

第11回

因子分析 (改訂版)

因子分析のモデル

因子分析モデル

y_{ij} , $i = 1, 2, \dots, n$, $j = 1, 2, \dots, p$ を観測される p 変量のデータとする。因子分析モデルは、このデータに以下のモデルを当てはめる。

$$y_{ij} = \mu_j + \sum_{a=1}^r f_{ia}\lambda_{ja} + e_{ij} \quad (122)$$

ただし、 μ_j は変数 j の平均、 f_{ia} は個体 i の第 a 因子のスコア、 λ_{ja} は変数 j の因子 a への因子負荷 (因子パターン)、 e_{ij} は残差である。

因子スコア f_{ia} , $a = 1, 2, \dots, r$ は観測されない r 次元の潜在変数であり、多くの場合、単に因子と呼ばれるが、通常、各因子の平均はゼロ、分散は 1 が仮定されている。また、因子間の相関に関しては、無相関とする場合を直交因子、相関を認める場合を斜交因子と呼ぶ。

3

データ行列の各列

ここで、 y_{ij} を $n \times p$ の行列の形に並べたものを \mathbf{Y} とし、その各列を $\mathbf{y}_{(j)}$, $j = 1, 2, \dots, p$ とする。また、 e_{ij} を $n \times p$ の行列の形に並べたものを \mathbf{E} とし、その各列を $\mathbf{e}_{(j)}$, $j = 1, 2, \dots, p$ とする。また、因子 f_{ia} を $n \times r$ の行列の形に並べたものを \mathbf{F} 、 λ_{ja} , $a = 1, 2, \dots, r$ を $r \times 1$ のベクトルの形に並べたものを $\boldsymbol{\lambda}_j$ とすれば、

$$\mathbf{y}_{(j)} = \mu_j \mathbf{1} + \mathbf{F}\boldsymbol{\lambda}_j + \mathbf{e}_{(j)} \quad (123)$$

と書けるが、この表現から因子分析は観測される各変数を共通の説明変数行列 \mathbf{F} へ回帰する重回帰分析モデルであり、因子負荷ベクトル $\boldsymbol{\lambda}_j$ は回帰係数ベクトル、 μ_j は回帰の切片であることが分かる。ただし、説明変数行列 \mathbf{F} は観測されていない。

4

上記の重回帰モデルをまとめて書けば

$$Y = \mathbf{1}\mu' + F\Lambda' + E \quad (124)$$

を得るが、これは多変量重回帰モデルである。ただし、 Λ は λ_j の転置を各行に持つ $p \times r$ の回帰係数行列であり、因子負荷行列または因子パターン行列と呼ばれるものである。

この場合、因子の平均と分散共分散行列に関する制約は

$$\mathbf{1}'F = \mathbf{0}, \quad \text{and} \quad \frac{1}{n}F'F = I \quad (125)$$

となる。

なお、因子分析モデルが観測される変数 Y の潜在変数からなる説明変数行列 F への回帰モデルであることを強調するために、上記のモデルを

$$Y = \mathbf{1}\mu' + X\beta' + E \quad (126)$$

と書くことがあるが、その場合、説明変数行列は F の代わりに X と記され、回帰係数行列は Λ の代わりに β と記されている。したがって、 β 行列の第 j 行の転置が β_j であり λ_j に相当する。

データ行列の各行

また、Eq.(124) に示される Y 行列の各行に着目すれば、その第 i 行の転置は

$$y_i = \mu + \Lambda f_i + e_i \quad (127)$$

と書くことが出来る。ただし、 $y_i, i = 1, 2, \dots, n$ は Y 行列の第 i 行の転置を示す $p \times 1$ のベクトル、 μ は $\mu_j, j = 1, 2, \dots, p$ を縦に並べた $p \times 1$ のベクトル、 $f_i, i = 1, 2, \dots, n$ は F 行列の第 i 行を転置した $r \times 1$ のベクトル、 $e_i, i = 1, 2, \dots, n$ は E 行列の第 i 行を転置した $p \times 1$ のベクトルである。

この場合、平均や分散共分散行列の制約は

$$\frac{1}{n} \sum_{i=1}^n f_i = \mathbf{0} \quad \text{and} \quad \frac{1}{n} \sum_{i=1}^n f_i' f_i = \mathbf{I} \quad (128)$$

となる。

7

母数モデルとランダムモデル（変数モデル）

ここで、潜在変数である因子を未知の定数と見なすか確率変数と見なすかにより因子分析モデルを母数モデルとランダムモデルに分類することが出来る。通常、母数モデルにおいては Eq.(123) や Eq.(124) そして Eq.(126) の表現を用いることが多く、ランダムモデルに於いては Eq.(127) の表現が用いられる。

8

母数モデルの母数（パラメタ）は、平均ベクトル μ 、因子スコア行列 F 、因子パタン行列 Λ 、（ならびに誤差の分散）であり、ランダムモデルの母数は平均ベクトル μ 、因子パタン行列 Λ 、ならびに誤差の分散である。また、母数モデルに於いては、最小2乗法を用いて母数を推定することが多いが、誤差分散は等しいと言うことを仮定すれば最尤解を用いることも出来る。他方、ランダムモデルにおいては誤差ベクトル e_i と因子スコアベクトル f_i に標準正規分布を仮定した上で y_i の周辺分布を導き、それを利用した最尤解や最小2乗解が用いられる。

9

モデルの不定性：原点の不定性

原点の不定性

$$F^{(*)} = F + \mathbf{1}m' \quad \text{and} \quad \mu^{(*)} = \mu - m \quad (129)$$

