}_{ej}\, \dot{r}_{ek}}{\sum_{j,k} \dot{r}_{jk}}$$

によって group factor の係数 a_{es} が求まる訳である。但し上の二式共 $\sum_{j,k}$ は $j < k,\ j,k \neq e,\ j,k \in G_s$ となるやうなすべての j, k に対する和を表はす。

かうして $a_{e,s}$ $(e \in G_s,\ s = 1, 2, \cdots, m)$ を求めた後は $r_{j,k}$ の剰余（final residual）は

$$\bar{r}_{j,k} = \begin{cases} \dot{r}_{jk} - a_{js}\, a_{ks} & (j,k \in G_s \text{ のとき}) \\ \dot{r}_{jk} & (j \in G_s,\ k \in G_t,\ s \neq t \text{ のとき}) \end{cases}$$

となる。最初に假定した bi-factor pattern の形が正しけ

れば，$\bar{r}_{jk}$ は sampling error を除いて 0 に等しい訳である。若しさうなってゐなければ pattern の形に不合理があつたのであって，その場合にはそれに相当する補正を施さなくてはならない。

以上に於て屢々"sampling error を除いて"云々といふ言葉を用ひて来たが，その為には厳密にいへば問題になっている夫々の量の sampling variance を用ひて有意検定を行ふべきである。大標本に対してこれらの公式を挙げて置かう。

先づ

$$a_{e0}^2 = t_{jk} = t,\ r_{ej} = a,\ r_{ek} = b,\ r_{jk} = c,$$

とすれば

$$t = \frac{ab}{c}$$

であるから (3) を用ひて a_{e0} を出すとき，その sampling variance の公式を求めよう。その為に両辺の対数をとれば

$$\log t = \log a + \log b - \log c$$

両辺の微分をとって

$$\frac{dt}{t} = \frac{da}{a} + \frac{db}{b} - \frac{dc}{c}$$

両辺を平方してその平均値をとれば

$$\frac{\sigma_t^2}{t^2} = \frac{\sigma_a^2}{a^2} + \frac{\sigma_b^2}{b^2} + \frac{\sigma_c^2}{c^2} + 2\left(\frac{\sigma_a\sigma_b}{ab}r_{ab} - \frac{\sigma_a\sigma_c}{ac}r_{ac} - \frac{\sigma_b\sigma_c}{bc}r_{bc}\right)$$

但しここで，σ_t^2，σ_a^2，----は夫々 t，a，----の sampling variance であり，$r_{a.b}$ は a と b との相関係数である。

所が既に次の公式はわかつている。

$$\sigma_a^2 = \frac{(1-a^2)^2}{N} \qquad \text{etc.}$$

$$r_{ab} = c - \frac{ab(1-a^2-b^2-c^2+2abc)}{2(1-a^2)(1-b^2)} \quad \text{etc.}$$

(Peason-Filon，公式)

そこでこれ等の公式を用ひれば

$$N\sigma_t^2 = t\left[t(1-2t)\left(\frac{1}{a^2}+\frac{1}{b^2}\right)+\frac{t}{c^2}+2(1-t)(1-2t)+t\right]$$

又 $d\sqrt{t} = \dfrac{dt}{2\sqrt{t}}$ であるから

$$4N\sigma_{\sqrt{t}}^2 = \frac{1}{t}N\sigma_t^2 = 2+\frac{a}{bc}+\frac{b}{ac}-2\left(\frac{a}{c}\right)^2-2\left(\frac{b}{c}\right)^2 + \frac{ab}{c^3}-5\frac{ab}{c}+4\left(\frac{ab}{c}\right)^2$$

同じ $a_e = \sqrt{t}$ を求めるにも j, k を色々に変へる事によつて a, b, c の取り方は色々あるが，あらゆる可能な取り方に対する a, b, c の平均値を夫々 $\bar{a}, \bar{b}, \bar{c}$ で表はせば，Tailor展開の第一項のみを用いる事によつて

$$(4) \qquad \sigma_{a_{e0}}^2 = \frac{1}{4N}\left(2+\frac{\bar{a}}{\bar{b}\bar{c}}+\frac{\bar{b}}{\bar{a}\bar{c}}-\frac{2\bar{a}^2}{\bar{c}^2}-\frac{2\bar{b}^2}{\bar{c}^2}+\frac{\bar{a}\bar{b}}{\bar{c}^3}-\frac{5\bar{a}\bar{b}}{\bar{c}}+\frac{4\bar{a}^2\bar{b}^2}{\bar{c}^2}\right)$$

