

第 5 章

因子分析

5.1 因子分析の概略

まず，因子分析ではどのようなことが出来るのかについて考えてみよう．まず以下のデータを良く観察して頂きたい．

例) 以下の表は新卒生に対して，就職先を選ぶ際に重視することをテーマに行ったアンケートの結果である．以下のアンケートから，結局新卒生は何を基準に会社を判断していると言えるのか分析したい．

	有名である	教育の充実	若手重視	専門技能の獲得	伝統がある	産休制度	将来性
A	2	5	1	5	1	5	2
B	3	4	2	5	3	4	1
C	4	1	1	2	5	3	2
D	4	1	3	2	5	3	4
E	1	2	5	1	2	1	4
F	5	1	1	1	4	2	2
G	4	1	1	2	3	2	2
H	3	3	3	4	4	5	4
I	3	2	4	3	5	3	5
J	3	1	2	2	4	3	3
K	2	2	3	2	3	1	1
L	4	3	2	3	5	3	3
M	2	1	1	2	3	3	1
N	3	1	1	1	4	2	2
O	3	3	2	4	4	5	3

問題 29 以上のアンケートから、各要因の背後に潜む共通因子を推測せよ。

問題 30 各変数の相関係数を求め、相関行列を求めよ。またその結果から何が言えそうか考えて見なさい。

R

#データの読み込み

```
> d <- read.csv("shinsotsu.csv", fileEncoding="cp932")  
> head(d)  
> str(d)
```

