

因子分析と共分散構造分析

前川眞一

20021204,060125,071114,100113,14,15,27,29, 20160107,0510,12,13,14
,0915c,22,23,25,26redcross,27,1011dnc,12,15c,16c,17c,18,21dnc,24,26,27dnc,28,
1101,20170122,23,0202active,15,0424,25dnc,28dnc,0501,06cot,13x, 0612 , 20210630,0701,20220725,26dnc
2022. 7. 27

目次

1	はじめに	2
1.1	因子分析のモデル	2
1.2	モデルの不定性	3
1.2.1	原点の不定性	3
1.2.2	単位の不定性	3
1.2.3	回転の不定性	4
2	記述的因子分析	4
2.1	母数モデル	4
2.1.1	最小2乗解	6
2.1.2	最尤解	10
2.2	ランダムモデル	12
2.2.1	最小2乗解	14
2.2.2	最尤解	17
2.2.3	Joereskog の方法 [Joe]	22
2.2.4	誤差分散に構造を仮定した場合の解法	26
2.3	因子分析の解析解のまとめ	28
3	確認的因子分析	30
3.1	EM アルゴリズムによる方法	30
3.2	因子の平均と相関行列の推定	31
4	共分散構造分析	33
4.1	モデルと共分散構造	33
4.1.1	因子分析 (多因子)	33
4.1.2	因子分析 (1 因子)	33
4.1.3	2 次因子分析 (多因子)	34
4.1.4	多重指標モデル	34
4.1.5	一般形 (RAM)	35
4.1.6	モデルの拡張	39

A	付録 [appendix]	40
1.1	偶に使う公式集	40
1.2	逆行列の公式 [matrixinversionformulae]	41
1.3	確率変数の線形結合 [linearcomb]	44

1 はじめに

1.1 因子分析のモデル

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。

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

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。

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 λ_j とすれば、

$$\mathbf{y}_{(j)} = \mu_j \mathbf{1} + \mathbf{F} \lambda_j + \mathbf{e}_{(j)} \quad [\text{fayj}] \quad (2)$$

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 λ_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。

上記の重回帰モデルをまとめて書けば

$$\mathbf{Y} = \mathbf{1} \boldsymbol{\mu}' + \mathbf{F} \boldsymbol{\Lambda}' + \mathbf{E} \quad [\text{faY}] \quad (3)$$

を得るが、これは多変量重回帰モデルである。ただし、 $\boldsymbol{\Lambda}$ 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

$$\mathbf{1}' \mathbf{F} = \mathbf{0}, \quad \text{and} \quad \frac{1}{n} \mathbf{F}' \mathbf{F} = \mathbf{I} \quad (4)$$

となる。

なお、因子分析モデルが観測される変数 \mathbf{Y} の潜在変数からなる説明変数行列 \mathbf{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

$$\mathbf{Y} = \mathbf{1} \boldsymbol{\mu}' + \mathbf{X} \boldsymbol{\beta}' + \mathbf{E} \quad [\text{faXB}] \quad (5)$$

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

また、Eq.(3) [faY] に示される \mathbf{Y} 行列の各行に着目すれば、その第 i 行の転置は

$$\mathbf{y}_i = \boldsymbol{\mu} + \boldsymbol{\Lambda} \mathbf{f}_i + \mathbf{e}_i \quad [\text{fayi}] \quad (6)$$

と書くことが出来る。ただし、 $\mathbf{y}_i, i = 1, 2, \dots, n$ は \mathbf{Y} 行列の第 i 行の転置を示す $p \times 1$ のベクトル、 $\boldsymbol{\mu}$ は $\mu_j, j = 1, 2, \dots, p$ を縦に並べた $p \times 1$ のベクトル、 $\mathbf{f}_i, i = 1, 2, \dots, n$ は \mathbf{F} 行列の第 i 行を転置した $r \times 1$ のベクトル、 $\mathbf{e}_i, i = 1, 2, \dots, n$ は \mathbf{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

$$\frac{1}{n} \sum_{i=1}^n \mathbf{f}_i = \mathbf{0} \quad \text{and} \quad \frac{1}{n} \sum_{i=1}^n \mathbf{f}_i' \mathbf{f}_i = \mathbf{I} \quad (7)$$

となる。

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.(2)[faYj] や Eq.(3)[faY] そして Eq.(5)[faXB] 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.(6)[fayi] の表現が用いられる。

母数モデルの母数（パラメタ）は、平均ベクトル $\boldsymbol{\mu}$ 、因子スコア行列 \mathbf{F} 、因子パタン行列 $\boldsymbol{\Lambda}$ 、（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 $\boldsymbol{\mu}$ 、因子パタン行列 $\boldsymbol{\Lambda}$ 、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 \mathbf{e}_i と因子スコアベクトル \mathbf{f}_i に標準正規分布を仮定した上で \mathbf{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。

1.2 モデルの不定性

1.2.1 原点の不定性

$$\mathbf{F}^{(*)} = \mathbf{F} + \mathbf{1}\mathbf{m}' \quad \text{and} \quad \boldsymbol{\mu}^{(*)} = \boldsymbol{\mu} - \mathbf{m} \quad (8)$$

を定義する、これより

$$\mathbf{F} = \mathbf{F}^{(*)} - \mathbf{1}\mathbf{m}' \quad (9)$$

を Eq.(3)[faY] に代入すると

$$\mathbf{Y} = \mathbf{1}\boldsymbol{\mu}' + \mathbf{F}\boldsymbol{\Lambda}' + \mathbf{e} \quad (10)$$

$$= \mathbf{1}\boldsymbol{\mu}' + (\mathbf{F}^{(*)} - \mathbf{1}\mathbf{m}')\boldsymbol{\Lambda}' + \mathbf{e} \quad (11)$$

$$= \mathbf{1}(\boldsymbol{\mu}' - \mathbf{m}') + \mathbf{F}^{(*)}\boldsymbol{\Lambda}' + \mathbf{e} \quad (12)$$

$$= \mathbf{1}\boldsymbol{\mu}^{(*)}' + \mathbf{F}^{(*)}\boldsymbol{\Lambda}' + \mathbf{e} \quad (13)$$

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。

1.2.2 単位の不定性

$$\mathbf{F}^{(*)} = \mathbf{F}\mathbf{D} \quad \text{and} \quad \boldsymbol{\Lambda}^{(*)} = \boldsymbol{\Lambda}\mathbf{D}^{-1} \quad (14)$$

ただし D は対角行列、を定義する、これより

$$F = F^{(*)} D^{-1} \quad (15)$$

を Eq.(3)[faY] に代入すると

$$Y = \mathbf{1}\mu' + F\Lambda' + e \quad (16)$$

$$= \mathbf{1}\mu' + F^{(*)} D^{-1} \Lambda' + e \quad (17)$$

$$= \mathbf{1}\mu' + F^{(*)} (\Lambda D^{-1})' + e \quad (18)$$

$$= \mathbf{1}\mu' + F^{(*)} \Lambda^{(*)}' + e \quad (19)$$

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。

1.2.3 回転の不定性

$$F^{(*)} = FT \quad \text{and} \quad \Lambda^{(*)} = \Lambda T'^{-1} = \Lambda T^{-1'} \quad (20)$$

ただし T は正則行列、を定義する、これより

$$F = F^{(*)} T^{-1} \quad (21)$$

を Eq.(3)[faY] に代入すると

$$Y = \mathbf{1}\mu' + F\Lambda' + e \quad (22)$$

$$= \mathbf{1}\mu' + F^{(*)} T^{-1} \Lambda' + e \quad (23)$$

$$= \mathbf{1}\mu' + F^{(*)} \Lambda^{(*)}' + e \quad (24)$$

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。

この不定性を取り除くためには、例えば

$$\Lambda' \Lambda = \Gamma = \text{対角行列} \quad (25)$$

という制約を置けば良い。なぜなら

$$\Lambda^{(*)}' \Lambda^{(*)} = T^{-1} \Lambda' \Lambda T'^{-1} = T^{-1} \Gamma T'^{-1} \quad (26)$$

は対角になるとは限らないから。

2 記述的因子分析

2.1 母数モデル

本稿で取り扱うモデルは p 個の基準変数 $\mathbf{y}_{(j)}, j = 1, 1, \dots, p$ のそれぞれを、共通の $n \times r$ の説明変数行列 X へ回帰するというモデルである。ただし、 X は観測されていない。言い換えれば基準変数を回帰モデル（線形モデル）で説明するような、少数の説明変数群を求める。このような説明変数は潜在変数 latent variables と呼ばれている。

$n \times p$ のデータ行列を Y とし、その第 i 行を転置した $p \times 1$ のベクトルを y_i 、その第 j 列を表す $n \times 1$ のベクトルを $y_{(j)}$ とする。 $n \times p$ の残差行列 E も同様に定義し、 e_i をその第 i 行の転置、 $e_{(j)}$ をその第 j 列とする。また、 $n \times 1$ の全ての要素が 1 であるベクトルを $\mathbf{1}$ 、各基準変数の平均を表す $p \times 1$ のベクトルを μ 、 $\beta'_j, j = 1, 2, \dots, p$ をその行に持つ $p \times r$ の回帰係数行列を β とする。これらを用いると、上記の回帰モデルは

$$Y = \mathbf{1}\mu' + X\beta' + E \quad [\text{mat}] \quad (27)$$

と表現される。これは、その各列を見れば

$$y_{(j)} = \mu_j \mathbf{1} + X\beta_j + e_{(j)} \quad [\text{col}] \quad (28)$$

を、また、各行を見れば

$$y_i = \mu + \beta x_i + e_i \quad [\text{row}] \quad (29)$$

を表している。また、誤差に関しては

$$\mathcal{E}(e_{(j)}) = \mathbf{0}, \quad \text{var}(e_{(j)}) = \psi_j \mathbf{I} \quad (30)$$

$$\mathcal{E}(e_i) = \mathbf{0}, \quad \text{var}(e_i) = \Psi \quad (\text{diagonal}) \quad (31)$$

を仮定する。

なお、 X は潜在変数であるため、因子分析モデルにはパラメタの不定性が存在する。すなわち、あるパラメタのセット (X, β, μ) が Eq.(27) [mat] を満たす場合、以下のように定義される新しいパラメタのセット (X^*, β^*, μ^*) を用いても、因子分析モデルが成り立つことを示すことが出来る。ただし、 T を任意の $r \times r$ の正則な行列、 c を $p \times 1$ のベクトルとして、

$$X^* = (X - \mathbf{1}c')T, \quad \beta^* = \beta T'^{-1}, \quad \mu^* = \beta c + \mu \quad [\text{newparam}] \quad (32)$$

である。なぜなら、Eq.(32) [newparam] より

$$X = X^* T^{-1} + \mathbf{1}c', \quad \beta = \beta^* T', \quad \mu = \mu^* - \beta^* T'c \quad (33)$$