この式が実際に用ひられるのであるが，j, k の変化に対する t の変化は非常に取扱ひにくいものであるので，かように t の

分散の平均を用いた訳である。

bi-factor solution を得るときには，*general factor residual* が問題になるが，それは

$$\dot{r}_{jk} = r_{jk} - a_{j0}\, a_{k0}$$

であるから

$$d\dot{r}_{jk} = dr_{jk} - a_{j0}\, da_{k0} - a_{k0}\, da_{j0}$$

故に
$$\sigma^2_{\dot{r}_{jk}} = \sigma^2_{r_{jk}} + a^2_{j0}\, d\sigma^2_{a_{k0}} + a^2_{k0}\sigma^2_{a_{j0}} + 2\{a_{j0}a_{k0}\, cov(a_{j0}, a_{k0}) - \cdots\}$$

となるが，ここで $2\{\ \ \}$ の項を無視すれば次の近似式が得られる。

(5) $$\sigma^2_{\dot{r}_{jk}} = \sigma^2_{r_{jk}} + a^2_{j0}\sigma^2_{a_{k0}} + a^2_{k0}\sigma^2_{a_{j0}}$$

次に同じ群に属する二つの変数の間で相関係数の *final residual* をつくれば

$$\bar{r}_{jk} = r_{jk} - a_{j0}\, a_{k0} - a_{js}\, a_{ks} \qquad (j, k \in G_s)$$

これの分散を求めれば，前と同様の近似法を用いて

(6) $$\sigma^2_{\bar{r}_{jk}} = \sigma^2_{r_{jk}} + a^2_{j0}\sigma^2_{a_{k0}} + a^2_{k0}\sigma^2_{a_{j0}} + a^2_{js}\sigma^2_{a_{ks}} + a^2_{ks}\sigma^2_{a_{js}}$$

観測より計算された相関係数がすべて ρ に等しいと看做すこ

との出来るときには，(4)(5)は更に簡単に次のやうになる。

$$\sigma^2_{a_{e0}} = \frac{1}{4N}\left(\frac{3}{\rho} - 2 - 5\rho + 4\rho^2\right)$$

$$\sigma^2_{\dot{r}_{jk}} = \frac{(1-r_{jk}^2)^2}{N} + \frac{a_{j0}^2 + a_{k0}}{4N}\left(\frac{3}{\rho} - 2 - 5\rho + 4\rho^2\right) \fallingdotseq \frac{(1-\rho)^2(5+8\rho+2\rho^2)}{2N}$$

更に同一の群の中で $\dot{r}_{jk}$ がすべて ρ_1 に等しいと看做せるならば (6) は

$$\sigma^2_{\bar{r}_{jk}} = \frac{(1-\rho)^2(5+8\rho+2\rho^2)}{2N} + \frac{\rho_1}{2N}\left(\frac{3}{\rho_1} - 2 - 5\rho_1 + 4\rho_1^2\right)$$

となる。

かようにして計算された種々の分散の式が剰余の有意検定に用ひられるのである。

3. Principal Factor Pattern

これを求めるには先づ communalities を求めなければならないが，その求め方は既に述べた所である。

これを求めて $h_j^2 = r_{jj}$ $(j = 1, 2, \cdots, n)$ と置けば

$$z_j = a_{j1}F_1 + a_{j2}F_2 + \cdots + a_{jm}F_m + a_j U_j \quad (j=1,2,\cdots,n)$$

により，sampling error を除いて

$$r_{jk} = r'_{jk} = \sum_{t=1}^{m} a_{jt}a_{kt} \quad (j, k = 1, 2, \cdots, n)$$

となる。

r_{jk} は與へられてゐて，a_{jt} $(j = 1, 2, \cdots n;\ t = 1, 2, \cdots m)$ を

求めるには　先づ

$$A_1 = a_{11}^2 + a_{21}^2 + --- + a_{n1}^2$$

と置いて A_1 を最大にすることを考へる。その為には La-grange の multiplier μ_{jk} ($j,k = 1,2,---n, \mu_{jk} = \mu_{kj}$) を使って

$$2T = A_1 - \sum_{j,k=1}^{n} \mu_{jk} r_{jk} = A_1 - \sum_{j,k=1}^{n} \sum_{t=1}^{m} \mu_{jk} a_{jt} a_{kt}$$