#相関のチェック

```
> cor(d[, -1])
```

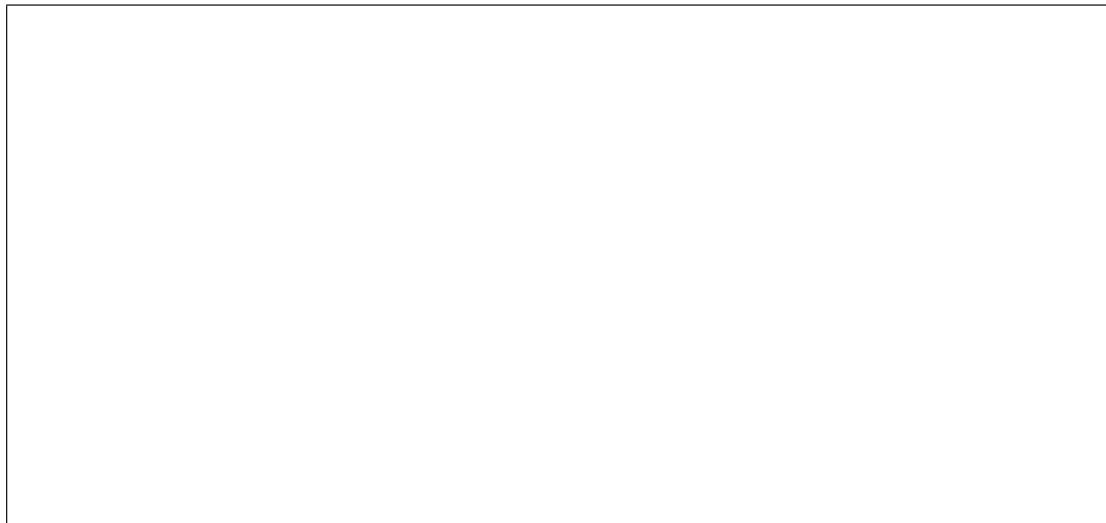
5.2 因子分析とパス図

因子分析 (factor analysis)も重回帰分析、主成分分析などと同じ多変量解析と呼ばれる統計学的分析手法の一つである。

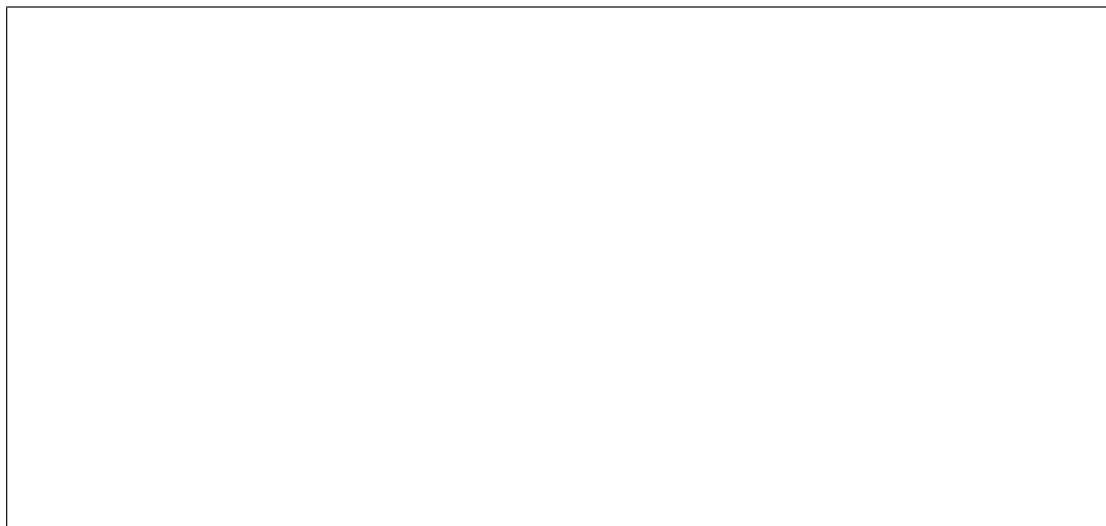
因子分析はいくつかの要因の背後に存在する共通要因 (=共通因子 (common facotor)) を発見するための手法であると言われる。この章では、因子分析を学んでいくわけだが、今一度これらの違いについて説明しておこう。

今までも、何度か説明のために書いてきたが、統計学的手法の違いを理解するためのパス図 (path diagram)というものを使って説明する。

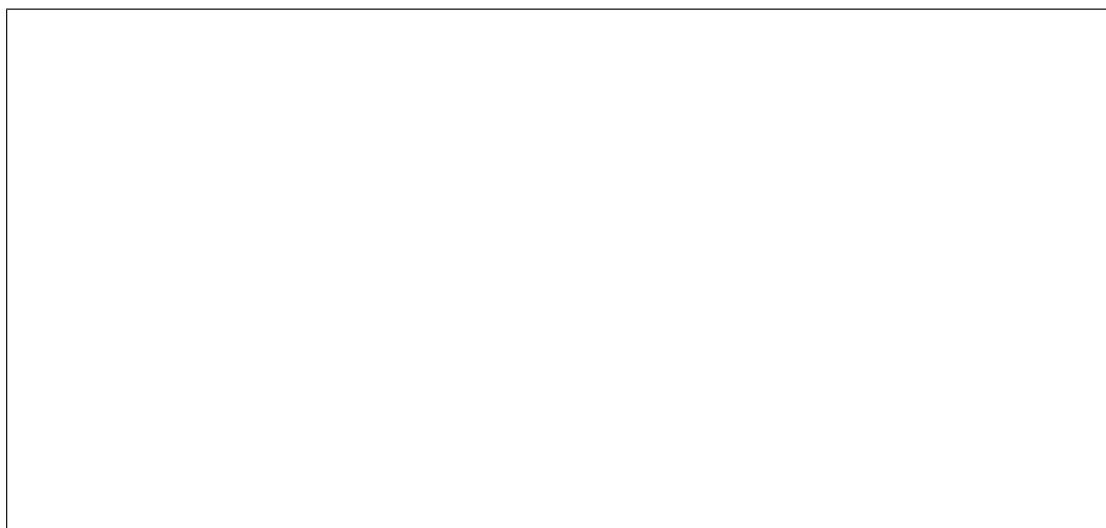
問題 31 重回帰分析のパス図を完成させよ.



問題 32 主成分分析のパス図を完成させよ.



では肝心の因子分析のパス図はどうなっているかというと、以下のようなになる.



どうだろうか。主成分分析と因子分析は非常に良く似ているが、その違いはどこにあるだろうか。

問題 33 パス図上での主成分分析と因子分析の類似点と相違点について話し合いなさい。

5.3 因子分析をする際に知っておくこと

では因子分析を実際に行う前に、因子分析を行うときの注意点を確認しておこう。

1. 因子の順番は意味を持たない

因子分析でも、第1因子、第2因子など因子に番号をつけて呼ぶ事があるが、それは順序を意味しない。

2. 共通因子の個数をあらかじめ想定しておかなければならない

分析をはじめる前にいくつかの因子からデータが説明されるか決めておかなければならない。もし決められない場合は、手当たり次第にやる必要がある。

3. 共通因子の数をいくつにしても良い結果が出る場合は、分析者の判断が優先する

2つの因子を想定しても、3つの因子を想定しても分析が納得いく結果に落ち着いてしまう事がある。そのようなときには数学的に良否を判断する事が出来ない。そういうときには分析者が自身でより納得のいくモデルを選び、なぜそのモデルを選んだか説明しなければならない。