であるが、これを Eq.(27) [mat] に代入すると

$$Y = \mathbf{1}(\mu^* - \beta^* T'c)' + (X^* T^{-1} + \mathbf{1}c')(\beta^* T')' + E \quad (34)$$

$$= \mathbf{1}\mu^{*'} + X^*\beta^{*'} + E \quad [\text{mat2}] \quad (35)$$

となり、これもまた因子分析モデルとなっているからである。

したがって、因子分析モデルのパラメタを一意に定めるためには、上式に含まれる c と T で表現されるモデルの不定性を解消しなければならないが、これは、パラメタに $p + r^2$ の制約条件を課すことによって実現される。まず、 X 行列の列平均が全て 0 であるという制約

$$\frac{1}{n} \mathbf{1}' X = \mathbf{0}' \quad [\text{const}] \quad (36)$$

を置くことにより c に含まれる r 個の不定性を解消する。この制約により μ が Y 行列の列平均を表すパラメタという意味を持つ。次に、

$$\frac{1}{n} X' X = I \quad [\text{diag}] \quad (37)$$

と置くことにより、 \mathbf{T} に含まれる不定性のうち $r(r+1)/2$ 個の部分解消する。なぜなら、上式は対称であるため、実質的に、 r 個の対角要素が全て 1、かつ、 $r(r-1)/2$ 個の非対角要素が 0 という制約を掛けているからである。この制約は、因子スコア行列の各列の分散が 1、そして各列が直交していることを意味している。そして、最後に

$$\beta' \beta = \mathbf{I} \quad \text{ただし } \mathbf{I} \text{ は正の対角要素を持つ対角行列} \quad [\text{rotate}] \quad (38)$$

という制約、すなわち、対称行列 $\beta' \beta$ の非対角要素が全て 0 であるという制約を用いて、最後の $r(r-1)/2$ 個の制約とする。

なお、モデルの制約としては Eq.(36) [const] と Eq.(37) [diag] のみを考え、残った $r(r-1)/2$ 個の不定性を因子の直交回転という形で解消する方法も存在する。より一般的には、 Φ をその対角要素が全て 1 であるような既知の r 次元の対称行列として、Eq.(38) [rotate] の代わりに

$$\frac{1}{n} \mathbf{X}' \mathbf{X} = \Phi \quad [\text{Phi}] \quad (39)$$

という $r(r+1)/2$ 個の制約を課したものを斜交回転と呼ぶ。

2.1.1 最小2乗解

データとモデルの値の残差の2乗和

$$\text{RSS} = \sum_{i=1}^N \sum_{j=1}^p (y_{ij} - \hat{y}_{ij})^2 \quad (40)$$

$$= \text{tr} \left((\mathbf{Y} - \hat{\mathbf{Y}})' (\mathbf{Y} - \hat{\mathbf{Y}}) \right) \quad [\text{rss}] \quad (41)$$

ただし、

$$\hat{y}_{ij} = \mu_j + \sum_{a=1}^r x_{ia} \beta_{ja} \quad \text{or} \quad \hat{\mathbf{Y}} = \mathbf{1} \boldsymbol{\mu}' + \mathbf{X} \boldsymbol{\beta}' \quad (42)$$

を最小とするパラメタ $(\boldsymbol{\mu}, \boldsymbol{\beta}, \mathbf{X})$ の値は、

$$\frac{\partial \text{RSS}}{\partial \boldsymbol{\mu}} = 2((\mathbf{Y} - \mathbf{X} \boldsymbol{\beta}')' \mathbf{1} + N \boldsymbol{\mu}) = 0, \quad [\text{DRssDmu}] \quad (43)$$

$$\frac{\partial \text{RSS}}{\partial \boldsymbol{\beta}} = 2(\boldsymbol{\beta} \mathbf{X}' \mathbf{X} + \boldsymbol{\mu} \mathbf{1}' \mathbf{X} - \mathbf{Y}' \mathbf{X}) = 0 \quad [\text{DRssDBeta}] \quad (44)$$

$$\frac{\partial \text{RSS}}{\partial \mathbf{X}} = 2(-\mathbf{Y} \boldsymbol{\beta} + \mathbf{1} \boldsymbol{\mu}' \boldsymbol{\beta} + \mathbf{X} \boldsymbol{\beta}' \boldsymbol{\beta}) = 0 \quad [\text{DRssDX}] \quad (45)$$

の解として与えられる。 $\boldsymbol{\mu}$ に関する偏微分から、Eq.(36) [const] を考慮すると

$$\boldsymbol{\mu} = \frac{1}{n} \mathbf{Y}' \mathbf{1} \quad [\text{muhat}] \quad (46)$$

を得る。すなわち $\boldsymbol{\mu}$ はデータ行列の各列の平均値となる。この $\boldsymbol{\mu}$ の推定値を Eq.(44) [DRssDBeta] と Eq.(45) [DRssDX] に代入しても良いが、ここでは、以下のようにして RSS を書き直してみる。

まず、以下のように中心化行列 \mathbf{J} を定義する。

$$\mathbf{J} = \mathbf{I} - \frac{1}{n} \mathbf{1} \mathbf{1}' \quad [\mathbf{J}] \quad (47)$$

この J 行列は、対称であり、以下のような性質を持っている。

$$\mathbf{1}'J = \mathbf{0}', \quad J\mathbf{1} = \mathbf{0}, \quad JJ = J \quad [JJ] \quad (48)$$

したがって、制約条件 Eq.(36) [const] から

$$JX = X \quad (49)$$

であることに注意すると、

$$\text{RSS} = \text{tr} \left((\mathbf{Y} - \hat{\mathbf{Y}})'(\mathbf{Y} - \hat{\mathbf{Y}}) \right) \quad (50)$$

$$= \text{tr} \left((\mathbf{Y} - \hat{\mathbf{Y}})'(J + \frac{1}{n}\mathbf{1}\mathbf{1}')(\mathbf{Y} - \hat{\mathbf{Y}}) \right) \quad (51)$$

$$= \text{tr} \left((\mathbf{Y} - \hat{\mathbf{Y}})'J(\mathbf{Y} - \hat{\mathbf{Y}}) \right) + \frac{1}{n}\text{tr} \left((\mathbf{Y} - \hat{\mathbf{Y}})'\mathbf{1}\mathbf{1}'(\mathbf{Y} - \hat{\mathbf{Y}}) \right) \quad (52)$$

$$= \text{tr} \left((JY - J\hat{Y})'(JY - J\hat{Y}) \right) + \frac{1}{n}\text{tr} \left((\mathbf{Y}'\mathbf{1} - \hat{\mathbf{Y}}'\mathbf{1})(\mathbf{1}'\mathbf{Y} - \mathbf{1}'\hat{\mathbf{Y}}) \right) \quad [\text{rssJ}] \quad (53)$$

となるが、

$$J\hat{Y} = J(\mathbf{1}\mu' + X\beta') = JX\beta' = X\beta' \quad (54)$$

$$\mathbf{1}'\hat{Y} = \mathbf{1}'(\mathbf{1}\mu' + X\beta') = n\mu' \quad (55)$$

であり、Eq.(53) [rssJ] の第 1 項には μ は含まれない。したがって、 X や β の値にかかわらず、Eq.(46) [muhat] で求められた μ の値を用いれば Eq.(53) [rssJ] の第 2 項は常に 0 となり、

$$\text{RSS} = \text{tr} \left((JY - X\beta')'(JY - X\beta') \right) \quad [\text{rss2}] \quad (56)$$

を得る。

したがって、RSS の偏微分は

$$\frac{\partial \text{RSS}}{\partial \beta} = 2(\beta X'X - Y'JX) = \mathbf{O} \quad [\text{DRssDBeta1}] \quad (57)$$

$$\frac{\partial \text{RSS}}{\partial X} = 2(-JY\beta + X\beta'\beta) = 0 \quad [\text{DRssDX1}] \quad (58)$$

と簡単に書くことができる。

2.1.1.1 交互最小2乗法 ALS: Alternating Least Squares による解法

Eq.(57) [DRssDBeta1] と Eq.(58) [DRssDX1] をそれぞれ \mathbf{O} と置いたものから条件付きの最小2乗解を求めると

$$\beta = Y'JX(X'X)^{-1} \quad [\text{LSBeta}] \quad (59)$$

$$X = JY\beta(\beta'\beta)^{-1} \quad [\text{LSX}] \quad (60)$$

となるが、ある初期値から初めて、これをを繰り返すと最小2乗解が得られる。

ただ、上記の繰り返しの結果得られる解は、解の制約条件 Eq.(37) [diag]-Eq.(38) [rotate] を満たしているとは限らないので、以下の様に、繰り返しの最後に得られた解を調整する必要がある。なお、この調整は、解を一意に定めるための調整であり、そのために RSS の値が変化することはない。なお、Eq.(36) [const] は、Eq.(60) [LSX] の形から、 Y として中心化されたものを用いれば常に満たされる。

[制約条件について]

Eq.(37) [diag] の形は、正則な行列 T を用いて

$$X_* = XT \quad (61)$$

という形で X を基準化するとして

$$\frac{1}{n} X_*' X_* = I \quad (62)$$

となるように T を定めれば良い。また、この変換の後に、

$$\beta_* = \beta(T^{-1})' \quad (63)$$

で β の値も調節する必要がある。この変換では \hat{Y} の値は

$$\hat{Y} = X_* \beta_*' = X T T^{-1} \beta' = X \beta' \quad (64)$$

となり不変であるため、RSS の値は変化しない。

また、Eq.(38) [rotate] に関しては、直交行列 Q を用いて

$$\beta_{**} = \beta_* Q \quad (65)$$

という形で β_* を基準化するとして

$$\beta_{**}' \beta_{**} = \text{対角} \quad (66)$$

となるように Q を定めれば良い。また、この変換の後に、

$$X_{**} = X_* Q \quad (67)$$

で X_* の値も再度調節する必要があるが、 Q が直交行列であるため、

$$\frac{1}{n} X_{**}' X_{**} = \frac{1}{n} Q' X_*' X_* Q = Q' I Q = I \quad (68)$$

となり、Eq.(37) [diag] を満たす。また、この変換では \hat{Y} の値は

$$\hat{Y} = X_{**} \beta_{**}' = X_* Q Q' \beta_*' = X_* \beta_*' = X \beta' \quad (69)$$

となり不変であるため、RSS の値は変化しない。

なお、 T 行列は、 $\frac{1}{n} X' X$ のスペクトル分解を

$$\frac{1}{n} X' X = R \Delta R', \quad R' R = I \quad \text{and} \quad \Delta = \text{対角} \quad (70)$$

とすれば、

$$T = R \Delta^{-\frac{1}{2}} \quad (71)$$

で与えられる。また、 Q 行列は、 $\beta_*' \beta_*$ のスペクトル分解を

$$\beta_*' \beta_* = Q \Delta^* Q', \quad Q' Q = I \quad \text{and} \quad \Delta^* = \text{対角} \quad (72)$$

とした時に、固有ベクトルからなる正規直交行列 Q となる。

2.1.1.2 固有値分解を利用する解

Eq.(56)[rss2] を Eq.(37)[diag] ならびに Eq.(38)[rotate] の制約の下で最小化することを考える。いま、Eq.(37)[diag] が成り立っているとして、Eq.(59)[LSBeta] を書き直し、Eq.(60)[LSX] を代入し、Eq.(38)[rotate] を使うと

$$\beta = Y' J X (X' X)^{-1} = J Y' X (nI)^{-1} \quad (73)$$

$$= \frac{1}{n} Y' J Y \beta (\beta' \beta)^{-1} \quad (74)$$

$$= S \beta \Gamma^{-1} \quad (75)$$

を得る。ただし、 S は

$$S = \frac{1}{n} Y' J Y \quad [S] \quad (76)$$

で定義される基準変数 Y の間の分散共分散行列である。

ここで、上式の両辺の右から Γ を掛け、左右を入れ替えると

$$S \beta = \beta \Gamma \quad (77)$$

となるが、この式は 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の列を定数倍したものと なっていることがわかる。したがって、 S の r 個の固有値と固有ベクトルを

$$S P = P \Gamma, \quad P' P = I \quad \text{and} \quad \Gamma = \text{対角} \quad [\text{eigen} - S0] \quad (78)$$