を定義する、これより

$$F = F^{(*)} - \mathbf{1}m' \quad (130)$$

を Eq.(124) に代入すると

$$Y = \mathbf{1}\mu' + F\Lambda' + e \quad (131)$$

$$= \mathbf{1}\mu' + (F^{(*)} - \mathbf{1}m')\Lambda' + e \quad (132)$$

$$= \mathbf{1}(\mu' - m') + F^{(*)}\Lambda' + e \quad (133)$$

$$= \mathbf{1}\mu^{(*)}' + F^{(*)}\Lambda' + e \quad (134)$$

となり、星付きの変数に於いても因子分析モデルが成り立つ。

$$\mathbf{1}'F = 0$$

10

モデルの不定性：単位の不定性

単位の不定性

$$\mathbf{F}^{(*)} = \mathbf{F}\mathbf{D} \quad \text{and} \quad \mathbf{\Lambda}^{(*)} = \mathbf{\Lambda}\mathbf{D}^{-1} \quad (135)$$

ただし \mathbf{D} は対角行列、を定義する、これより

$$\mathbf{F} = \mathbf{F}^{(*)}\mathbf{D}^{-1} \quad (136)$$

を Eq.(124) に代入すると

$$\mathbf{Y} = \mathbf{1}\boldsymbol{\mu}' + \mathbf{F}\boldsymbol{\Lambda}' + \mathbf{e} \quad (137)$$

$$= \mathbf{1}\boldsymbol{\mu}' + \mathbf{F}^{(*)}\mathbf{D}^{-1}\boldsymbol{\Lambda}' + \mathbf{e} \quad (138)$$

$$= \mathbf{1}\boldsymbol{\mu}' + \mathbf{F}^{(*)}(\mathbf{\Lambda}\mathbf{D}^{-1})' + \mathbf{e} \quad (139)$$

$$= \mathbf{1}\boldsymbol{\mu}' + \mathbf{F}^{(*)}\boldsymbol{\Lambda}^{(*)}' + \mathbf{e} \quad (140)$$

となり、星付きの変数に於いても因子分析モデルが成り立つ。

$$\frac{1}{n}\mathbf{F}'\mathbf{F} = \mathbf{I}$$

11

モデルの不定性：回転に関する不定性

回転の不定性

$$\mathbf{F}^{(*)} = \mathbf{F}\mathbf{T} \quad \text{and} \quad \mathbf{\Lambda}^{(*)} = \mathbf{\Lambda}\mathbf{T}'^{-1} = \mathbf{\Lambda}\mathbf{T}^{-1'} \quad (141)$$

ただし \mathbf{T} は正則行列、を定義する、これより

$$\mathbf{F} = \mathbf{F}^{(*)}\mathbf{T}^{-1} \quad (142)$$

を Eq.(124) に代入すると

$$\mathbf{Y} = \mathbf{1}\boldsymbol{\mu}' + \mathbf{F}\boldsymbol{\Lambda}' + \mathbf{e} \quad (143)$$

$$= \mathbf{1}\boldsymbol{\mu}' + \mathbf{F}^{(*)}\mathbf{T}^{-1}\boldsymbol{\Lambda}' + \mathbf{e} \quad (144)$$

$$= \mathbf{1}\boldsymbol{\mu}' + \mathbf{F}^{(*)}\boldsymbol{\Lambda}^{(*)}' + \mathbf{e} \quad (145)$$

となり、星付きの変数に於いても因子分析モデルが成り立つ。

12

回転に関する不定性

この不定性を取り除くためには、例えば

$$\Lambda' \Lambda = \Gamma = \text{対角行列} \quad (146)$$

という制約を置けば良い。なぜなら

$$\Lambda^{(*)'} \Lambda^{(*)} = \mathbf{T}^{-1} \Lambda' \Lambda \mathbf{T}^{-1} = \mathbf{T}^{-1} \Gamma \mathbf{T}^{-1} \quad (147)$$

は対角になるとは限らないから。

■ 母数モデル

母数モデル

$$Y = \mathbf{1}\mu' + X\beta' + E$$

$$y_{(j)} = \mu_j \mathbf{1} + X\beta_j + e_{(j)}$$

$$y_i = \mu + \beta x_i + e_i$$

$$\mathcal{E}(e_{(j)}) = \mathbf{0}, \quad \text{var}(e_{(j)}) = \psi_j I$$

$$\mathcal{E}(e_i) = \mathbf{0}, \quad \text{var}(e_i) = \Psi \quad (\text{diagonal})$$

制約条件: 直交因子モデル

$$\frac{1}{n} \mathbf{1}' X = \mathbf{0}'$$

$$\frac{1}{n} X' X = I$$

15

母数モデルの最小2乗解

データとモデルの値の残差の2乗和

$$\begin{aligned} \text{RSS} &= \sum_{i=1}^N \sum_{j=1}^p (y_{ij} - \hat{y}_{ij})^2 \\ &= \text{tr} \left((Y - \hat{Y})'(Y - \hat{Y}) \right) \end{aligned}$$

ただし、

$$\hat{y}_{ij} = \mu_j + \sum_{a=1}^r x_{ia} \beta_{ja} \quad \text{or} \quad \hat{Y} = \mathbf{1}\mu' + X\beta'$$

を最小とするパラメタ (μ, β, X) の値は、

$$\frac{\partial \text{RSS}}{\partial \mu} = 2((Y - X\beta')'\mathbf{1} + N\mu) = 0,$$

$$\frac{\partial \text{RSS}}{\partial \beta} = 2(\beta X' X + \mu \mathbf{1}' X - Y' X) = 0$$

$$\frac{\partial \text{RSS}}{\partial X} = 2(-Y\beta + \mathbf{1}\mu'\beta + X\beta'\beta) = 0$$