4. 目的変数の個数までしか因子数を想定できない

もし、アンケートの項目が7項目しかないなら、共通因子7個までしか分析にはかけられない。たとえもし現実にはより多くの因子が関係していたとしてもである。なので普通は、共通因子として考えるには値しないような影響の小さい因子を排除して、なるべく少ない因子で多くの事を説明できるようなモデルを構築する事が目的となる。

5. 誤差を独自因子と呼ぶ

パス図の誤差を独自因子 (uniqueness)と呼び、その要因が独自に持っている性質だと解釈する事がある。これは共通因子によって説明できなかった部分と解釈できる。

6. データは標準化してから用いる

変数ごとに標準化して、変数のデータサイズ等の影響を排除してから行う。

7. 共通因子の意味は自分で与えるものである。

因子分析は、分析者が想定していなかった共通因子をあぶり出してくれるような魔法のような手法ではない。共通因子はあくまで、経験や理論、もしくはカンなどから分析者自身が想定するものである。

5.4 因子分析の数学的構造

次に因子分析の数学的構造を確認しておこう。いま、 i 個の要因 x_i から共通因子 m 個、すなわち f_m が抽出されたとする。このとき因子分析のモデル式は以下のようになる。

$$\begin{cases} x_1 = a_{11}f_1 + a_{12}f_2 + \cdots + a_{1m}f_m + \varepsilon_1 \\ x_2 = a_{21}f_1 + a_{22}f_2 + \cdots + a_{2m}f_m + \varepsilon_2 \\ \vdots \\ x_i = a_{i1}f_1 + a_{i2}f_2 + \cdots + a_{im}f_m + \varepsilon_i \end{cases}$$

またこの式は行列を用いて表現する事が出来る。

問題 34 上記の方程式を行列表示しなさい。

因子分析の主な目的は、各共通因子 f の係数 a を決定する事にある。この a のことを 因子負荷量 (factor loading) といい、この値によって、共通因子が書く要因をどれだけ説明できるかを見る事が出来る。この a の値は負になる事もある。

また、先ほど行列表記した場合に出てきた、因子負荷量が並んでいる行列の事を 因子負荷行列 (factor loading matrix) もしくは 因子パターン行列 (factor pattern) という。この因子負荷量を計算する方法はいくつかあり、その中でも特に有名なものとしては 主因子法 (principal factor method) と 最尤法 (maximum likelihood method) が挙げられる。この講義では、主因子法について説明を行うこととする。

では、因子分析の数学的構造について説明しよう。かなり難しい部分もあるが、一度に全て分からなくても良いので、なるべく数式を追って見てほしい。

その前に因子分析の手順について確認しておくことにする。これは後に再度登場する。

一般的に最尤法のが良い結果が得られることが多いと言われており、主因子法よりも最尤法を用いることが多いです。しかしながら、いつでも最尤法を用いれば良いというのは間違いで、最尤法はその前提として、母集団における各変数の確率分布が多変量正規分布に従う事を要求しているため、それが満たされていないデータに関しては、最尤法を使わないほうが良い推定となることが知られています。また、ここであげた以外にも様々な方法があります。特に R の psych パッケージによる fa 関数では、様々な推定法を選ぶことができます

因子分析の手順

【STEP1】回転前の因子負荷量を求める。

主因子法 or 最尤法

【STEP2】回転後の因子負荷量を求める。

バリマックス回転 or プロマックス回転

【STEP3】各共通因子の意味を解釈する。

【STEP4】分析結果の精度を確認する。

【STEP5】因子得点を求め、各個体の特徴を把握する。

因子分析の手順は以上である。具体的な例を用いて数学的構造を確認する事にしよう。

【事例】

以下の表は杉並区にある有名カレー店「キュウリ」のお客様アンケートを集計したものである。今この結果に対して、共通因子を2つと仮定して因子分析をしたい。

	外観	店内の雰囲気	店員の対応	カレーの味	カレーの価格	食器
A	5	5	5	4	4	2
B	5	4	5	2	2	2
C	4	4	4	4	4	4
D	2	3	4	3	3	3
E	3	3	3	3	4	1
F	5	4	5	3	2	3
G	5	5	5	4	5	5
H	3	1	2	5	4	4
I	4	1	3	3	2	3
J	1	2	2	2	2	2
K	3	2	3	1	1	1
L	4	3	4	4	3	4
M	3	2	3	4	5	5
N	4	3	4	5	4	5
O	2	2	3	5	5	4
平均	3.5	2.9	3.7	3.5	3.3	3.2
標準偏差	1.5	1.4	1.4	1.4	1.5	1.5