とすれば、 β は固有ベクトル行列 P を右から対角行列（それを Δ と置く。）を

$$\beta = P \Delta \quad (79)$$

という形で掛けてスケールリングしたものとなる。

なお、上記の固有値、固有ベクトルを用いて、 S 行列の r 次のスペクトル分解を

$$S \approx P \Gamma P' \quad [\text{eigen} - S] \quad (80)$$

と記すことがある。

そして、Eq.(38)[rotate] より

$$\beta' \beta = \Delta P' P \Delta = \Delta^2 = \Gamma \quad (81)$$

であるため、

$$\Delta = \Gamma^{\frac{1}{2}} \quad (82)$$

すなわち

$$\beta = P \Gamma^{\frac{1}{2}} \quad [\text{LSBeta} - \text{eig}] \quad (83)$$

を得る。

また、このとき、Eq.(60)[LSX] より

$$\mathbf{X} = \mathbf{JY}\boldsymbol{\beta}(\boldsymbol{\beta}'\boldsymbol{\beta})^{-1} \quad (84)$$

$$= \mathbf{JY}\boldsymbol{\beta}\boldsymbol{\Gamma}^{-1} \quad (85)$$

$$= \mathbf{JY}\mathbf{P}\boldsymbol{\Gamma}^{-\frac{1}{2}} \quad [\text{LSX} - \text{eig}] \quad (86)$$

であるため、Eq.(37)[diag] が成り立つことが以下のように示せる。

$$\frac{1}{n} \mathbf{X}'\mathbf{X} = \frac{1}{n} \boldsymbol{\Gamma}^{-\frac{1}{2}} \mathbf{P}'\mathbf{Y}'\mathbf{JY}\mathbf{P}\boldsymbol{\Gamma}^{-\frac{1}{2}} \quad (87)$$

$$= \boldsymbol{\Gamma}^{-\frac{1}{2}} \mathbf{P}' \mathbf{S}\mathbf{P} \boldsymbol{\Gamma}^{-\frac{1}{2}} \quad (88)$$

$$= \boldsymbol{\Gamma}^{-\frac{1}{2}} \mathbf{P}' \mathbf{P}\boldsymbol{\Gamma} \boldsymbol{\Gamma}^{-\frac{1}{2}} \quad (89)$$

$$= \mathbf{I} \quad (90)$$

当然であるが、交互最小自乗法を用いて得られる解を基準化したものと分散共分散行列のスペクトル分解を用いて得られる解は符号を除いて同一である。

[$\boldsymbol{\Gamma}$ の対角化に関して]

いま、 \mathbf{M} と $\boldsymbol{\beta}$ が与えられているとする。このとき、以下のスペクトル分解を

$$\boldsymbol{\Gamma} = \boldsymbol{\beta}'\boldsymbol{\beta} - \mathbf{M} = \mathbf{Q}\boldsymbol{\Delta}\mathbf{Q}' \quad (91)$$

の \mathbf{Q} 行列を用いて $\boldsymbol{\Gamma}$ を以下のように変換すると

$$\mathbf{Q}'\boldsymbol{\Gamma}\mathbf{Q} = \mathbf{Q}'(\boldsymbol{\beta}'\boldsymbol{\beta} - \mathbf{M})\mathbf{Q} = (\boldsymbol{\beta}\mathbf{Q})'(\boldsymbol{\beta}\mathbf{Q}) - \mathbf{Q}'\mathbf{M}\mathbf{Q} = \boldsymbol{\Delta} \quad (92)$$

結果は対角行列となる。これは、 $\boldsymbol{\beta}$ が回転の不定性を利用して直交回転され、ラグランジュの乗数行列は再定義されたことを示している。

2.1.2 最尤解

Eq.(28)[col] を $\mathbf{y}_{(j)}$ の \mathbf{X} への回帰モデルとみれば、以下のような分布の仮定を置くことが出来る。

$$\mathbf{y}_{(j)} \sim N(\mu_j + \mathbf{X}\boldsymbol{\beta}_j, \psi_j) \quad (93)$$

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

$$f(\mathbf{Y}|\mathbf{X}, \boldsymbol{\beta}, \boldsymbol{\Psi}, \boldsymbol{\mu}) = \left(\frac{1}{\sqrt{(2\pi)^p |\boldsymbol{\Psi}|}} \right)^n \exp \left(-\frac{1}{2} \text{tr} \left(\boldsymbol{\Psi}^{-1} (\mathbf{Y} - \mathbf{1}\boldsymbol{\mu}' - \mathbf{X}\boldsymbol{\beta}')' (\mathbf{Y} - \mathbf{1}\boldsymbol{\mu}' - \mathbf{X}\boldsymbol{\beta}') \right) \right) \quad (94)$$

であるが、説明変数行列である \mathbf{X} が未知であるため、少なくとも一つの ψ_j を 0 に近づけることにより、 f の値を限りなく上昇させることが出来るため、最尤解は存在しない。しかし、 \mathbf{C} を所与の行列として、誤差分散が以下のような形

$$\boldsymbol{\Psi} = \boldsymbol{\psi}\mathbf{C} \quad [\text{psiC}] \quad (95)$$

を仮定した場合には最尤解が存在する。

いま、 μ を \bar{y} で推定するとし、Eq.(56) [rss2] を用いて尤度関数を書くと

$$L \propto \left(\frac{1}{\sqrt{(2\pi)^p |\Psi|}} \right)^n \exp \left(-\frac{1}{2} \text{tr} \left(\Psi^{-1} (\mathbf{JY} - \mathbf{X}\beta')' (\mathbf{JY} - \mathbf{X}\beta') \right) \right) \quad (96)$$

$$\propto (\psi^p)^{-\frac{n}{2}} \exp \left(-\frac{1}{2\psi} \text{tr} \left(\mathbf{C}^{-1} (\mathbf{JY} - \mathbf{X}\beta')' (\mathbf{JY} - \mathbf{X}\beta') \right) \right) \quad [\text{fixedC}] \quad (97)$$

となるが、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

$$\text{RSS}_C = \text{tr} \left(\mathbf{C}^{-1} (\mathbf{JY} - \mathbf{X}\beta')' (\mathbf{JY} - \mathbf{X}\beta') \right) \quad (98)$$

$$= \text{tr} \left(\mathbf{C}^{-\frac{1}{2}} (\mathbf{JY} - \mathbf{X}\beta')' (\mathbf{JY} - \mathbf{X}\beta') \mathbf{C}^{-\frac{1}{2}} \right) \quad (99)$$

$$= \text{tr} \left(\left(\mathbf{JY} \mathbf{C}^{-\frac{1}{2}} - \mathbf{X}\beta' \mathbf{C}^{-\frac{1}{2}} \right)' \left(\mathbf{JY} \mathbf{C}^{-\frac{1}{2}} - \mathbf{X}\beta' \mathbf{C}^{-\frac{1}{2}} \right) \right) \quad (100)$$

$$= \text{tr} \left((\mathbf{JY}^\dagger - \mathbf{X}\beta^\dagger)' (\mathbf{JY}^\dagger - \mathbf{X}\beta^\dagger) \right) \quad (101)$$

となる。ただし、

$$\mathbf{Y}^\dagger = \mathbf{Y} \mathbf{C}^{-\frac{1}{2}} \quad \text{and} \quad \beta^\dagger = \mathbf{C}^{-\frac{1}{2}} \beta \quad (102)$$

である。したがって、Eq.(97) [fixedC] を最大とする母数 β^\dagger は、前節の議論から、 $\mathbf{C}^{-\frac{1}{2}}$ を用いてスケールリングされた分散共分散行列の r 次元のスペクトル分解を

$$\mathbf{S}^\dagger = \frac{1}{n} \mathbf{Y}^{\dagger'} \mathbf{JY}^\dagger = \mathbf{C}^{-\frac{1}{2}} \mathbf{S} \mathbf{C}^{-\frac{1}{2}} \approx \mathbf{P}^\dagger \mathbf{\Gamma}^\dagger \mathbf{P}^{\dagger'} \quad [\text{sd2}] \quad (103)$$

とすれば、

$$\beta^\dagger = \mathbf{C}^{-\frac{1}{2}} \beta = \mathbf{P}^\dagger \mathbf{\Gamma}^{\dagger \frac{1}{2}} \quad (104)$$

で与えられるため

$$\beta = \mathbf{C}^{\frac{1}{2}} \mathbf{P}^\dagger \mathbf{\Gamma}^{\dagger \frac{1}{2}} \quad [\text{MLf - Beta}] \quad (105)$$

となる。また、 \mathbf{X} の推定値は

$$\mathbf{X} = \mathbf{JY}^\dagger \mathbf{P}^\dagger \left(\mathbf{\Gamma}^\dagger \right)^{-\frac{1}{2}} \quad [\text{MLf - X}] \quad (106)$$

であり、 ψ の推定値は

$$\psi = \frac{1}{np} \text{RSS}_C = \frac{1}{p} \left(\text{tr} \left(\mathbf{S}^\dagger \right) - \text{tr} \left(\mathbf{\Gamma}^\dagger \right) \right) \quad [\text{MLf - psi}] \quad (107)$$

となる。なお、上式の導出に際しては、Eq.(103) [sd2]、Eq.(105) [MLf-Beta] や Eq.(106) [MLf-X] から導かれる、

$$\mathbf{X}\beta^{\dagger'} = \mathbf{JY}^\dagger \mathbf{P}^\dagger \mathbf{P}^{\dagger'}, \quad \beta^{\dagger'} \beta^\dagger = \mathbf{\Gamma}^\dagger, \quad \mathbf{X}'\mathbf{X} = n\mathbf{I}, \quad \mathbf{P}^{\dagger'} \mathbf{S}^\dagger \mathbf{P}^\dagger = \mathbf{\Gamma}^\dagger \quad (108)$$

の関係を用いた。すなわち

$$\text{RSS}_C = \text{tr}(\mathbf{Y}'\mathbf{J}\mathbf{Y}^\dagger) - 2\text{tr}(\mathbf{Y}'\mathbf{J}\mathbf{X}\boldsymbol{\beta}^\dagger) + \text{tr}(\boldsymbol{\beta}^\dagger\mathbf{X}'\mathbf{X}\boldsymbol{\beta}^\dagger) \quad (109)$$

$$= n\text{tr}(\mathbf{S}^\dagger) - 2\text{tr}(\mathbf{Y}'\mathbf{J}\mathbf{Y}^\dagger\mathbf{P}^\dagger\mathbf{P}^\dagger) + n\text{tr}(\boldsymbol{\Gamma}^\dagger) \quad (110)$$

$$= n(\text{tr}(\mathbf{S}^\dagger) - 2\text{tr}(\mathbf{S}^\dagger\mathbf{P}^\dagger\mathbf{P}^\dagger) + \text{tr}(\boldsymbol{\Gamma}^\dagger)) \quad (111)$$

$$= n(\text{tr}(\mathbf{S}^\dagger) - 2\text{tr}(\mathbf{P}^\dagger\mathbf{S}^\dagger\mathbf{P}^\dagger) + \text{tr}(\boldsymbol{\Gamma}^\dagger)) \quad (112)$$

$$= n(\text{tr}(\mathbf{S}^\dagger) - \text{tr}(\boldsymbol{\Gamma}^\dagger)) \quad (113)$$