の解として与えられる。

16

$$\boldsymbol{\mu} = \frac{1}{n} \mathbf{Y}' \mathbf{1}$$

中心化行列 J を定義する。

$$J = I - \frac{1}{n} \mathbf{1} \mathbf{1}' \quad \mathbf{1}' J = \mathbf{0}', \quad J \mathbf{1} = \mathbf{0}, \quad J J = J$$

$$\begin{aligned} \text{RSS} &= \text{tr} \left((\mathbf{Y} - \hat{\mathbf{Y}})' (\mathbf{Y} - \hat{\mathbf{Y}}) \right) \\ &= \text{tr} \left((\mathbf{Y} - \hat{\mathbf{Y}})' \left(J + \frac{1}{n} \mathbf{1} \mathbf{1}' \right) (\mathbf{Y} - \hat{\mathbf{Y}}) \right) \\ &= \text{tr} \left((\mathbf{Y} - \hat{\mathbf{Y}})' J (\mathbf{Y} - \hat{\mathbf{Y}}) \right) + \frac{1}{n} \text{tr} \left((\mathbf{Y} - \hat{\mathbf{Y}})' \mathbf{1} \mathbf{1}' (\mathbf{Y} - \hat{\mathbf{Y}}) \right) \\ &= \text{tr} \left((\mathbf{JY} - \mathbf{J}\hat{\mathbf{Y}})' (\mathbf{JY} - \mathbf{J}\hat{\mathbf{Y}}) \right) + \frac{1}{n} \text{tr} \left((\mathbf{Y}' \mathbf{1} - \hat{\mathbf{Y}}' \mathbf{1}) (\mathbf{1}' \mathbf{Y} - \mathbf{1}' \hat{\mathbf{Y}}) \right) \end{aligned}$$

第 1 項には $\boldsymbol{\mu}$ は含まれない。

17

$$\text{RSS} = \text{tr} \left((\mathbf{JY} - \mathbf{X}\boldsymbol{\beta}')' (\mathbf{JY} - \mathbf{X}\boldsymbol{\beta}') \right)$$

$$\frac{\partial \text{RSS}}{\partial \boldsymbol{\beta}} = 2(\boldsymbol{\beta} \mathbf{X}' \mathbf{X} - \mathbf{Y}' \mathbf{J} \mathbf{X}) = \mathbf{0}$$

$$\frac{\partial \text{RSS}}{\partial \mathbf{X}} = 2(-\mathbf{JY} \boldsymbol{\beta} + \mathbf{X} \boldsymbol{\beta}' \boldsymbol{\beta}) = \mathbf{0}$$

18

交互最小2乗法による解

$$\begin{aligned}\beta &= Y' J X (X' X)^{-1} \\ X &= J Y \beta (\beta' \beta)^{-1}\end{aligned}$$

回転の不定性を取り除くために $\beta' \beta$ が対角になるように回転する。

19

固有分解を用いた解

$$\frac{1}{n} \mathbf{1}' X = 0' \longrightarrow J X = X \quad \frac{1}{n} X' X = I$$

$$\begin{aligned}\beta &= Y' J X (X' X)^{-1} = J Y' X (nI)^{-1} \\ &= \frac{1}{n} Y' J Y \beta (\beta' \beta)^{-1} \\ &= S \beta \Gamma^{-1} \quad \Gamma = (\beta' \beta)\end{aligned}$$

ただし、 S は

$$S = \frac{1}{n} Y' J Y$$

Y の間の分散共分散行列

20

上式の両辺の右から Γ を掛け、左右を入れ替えると

$$S\beta = \beta\Gamma$$

もしも Γ が対角行列なら（回転の不定性）

Γ が S の固有値からなる対角行列、 β が S の固有ベクトル行列、

$$SP = P\Gamma, \quad P'P = I \quad \text{and} \quad \Gamma = \text{対角}$$

$$S \approx P\Gamma P'$$

$$\beta = P\Gamma^{\frac{1}{2}}$$

$$\beta'\beta = \Gamma \quad \text{を満たす。}$$

21

$$\begin{aligned} X &= JY\beta(\beta'\beta)^{-1} \\ &= JY\beta\Gamma^{-1} \\ &= JYP\Gamma^{-\frac{1}{2}} \end{aligned}$$

$$\begin{aligned} \frac{1}{n} X'X &= \frac{1}{n} \Gamma^{-\frac{1}{2}} P'Y'JYP\Gamma^{-\frac{1}{2}} \\ &= \Gamma^{-\frac{1}{2}} P' SP \Gamma^{-\frac{1}{2}} \\ &= \Gamma^{-\frac{1}{2}} P' P\Gamma \Gamma^{-\frac{1}{2}} \\ &= I \end{aligned}$$

22

R code PCA by ALS

■ 交互最小2乗法による主成分分析（因子分析の母数モデル）

```
for( iter in 1:maxiter ){  
  
  # full conditional LSE of A given F  
  A <- t( solve(t(F)%*%F)%*%t(F)%*%Y )  
  # full conditional LSE of F given A  
  F <- t( solve(t(A)%*%A)%*%t(A)%*%t(Y) )  
  
  # evaluate fit  
  rmse <- sqrt( sum( (Y-F%*%t(A))^2 )/nrow(Y)/ncol(Y) )  
  rmseimp <- rmsep-rmse  
  maxad <- max(abs(F-Fp),abs(A-Ap))  
  
  if( maxad <= eps ) break  
  
  # next iteration  
  rmsep <- rmse  
  Fp <- F; Ap <- A  
  
} # end of main iteration loop
```

23

FとAを分割せずにNRを使う (PCAbyNRGN.R)

```
rssFA <- function( vecFA, ndim=2, Y=0 ){  
  # recover F and A from vecFA and evaluate rss  
  n=nrow(Y); nvar=ncol(Y)  
  F=matrix(vecFA[1:(n*ndim)],n,ndim)  
  A=matrix(vecFA[-(1:(n*ndim))],nvar,ndim)  
  rss=ssq(Y-F%*%t(A))  
  return( rss )  
} # end of rssFA  
  
ndim=2  
F0=QRGS( Y, orth=2 )$Q[,1:ndim]*sqrt(n)  
A0=t( solve(t(F0)%*%F0)%*%t(F0)%*%Y )  
vecFA0=c(c(F0),c(A0))  
cat("¥n¥n**** Startin PCA by NR ****¥n¥n")  
resNR=NR( vecFA0, rssFA, print=2, Y=Y, ndim=ndim )  
vecFANR=resNR$par  
F=matrix(vecFANR[1:(n*ndim)],n,ndim)  
A=matrix(vecFANR[-(1:(n*ndim))],nvar,ndim)  
resnorm=normalize( F,A )  
FNR=resnorm$F; ANR=resnorm$A  
rssNR=ssq(Y-FNR%*%t(ANR))  
Print(rss1,rssNR)  
Print(A1,ANR)          resnlm=nlminb( vecFA0, rssFA  
                        , control=list(trace=1), Y=Y, ndim=ndim )
```