と置く。

これを， a_{jt} ($j = 1,2,---n$; $t = 1,2,---m$) について偏微分すれば

(7)
$$\frac{\partial T}{\partial a_{j1}} = a_{j1} - \sum_{k=1}^{n} \mu_{jk} a_{k1} = 0$$

$$\frac{\partial T}{\partial a_{jt}} = -\sum_{k=1}^{n} \mu_{jk} a_{kt} = 0 \qquad (t \neq 1)$$

或は纏めて
$$\frac{\partial T}{\partial a_{jt}} = \delta_{1t} a_{j1} - \sum_{k=1}^{n} \mu_{jk} a_{kt} = 0 \quad (t = 1,2,--- m)$$

この式の両辺に a_{j1} を掛けて j について加へれば

$$\delta_{1t} \sum_{j=1}^{n} a_{j1}^2 - \sum_{j=1}^{n} \sum_{k=1}^{n} \mu_{jk} a_{j1} a_{kt} = 0$$

次に A_1 の即ち $\sum_{j=1}^{n} a_{j1}^2$ の最大値を λ_1 と置けば (7) により

$$\sum_{j=1}^{n} \mu_{jk} a_{j1} = a_{k1}$$

であるから

$$\delta_{1t}\lambda - \sum_{k=1}^{n} a_{k1} a_{kt} = 0$$

この式の両辺に a_{jt} を掛けて t について加へれば

$$\lambda_1 a_{j1} - \sum_{k=1}^{n} \left(\sum_{t=1}^{m} a_{jt} a_{kt} \right) a_{k1} = 0$$

即ち

$$\lambda_1 a_{j1} - \sum_{k=1}^{n} r_{jk} a_{k1} = 0 \quad (j=1,2,\cdots n)$$

依って λ_1 は matrix R の特有方程式 $|\lambda_1 \delta_{jk} - r_{jk}| = 0$ の最大根であって、これが $A_1 = \sum_{j=1}^{n} a_{j1}^2$ の最大値を与える事になる。そして $a_{11}, a_{21}, \cdots a_{n1}$ は $\lambda_1 = \sum_{j=1}^{n} a_{j1}^2$ を満足する所の λ_1 に対する固有 vector である。

$R = (r_{jk})$ の如き対称行列に対する固有 vector 並に固有値を求める事は周知の通りであるが、一応述べて置かう。

先づ任意の vector $\alpha = \begin{pmatrix} \alpha_1 \\ \alpha_2 \\ \vdots \\ \alpha_n \end{pmatrix}$ より始めて $R\alpha$, $R^2\alpha$, $R^3\alpha \cdots$ を順次作って行けば、これ等の vectors の方向は段々一定して行く訳である。そこで $R\alpha$, $R^2\alpha$, $R^3\alpha \cdots$ に於て絶対値を一定にして置くとか、第一座標を一定にして置くやうに normalize して置けば、すべての座標は一定の極限に達する筈である。その極限の vector を $\bar{\alpha} = \begin{pmatrix} \bar{\alpha}_1 \\ \bar{\alpha}_2 \\ \bar{\alpha}_n \end{pmatrix}$ と置けば

$$R\bar{\alpha} = \lambda_1 \bar{\alpha}$$

は二つの common factors を F_1, F_2 を用ひてあるものとし，(1, 2, 3, 4)，(5, 6, 7, 8) といふ組が二つの群になってゐるとする。このときその pattern を，

$$Z_j = a_{j1}F_1 + a_{j2}F_2 + a_j U_j \quad (j = 1, 2, \cdots, 8)$$

とすれば 8ヶの点 $P_j(a_{j1}, a_{j2})$ が出来るが，そのとき，P_1, P_2, P_3, P_4 の重心を通る線，P_5, P_6, P_7, P_8 の重心を通る線を二つの新しい斜交座標軸にとる。これに相当する factors を γ_1, γ_2，それ等の方向余弦を夫々 $(\lambda_{11}, \lambda_{21})$, $(\lambda_{12}, \lambda_{22})$ とすれば

$$r_{\gamma_1\gamma_2} = \lambda_{11}\lambda_{12} + \lambda_{21}\lambda_{22}$$

これによって γ_1, γ_2 の間の相関係数が求まった訳である。

次に factor structure は

$$(r_{j\gamma_1}\ r_{j\gamma_2}) = (a_{j1}\ a_{j2})\begin{pmatrix}\lambda_{11} & \lambda_{12} \\ \lambda_{21} & \lambda_{22}\end{pmatrix}$$

$$(j = 1, 2, \cdots, 8)$$

なる式の右辺より計算して求められる。factor pattern の係数 b_{j1}, b_{j2} は

$$\begin{cases} r_{j\gamma_1} = b_{j1} + b_{j2}\, r_{\gamma_1\gamma_2} \\ r_{j\gamma_2} = b_{j1}\, r_{\gamma_1\gamma_2} + b_{j2} \end{cases} \quad (j = 1, 2, \cdots, 8)$$