5.5 回転前の因子負荷量を求める。

まず、【STEP1】の回転前の因子負荷量を求めるというところから解説していこう。

[1] 変数ごとに標準化する。

各変数ごとに標準化する。このとき用いられる標準偏差は不変分散から導く。詳細は省く

が, 平均 0, 標準偏差 1 になる事を確認しておこう.

R

```
#Curry データの読み込み
> d2 <- read.csv("Curry.csv", fileEncoding="cp932")
> head(d2)
> str(d2)

#標準化
> d2.s <- scale(d2[, -1])
```

[2] 標準化後のデータが以下の数式に従っていることを仮定する.

例えば, A さんのデータをパス図で表すと, 以下である.

[3] 各共通因子と独自因子の相関係数と独自因子同士の相関係数は0であると仮定する。すなわちこの場合 27 通りの組み合わせが考えられる。

問題 35 上の文章中の 27 通りを確認せよ。

単相関係数が0である事を仮定するという事は、以下の問題にあるような数式を条件とする事と同値である。

問題 36 一例として f_1 と e_3 の単相関係数が0であると仮定した場合に導かれる条件式を記述せよ。

[4] 各共通因子同士の相関係数は0であると仮定する。

問題 37 f_1 と f_2 の相関係数が0という事を数式として表しなさい。

【補足】 各共通因子間の相関係数が 0 であると仮定する, 今回のようなモデルを直交因子 (orthogonal factor)モデルという. この仮定を置かないモデルの事を斜交因子 (oblique factor)モデルという.

直交因子モデルの仮定は一般に受け入れがたいものだが, 斜交因子モデルはあまりにも計算が煩雑になりすぎるので, 今までは直交因子モデルが用いられてきた. しかしながら, コンピュータの計算性能の進歩により, 斜交因子モデルが一般に浸透しつつある.

[5] 標準化されたデータ間の単相関係数を今までの議論を用いて書き換える.

[6] 同一データの単相関係数を今までの議論を用いて書き換える.

[7] [5], [6] より相関行列を以下のように書き換える.

[8] [7] で導かれた式を変形する. なおこのときの対角線上の値を 共通性 (commonality) と呼ぶ.

ここまでいまいち何を目的としているか見えなかったところがあると思う. ここで因子パターン行列 (右辺) と相関行列っぽいもの (左辺) の関係が論じられる. すなわちもし相関行列 (っぽいもの) をこの式の右辺と同じように分解できたならば, それは因子パターン行列として認められるということである.

[9] データから実際の相関行列を求め、そこから対角成分に個別因子の分散が並んだ行列を引く.

R

```
#相関行列を確認  
> cor(d2.s)
```

[10] [9] において求めた行列の対角成分の値をいくつかに仮定する. ここでは重回帰式の寄与率を用いる.

【補足】 [10] で行ったのは，いくつかある有名な方法の一つである．

[11] [10] で求められた行列の固有値と固有ベクトルを求める．

[12] 最大から 3 番目以降の固有値は 0 だと見なす．すると以下の関係が成り立つ．

[13] 求められた行列の対角成分だけを，元の行列に戻し，その固有値と固有ベクトルを求める．

[14] [11]~[13] を行列の対角成分のどれかが 1 を超えるまで繰り返す.

[15] 最後の繰り返し (1 を超えてしまう前) における固有値の平方根×固有ベクトルが因子負荷行列を作ると解釈する.

R

```
#factanal 関数による因子分析 2 因子 (無回転)
> fa2 <- factanal(d[, -1], factors=2, rotation="none")
> print(fa2$loadings, sort=T) #因子負荷量が大きい順に並べ替え
```

ここで R の `factanal` 関数を用いています. `factanal` 関数は主因子法ではなく, 最尤法を用いた関数なので本当はここで使うのは反則なのですが, この例では, 結果が大きく変わらないので, あまり気にせずにいきましょう.

[16] 因子負荷行列を点グラフ化する.