vecFA 以外の定数
ndim, Y
は ... の引数として渡す。

24

FとAを分割せずにGNを使う (PCAbyNRGN.R)

```
ymyhatFA <- function( vecFA, ndim=2, Y=0 ){
  # recover F and A from vecFA and evaluate vec(Y-F%*%t(A))
  n=nrow(Y); nvar=ncol(Y)
  F=matrix(vecFA[1:(n*ndim)],n,ndim)
  A=matrix(vecFA[-(1:(n*ndim))],nvar,ndim)
  resid=c(Y-F%*%t(A))
  return( resid )
} # end of ymyhatFA

cat("¥n¥n**** Startin PCA by GN ****¥n¥n")
resGN=GN( vecFA0, ymyhatFA, print=2, Y=Y, ndim=ndim )
vecFAGN=resGN$par
F=matrix(vecFAGN[1:(n*ndim)],n,ndim)
A=matrix(vecFAGN[-(1:(n*ndim))],nvar,ndim)
resnorm=normalize( F,A )
FGN=resnorm$F; AGN=resnorm$A
rssGN=ssq(Y-FGN%*%t(AGN))
Print(rss1,rssGN)
Print(A1,AGN)

library(nlsr)
yhatmyFA <- function( vecFA, ndim=2, Y=0 ){
  return( -ymyhatFA( vecFA, ndim=ndim, Y=Y ) )
} # end of yhatmyFA
resnlsr=nlsr( vecFA0, yhatmyFA, Y=Y, ndim=ndim
, control=list(femax=500) )
```

25

母数モデルの最尤解は存在しない

$$\mathbf{y}_{(j)} \sim N(\mu_j + \mathbf{X}\boldsymbol{\beta}_j, \psi_j) \quad (67)$$

p 個の $\mathbf{y}_{(j)}$ が独立である場合、データ \mathbf{Y} を得たときのパラメタ $\{\mathbf{X}, \boldsymbol{\beta}, \boldsymbol{\Psi}, \boldsymbol{\mu}\}$ の尤度は

$$f(\mathbf{Y}|\mathbf{X}, \boldsymbol{\beta}, \boldsymbol{\Psi}, \boldsymbol{\mu}) = \left(\frac{1}{\sqrt{(2\pi)^p |\boldsymbol{\Psi}|}} \right)^n \exp \left(-\frac{1}{2} \text{tr} \left(\boldsymbol{\Psi}^{-1} (\mathbf{Y} - \mathbf{1}\boldsymbol{\mu}' - \mathbf{X}\boldsymbol{\beta}')' (\mathbf{Y} - \mathbf{1}\boldsymbol{\mu}' - \mathbf{X}\boldsymbol{\beta}') \right) \right) \quad (68)$$

この尤度関数は一つの ψ_j を 0 に近づけることにより
いくらでも大きくすることが出来る。

誤差分散に構造を仮定した場合の最尤解

$\Psi = \psi C$ ただし C は既知。

$$\begin{aligned} L &\propto \left(\frac{1}{\sqrt{(2\pi)^p |\Psi|}} \right)^n \exp \left(-\frac{1}{2} \text{tr} \left(\Psi^{-1} (JY - X\beta)' (JY - X\beta) \right) \right) \\ &\propto (\psi^p)^{-\frac{n}{2}} \exp \left(-\frac{1}{2\psi} \text{tr} \left(C^{-1} (JY - X\beta)' (JY - X\beta) \right) \right) \end{aligned}$$

指数関数の中の $\text{tr}()$ で表現されている部分を RSS_C とすると

$$\begin{aligned} \text{RSS}_C &= \text{tr} \left(C^{-1} (JY - X\beta)' (JY - X\beta) \right) \\ &= \text{tr} \left(C^{-\frac{1}{2}} (JY - X\beta)' (JY - X\beta) C^{-\frac{1}{2}} \right) \\ &= \text{tr} \left(\left(JYC^{-\frac{1}{2}} - X\beta' C^{-\frac{1}{2}} \right)' \left(JYC^{-\frac{1}{2}} - X\beta' C^{-\frac{1}{2}} \right) \right) \\ &= \text{tr} \left((JY^\dagger - X\beta^\dagger)' (JY^\dagger - X\beta^\dagger) \right) \end{aligned}$$

27

誤差分散に構造を仮定した場合の最尤解

$$Y^\dagger = YC^{-\frac{1}{2}} \quad \text{and} \quad \beta^\dagger = C^{-\frac{1}{2}}\beta$$

$$S^\dagger = \frac{1}{n} Y^{\dagger'} J Y^\dagger = C^{-\frac{1}{2}} S C^{-\frac{1}{2}} \approx P^\dagger \Gamma^\dagger P^{\dagger'}$$

$$\beta^\dagger = C^{-\frac{1}{2}}\beta = P^\dagger \Gamma^{\dagger \frac{1}{2}}$$

$$\beta = C^{\frac{1}{2}} P^\dagger \Gamma^{\dagger \frac{1}{2}}$$

$$X = JY^\dagger P^\dagger (\Gamma^\dagger)^{-\frac{1}{2}}$$

$$\psi = \frac{1}{np} \text{RSS}_C = \frac{1}{p} \left(\text{tr}(S^\dagger) - \text{tr}(\Gamma^\dagger) \right)$$

28

■ ランダムモデル

29

ランダムモデル

固体を表す添字 i は省略。

因子を与えたときの観測値の分布を以下の正規分布とする。

$$y_j | \mathbf{f} \sim N(\mu_j + \boldsymbol{\lambda}'_j \mathbf{f}, \psi_j), \quad j = 1, 2, \dots, p \quad (94)$$