なる聯立方程式を解く事によって得られる。これで完全なる oblique solution が得られたのである。

元来 factor solution の理想型といへば uni-factor pattern であって，oblique solution は，uni-factor pattern に近づくといふときに，初めて用ひて利があるのである。uni-factor pattern に近づかないものは，oblique solution を用ひる事によって，common factors の間に相関が出て来てかへって複雑になるだけである。

以上で factor solution の解法を終る事にする。かうして求めた factor F_s には，F_s を最もよく表はしている幾つかの変数をとってその名称を與へる。次に書物に載ってゐる例を挙げて置かう。

例．15才の女子305人に対する身体計測に出て来る8変数について，次に相関係数の表を掲げる（小数点以下3位迄）

	1	2	3	4	5	6	7	8
1. Height								
2. Arm span	846							
3. Length of forearm	805	881						
4 Length of lower leg	859	826	801					
5 Weight	473	376	380	436				
6. Bitrochanteric diameter	398	326	319	329	762			
7 Chest girth	301	277	237	327	730	583		
8 Chest width	382	415	345	365	629	577	539	

の関係が成立つ筈であるから，これより R の最大固有値 λ_1 を求める事が出來る。 $\bar{\alpha}$ に適当な数を掛けその絶対値の平方が λ_1 になるやうにして，それを $\begin{pmatrix} a_{11} \\ a_{21} \\ \vdots \\ a_{n1} \end{pmatrix}$ とすれば

$$R\begin{pmatrix} a_{11} \\ a_{21} \\ \vdots \\ a_{n1} \end{pmatrix} = \lambda_1 \begin{pmatrix} a_{11} \\ a_{21} \\ \vdots \\ a_{n1} \end{pmatrix}, \quad \lambda_1 = \sum_{j=1}^{n} a_{j1}^2$$

となるから，この $a_{11}, a_{21}, \cdots\cdots a_{n1}$ が求めるものなのである。 尚 $R\alpha, R^2\alpha, R^3\alpha, \cdots\cdots$ を順々に求めて行く代りに，$R^2, R^4, R^8 \cdots\cdots$ を計算して $R^2\alpha, R^4\alpha, R^8\alpha$ を求めて行けば計算は簡単で收斂は一層早いのである。

次に $a_{12}, a_{22}, \cdots\cdots, a_{n2}$ を求めるには前の R の代りに

$$R_1 = (\dot{r}_{jk}) = (r_{jk} - a_{j1}a_{k1})$$

を用ひて

$$\dot{r}_{jk} = \sum_{t=2}^{m} a_{jt}a_{kt}$$

の條件の下に

$$A_2 = a_{12}^2 + a_{22}^2 + \cdots\cdots + a_{n2}^2$$

を最大にする事を考へる訳であるから，今度は R_1 の最大固有値及びそれに対する固有 vector を求めればよい訳である。

そこで R_1^e を求める計算を行ふのであるが，これは前に求めた R^e を使って簡単に計算出來る。即ち $Q_1 = (a_{j1}a_{k1})$ と置けば $R_1 = R - Q_1$ であり

$$RQ_1 = Q_1R = \lambda_1Q_1, \qquad Q_1^2 = \lambda_1Q_1$$

なる関係が成立つから，一般に数学的帰納法法で

$$R^eQ_1 = RQ_1^e = \lambda_1^e Q_1,\quad Q_1^e = \lambda_1^{e-1} Q_1$$

となる。更に

$$R_1^2 = (R-Q_1)^2 = R^2 - 2RQ_1 + Q_1^2 = R^2 - \lambda_1 Q_1$$

或は一般に数学的帰納法で

$$R_1^e = R^e - \lambda_1^{e-1} Q_1$$

なることが証明出来る。

以下の計算も全く同様に行く。

かような計算によって得られる固有値は順次小さくなって行くから，幾つかの common factors の係数を求めて行く中に，現はれる係数 $a_{1t}, a_{2t}, \text{-----}, a_{nt}$ は皆 insignificant になってしまふ。そこで common factors への分解は止めてあとは unique factor に繰入れるといふのが principal factor pattern の解法である。