である。

なお、

$$\mathbf{R} = \mathbf{C}^{-1}\boldsymbol{\beta}(\boldsymbol{\beta}'\mathbf{C}^{-1}\boldsymbol{\beta})^{-1} \quad (114)$$

として

$$\mathbf{Y}'\mathbf{J}\mathbf{X} = \mathbf{Y}'\mathbf{J}\mathbf{Y}\mathbf{R} = n\mathbf{S}\mathbf{R} \quad (115)$$

$$\mathbf{X}'\mathbf{X} = \mathbf{R}'\mathbf{Y}'\mathbf{J}\mathbf{Y}\mathbf{R} = n\mathbf{R}'\mathbf{S}\mathbf{R} \quad (116)$$

であるから、尤度関数の中に現れる残差積和行列は

$$(\mathbf{J}\mathbf{Y} - \mathbf{X}\boldsymbol{\beta}')'(\mathbf{J}\mathbf{Y} - \mathbf{X}\boldsymbol{\beta}') = n(\mathbf{S} - \mathbf{S}\mathbf{R}\boldsymbol{\beta}' - \boldsymbol{\beta}\mathbf{R}'\mathbf{S} + \boldsymbol{\beta}\mathbf{R}'\mathbf{S}\mathbf{R}\boldsymbol{\beta}') \quad [\text{RssbyP2}] \quad (117)$$

と \mathbf{S} 行列を用いて書くことが出来るため、 $\frac{-2}{n}$ 倍の対数尤度関数は

$$\ln L = p \log(2\pi) + \log |\psi \mathbf{C}| + \text{tr}(\psi \mathbf{C}^{-1}(\mathbf{J}\mathbf{Y} - \mathbf{X}\boldsymbol{\beta}')'(\mathbf{J}\mathbf{Y} - \mathbf{X}\boldsymbol{\beta}')) \quad (118)$$

$$= p \log(2\pi) + \sum_{j=1}^p \log(c_{jj}) + p \log(\psi) + \frac{1}{\psi} \text{tr}(\mathbf{C}^{-1}(\mathbf{S} - 2\mathbf{S}\mathbf{R}\boldsymbol{\beta}' + \boldsymbol{\beta}\mathbf{R}'\mathbf{S}\mathbf{R}\boldsymbol{\beta}')) \quad [\text{RssbyS2}] \quad (119)$$

となる。なお、この形は、 \mathbf{R} の $\boldsymbol{\beta}$ への依存性を無視すれば、 \mathbf{X} を所与の場合の $\boldsymbol{\beta}$ と ψ の対数尤度関数となっている。

2.2 ランダムモデル

上記のモデルでは、因子スコア行列 \mathbf{X} はモデルの母数として取り扱われていた。したがって、個体数を増せば増やすだけ、推定すべき母数の数が増えることになる。そこで、因子を確率変数と考え、モデルの母数から除外すれば、因子パターン $\boldsymbol{\beta}$ に関して効率の良い推定が行える。

なお、因子分析においては、ランダムモデルが普通であり、ノーテーションとしては、データ行列 \mathbf{Y} の任意の一行の転置を \mathbf{y} と表し、その要素を $y_j, j = 1, 2, \dots, p$ と記す。また、因子スコアに \mathbf{f} を、因子パターン(回帰係数、因子負荷)に $\boldsymbol{\Lambda}$ を用い、 \mathbf{f} は \mathbf{X} 行列の任意の一行の転置に相当する。すなわち $\boldsymbol{\Lambda} = \boldsymbol{\beta}, \mathbf{F} = \mathbf{X}$ である。

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。

$$y_j | \mathbf{f} \sim N(\mu_j + \boldsymbol{\lambda}_j' \mathbf{f}, \psi_j), \quad j = 1, 2, \dots, p \quad (120)$$

また、これらをまとめて、

$$\mathbf{y}|\mathbf{f} \sim N(\boldsymbol{\mu} + \boldsymbol{\Lambda}\mathbf{f}, \boldsymbol{\Psi}) \quad [\text{randomy}] \quad (121)$$

と記す。ただし、通常は、 $\boldsymbol{\Psi}$ は、条件付き独立（局所独立）を仮定して ψ_j を対角要素に持つ対角行列とする。これは

$$\mathbf{y} = \boldsymbol{\mu} + \boldsymbol{\Lambda}\mathbf{f} + \mathbf{e}, \quad \text{and} \quad \mathbf{e} \sim N(\mathbf{0}, \boldsymbol{\Psi}) \quad [\text{ybyf}] \quad (122)$$

と同値である。

また、 \mathbf{f} の分布（事前分布、周辺分布）を以下の正規分布とする。

$$\mathbf{f} \sim N(\boldsymbol{\mu}_0, \boldsymbol{\Phi}) \quad [\text{randomf}] \quad (123)$$

すると、付録の公式から、 \mathbf{y} の周辺分布は、

$$\mathbf{y} \sim N(\boldsymbol{\mu} + \boldsymbol{\Lambda}\boldsymbol{\mu}_0, \boldsymbol{\Lambda}\boldsymbol{\Phi}\boldsymbol{\Lambda}' + \boldsymbol{\Psi}) \quad [\text{pdfy}] \quad (124)$$

という正規分布になる。ただし、ふつうは、原点と単位の不定性を除くため、

$$\boldsymbol{\mu}_0 = \mathbf{0}, \quad \boldsymbol{\Phi} = \mathbf{I} \quad [\text{manddrand}] \quad (125)$$

とする場合が多いが、その時は、

$$\mathbf{y} \sim N(\boldsymbol{\mu}, \boldsymbol{\Lambda}\boldsymbol{\Lambda}' + \boldsymbol{\Psi}) \quad (126)$$

である。

なお、Eq.(121)[randomy] と Eq.(123)[randomf] に示す正規分布の仮定を、その平均と分散共分散行列のみの仮定

$$\mathcal{E}(\mathbf{y}|\mathbf{f}) = \boldsymbol{\mu} + \boldsymbol{\Lambda}\mathbf{f} \quad \text{and} \quad \mathcal{D}(\mathbf{y}|\mathbf{f}) = \boldsymbol{\Psi} \quad [\text{condmandd}] \quad (127)$$

$$\mathcal{E}(\mathbf{f}) = \boldsymbol{\mu}_0 \quad \text{and} \quad \mathcal{D}(\mathbf{f}) = \boldsymbol{\Phi} \quad [\text{muandPhi}] \quad (128)$$

に緩めた場合でも、次節に示す公式より、 \mathbf{y} の周辺分布の平均と分散共分散行列行列が

$$\mathcal{E}(\mathbf{y}) = \boldsymbol{\mu} + \boldsymbol{\Lambda}\boldsymbol{\mu}_0 \quad (129)$$

$$\mathcal{D}(\mathbf{y}) \equiv \boldsymbol{\Sigma} = \boldsymbol{\Lambda}\boldsymbol{\Phi}\boldsymbol{\Lambda}' + \boldsymbol{\Psi} \quad [\text{Sigma}] \quad (130)$$

という同じ構造を持つことが分かる。なお、Eq.(130)[Sigma] で示される \mathbf{y} の分散共分散行列 $\boldsymbol{\Sigma}$ の構造を、因子分析モデルの基本モデルと考えることもできる。

このようにして導かれた観測される変数 \mathbf{y} の分散共分散行列の構造から、それを規定するパラメタ $\boldsymbol{\Lambda}, \boldsymbol{\Psi}$ を推定するには、たとえば、最尤法や最小2乗法を持ちいれればよい。

しかし、母数モデルと同じく、ランダムモデルにおいても因子の不定性が存在する。すなわち、Eq.(136)[muandPhi] に示す因子の平均と相関行列（分散共分散行列）を指定するだけでは、因子分析のランダムモデルの母数 $(\boldsymbol{\mu}, \boldsymbol{\Lambda}, \boldsymbol{\Psi})$ は一意には定まらない。なぜなら、いま、因子 \mathbf{f} に関して因子分析モデル Eq.(127)[condmandd]-Eq.(130)[Sigma] が成立しているとする。その時

$$\mathbf{g} = \mathbf{T}'(\mathbf{f} - \mathbf{c}), \quad \boldsymbol{\Lambda}_g = \boldsymbol{\Lambda}\mathbf{T}'^{-1}, \quad \boldsymbol{\mu}_g = \boldsymbol{\Lambda}\mathbf{c} + \boldsymbol{\mu} \quad [\mathbf{g}] \quad (131)$$

で定義される新しい因子 g と平均 μ_g 並びに因子パターン行列 Λ_g を用いれば、

$$y = \mu_g + \Lambda_g g + e \quad [\text{ybyg}] \quad (132)$$

となり、これもまた因子分析モデルとなっている。なぜなら、Eq.(131) [g] より

$$f = T'^{-1}g + c, \quad \Lambda = \Lambda_g T', \quad \mu = \mu_g - \Lambda_g T' c \quad (133)$$

であるが、これを Eq.(122) [ybyf] に代入すると

$$y = \mu_g - \Lambda_g T' c + \Lambda_g T' (T'^{-1}g + c) + e \quad (134)$$

$$= \mu_g + \Lambda_g g + e \quad (135)$$

となるからである。また、

$$\mathcal{E}(g) = T'(\mu_0 - c) \quad \text{and} \quad \mathcal{D}(g) = T' \Phi T \quad [\text{muandPhi}] \quad (136)$$

であるから

$$\mathcal{E}(y) = \mathcal{E}(\mu_g + \Lambda_g g + e) \quad (137)$$

$$= \Lambda c + \mu + \Lambda_g T'(\mu_0 - c) \quad (138)$$

$$= \Lambda c + \mu + \Lambda T'^{-1} T'(\mu_0 - c) \quad (139)$$

$$= \mu + \Lambda \mu_0 \quad (140)$$

$$\mathcal{D}(y) = \mathcal{D}(\mu_g + \Lambda_g g + e) \quad (141)$$

$$= \mathcal{D}(\Lambda_g g) + \mathcal{D}(e) \quad (142)$$

$$= \Lambda_g \mathcal{D}(g) \Lambda_g' + \Psi \quad (143)$$

$$= \Lambda T'^{-1} T' \Phi T T^{-1} \Lambda' + \Psi \quad (144)$$

$$= \Lambda \Phi T \Lambda' + \Psi \quad (145)$$

である。

T の選択

2.2.1 最小2乗解

最小2乗法の場合、 S を観測された y の分散共分散行列とし、最小2乗基準を

$$\text{RSS} = \text{tr}((S - \Sigma)'(S - \Sigma)) \quad [\text{rssS}] \quad (146)$$

$$= \sum_{i=1}^p \sum_{j=1}^p (s_{ij} - \sigma_{ij})^2 \quad (147)$$

と定めれば、そのパラメタに関する偏微分は

$$\frac{\partial \text{RSS}}{\partial \Lambda} = 4(-S\Lambda + \Psi\Lambda + \Lambda\Lambda'\Lambda) \quad [\text{drssSdLambda0}] \quad (148)$$

$$= -4(S - \Sigma)\Lambda \quad [\text{drssSdLambda}] \quad (149)$$

$$\frac{\partial \text{RSS}}{\partial \Psi} = -2\text{diag}(\mathbf{S} - \mathbf{\Lambda}\mathbf{\Lambda}' - \Psi) \quad [\text{drssSdPsi}] \quad (150)$$