また、これらをまとめて、

$$\mathbf{y} | \mathbf{f} \sim N(\boldsymbol{\mu} + \boldsymbol{\Lambda} \mathbf{f}, \boldsymbol{\Psi}) \quad (95)$$

$\boldsymbol{\Psi}$ は、条件付き独立（局所独立）を仮定して対角行列とする。

$$\mathbf{y} = \boldsymbol{\mu} + \boldsymbol{\Lambda} \mathbf{f} + \mathbf{e}, \quad \text{and} \quad \mathbf{e} \sim N(\mathbf{0}, \boldsymbol{\Psi}) \quad \text{と同値} \quad (96)$$

\mathbf{f} の分布（事前分布、周辺分布）を以下の正規分布とする。

$$\mathbf{f} \sim N(\boldsymbol{\mu}_0, \boldsymbol{\Phi}) \quad (97)$$

30

周辺分布

y の周辺分布

$$f(y|\Lambda, \Psi) = \int_F f(y|f, \Lambda, \Psi) \times h(f|\mu_0, \Phi) df$$

ただし $f()$ も $h()$ も正規分布の密度関数。

$$y \sim N(\mu + \Lambda\mu_0, \Lambda\Phi\Lambda' + \Psi) \quad (98)$$

ふつうは、原点と単位の不定性を除くため、

$$\mu_0 = 0, \quad \Phi = I$$

$$y \sim N(\mu, \Lambda\Lambda' + \Psi) \quad (100)$$

31

周辺分布

(95) 式 と (97) 式 に示す正規分布の仮定 \longrightarrow

平均と分散共分散行列のみの仮定

$$\mathcal{E}(y|f) = \mu + \Lambda f \quad \text{and} \quad \mathcal{D}(y|f) = \Psi$$

$$\mathcal{E}(f) = \mu_0 \quad \text{and} \quad \mathcal{D}(f) = \Phi$$

y の周辺分布の平均と分散共分散行列:

$$\mathcal{E}(y) = \mu + \Lambda\mu_0 \quad (103)$$

$$\mathcal{D}(y) \equiv \Sigma = \Lambda\Phi\Lambda' + \Psi \quad (104)$$

μ を $\mu = \frac{1}{n}Y'\mathbf{1}$ で推定する。

■ ランダムモデルの最小2乗解

33

ランダムモデルの最小2乗解

データに以下の構造を持った分散共分散行列を当てはめる。

$$\Sigma = \Lambda\Lambda' + \Psi$$

最小2乗基準

$$\text{RSS} = \text{tr}((\mathbf{S} - \Sigma)'(\mathbf{S} - \Sigma)) \quad (120)$$

$$= \sum_{i=1}^p \sum_{j=1}^p (s_{ij} - \sigma_{ij})^2 \quad (121)$$

34

- R の最小2乗法の関数： GN , nlfb in nlsvr package
 - ◆ データ Y とモデル Yhat のずれを最小とするように Yhat に含まれるパラメタを推定する。
 - ◆ 関数には、第1引数としてパラメタをまとめたベクトルを、第2引数として Y-Yhat または Yhat-Y を計算する関数を渡す。
- R の最適化の関数： NR, nlminb
 - ◆ 負の対数尤度関数もしくは最小2乗基準を最小とするようなパラメタを推定する。
 - ◆ 関数には、第1引数としてパラメタをまとめたベクトルを、第2引数と最小化されるべき関数を渡す。

35

R code: LS 全てを GNで (FALSbyGN)

```

residfa <- function( param ){
  # residual of random fa model: to be used by GN
  # vecS - vecSigma
  #

  Lambda=matrix(param[1:(nvar*ndim)], nvar)
  psi=param[-(1:(nvar*ndim))]
  Sigma=Lambda%*%t(Lambda)+diag(psi)

  resid=vecS-Sigma[locLH]

  return( resid )
} # end of residfa

# initial by pca
ev=eigen( S )
Lambda0=ev$vectors[, 1:ndim]%*%diag(sqrt(ev$values[1:ndim]))
psi0=diag(S-Lambda0%*%t(Lambda0))

# LS fa by nlsvr with constraints
param0=c(Lambda0, psi0)
resNlfb=nlfb( param0, residfa2, control=list(femax=100) )

# LS fa by GN with reparametrization
param0=c(Lambda0, psi0)
resGN=GN( param0, residfa, maxiter=500 )

```

param 以外の定数
nvar, ndim, vecS, locLH
は residfa を定義した環境の中にある。

数値例は後ほど

36

R code: LS 全てを NR で (FALSbyNR)

```

rssfa <- function( param ){
  # evaluation of LS criterion function
  # param=c( c(Lambda), psi )
  #
  # S, ndim and nvar is from the parent env.
  #
  Lambda=matrix(param[1:(nvar*ndim)], nvar)
  psi=param[-(1:(nvar*ndim))]
  Sigma=Lambda%*%t(Lambda)+diag(psi)

  rss=sum((S-Sigma)^2)          rss=sum((vecS-Sigma[locLH])^2)

  return( rss )
} # end of rssfa

param0=c(Lambda0, psi0)
# LS fa by nlminb with constraints
resnlminb=nlminb( param0, rssfa, control=list(trace=0) )

# LS fa by NR with reparametrization
resNR=NR( param0, rssfa, maxiter=50 )