R

```
#因子負荷量を plot してみる（無回転）
> plot(fa2$loadings[,1],fa2$loadings[,2],
+      xlim=c(0,1), ylim=c(-1,1))
> abline(a=0, b=0, lty=2, col=2) #この2行は x 軸, y 軸を書いているだけ
> abline(v=0, lty=2, col=2)
```

今回は、いきなり2因子モデルに挑戦したので、とても難しいと感じたかもしれない。1因子であればもう少し簡単に説明できるのだが、現実には2因子以上の分析がほとんどなので、最初から2因子モデルを例を解説した。

5.6 回転後の因子負荷行列を求める

前節では、回転前の因子負荷を求めた。それによれば、因子負荷行列は以下のようになっていた。

少しだけ前節でRを使って求めたものと違うかもしれませんが、それは主因子法と最尤法の違いです。

$$\begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \\ a_{31} & a_{32} \\ a_{41} & a_{42} \\ a_{51} & a_{52} \\ a_{61} & a_{62} \end{bmatrix} = \begin{bmatrix} 0.67 & -0.41 \\ 0.74 & -0.49 \\ 0.80 & -0.57 \\ 0.57 & 0.75 \\ 0.55 & 0.66 \\ 0.51 & 0.60 \end{bmatrix}$$

問題 38 これをパス図を書いて表してみよ。

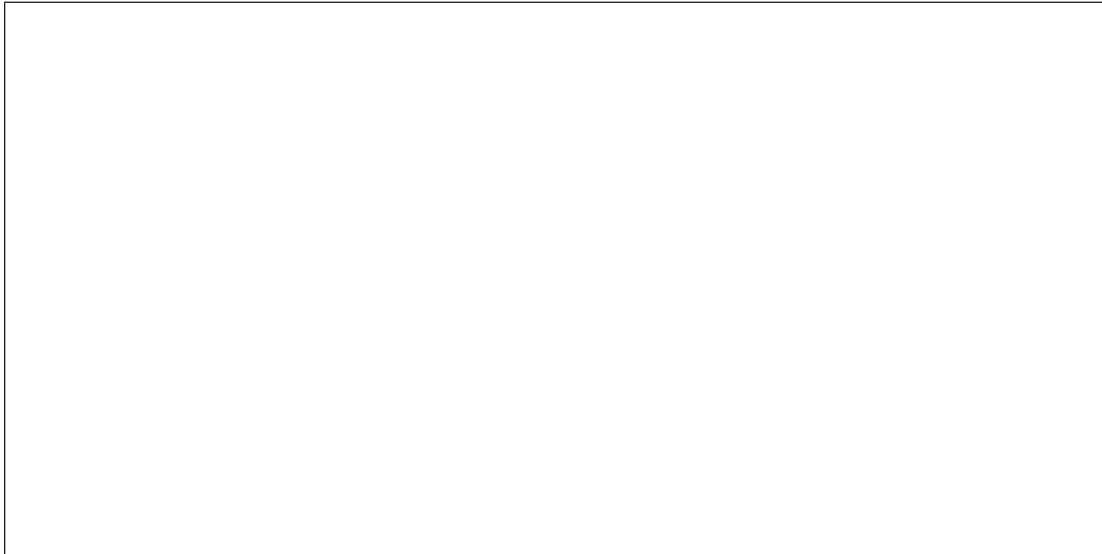
次は【2】回転後の因子負荷量を求めるという手順だが、そもそもなぜ回転などという事をしなければならないのだろうか？ それについて簡単に説明しておこう。

おおざっぱに言うと、回転とは各因子の意味をとらえやすくするための作業なのである。

どういう事が説明するために、因子負荷行列を座標とみて、それぞれのデータごとに x-y 平面上に書き起こしてみる。

問題 39 因子負荷行列の1列目を f_1 成分、2列目を f_2 成分として、各質問（各質問を標準化したデータ）の座標を図示せよ。図示した後に、本来はどの軸ではかったら一番データ

が見やすいのかについて考えなさい。



R

```
#因子負荷量を plot してみる（無回転 再掲）
> plot(fa2$loadings[,1],fa2$loadings[,2],
+      xlim=c(0,1), ylim=c(-1,1))
> abline(a=0, b=0, lty=2, col=2) #この 2 行は x 軸, y 軸を書いているだけ
> abline(v=0, lty=2, col=2)

#factanal 関数による因子分析 (varimax 回転)
> fa2.vari <- factanal(d[, -1], factors=2, rotation="varimax")
> print(fa2.vari$loadings, sort=T)

#因子負荷量を plot してみる（回転後）
> plot(fa2.vari$loadings[,1],fa2.vari$loadings[,2],
+      xlim=c(0,1), ylim=c(-1,1))
> abline(a=0, b=0, lty=2, col=2)
> abline(v=0, lty=2, col=2)

#詳細
> print(fa2, sort=T)
```