となる。

したがって、Eq.(37)[diag] に相当する制約条件

$$\mathbf{\Lambda}'\mathbf{\Lambda} = \mathbf{\Gamma} \quad (\text{対角}) \quad (151)$$

を用いれば、Eq.(148)[drssSdLambda0] を \mathbf{O} と置いた方程式から

$$(\mathbf{S} - \Psi)\mathbf{\Lambda} = \mathbf{\Lambda}\mathbf{\Gamma} \quad [\text{normaleqLS}] \quad (152)$$

を得るが、これは $\mathbf{\Gamma}$ が $\mathbf{S} - \Psi$ 行列の固有値を対角行列に持つ行列であること、また $\mathbf{\Lambda}$ が、それぞれの長さの2乗が固有値の値に基準化された固有ベクトルを列に持つ行列であることを示している。すなわち、 $\mathbf{S} - \Psi$ の r 次元のスペクトル分解（固有値分解）を

$$\mathbf{S} - \Psi \approx \mathbf{P}\mathbf{\Gamma}\mathbf{P}' \quad [\text{eigen}] \quad (153)$$

とすれば、

$$\mathbf{\Lambda} = \mathbf{P}\mathbf{\Gamma}^{1/2} \quad [\text{LambdaLS}] \quad (154)$$

である。

$(\mathbf{S} - \Psi)$ は $p \times p$ の対称行列であるため、それは p 個の固有値と固有ベクトルを持つが、その中からどの固有ベクトルを選べば良いかは、RSS を固有値を用いて書き表せば解る。以下の記述では簡便性を考慮して $\mathbf{A} = \mathbf{S} - \Psi$ と置くが、 $\mathbf{\Lambda}'\mathbf{A}\mathbf{\Lambda} = \mathbf{\Gamma}^2$ であるため、

$$\text{RSS} = \text{tr} \left[(\mathbf{A} - \mathbf{\Lambda}\mathbf{\Lambda}')'(\mathbf{A} - \mathbf{\Lambda}\mathbf{\Lambda}') \right] = \text{tr}(\mathbf{A}'\mathbf{A}) - 2\text{tr}(\mathbf{A}'\mathbf{\Lambda}\mathbf{\Lambda}') + \text{tr}(\mathbf{\Lambda}\mathbf{\Lambda}'\mathbf{\Lambda}\mathbf{\Lambda}') \quad (155)$$

$$= \text{tr}(\mathbf{A}'\mathbf{A}) - 2\text{tr}(\mathbf{\Lambda}'\mathbf{A}'\mathbf{\Lambda}) + \text{tr}(\mathbf{\Lambda}\mathbf{\Gamma}\mathbf{\Lambda}') = \text{tr}(\mathbf{A}'\mathbf{A}) - 2\text{tr}(\mathbf{\Gamma}^2) + \text{tr}(\mathbf{\Gamma}^2) \quad (156)$$

$$= \text{tr}(\mathbf{A}'\mathbf{A}) - 2\text{tr}(\mathbf{\Gamma}^2) \quad [\text{rssbygamma}] \quad (157)$$

となり、 Ψ が所与の場合にこれを最小とするためには、 $\text{tr}(\mathbf{\Gamma}^2)$ を最大とすれば良く、そのためには固有値の大きい方から r 個を $\mathbf{\Gamma}$ として用いれば良く、固有ベクトルもこれらの固有値に対応するものを選べば良いことが解る。

なお、 Ψ の推定値が何らかの形で与えられている場合、それを用いて上記の解を求める方法を主因子解と呼ぶが、Eq.(150)[drssSdPsi] を \mathbf{O} と置くことにより、 $\mathbf{\Lambda}$ が所与の場合の Ψ の最小2乗解は

$$\Psi = \text{diag}(\mathbf{S} - \mathbf{\Lambda}\mathbf{\Lambda}') \quad [\text{PsiLS}] \quad (158)$$

となる。そして、この式を用いて Ψ の推定値をアップデートしながら $(\mathbf{S} - \Psi)$ 行列の固有値固有ベクトルを求めていく方法を繰り返しのある主因子法（minres 法もしくは refactoring 法（反復解法））と呼ぶ。

因子分析のランダムモデルの解を最小2乗法を用いて推定する方法は、上記の方法以外に Joereskog (1977) が示した RSS を Ψ のみの関数として取り扱う方法があるが、その詳細は第 2.2.3 節に示した Joereskog の最尤解の解説を参照されたい。

なお、因子間相関がある場合、最小2乗基準は Eq.(130)[Sigma] と Eq.(146)[rssS] により

$$\text{RSS} = \text{tr} \left((\mathbf{S} - \mathbf{\Lambda}\mathbf{\Phi}\mathbf{\Lambda}' - \Psi)'(\mathbf{S} - \mathbf{\Lambda}\mathbf{\Phi}\mathbf{\Lambda}' - \Psi) \right) \quad [\text{rssS2}] \quad (159)$$

となるが、そのパラメタに関する偏微分は

$$\frac{\partial \text{RSS}}{\partial \Lambda} = 4(-S\Lambda\Phi + \Psi\Lambda\Phi + \Lambda\Phi\Lambda'\Lambda\Phi) \quad (160)$$

$$= 4(-S + \Psi + \Lambda\Phi\Lambda')\Lambda\Phi \quad [\text{drssSdLambda1}] \quad (161)$$

$$= -4(S - \Sigma)\Lambda\Phi \quad (162)$$

となる。したがって、正規方程式は、Eq.(161)[drssSdLambda1] より、

$$(S - \Psi)\Lambda\Phi = \Lambda\Phi\Lambda'\Lambda\Phi \quad (163)$$

となるが、 Φ は正則であるため、上式は、その両辺の右から $\Phi^{-\frac{1}{2}}$ 掛けたもの、

$$(S - \Psi)\Lambda\Phi^{\frac{1}{2}} = \Lambda\Phi^{\frac{1}{2}}\Phi^{\frac{1}{2}}\Lambda'\Lambda\Phi^{\frac{1}{2}} \quad (164)$$

と同値である。そして、

$$\Lambda^* = \Lambda\Phi^{\frac{1}{2}} \quad \text{and} \quad \Gamma^* = \Lambda^{*\prime}\Lambda^* = (\text{対角}) \quad (165)$$

を用いると、最終的に、正規方程式は

$$(S - \Psi)\Lambda^* = \Lambda^*\Lambda^{*\prime}\Lambda^* = \Lambda^*\Gamma^* \quad [\text{normaleqLS2}] \quad (166)$$

と書くことが出来る。したがって、これを Eq.(152)[normaleqLS] と比べることにより、 Φ を指定したモデルの最小2乗解は、Eq.(154)[LambdaLS] で得られる Λ 行列を Λ^* とすれば、それを

$$\Lambda = \Lambda^*\Phi^{-\frac{1}{2}} \quad [\text{LambdaLS2}] \quad (167)$$

とスケーリングしたものとなる。また、 Ψ の最小2乗解は

$$\Psi = \text{diag}(S - \Lambda\Phi\Lambda') \quad [\text{PsiLS2}] \quad (168)$$

で与えられる。

なお、 Λ や Φ の要素に制約が無い場合、上記の Φ を指定した最小2乗解は通常の ($\Phi = I$ を指定した) 直交解をその因子間相関が Φ となるように回転して得られる解と等しくなり、最小2乗基準の値も変わらない。しかし、 Λ や Φ の要素に制約がある場合には、どのような Φ を用いるかにより、その解は異なってくる可能性がある。この問題は、一般的には、次節で述べる確認的因子分析の問題であるが、例えば、 Ψ 行列を既知として Λ 行列のみを推定する場合、どのような Φ を用いるかにより、得られる Λ 行列の値は異なると考えられる。

最後に、様々な解を用いて残差の2乗和のパラメタに関する偏微分、すなわち gradient を計算する必要がある場合、もしも RSS が S 等の対称性を考慮して残差行列の対角要素を含む下半分もしくは上半分の和として定義されている場合には、

$$\frac{\partial \text{RSS}}{\partial \Lambda} = -2 (\text{diag}(S - \Sigma) + S - \Sigma) \Lambda \quad [\text{drssSdLambdaSym}] \quad (169)$$

となる。なぜなら、Eq.(146)[rsslS] は

$$\text{RSS} = \sum_{i=1}^p (s_{ii} - \sigma_{ii})^2 + 2 \sum_{i>j}^p (s_{ij} - \sigma_{ij})^2 \quad (170)$$

と分解されるが、残差行列の下半分の和は

$$\text{RSS1} = \sum_{i=1}^p (s_{ii} - \sigma_{ii})^2 + \sum_{i>j}^p (s_{ij} - \sigma_{ij})^2 = \sum_{i=1}^p (s_{ii} - \sigma_{ii})^2 + \frac{1}{2} (\text{RSS} - \sum_{i=1}^p (s_{ii} - \sigma_{ii})^2) \quad (171)$$

$$= \frac{1}{2} (\sum_{i=1}^p (s_{ii} - \sigma_{ii})^2 + \text{RSS}) \quad (172)$$

であり、

$$\frac{\partial \sum_{i=1}^p (s_{ii} - \sigma_{ii})^2}{\partial \mathbf{\Lambda}} = -4 \text{diag}(\mathbf{S} - \mathbf{\Sigma})\mathbf{\Lambda} \quad (173)$$

となるからである。なお、Eq.(150) [drssSdPsi] はこの場合でも変わらない。

2.2.2 最尤解

最尤法の場合は、 \mathbf{y}_i , $i = 1, 2, \dots, n$, が Eq.(124) [pdfy] に示す多変量正規分布をすることを仮定して、その同時分布が、 $\boldsymbol{\mu}_0 = \mathbf{0}$ を仮定すると

$$f(\mathbf{Y}|\mathbf{\Lambda}, \mathbf{\Psi}, \boldsymbol{\mu}) = \prod_{i=1}^n \frac{1}{\sqrt{(2\pi)^p |\boldsymbol{\Sigma}|}} \exp\left(-\frac{1}{2}(\mathbf{y}_i - \boldsymbol{\mu})' \boldsymbol{\Sigma}^{-1} (\mathbf{y}_i - \boldsymbol{\mu})\right) \quad (174)$$

$$= \left(\frac{1}{\sqrt{(2\pi)^p |\boldsymbol{\Sigma}|}}\right)^n \exp\left(-\frac{1}{2} \text{tr}\left(\boldsymbol{\Sigma}^{-1} \sum_{i=1}^n (\mathbf{y}_i - \boldsymbol{\mu})(\mathbf{y}_i - \boldsymbol{\mu})'\right)\right) \quad (175)$$

$$= \left(\frac{1}{\sqrt{(2\pi)^p |\boldsymbol{\Sigma}|}}\right)^n \exp\left(-\frac{1}{2} \text{tr}\left(\boldsymbol{\Sigma}^{-1} (\mathbf{Y} - \mathbf{1}\boldsymbol{\mu}')' (\mathbf{Y} - \mathbf{1}\boldsymbol{\mu}')\right)\right) \quad (176)$$