4. Multiple-factor Pattern

これは初めに求めた或る orthagonal solution により，factors の間の直交変換を行って求めるのである。その場合成可く多くの係数を insignifi-cant にするやうに順次直交変換を進めて行けばよい。5 Oblique Pattern これもやはり最初に orthogonal pattern を求めてその factors の間に適当な座標変換を行って求めるのである その orthogonal pattern としては、変数の類別が一目瞭然たるものがよいが変数が多いときには更に B－係数の方法によって群別を確める。

簡単の為に，最初に求めた orthogonal pattern に於て

B (1, 2, 3, 4) = 235, B (5, 6, 7, 8) = 179 で夫々一つの群をなしてゐる。

この data に対して，種々の factor solution の結果 (common factors の係数を小数点以下 3 位迄) を與えて置かう。

1. Bi-factor pattern

	F_0	F_1	F_2
1	691	614	
2	591	740	
3	581	704	
4	598	652	
5	694		623
6	611		560
7	562		453
8	596		473

F_0 = general physical growth factor

F_1 = lankiness factor

F_2 = stockiness factor

2. Principal factor pattern

	F_1	F_2
1	858	−328
2	849	−414
3	810	−412
4	825	−339

5	7 4 7	5 6 1
6	6 3 7	5 0 7
7	5 6 1	4 8 8
8	6 1 9	3 7 1

F_1 = general growth factor

F_2 = body-type factor

3. Multiple-factor pattern

	M_1	M_2
1	8 8 2	2 6 1
2	9 2 2	1 9 9
3	8 9 2	1. 7 1
4	8 6 1	2 3 6
5	2 5 2	8 9 9
6	1 9 3	7 9 0
7	1 4 4	7 2 7
8	2 5 8	6 8 0

M_1 に対しては Z_5, Z_6, Z_7, Z_8 の係数が

M_2 に対しては Z_1, Z_2, Z_3, Z_4 の係数が

比較的に小さい。

4. Oblique solution

	structure		pattern	
	$r_{j\gamma_1}$	$r_{j\gamma_2}$	γ_1	γ_2
1	9 1 9	4 8 4	8 9 4	0 5 1

2	943	435	956	-027
3	907	399	932	-052
4	893	454	879	029
5	455	932	005	930
6	374	813	-025	825
7	312	740	-060	769
8	412	724	080	685

Z_5, Z_6, Z_7, Z_8 に対する γ_1 の係数，Z_1, Z_2, Z_3, Z_4 に対する γ_2 の係数は何れも非常に小さくて0と見做すことが出来る。

次に factors を求める為に行列記号を導入して置く。

$$Z = \begin{pmatrix} Z_1 \\ \vdots \\ Z_n \end{pmatrix}, \quad F = \begin{pmatrix} F_1 \\ \vdots \\ F_m \\ U_1 \\ \vdots \\ U_n \end{pmatrix}, \quad A = \begin{pmatrix} a_{11} & \cdots & a_{1m} \\ a_{21} & \cdots & a_{2m} \\ \cdots & \cdots & \cdots \\ a_{n1} & \cdots & a_{nm} \end{pmatrix}$$

$$U = \begin{pmatrix} a_1 & 0 & \cdots & 0 \\ 0 & a_2 & \cdots & 0 \\ & & & \\ 0 & 0 & \cdots & a_n \end{pmatrix}$$

このとき pattern は

$$Z = MF \quad \text{但し} \quad M = (A, U)$$

によって表はされ，structure は

$$S = \begin{pmatrix} t_{11} & \cdots & t_{1m} & a_1 & 0 & \cdots & 0 \\ t_{21} & \cdots & t_{2m} & 0 & a_2 & \cdots & 0 \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ t_{n1} & \cdots & t_{nm} & 0 & 0 & \cdots & a_n \end{pmatrix} = (T, U)$$

$$但し \begin{cases} t_{js} = r_{Z_j F_s} \\ a_j = r_{Z_j U_j} \end{cases}$$

となる。

F_s を $Z_1, Z_2 \cdots Z_n$ の一次形式

$$\overline{F}_s = \beta_{s1} Z_1 + \beta_{s2} Z_2 + \cdots + \beta_{sn} Z_n \qquad (S=1,2,\cdots m)$$