上の例はちょうど、第一因子と第二因子が 90° になるような座標になっているが、毎回こうなるとは限らない。だから、x-y 軸を別々に回転させるということが必要になるかもしれない。今回やったような回転を、直交回転 (orthogonal rotation)といい、x-y 軸を別々に動かすような回転を斜行回転 (oblique rotation)という。

直交回転と斜行回転にはいくつかのやり方が存在するが、一般に直交回転は

「バリマックス法 (varimax rotation)」斜行回転は「プロマックス法 (promax rotation)」と呼ばれる回転が使われる。

今では、あまりにバリマックス回転とプロマックス回転が使われすぎているため、その2つ以外の回転法を使うには勇気がいります。バリマックス回転とか、プロマックス回転させましたというとも何も突っ込まれないのですが、それ以外の回転法をとると、なぜその回転を用いたのですかと質問されます（笑）直交回転には「コーティマックス法」「バイコーティマックス法」、斜行回転には「コーティミン法」「バイコーティミン法」「コバリミン法」などがあります。

今回はこのバリマックス法について、解説していく事にしよう。まずは因子負荷行列を以下のように書き換える。

問題 40 この書き換えの意味を考えなさい。

（標準化）バリマックス法

回転角 θ は以下の条件により決定する。

$$\left(\frac{b_{11}^2}{b_{11}^2 + b_{12}^2} - \frac{\frac{b_{11}^2}{b_{11}^2 + b_{12}^2} + \cdots + \frac{b_{61}^2}{b_{61}^2 + b_{62}^2}}{6} \right)^2 + \cdots + \left(\frac{b_{12}^2}{b_{11}^2 + b_{12}^2} - \frac{\frac{b_{12}^2}{b_{11}^2 + b_{12}^2} + \cdots + \frac{b_{62}^2}{b_{61}^2 + b_{62}^2}}{6} \right)^2$$

この数式の値を最大にする角度 θ をバリマックス法における回転角とする。

これを手計算でするのは、いささか無理があるのでコンピュータを用いて計算すると、今回の例では約 -36° 回転させれば、最適になる事がわかる。

問題 41 因子負荷行列をバリマックス法を用いて回転させ、回転後の因子負荷量を求めなさい。ただし、 $\sin(-36^\circ) = -0.5878$ と $\cos(-36^\circ) = 0.8090$ は既知のものとして良い。

ここで回転行列が登場します。

5.7 最尤法*

ここまで因子負荷量の計算方法として主因子法を，因子負荷量の回転法として，バリマックス回転を説明したわけだが，現在多くの解析は最尤法とプロマックス法という方法で行われている．最尤法は数学的に高度な計算が必要とされるので，詳しくは踏み込まない事にするが，概要だけ紹介しておこう．

最尤法（概要）

まず，行列 X を以下のように定義する．

$X =$

$$\begin{bmatrix} r_{11} & r_{12} & \cdots & r_{16} \\ r_{21} & r_{22} & \cdots & r_{26} \\ \vdots & \vdots & \ddots & \vdots \\ r_{61} & r_{62} & \cdots & r_{66} \end{bmatrix} \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \\ \vdots & \vdots \\ a_{61} & a_{62} \end{bmatrix} \begin{bmatrix} a_{11} & a_{21} & \cdots & a_{61} \\ a_{12} & a_{22} & \cdots & a_{62} \end{bmatrix} + \begin{bmatrix} d_1^2 & 0 & \cdots & 0 \\ 0 & d_2^2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & d_6^2 \end{bmatrix}^{-1}$$

次に以下のように L を定義する．

$$L = \text{目的変数の個数} + \log(\det X) - X \text{ の対角成分の和} \quad (\det X = X \text{ の行列式})$$

この L が最大になるような因子負荷行列 $\begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \\ \vdots & \vdots \\ a_{61} & a_{62} \end{bmatrix}$ を最尤法の解とする．

Rにおけるfactanal関数もこちらの最尤法を用いている。気軽に使っても問題ないが、しかし、最尤法を使うときに注意しないといけない前提があるので、それだけでも覚えておこう。それは、「母集団におけるデータが多変量正規分布に従っている」という前提である。この前提がないようなデータに最尤法を用いると、思わぬ間違いを犯すので注意しておこう。