である。ただし、 \mathbf{Y} は \mathbf{y}_i の転置を各行に持つ $n \times p$ の行列である。また、Eq.(47) [J] 式の中心化行列を用いると

$$(\mathbf{Y} - \mathbf{1}\boldsymbol{\mu}')' (\mathbf{Y} - \mathbf{1}\boldsymbol{\mu}') = (\mathbf{Y} - \mathbf{1}\boldsymbol{\mu}')' \left(\frac{1}{n} \mathbf{1}\mathbf{1}' + \mathbf{J}\right) (\mathbf{Y} - \mathbf{1}\boldsymbol{\mu}') \quad (177)$$

$$= (\mathbf{Y} - \mathbf{1}\boldsymbol{\mu}')' \left(\frac{1}{n} \mathbf{1}\mathbf{1}'\right) (\mathbf{Y} - \mathbf{1}\boldsymbol{\mu}') + (\mathbf{Y} - \mathbf{1}\boldsymbol{\mu}')' \mathbf{J} (\mathbf{Y} - \mathbf{1}\boldsymbol{\mu}') \quad (178)$$

$$= n(\bar{\mathbf{y}} - \boldsymbol{\mu})(\bar{\mathbf{y}} - \boldsymbol{\mu})' + n\mathbf{S} \quad [\text{sscp}] \quad (179)$$

を得るが、 \mathbf{S} は Eq.(76) [S] で定義されたデータの分散共分散行列である。また、Eq.(48) [JJ] に示した \mathbf{J} の性質と

$$\left(\frac{1}{n} \mathbf{1}\mathbf{1}'\right) \times \left(\frac{1}{n} \mathbf{1}\mathbf{1}'\right) = \left(\frac{1}{n} \mathbf{1}\mathbf{1}'\right) \quad (180)$$

を用いた。

したがって、対数尤度関数は

$$\log f(\mathbf{Y}|\mathbf{\Lambda}, \mathbf{\Psi}, \boldsymbol{\mu}) = -\frac{1}{2}n \log(2\pi)^p - \frac{1}{2}n \log |\boldsymbol{\Sigma}| - \frac{1}{2}n \text{tr}\left(\boldsymbol{\Sigma}^{-1} \mathbf{S}\right) - \frac{1}{2}n \text{tr}\left(\boldsymbol{\Sigma}^{-1} (\bar{\mathbf{y}} - \boldsymbol{\mu})(\bar{\mathbf{y}} - \boldsymbol{\mu})'\right) \quad (181)$$

となるが、 μ を \bar{y} で推定するとすれば、 Σ を含まない部分を c とすれば

$$\ln L = -\frac{n}{2} \left(\log |\Sigma| + \text{tr} \left(\Sigma^{-1} \mathbf{S} \right) \right) + c \quad [\ln L] \quad (182)$$

となる。したがって、最尤解は、Eq.(182) [lnL] を最大とするパラメタ Λ と Ψ の値となる。

ただし、歴史的経緯から、上記の対数尤度ではなく、以下に示す対数尤度比 F を最小化することを行う場合が多い。すなわち、 Σ に如何なる構造も考えない場合の最尤解は

$$\Sigma = \mathbf{S} \quad (183)$$

であるが、その場合の最大化された対数尤度は

$$\ln L_1 = -\frac{n}{2} \left(\log |\mathbf{S}| + \text{tr} \left(\mathbf{S}^{-1} \mathbf{S} \right) \right) + c = -\frac{n}{2} (\log |\mathbf{S}| + p) + c \quad (184)$$

となるため、対数尤度比を $-\frac{2}{n}$ したものは、定数部分を除いて

$$-\frac{2}{n} \log \frac{L}{L_1} \equiv F = -\frac{2}{n} (\ln L - \ln L_1) = \log |\Sigma| + \text{tr} \left(\Sigma^{-1} \mathbf{S} \right) - \log |\mathbf{S}| - p \quad [\text{F0}] \quad (185)$$

$$= \text{tr} \left(\Sigma^{-1} \mathbf{S} \right) - \log |\Sigma^{-1} \mathbf{S}| - p \quad [\text{F}] \quad (186)$$

と書き表される。因子分析のランダムモデルでは、この関数を最小とするような母数を求めることになる。

2.2.2.1 Lawley の方法

Eq.(186) [F] の偏微分、すなわち、 $-\frac{2}{n}$ 倍の対数尤度比 Eq.(185) [F0] の偏微分は

$$\frac{\partial F}{\partial \Lambda} = 2 \Sigma^{-1} \Lambda \Phi - 2 \Sigma^{-1} \mathbf{S} \Sigma^{-1} \Lambda \Phi \quad (187)$$

$$= 2 \Sigma^{-1} (\Sigma - \mathbf{S}) \Sigma^{-1} \Lambda \Phi \quad [\text{dFdLambda}] \quad (188)$$

$$\frac{\partial F}{\partial \Psi} = \text{diag} \left(\Sigma^{-1} - \Sigma^{-1} \mathbf{S} \Sigma^{-1} \right) \quad (189)$$

$$= \text{diag} \left(\Sigma^{-1} (\Sigma - \mathbf{S}) \Sigma^{-1} \right) \quad [\text{dFdPsi}] \quad (190)$$

となるが、 Σ と Φ は正則と仮定されるので、Eq.(188) [dFdLambda] を \mathbf{O} と置いた方程式は、その両辺の左から Σ 、右から Φ^{-1} を掛けることにより

$$\Lambda = \mathbf{S} \Sigma^{-1} \Lambda \quad [\text{dFdLambda0}] \quad (191)$$

となる。ここで、逆行列の公式

$$(\mathbf{A} + \mathbf{BCD})^{-1} = \mathbf{A}^{-1} - \mathbf{A}^{-1} \mathbf{B} (\mathbf{C}^{-1} + \mathbf{DA}^{-1} \mathbf{B})^{-1} \mathbf{DA}^{-1} \quad [\text{mSI9}] \quad (192)$$

を使うと

$$\Sigma^{-1} = \Psi^{-1} - \Psi^{-1} \Lambda (\Phi^{-1} + \Lambda' \Psi^{-1} \Lambda)^{-1} \Lambda' \Psi^{-1} \quad [\text{bif}] \quad (193)$$

を得るが、その両辺の右から Λ を掛けることにより得られる

$$\Sigma^{-1} \Lambda = \Psi^{-1} \Lambda (\Phi^{-1} + \Lambda' \Psi^{-1} \Lambda)^{-1} \Phi^{-1} \quad [\text{bfi-lp}] \quad (194)$$

を Eq.(191) [dFdLambda0] に代入すると

$$\Lambda = S\Psi^{-1}\Lambda(\Phi^{-1} + \Lambda'\Psi^{-1}\Lambda)^{-1}\Phi^{-1} \quad (195)$$

を得るが、以下、表記の煩雑さを避けるための、

$$D = \Lambda'\Psi^{-1}\Lambda \quad [D] \quad (196)$$

と置くと

$$\Lambda = S\Psi^{-1}\Lambda(\Phi^{-1} + D)^{-1}\Phi^{-1} \quad (197)$$

である。(逆行列の公式の導出は節末の中を参照のこと。)

これを以下の様に変形する。まず、右辺右端の Φ をその平方根 $\Phi^{\frac{1}{2}}$ の積で書き、上式の両辺の右側から $\Phi^{\frac{1}{2}}$ を掛けると

$$\Lambda\Phi^{\frac{1}{2}} = S\Psi^{-1}\Lambda(\Phi^{-1} + D)^{-1}\Phi^{-\frac{1}{2}} \quad (198)$$

を得るが、右辺中央に現れる Λ に $\Phi^{\frac{1}{2}}$ を掛けた形を作るためにさらに以下の様に変形する。

$$\Lambda\Phi^{\frac{1}{2}} = S\Psi^{-1}\Lambda(\Phi^{-1} + D)^{-1}\Phi^{-\frac{1}{2}} \quad (199)$$

$$= S\Psi^{-1}\Lambda(\Phi^{\frac{1}{2}}(\Phi^{-1} + D))^{-1} \quad (200)$$

$$= S\Psi^{-1}\Lambda(\Phi^{-\frac{1}{2}} + \Phi^{\frac{1}{2}}D)^{-1} \quad (201)$$

$$= S\Psi^{-1}\Lambda\Phi^{\frac{1}{2}}\Phi^{-\frac{1}{2}}(\Phi^{-\frac{1}{2}} + \Phi^{\frac{1}{2}}D)^{-1} \quad (202)$$

$$= S\Psi^{-1}\Lambda\Phi^{\frac{1}{2}}((\Phi^{-\frac{1}{2}} + \Phi^{\frac{1}{2}}D)\Phi^{\frac{1}{2}})^{-1} \quad (203)$$

$$= S\Psi^{-1}\Lambda\Phi^{\frac{1}{2}}(I + \Phi^{\frac{1}{2}}D\Phi^{\frac{1}{2}})^{-1} \quad (204)$$

$$= S\Psi^{-1}\Lambda\Phi^{\frac{1}{2}}(I + \Delta)^{-1} \quad [LP] \quad (205)$$

ただし、

$$\Delta = \Phi^{\frac{1}{2}}D\Phi^{\frac{1}{2}} = \Phi^{\frac{1}{2}}\Lambda'\Psi^{-1}\Lambda\Phi^{\frac{1}{2}} \quad [Delta] \quad (206)$$

と置いた。また、Eq.(205) [LP] の両辺の左から $\Psi^{-\frac{1}{2}}$ を、右から $I + \Delta$ を掛けると

$$\Psi^{-\frac{1}{2}}\Lambda\Phi^{\frac{1}{2}}(I + \Delta) = \Psi^{-\frac{1}{2}}S\Psi^{-1}\Lambda\Phi^{\frac{1}{2}} \quad (207)$$

$$= \Psi^{-\frac{1}{2}}S\Psi^{-\frac{1}{2}}\Psi^{-\frac{1}{2}}\Lambda\Phi^{\frac{1}{2}} \quad (208)$$

$$= S^*\Psi^{-\frac{1}{2}}\Lambda\Phi^{\frac{1}{2}} \quad (209)$$

を得る。ただし、

$$S^* = \Psi^{-\frac{1}{2}}S\Psi^{-\frac{1}{2}} \quad [Sstar] \quad (210)$$

と置いている。

そして、 Δ が対角行列であると仮定すると、(このことは、因子の軸をある方向に向けることに等しく、尤度関数の値に影響を与えるものではない) 上式の左右を入れ替えた式

$$S^* \left(\Psi^{-\frac{1}{2}}\Lambda\Phi^{\frac{1}{2}} \right) = \left(\Psi^{-\frac{1}{2}}\Lambda\Phi^{\frac{1}{2}} \right) (I + \Delta) \quad [eigenML] \quad (211)$$