```

数値例は後ほど

37

交互最小2乗法：Minres etc 一次微分を利用

$$\text{RSS} = \text{tr}((\mathbf{S} - \mathbf{\Sigma})'(\mathbf{S} - \mathbf{\Sigma})) \quad (120)$$

$$= \sum_{i=1}^p \sum_{j=1}^p (s_{ij} - \sigma_{ij})^2 \quad (121)$$

$$\frac{\partial \text{RSS}}{\partial \mathbf{\Lambda}} = 4(-\mathbf{S}\mathbf{\Lambda} + \mathbf{\Psi}\mathbf{\Lambda} + \mathbf{\Lambda}\mathbf{\Lambda}'\mathbf{\Lambda}) \quad (122)$$

$$= -4(\mathbf{S} - \mathbf{\Sigma})\mathbf{\Lambda} \quad (123)$$

$$\frac{\partial \text{RSS}}{\partial \mathbf{\Psi}} = -2\text{diag}(\mathbf{S} - \mathbf{\Lambda}\mathbf{\Lambda}' - \mathbf{\Psi}) \quad (124)$$

38

Minres, Refactoring, 繰り返しのある主因子法

(122) 式を O と置いた方程式から

$$(S - \Psi)\Lambda = \Lambda\Gamma$$

$$S - \Psi \approx P\Gamma P'$$

$$\Lambda'\Lambda = \Gamma \quad (\text{対角})$$

$$\Lambda = P\Gamma^{1/2}$$

(128)

Ψ が既知なら主因子法

(124) 式を O と置くことにより

$$\Psi = \text{diag}(S - \Lambda\Lambda')$$

(129)

ALS

繰り返しのある主因子法 (minres 法もしくは refactoring 法 (反復解法))³⁹

R code: LS Minres etc (FALSMLbyALT.R)

```
for( llll in 1:maxiter ){  
  
  # Update Lambda  
  ev=eigen( S-diag(psi) )  
  P=ev$vectors[, 1:ndim];   Gamma=ev$values[1:ndim]  
  Lambda=P*%*%diag(sqrt(Gamma))  
  
  # Update psi  
  psi=diag(S-Lambda*%*%t(Lambda))  
  
  # check convergence  
  param0=c(c(Lambda), psi)  
  crit=critfa(param0, method=method); critimp=(critp-crit)  
  maxad=max(abs(abs(Lambda)-abs(Lambda)), abs(abs(psi)-abs(psi)))  
  
  if( maxad <= epsd ) break  
  
  # next iteration  
  Lambda=Lambda; psip=psi; critp=crit  
  
} # end of llll loop
```

数値例は後ほど

Λ を消すと RSS は ψ だけの関数となる。

(128) 式の Ψ を与えたときの Λ の条件付き最小2乗解を用いて RSS から Λ を消去することが出来る。(128) 式で与えられる Λ や P や Γ が Ψ に依存することを明記して

$$\Lambda(\Psi) = P(\Psi)\Gamma(\Psi)^{1/2} \quad \text{ただし} \quad (S - \Psi)P(\Psi) = P(\Psi)\Gamma(\Psi)$$

と書き、この表現を用いた Σ を

$$\Sigma(\Psi) = \Lambda(\Psi)\Lambda(\Psi)' + \Psi$$

とすれば (120) は

$$\text{RSS} = \text{tr} \left[(S - \Sigma(\Psi))' (S - \Sigma(\Psi)) \right]$$

と書けるため、これを Ψ の対角要素の関数と見なして Newton-Raphson 法等で最小値を求めることが出来る。

(実際はもう少し簡単にしてから使う方が多い。)

41

Ψ の関数としての Rss

```
rss <- function( psi ){  
  #  $\Psi$  が所与の場合の RSS の計算  
  # 引数 psi は  $\Psi$  の対角要素をベクトルにしたもの。  
  #  $S$  はこの関数が定義された環境に存在するものとする。  
  
   $\Psi$  <- diag(psi)  
  
   $S - \Psi$  の固有値  $\Gamma(\Psi)$  と固有ベクトル  $P(\Psi)$  を求める。  
   $\Lambda(\Psi) = P(\Psi)\Gamma(\Psi)^{\frac{1}{2}}$  を計算する。  
  RSS =  $\| (S - \Lambda(\Psi)\Lambda(\Psi)' - \Psi) \|^2$  を計算する。  
  
  return( RSS )  
} # end of rss
```

42

R code: LS psi only by NR (FALSbyJo.R)

```
critfapsi <- function( psi, uselogpsiasparam=0, method="ML" ){
# evaluation of ML or LS criterion function to be used by nlminb or NR
# param=psi
# 20210707
#
# S, ndim and nvar is from the parent env.
#
# psi or log(psi)
# If the parameter is log(psi), recover original as exp(log(psi)).
if( uselogpsiasparam ) psi=exp(psi)

psi=c(psi)

ev=eigen( S-diag(psi) )
P=ev$vector[, 1:ndim]
Gamma=ev$values[1:ndim]
Lambda=P%*%diag(sqrt(Gamma))
Sigma=Lambda%*%t(Lambda)+diag(psi)

crit=sum((S-Sigma)^2)

return( crit )

} # end of critfapsi

resnlminb=nlminb( psi0, critfapsi, control=list(trace=0) )
resNR=NR( psi0, critfapsi, maxiter=50)
```

数値例は後ほど

43

■ ランダムモデルの最尤解

44

(周辺) 最尤解

$\mathbf{y}_i, i = 1, 2, \dots, n$, が (98) 式 に示す多変量正規分布: $\boldsymbol{\mu}_0 = 0$

$$\begin{aligned} f(\mathbf{Y}|\boldsymbol{\Lambda}, \boldsymbol{\Psi}, \boldsymbol{\mu}) &= \prod_{i=1}^n \frac{1}{\sqrt{(2\pi)^p |\boldsymbol{\Sigma}|}} \exp\left(-\frac{1}{2}(\mathbf{y}_i - \boldsymbol{\mu})' \boldsymbol{\Sigma}^{-1} (\mathbf{y}_i - \boldsymbol{\mu})\right) \\ &= \left(\frac{1}{\sqrt{(2\pi)^p |\boldsymbol{\Sigma}|}}\right)^n \exp\left(-\frac{1}{2} \text{tr}\left(\boldsymbol{\Sigma}^{-1} \sum_{i=1}^n (\mathbf{y}_i - \boldsymbol{\mu})(\mathbf{y}_i - \boldsymbol{\mu})'\right)\right) \\ &= \left(\frac{1}{\sqrt{(2\pi)^p |\boldsymbol{\Sigma}|}}\right)^n \exp\left(-\frac{1}{2} \text{tr}\left(\boldsymbol{\Sigma}^{-1} (\mathbf{Y} - \mathbf{1}\boldsymbol{\mu}')' (\mathbf{Y} - \mathbf{1}\boldsymbol{\mu}')\right)\right) \end{aligned}$$