によって推定するが，その場合 $\sum_{i=1}^{N} (F_{si} - \overline{F}_{si})$ を最小にするやうな係数 $\beta_{s1}, \beta_{s2}, \cdots \beta_{sn}$ は次の聯立一次方程式から求められる。

$$\begin{cases} \beta_{s1} + r_{12}\beta_{s2} + \cdots + r_{1n}\beta_{sn} = t_{1s} \\ r_{21}\beta_{s1} + \beta_{s2} + \cdots + r_{2n}\beta_{sn} = t_{2s} \\ \cdots\cdots\cdots\cdots \\ r_{n1}\beta_{s1} + r_{n2}\beta_{s2} + \cdots + \beta_{sn} = t_{ns} \end{cases}$$

故に

$$R = \begin{pmatrix} 1 & r_{12} & \cdots & r_{1n} \\ r_{21} & 1 & \cdots & r_{2n} \\ \cdots & \cdots & \cdots & \cdots \\ r_{n1} & r_{n2} & & 1 \end{pmatrix}, \quad t = \begin{pmatrix} t_{1s} \\ t_{2s} \\ \vdots \\ t_{ns} \end{pmatrix}$$

と置くとき

$$\overline{F}_s = t'_s R^{-1} Z$$

と表はされる。

若しも $\overline{F}_s$ の代りに $\overline{U}_j$ を出すのであれば，t'_s の代りに $(0 \cdots 0 a_j 0 \cdots 0)$ を用いればよい。

更に，

$$\overline{f} = \begin{pmatrix} \overline{F}_1 \\ \vdots \\ \overline{F}_m \end{pmatrix}, \qquad \overline{F} = \begin{pmatrix} \overline{F}_1 \\ \vdots \\ \overline{F}_m \\ \overline{U}_1 \\ \vdots \\ \overline{U}_n \end{pmatrix}$$

と置けば，上の式は纏めて

$$\overline{f} = T' R^{-1} Z, \quad \overline{F} = S' R' Z$$

特に factors が互に無相関のときは $S = M$ で

$$\overline{F} = M' R^{-1} Z$$

となるが，一般には

$$\phi = \begin{pmatrix} 1 & r_{F_1F_2} & \cdots & r_{F_1F_m} \\ r_{F_2F_1} & 1 & \cdots & r_{F_2F_n} \\ \cdots & \cdots & \cdots & \cdots \\ r_{F_mF_1} & r_{F_mF_2} & \cdots & 1 \end{pmatrix}, \quad \Phi = \begin{pmatrix} \phi & 0 \\ 0 & E_m \end{pmatrix}$$

（但し E_n は n 次の単位行列）

と置くとき　$S = M\Phi$ となるから

$$\overline{F} = \Phi M' R^{-1} Z$$

若しも common factors のみを問題にするなら

$$\overline{f} = \phi A' R^{-1} Z$$

であり common factor の間で互に無相関のときには

$$\bar{f} = A' R^{-1} Z$$

となる。

扨，$\bar{F}_s$ といふ推定値の正確さは何によって計られるであらうか。それに対して一通りの解答を與えるものは F_s と $\bar{F}_s$ との間の相関係数，即ち F_s と $Z_1, Z_2, \cdots Z_n$ との間の重相関係数である。これを R_s として，次に R_s を求めよう。先づ

$$\sum_i (F_{si} - \bar{F}_{si}) Z_{ji} = 0 \qquad \left(\begin{array}{l} s = 1, 2, \cdots, m \\ j = 1, 2, \cdots, n \end{array} \right)$$

故に

$$\sum_i (F_{si} - \bar{F}_{si}) \bar{F}_{si} = 0$$

即ち

$$\sum_i F_{si} \bar{F}_{si} = \sum_i \bar{F}_{si}^2 \qquad \sigma_{\bar{F}_s} r_{F_s \bar{F}_s} = \sigma_{\bar{F}_s}^2$$

$$R_s^2 = r_{F_s \bar{F}_s}^2 = \sigma_{\bar{F}_s}^2 \doteq \frac{1}{N} \sum_i F_{si} \bar{F}_{si}$$

$$= \frac{1}{N} \left(\beta_{s1} \sum_i Z_{1i} F_{si} + \beta_{s2} \sum_i Z_{2i} F_{si} + \cdots + \beta_{sn} \sum_i Z_{ni} F_{si} \right)$$

$$= \beta_{s1} t_{1s} + \beta_{s2} t_{2s} + \cdots + \beta_{sn} t_{ns}$$

以上によって "Factor Analysis" の解説を終る事にする。全般的にいって数学的には非常に粗雑であるが，匆々の間に書いた原稿なので一々検討してみる暇を持たなかつた。それ等の点は読者諸賢の御宥恕を願ふことにしてここで筆を擱く事にする。