最近では、主因子法より最尤法がなんでも良いという噂？が、まことしやかに流れているが、前提を必要としないという点で、主因子法は最尤法より優れていると言っても良いだろう。

5.8 プロマックス法*

本講義では、直交回転の一種であるバリマックス回転について詳しく説明したのだが、現在では斜行回転のプロマックス回転で、回転後の因子負荷量を計算する事が一般的になってきた。プロマックス法と呼ばれる回転の方法について、これも概要のみだが紹介。

プロマックス法（概要）

- 1) バリマックス回転をする。
- 2) 経験や理論から、「母集団の真の姿はこんなふうになっているに違いない」という、「真の因子負荷行列」を想定する。この「真の因子負荷行列」のことをターゲット行列 (target matrix) という。
- 3) 2で想定したターゲット行列に出来るだけ近づくように1)における軸を回転する。

問題 42 プロマックス法の概要を読んで疑問に思うところはないか。

さて、われわれは母集団の状況がわからないから、統計学なんてものをやっているわけで、2)の「母集団の真の姿はこんな風になっているに違いない」なんていうのは、統計学の存在意義そのものが疑われてしまうような暴挙にも等しいわけである。したがってターゲット行列なんて想定できるわけが無いのだが、ターゲット行列を以下の式で求める事が可能になってしまうのである。

ターゲット行列

今ターゲット行列を $C = (c_{ij})$, バリマックス法における回転後の因子負荷行列を $B = (b_{ij})$ とする. そのときターゲット行列の各成分は下の式によって求められる.

$$c_{ij} = \frac{\sqrt{\sum_{j=1}^n b_{ij}}}{b_{ij}} \times \left| \frac{b_{ij}}{\sqrt{\sum_{j=1}^n b_{ij}}} \right|^{\kappa+1}$$

またこの κ は分析者自身が勝手に決めるものである. (2,3,4, あたりにするのが普通, 実質的には統計ソフトがエラーが出ないように勝手に決めてくれます.)

プロマックス法の妥当性については, 未だに議論の余地が多く残されています. ターゲット行列や κ の値の決め方にかなり不思議なところがあります. バリマックス法を改良したとして, 広く用いられている回転法ですが, 手放しにプロマックス法を賞賛していいというわけでもないとは私は考えます.

5.9 factanal 関数を用いた因子分析の手順

さて, だいぶ因子分析の数学的構造についての説明が長くなってしまったが, ここからコンピュータを用いつつ, 実際の運用を学んでいくことにしよう. まず, 手始めに以前にも紹介した因子分析の手順を再確認しておくことにしよう.

因子分析の手順 (再掲)

【STEP1】 回転前の因子負荷量を求める.

主因子法 or 最尤法

【STEP2】 回転後の因子負荷量を求める.

バリマックス回転 or プロマックス回転

【STEP3】 各共通因子の意味を解釈する.

【STEP4】 分析結果の精度を確認する.

【STEP5】 因子得点を求め, 各個体の特徴を把握する.

以前に登場した事例を用いて、具体的に因子分析を進めていこう。

例) 以下の表は新卒生に対して、就職先を選ぶ際に重視することをテーマに行ったアンケートの結果である。以下のアンケートから、結局新卒生は何を基準に会社を判断していると言えるのか分析したい。(再掲)

	有名である	教育の充実	若手重視	専門技能の獲得	伝統がある	産休制度	将来性
A	2	5	1	5	1	5	2
B	3	4	2	5	3	4	1
C	4	1	1	2	5	3	2
D	4	1	3	2	5	3	4
E	1	2	5	1	2	1	4
F	5	1	1	1	4	2	2
G	4	1	1	2	3	2	2
H	3	3	3	4	4	5	4
I	3	2	4	3	5	3	5
J	3	1	2	2	4	3	3
K	2	2	3	2	3	1	1
L	4	3	2	3	5	3	3
M	2	1	1	2	3	3	1
N	3	1	1	1	4	2	2
O	3	3	2	4	4	5	3

R で因子分析をする際には、factanal 関数 (factanal function)を用いるのが一般的である。この関数は `scores` オプションで、後述の因子得点 (factor score)の計算もすることが出来る。また `rotation` オプションでは回転法の指定も可能である。