$$\begin{aligned} \log f(\mathbf{Y}|\boldsymbol{\Lambda}, \boldsymbol{\Psi}, \boldsymbol{\mu}) &= \\ &-\frac{1}{2}n \log(2\pi)^p - \frac{1}{2}n \log |\boldsymbol{\Sigma}| - \frac{1}{2}n \text{tr}\left(\boldsymbol{\Sigma}^{-1} \mathbf{S}\right) - \frac{1}{2}n \text{tr}\left(\boldsymbol{\Sigma}^{-1} (\bar{\mathbf{y}} - \boldsymbol{\mu})(\bar{\mathbf{y}} - \boldsymbol{\mu})'\right) \end{aligned} \quad (147)$$

45

$\boldsymbol{\mu}$ を $\bar{\mathbf{y}}$ で推定 $\boldsymbol{\Sigma}$ を含まない部分を c

$$\log f(\mathbf{Y}|\boldsymbol{\Lambda}, \boldsymbol{\Psi}, \boldsymbol{\mu}) = \ln L = -\frac{n}{2} \left(\log |\boldsymbol{\Sigma}| + \text{tr}\left(\boldsymbol{\Sigma}^{-1} \mathbf{S}\right) \right) + c \quad (148)$$

これを最大とする母数は

$$F = -\frac{2}{n} (\ln L - \ln L_1) = \log |\boldsymbol{\Sigma}| + \text{tr}\left(\boldsymbol{\Sigma}^{-1} \mathbf{S}\right) - \log |\mathbf{S}| - p \quad (151)$$

を最小とする母数に等しい。

NR を使う?

46

Lawley の方法

$$\frac{\partial F}{\partial \Lambda} = 2 \Sigma^{-1} \Lambda \Phi - 2 \Sigma^{-1} S \Sigma^{-1} \Lambda \Phi \quad (152)$$

$$= 2 \Sigma^{-1} (\Sigma - S) \Sigma^{-1} \Lambda \Phi \quad (153)$$

$$\frac{\partial F}{\partial \Psi} = \text{diag} \left(\Sigma^{-1} - \Sigma^{-1} S \Sigma^{-1} \right) \quad (154)$$

$$= \text{diag} \left(\Sigma^{-1} (\Sigma - S) \Sigma^{-1} \right) \quad (155)$$

$$S^* = \Psi^{-\frac{1}{2}} S \Psi^{-\frac{1}{2}} \approx \Omega \Theta \Omega' \quad (176) \quad (177)$$

交互最尤解

$$\Lambda = \Psi^{1/2} \Omega (\Theta - I)^{1/2} \Phi^{-\frac{1}{2}} \quad (178)$$

$$\Psi = \text{diag} (S - \Lambda \Phi \Lambda') \quad (188)$$

47

R code: ML by Lawley (FAMLbyLawley.R)

```
for( llll in 1:maxiter ){
```

```
# Update Lambda
```

```
Psimh=diag(1/sqrt(psi))
```

```
ev=eigen( Psimh**S**Psimh )
```

```
Omega=ev$vectors[,1:ndim]; Theta=ev$values[1:ndim]
```

```
Lambda=sqrt(psi)*P**diag(sqrt(Theta-1), nrow=ndim)
```

```
# Update psi
```

```
psi=diag(S-Lambda**t(Lambda))
```

```
# check convergence
```

```
param0=c(c(Lambda), psi)
```

```
crit=critfa(param0, method=method); critimp=(critp-crit)
```

```
maxad=max(abs(abs(Lambda)-abs(Lambda)), abs(abs(psi)-abs(psi)))
```

```
if( maxad <= epsd ) break
```

```
# next iteration
```

```
Lambda=Lambda; psi=psi; critp=crit
```

```
} # end of llll loop
```

プログラムの構造的には
Minress と似ている。

48

Joereskog の方法

(178) 式を用いて Λ を消去することが出来る。

(178) 式で与えられる Λ や Ω や Θ が Ψ に依存することを明記して

$$\Lambda(\Psi) = \Psi^{\frac{1}{2}} \Omega(\Psi) (\Theta(\Psi) - I)^{1/2} \Phi^{-\frac{1}{2}}$$

と書き、この表現を用いた Σ を

$$\Sigma(\Psi) = \Lambda(\Psi) \Phi \Lambda(\Psi)' + \Psi$$

とすれば (151) は

$$F = \log |\Sigma(\Psi)| + \text{tr} \left(\Sigma(\Psi)^{-1} S \right) + \text{const}$$

と書けるため、これを Ψ の対角要素の関数と見なして Newton-Raphson 法等で最小値を求めることが出来る。

実際にはもう少し簡素化してから利用する。

49

Ψ の関数としての F

```
F <- function( psi ){  
  #  $\Psi$  が所与の場合の F の計算  
  # 引数 psi は  $\Psi$  の対角要素をベクトルにしたもの。  
  # S はこの関数が定義された環境に存在するものとする。  
  
   $\Psi <- \text{diag}(\text{psi})$   
  
   $\Psi^{-\frac{1}{2}} S \Psi^{-\frac{1}{2}}$  の固有値  $\Theta(\Psi)$  と固有ベクトル  $\Omega(\Psi)$  を求める。  
   $\Lambda(\Psi) = \Psi^{\frac{1}{2}} \Omega(\Psi) (\Theta(\Psi) - I)^{\frac{1}{2}}$  を計算する。  
   $\Sigma = \Lambda(\Psi) \Lambda(\Psi)' + \Psi$  を計算する。  
   $F = \log |\Sigma| + \text{tr} (\Sigma^{-1} S) - \log S - \text{nrow}(S)$  を計算する。  
  
  return( F )  
} # end of F
```