は $\Psi^{-\frac{1}{2}} \Lambda \Phi^{\frac{1}{2}}$ が S^* 行列の固有ベクトルからなる行列、 $I + \Delta$ がその固有値を対角に持つ対角行列であることを示している。したがって、 Ψ が所与の場合、 S^* の r 次元のスペクトル分解を

$$S^* \approx \Omega_r \Theta_r \Omega_r', \quad \Omega_r' \Omega_r = I_r \quad [\text{SstarDecomp}] \quad (212)$$

とすれば、 Λ は Ω_r の列を対角行列 Q を用いて定数倍したものをを用いて

$$\Lambda = \Psi^{1/2} \Omega_r Q \Phi^{-\frac{1}{2}} \quad (213)$$

となる。ただし、この Q は Eq.(206) [Delta] を満たさなければならないため

$$\Delta = \Phi^{\frac{1}{2}} \Lambda' \Psi^{-1} \Lambda \Phi^{\frac{1}{2}} = \Phi^{\frac{1}{2}} \Phi^{-\frac{1}{2}} Q \Omega_r' \Psi^{1/2} \Psi^{-1} \Psi^{1/2} \Omega_r Q \Phi^{-\frac{1}{2}} \Phi^{\frac{1}{2}} = Q^2 \quad (214)$$

となるが、これより

$$I_r + \Delta = I_r + Q^2 = \Theta_r \quad [\text{Theta}] \quad (215)$$

となり、

$$Q = (\Theta_r - I_r)^{-\frac{1}{2}} \quad (216)$$

を得る。すなわち

$$\Lambda = \Psi^{1/2} \Omega_r (\Theta_r - I_r)^{1/2} \Phi^{-\frac{1}{2}} \quad [\text{LambdaML}] \quad (217)$$

である。

これまでの議論で、 Λ は S^* の r 個の固有ベクトルを定数倍したものをを用いて算出されることがわかったが、 p 個ある固有ベクトルのうちどの固有値に対応するものを選べば良いかは不明である。それを知るためには対数尤度比に相当する Eq.(185) [F0] に出てくる $\text{tr}(\Sigma^{-1} S)$ や $|\Sigma^{-1} S|$ を固有値や固有ベクトルの関数として表現する必要があるが、そのためには以下の面倒な議論が必要となる。

いま、これらは、

$$\text{tr}(\Sigma^{-1} S) = \text{tr}\left(\Psi^{\frac{1}{2}} \Sigma^{-1} \Psi^{\frac{1}{2}} \Psi^{-\frac{1}{2}} S \Psi^{-\frac{1}{2}}\right) = \text{tr}\left(\Psi^{\frac{1}{2}} \Sigma^{-1} \Psi^{\frac{1}{2}} S^*\right) \quad (218)$$

$$|\Sigma^{-1} S| = \left|\Psi^{\frac{1}{2}} \Sigma^{-1} \Psi^{\frac{1}{2}} \Psi^{-\frac{1}{2}} S \Psi^{-\frac{1}{2}}\right| = \left|\Psi^{\frac{1}{2}} \Sigma^{-1} \Psi^{\frac{1}{2}} S^*\right| \quad (219)$$

と書くことが出来るため、 $\Psi^{\frac{1}{2}} \Sigma^{-1} \Psi^{\frac{1}{2}} S^*$ を S^* の固有値と固有ベクトルで表すことにする。

S^* のスペクトル分解を

$$S^* = \Omega \Theta \Omega', \quad \Omega' \Omega = \Omega \Omega' = I_p \quad (220)$$

とし、固有値と固有ベクトルを以下の様に二つのサブセットに分割する。

$$\Theta = \text{Diag}(\Theta_1, \Theta_2) \quad \text{and} \quad \Omega = [\Omega_1, \Omega_2] \quad (221)$$

ただし、 Θ_1, Θ_2 はそれぞれ r 次ならびに $p-r$ 次の対角行列、 Ω_1, Ω_2 はそれぞれ $p \times r$ ならびに $p \times (p-r)$ の行列であり、サブセットの選び方にかかわらず

$$\Omega_1' \Omega_1 = I_r, \quad \Omega_1' \Omega_2 = O, \quad \Omega_2' \Omega_2 = I_{p-r} \quad (222)$$

$$\Omega \Omega' = \Omega_1 \Omega_1' + \Omega_2 \Omega_2' = I_p \quad (223)$$

を満たす。また、Eq.(215) [Theta] は $r = p$ の場合にも成り立つので、全ての固有値は 1 より小さくなることはないことがわかる。

今ここで Λ を (Θ_1, Ω_1) を用いて算出したとすると、すなわち、Eq.(212) [SstarDecomp] に現れる (Θ_r, Ω_r) を (Θ_1, Ω_1) とすると、

$$\Psi^{-\frac{1}{2}} \Sigma \Psi^{-\frac{1}{2}} = \Psi^{-\frac{1}{2}} (\Lambda \Phi \Lambda' + \Psi) \Psi^{-\frac{1}{2}} \quad (224)$$

$$= \Psi^{-\frac{1}{2}} \left(\Psi^{1/2} \Omega_1 (\Theta_1 - I_r)^{1/2} \Phi^{-\frac{1}{2}} \Phi \Phi^{-\frac{1}{2}} (\Theta_1 - I_r)^{1/2} \Omega_1' \Psi^{1/2} + \Psi \right) \Psi^{-\frac{1}{2}} \quad (225)$$

$$= \Omega_1 (\Theta_1 - I_r) \Omega_1' + I_p = \Omega_1 \Theta_1 \Omega_1' - \Omega_1 \Omega_1' + I_p \quad (226)$$

$$= \Omega_1 \Theta_1 \Omega_1' - (I_p - \Omega_2 \Omega_2') + I_p = \Omega_1 \Theta_1 \Omega_1' + \Omega_2 \Omega_2' \quad (227)$$

$$= [\Omega_1, \Omega_2] \begin{bmatrix} \Theta_1 & \mathbf{O} \\ \mathbf{O} & I_{n-r} \end{bmatrix} [\Omega_1, \Omega_2]' \quad [\text{OTO}] \quad (228)$$

を得るが、 Ω は直交行列であり Θ_1 は対角行列であるため、

$$\Psi^{\frac{1}{2}} \Sigma^{-1} \Psi^{\frac{1}{2}} = \left(\Psi^{-\frac{1}{2}} \Sigma \Psi^{-\frac{1}{2}} \right)^{-1} = [\Omega_1, \Omega_2] \begin{bmatrix} \Theta_1^{-1} & \mathbf{O} \\ \mathbf{O} & I_{n-r} \end{bmatrix} [\Omega_1, \Omega_2]' \quad (229)$$

$$= \Omega_1 \Theta_1^{-1} \Omega_1' + \Omega_2 \Omega_2' \quad [\text{PhSPH}] \quad (230)$$

である。

この表現と、 S^* のスペクトル分解が

$$S^* = \Omega \Theta \Omega' = \Omega_1 \Theta_1 \Omega_1' + \Omega_2 \Theta_2 \Omega_2' \quad (231)$$

と書けることから、先に示した行列は

$$\Psi^{\frac{1}{2}} \Sigma^{-1} \Psi^{\frac{1}{2}} S^* = \left(\Omega_1 \Theta_1^{-1} \Omega_1' + \Omega_2 \Omega_2' \right) \left(\Omega_1 \Theta_1 \Omega_1' + \Omega_2 \Theta_2 \Omega_2' \right) \quad (232)$$

$$= \Omega_1 \Omega_1' + \Omega_2 \Theta_2 \Omega_2' \quad (233)$$

となるため、結局

$$\text{tr} \left(\Sigma^{-1} S \right) = r + \text{tr} (\Theta_2), \quad \text{and} \quad \left| \Sigma^{-1} S \right| = \prod_a^p \theta_{2a} \quad (234)$$

となり、Eq.(186) [F] の最小化すべき基準は S^* の固有値を使って

$$F = r + \sum_a \theta_{2a} - \sum_a \log \theta_{2a} \quad [\text{Fbytheta}] \quad (235)$$

と書くことが出来る。また、この関数は全ての θ_{2a} に関して区間 $1 \leq \theta_{2a}$ において単調増加であるため、この最小値を与える Θ_2 としては、 S^* の最も小さい固有値を下から $n - r$ 個選べば良いことになる。したがって、Eq.(217) [LambdaML] の (Θ_r, Ω_r) としては、 S^* の大きい方から r 個選んだ固有値とそれらに対応する固有ベクトルを用いれば良いことになる。

次に、 Ψ に関してであるが、Eq.(188) [dFdLambda] が \mathbf{O} の場合、すなわち Λ が Eq.(217) [LambdaML] で与えられている場合について考える。

初めに、Eq.(194) [bfi-1p] を Eq.(188) [dFdLambda] に代入して \mathbf{O} と置くと

$$\Sigma^{-1} (\Sigma - S) \Sigma^{-1} \Lambda \Phi = \Sigma^{-1} (\Sigma - S) \Psi^{-1} \Lambda (\Phi^{-1} + \Lambda' \Psi^{-1} \Lambda)^{-1} = \mathbf{O} \quad (236)$$

を得るが、 Σ と $\Phi^{-1} + \Lambda' \Psi^{-1} \Lambda$ の正則性により

$$(\Sigma - S) \Psi^{-1} \Lambda = \mathbf{O} \quad (237)$$

であることを準備しておく。

次に、Eq.(190)[dFdPsi] の $\text{diag}()$ の中の式の右側の Σ^{-1} に Eq.(193)[bif] を代入し、上式の間係を用いると

$$\Sigma^{-1}(\Sigma - S)\Sigma^{-1} = \Sigma^{-1}(\Sigma - S) \left(\Psi^{-1} - \Psi^{-1} \Lambda (\Phi^{-1} + \Lambda' \Psi^{-1} \Lambda)^{-1} \Lambda' \Psi^{-1} \right) \quad (238)$$

$$= \Sigma^{-1}(\Sigma - S) \Psi^{-1} - \Sigma^{-1}(\Sigma - S) \Psi^{-1} \Lambda (\Phi^{-1} + \Lambda' \Psi^{-1} \Lambda)^{-1} \Lambda' \Psi^{-1} \quad (239)$$

$$= \Sigma^{-1}(\Sigma - S) \Psi^{-1} \quad (240)$$

となるが、この Σ^{-1} にも、対称性を考慮して同様の操作を行うと、最終的に

$$\Sigma^{-1}(\Sigma - S)\Sigma^{-1} = \Psi^{-1}(\Sigma - S)\Psi^{-1} \quad [\text{Sigma2Psi}] \quad (241)$$

の間係を得る。したがって、 Ψ が対角行列であることから、Eq.(190)[dFdPsi] を \mathbf{O} と置いた方程式は、

$$\text{diag}(\Sigma - S) = \text{diag}(\Lambda \Phi \Lambda' + \Psi - S) = \mathbf{O} \quad (242)$$

となり、

$$\Psi = \text{diag}(S - \Lambda \Phi \Lambda') \quad [\text{PsiML}] \quad (243)$$