R

```
#3 因子モデル プロマックス回転・バートレット法で因子得点
fa3 <- factanal(d[, -1], factors=3,
                rotation="promax", scores="Bartlett")
print(fa3, sort=T)    #因子負荷量を並び替えて実行
```

この検定は通らない方が良いのだが、参考程度で良い。

【STEP1】～【STEP2】

まず、出力の `loadings` という部分であるが、これが因子負荷行列である。そしてこれは `rotation` オプションを指定したことにより、既に回転後のものとなっている。よって、**【STEP2】** 回転後の因子負荷行列を求めるという所まではここを見ればわかる。

この `loadings` の各項目の最大値を見てみよう。するときれいな階段状の構造が見えてくる。この構造をサーストンの階段構造 (Thurstone's staircase structure)といい、良い因子

分析の目安の一つとされている。

問題 43 サーストンの階段構造が見えるということはどういうことか。考えて見なさい。

【STEP3】

次は【STEP3】各共通因子の意味を解釈するという部分だが、ここが分析者が最も頭を働かせる部分かもしれない。例えば今【STEP2】において、第一因子は主に B,D,F を説明する因子であるということであろう。

問題 44 各因子に名前を付けなさい。

やってもらって分かったと思うが、この【STEP3】は分析者の主観が多分に影響する。この部分で、皆が納得できるような因子名つけるには、その対象についての知識が十分でないと難しい。【STEP2】までの因子分析は真の因子を魔法のように見つけ出してくれる方法というわけではなく、あくまで数学的な処理にすぎない。この【STEP3】で初めて因子に意味が与えられるのである。

【STEP4】

そして【STEP4】分析の精度を確認するに移ろう。ここでは3通りの見方を紹介する。

[1] 個別因子 (Uniqueness)を検討する。

まず、R の出力の Uniquenesses という部分を見てみよう。この値は各変数が共通因子によって説明されない部分がどの程度かを表している。因子分析は、主成分分析とは違って共通因子により説明されない部分がある分析である。因子分析の目的はもちろんいくつかの変数の総合値として得られる共通因子によってモデルを説明することにあるので、この個別性の値は低い方が望ましいのである。

[2] サーストンの階段構造の有無をみる。

因子負荷量の説明のときに見たように、サーストンの階段構造がキレイに現れているかどうかを判断の基準とするのも良い。これは階段構造が因子構造が単純で、混じりがない場合

に現れるためである。分析者にとってはありがたく、因子を解釈しやすいので、分析者による因子解釈のブレは少なくなる。

[3] 累積寄与率をみる。

最後に、累積寄与率 (cumulative variance)を見るという方法を紹介する。これは各因子がどの程度モデルへの説明力があるかということだと解釈してよい。この値が高いほど、共通因子に説明力があるという解釈ができる。

問題 45 以下の R コマンドをコンソールに入力し、その結果から累積寄与率と独自因子についての理解を深めよ。

R

#寄与率と独自性の関係

```
> mean(fa3$uniquenesses)+0.843 #0.843 は累積寄与率
```

```
> round(mean(FA3$uniquenesses)+0.843)
```

【STEP5】

最後に因子得点 (factor score)を計算しておこう。因子得点とは、各個体が持つ因子の値であるといってよい。これは因子分析の後の個別分析に属するもので、正式に因子分析の一部という訳ではないかもしれないが、各個体の属性把握には非常に良く使われるものである。

また、この因子得点を計算してから、因子得点を説明変数をして重回帰分析などの他の解析を行うことも考えられる。それゆえ、今回の分析ではこれも因子分析の手順として紹介することとする。

因子得点を計算するには、factanal 関数の scores オプションを有効にすればいい。今回は回帰法 (regression method)で計算させているが、その他にもバートレット法 (Bartlett method)やアンダーソンルービン法 (Anderson-Rubin method)などがある。

因子得点 (個別分析) は以下のコマンドで計算する。

R

```
#因子得点の確認  
> fa3$score
```

問題 46 同じデータについて、2 因子を仮定した場合の解析結果と今回の解析結果を見比べて、検討しなさい。

今回は R の標準の `factanal` 関数で解説をしましたが、より詳しい因子分析をするためには、`psych` パッケージに含まれる関数 `fa` の方がより多くのことが分かります。時間があれば講義でも触れますが、気になる方は調べてみても良いでしょう。

R

```
#2 因子との比較  
fa.2 <- factanal(d[, -1], factors=2,  
                 rotation="promax", scores="Bartlett")  
print(fa.2, sort=T)
```