50

R code: ML psi のみ (FALSbyJo.R)

```
critfapsi <- function( psi, uselogpsiasparam=0 ){
# evaluation of ML criterion function to be used by nlminb or NR
# S, logdets, ndim and nvar is from the parent env.

# psi or log(psi)
# If the parameter is log(psi), recover original as exp(log(psi)).
if( uselogpsiasparam ) psi=exp(psi)

# ML Lambda given psi
psi=c(psi)

Psimh=diag(1/sqrt(psi))
ev=eigen( Psimh %**% S %**% Psimh )
P=ev$vector[,1:ndim]
D=ev$values[1:ndim]
Lambda=sqrt(psi)*P %**% diag(sqrt(D-1), nrow=ndim) %**% invPsimh
Sigma=Lambda %**% Phi %**% t(Lambda) + diag(psi)
invSigma=solve(Sigma)

crit=tr(S %**% invSigma) + log(det(Sigma)) - logdets - nvar

return( crit )
} # end of critfapsi

# ML fa by nlminb with constraints
resnlminb <- nlminb( psi, critfapsi, control=list(trace=0)
, lower=1e-3, upper=1-1e-3 )

# ML fa by NR with reparametrization
resNR <- NR( log(psi), critfapsi, uselogpsiasparam=1,
maxiter=50 )
```

51

EMアルゴリズムによる解法

F を欠測したデータと見なす。

$$\Lambda = Y' J F^* (F^{*'} J F^* + n V^*)^{-1} \quad (211)$$

$$\Psi = \frac{1}{n} \text{diag} \left((Y - F^* \Lambda')' J (Y - F^* \Lambda') + n \Lambda V^* \Lambda' \right) \quad (212)$$

データとパラメタの値が所与の場合の F の条件付き分布

$$F^* = Y \Psi^{-1} \Lambda V^* \quad (202)$$

$$V^* = \Phi - \Phi \left(\Lambda' \Sigma^{-1} \Lambda \right) \Phi \quad (198)$$

$$= \Phi - \Phi \left(\Phi^{-1} - \Phi^{-1} (\Lambda' \Psi^{-1} \Lambda + \Phi^{-1})^{-1} \Phi^{-1} \right) \Phi \quad (199)$$

$$= (\Lambda' \Psi^{-1} \Lambda + \Phi^{-1})^{-1} \quad (200)$$

誤差分散が負にならない！

52

LS by Minress, ML by Lawley, ML by EM

- これらの3つの方法は基本的に同じ構造をしている。
 - ◆ Psi を所与として Lambda を求める。
 - ◆ Lambda を所与として Psi を求める。
- lazy.fa の efa 関数を使うとこれらの3つの方法での推定が可能。

53

- 誤差分散に構造を仮定した場合の解析解
因子分析と主成分分析

54

誤差分散に構造を仮定した場合の解析解のまとめ

$$\Psi = \psi I$$

$$\frac{1}{n} Y' J Y = S \approx P \Gamma P$$

ランダムモデルの最尤解 = ランダムモデルの最小2乗解

$$\Lambda = P(\Gamma - \psi I)^{\frac{1}{2}}$$
$$\psi = \frac{\text{tr}(S) - \text{tr}(\Gamma)}{p - r}$$

母数モデルの最尤解 = 母数モデルの最小2乗解 = 主成分分析

$$\beta = P \Gamma^{\frac{1}{2}}$$
$$\psi = \frac{\text{tr}(S) - \text{tr}(\Gamma)}{p}$$
$$X = J Y P \Gamma^{-\frac{1}{2}}$$

これでも歴とした因子分析！！
諸説有り。

55

■ 構造方程式モデル、共分散構造分析

56

- 構造方程式モデル、共分散構造分析に関しては
<http://mayekawa.in.coocan.jp/titech/docs/fa12.pdf>
 の p25 以降を参照のこと。

一般形 (RAM)

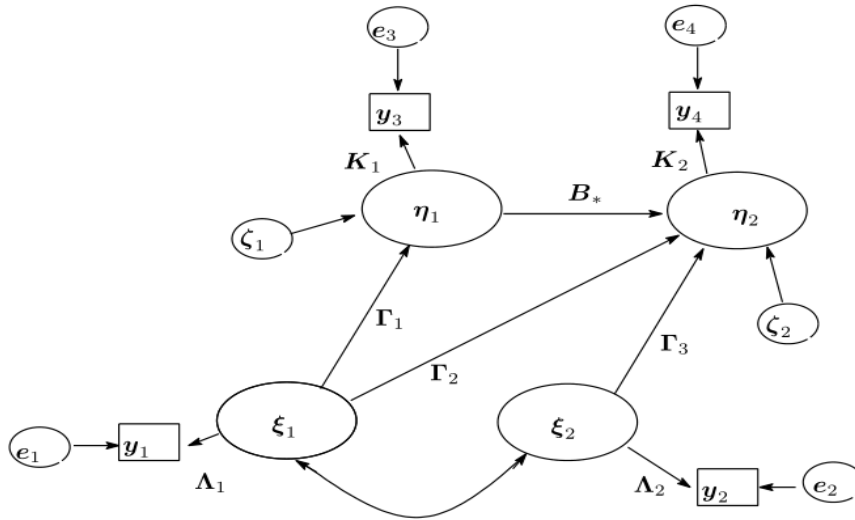


Figure1 Example of a SEM [Example1]

$$\mathcal{D}(\mathbf{y}) = \Psi + (\Lambda + K \cdot (I - B_{**})^{-1} \cdot \Gamma) \cdot \Phi \cdot (\Gamma^T \cdot (I - B_{**})^{-1} \cdot K^T + \Lambda^T) + K \cdot (I - B_{**})^{-1} \cdot \Delta \cdot (I - B_{**})^{-1} \cdot K^T$$