となる。ただし、この間係は Eq.(191)[dFdLambda0] が成り立っている場合に限りことに注意が必要である。

従って、最尤解は、 Ψ の初期値を与えられたときに、Eq.(210)[Sstar] を計算し、その固有値と固有ベクトルを用いて Eq.(217)[LambdaML] により Λ を更新し、その値を用いて Eq.(243)[PsiML] により Ψ の値を更新する、という繰り返し計算を行うことにより求めることができる。この繰り返しアルゴリズムは Laykey(1940) により提案された。(Lawkey & Maxwell, 1971 も参照)

2.2.3 Joereskog の方法 [Joe]

Eq.(217)[LambdaML] で与えられる Λ は、 Ψ を与えたときの条件付きの最尤解を与えるものであった。すなわち、式中の Ω や Θ は Ψ に依存する量であった。同様に、Eq.(154)[LambdaLS] や Eq.(167)[LambdaLS2] で与えられる最小2乗解も P や Γ を通して Ψ に依存している。すなわち、 Ψ が与えられれば、 S と Ψ の値を用いて定義される行列の固有値と固有ベクトルを計算することにより、 Λ が決まる。いま、この依存性を明らかにするために、それぞれを

$$\text{最尤解: } \Lambda(\Psi) = \Psi^{\frac{1}{2}} \Omega(\Psi) (\Theta(\Psi) - I)^{1/2} \Phi^{-\frac{1}{2}} \quad (244)$$

$$\text{最小2乗解: } \Lambda(\Psi) = P(\Psi) \Gamma(\Psi)^{\frac{1}{2}} \quad (245)$$

$$(246)$$

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 Eq.(146)[rssS] や Eq.(186)[F] は

$$F = \text{tr} \left(\Sigma(\Psi)^{-1} S \right) - \log \left| \Sigma(\Psi)^{-1} S \right| \quad \text{and} \quad \text{RSS} = \text{tr} \left[(S - \Sigma(\Psi))' (S - \Sigma(\Psi)) \right] \quad (247)$$

と書けるため、これを Ψ の対角要素のみの関数と見なして Newton-Raphson 法等で最小値を求めることが出来る。その際、数値微分を用いるのであればこの表現で十分であるが、Eq.(235) [Fbytheta] や Eq.(157) [rssbygamma] を用いれば F の Ψ に関する1次微分 (gradient) や2次微分 (Hessian) を解析的に求めることも可能である。ここでは説明を省略し 柳井他 (1990) が示した結果 (3.8 式と 3.40 式) のみを示す。

$$\frac{\partial F}{\partial \psi_j} = \frac{1}{\psi_j} \sum_{k=r+1}^p \omega_{jk}^2 (1 - \theta_k), \quad j = 1, 2, \dots, p \quad (248)$$

$$\frac{\partial \text{RSS}}{\partial \psi_j} = -2 \sum_{k=r+1}^p p_{jk}^2 \gamma_k, \quad j = 1, 2, \dots, p \quad (249)$$

0159.pdf eq(41)

2.2.3.1 EM アルゴリズムによる解法

最尤解を求めるために EM アルゴリズムを利用することができる。(Dempster, Laird, & Rubin, 1977 を参照。) EM アルゴリズムでは、 $f_i, i = 1, 2, \dots, n$, を欠測データを見なし、それと、実際に観測されるデータ $y_i, i = 1, 2, \dots, n$, を合わせたもの $(y_i, f_i), i = 1, 2, \dots, n$, を完全データと呼ぶ。EM アルゴリズムでは、E ステップで欠測データの観測されるデータを与えたときの条件付きの分布を導出し、それに関する完全データを与えたときの母数の対数尤度の期待値を計算する。次に、M ステップで、期待対数完全データ尤度を最大とするような母数を求める、という操作を繰り返すことにより、Eq.(182) [lnL] を最大化する、すなわち、Eq.(186) [F] を最小化する母数を推定する。

完全データを与えたときのパラメタの尤度、すなわち完全データ尤度は、Eq.(121) [randomy] と Eq.(123) [randomf] から

$$L(\Lambda, \Psi) \propto \prod_{i=1}^n g(y_i | f_i) h(f_i) \quad (250)$$

$$\propto \prod_{i=1}^n \frac{1}{\sqrt{(2\pi)^p |\Psi|}} \exp\left(-\frac{1}{2}(y_i - \mu - \Lambda f_i)' \Psi^{-1} (y_i - \mu - \Lambda f_i)\right) \times \frac{1}{\sqrt{(2\pi)^r |\Phi|}} \exp\left(-\frac{1}{2}(f_i - \mu_0)' \Phi^{-1} (f_i - \mu_0)\right) \quad [\text{cd1}] \quad (251)$$

となるが、その対数、すなわち、対数完全データ尤度は

$$\ln L(\Lambda, \Psi) = -\frac{1}{2} \left(n \log |\Psi| + \sum_{i=1}^n (y_i - \mu - \Lambda f_i)' \Psi^{-1} (y_i - \mu - \Lambda f_i) \right) + C \quad [\text{lcd1}] \quad (252)$$

となる。ただし C は母数を含まない定数である。

また、Eq.(251) [cd1] より、 y_i を与えたときの f_i の条件付き分布は、 f と y の同時分布を

$$\begin{pmatrix} f_i \\ y_i \end{pmatrix} \sim N \left(\begin{pmatrix} 0 \\ \mu \end{pmatrix}, \begin{bmatrix} \Phi & \Phi \Lambda' \\ \Lambda \Phi & \Sigma \end{bmatrix} \right) \quad (253)$$

とすれば、

$$\mathbf{f}_i | \mathbf{y}_i \sim N(\mathbf{f}_i^*, \mathbf{V}^*) \quad [\text{fgiveny}] \quad (254)$$

ただし、

$$\mathbf{V}^* = \Phi - \Phi \Lambda' \Sigma^{-1} \Lambda \Phi \quad [\text{Vstar}] \quad (255)$$

$$\mathbf{f}_i^* = \Phi \Lambda' \Sigma^{-1} \mathbf{y}_i \quad [\text{fstar}] \quad (256)$$

であるが、逆行列の公式

$$\mathbf{D}(\mathbf{A} + \mathbf{BCD})^{-1} \mathbf{B} = \mathbf{C}^{-1} - \mathbf{C}^{-1}(\mathbf{DA}^{-1} \mathbf{B} + \mathbf{C}^{-1})^{-1} \mathbf{C}^{-1} \quad [\text{mSI12}] \quad (257)$$

を用いると

$$\Lambda' \Sigma^{-1} \Lambda = \Phi^{-1} - \Phi^{-1}(\Lambda' \Psi^{-1} \Lambda + \Phi^{-1})^{-1} \Phi^{-1} \quad (258)$$

であるから

$$\mathbf{V}^* = \Phi - \Phi (\Lambda' \Sigma^{-1} \Lambda) \Phi \quad (259)$$

$$= \Phi - \Phi (\Phi^{-1} - \Phi^{-1}(\Lambda' \Psi^{-1} \Lambda + \Phi^{-1})^{-1} \Phi^{-1}) \Phi \quad (260)$$

$$= (\Lambda' \Psi^{-1} \Lambda + \Phi^{-1})^{-1} \quad [\text{Vstar2}] \quad (261)$$

となる。したがって、

$$\mathbf{f}_i^* = \mathbf{V}^* \Lambda' \Psi^{-1} \mathbf{y}_i \quad [\text{fstar2}] \quad (262)$$

であるが、これをまとめて $n \times r$ の行列 \mathbf{F}^* を

$$\mathbf{F}^* = \mathbf{Y} \Psi^{-1} \Lambda \mathbf{V}^* \quad [\text{Fstar}] \quad (263)$$

と書く。

次に、2次形式の期待値の公式

$$\mathcal{E}((\mathbf{x} - \mathbf{a})' \mathbf{B} (\mathbf{x} - \mathbf{a})) = (\mathcal{E}(\mathbf{x}) - \mathbf{a})' \mathbf{B} (\mathcal{E}(\mathbf{x}) - \mathbf{a}) + \text{tr}(\mathbf{B} \mathbf{D}(\mathbf{x})) \quad (264)$$

を用いると、対数完全データ尤度の期待値で母数を含む部分は

$$\begin{aligned} \mathcal{E} \ln L(\Lambda, \Psi) &= -\frac{1}{2} \left(n \log |\Psi| + \sum_{i=1}^n \left((\mathbf{y}_i - \boldsymbol{\mu} - \Lambda \mathbf{f}_i^*)' \Psi^{-1} (\mathbf{y}_i - \boldsymbol{\mu} - \Lambda \mathbf{f}_i^*) + \text{tr}(\Psi^{-1} \Lambda \mathbf{V}^* \Lambda') \right) \right) \\ & \quad [\text{elcd1}] \quad (265) \end{aligned}$$

$$= -\frac{1}{2} \left(n \log |\Psi| + \text{tr}(\Psi^{-1} (\mathbf{Y} - \mathbf{1} \boldsymbol{\mu}' - \mathbf{F}^* \Lambda')' (\mathbf{Y} - \mathbf{1} \boldsymbol{\mu}' - \mathbf{F}^* \Lambda')) \right) + n \text{tr}(\Psi^{-1} \Lambda \mathbf{V}^* \Lambda') \quad (266)$$

となるが、Eq.(179) [sscp] と同様の操作で上式の最初のトレースの中に現れる2次形式の部分は

$$(\mathbf{Y} - \mathbf{1} \boldsymbol{\mu}' - \mathbf{F}^* \Lambda')' (\mathbf{Y} - \mathbf{1} \boldsymbol{\mu}' - \mathbf{F}^* \Lambda') = \quad (267)$$

$$n \left(\frac{1}{n} (\mathbf{Y} - \mathbf{F}^* \Lambda')' \mathbf{1} - \boldsymbol{\mu}' \right) \left(\frac{1}{n} (\mathbf{Y} - \mathbf{F}^* \Lambda')' \mathbf{1} - \boldsymbol{\mu} \right) + (\mathbf{Y} - \mathbf{F}^* \Lambda')' \mathbf{J} (\mathbf{Y} - \mathbf{F}^* \Lambda') \quad (268)$$

と書けるため μ を

$$\mu = \frac{1}{n} (\mathbf{Y} - \mathbf{F}^* \mathbf{\Lambda}')' \mathbf{1} \quad (269)$$

として推定することにすれば、期待対数完全データ尤度は

$$E \ln L(\mathbf{\Lambda}, \mathbf{\Psi}) = -\frac{1}{2} \left(n \log |\mathbf{\Psi}| + \text{tr} \left(\mathbf{\Psi}^{-1} (\mathbf{Y} - \mathbf{F}^* \mathbf{\Lambda}')' \mathbf{J} (\mathbf{Y} - \mathbf{F}^* \mathbf{\Lambda}') \right) + n \text{tr} \left(\mathbf{\Psi}^{-1} \mathbf{\Lambda} \mathbf{V}^* \mathbf{\Lambda}' \right) \right) \quad [\text{elcd12}] \quad (270)$$

と書ける。したがって、M ステップは、上記の Eq.(270) [elcd12] を最大化する母数 $\mathbf{\Lambda}$ と $\mathbf{\Psi}$ を求めることになるが、その際、欠測データの条件付き分布のパラメタ \mathbf{F}^* と \mathbf{V}^* は定数として取り扱う。なお、この最大化されるべき関数の単純性が EM アルゴリズムの特徴である。

Eq.(270) [elcd12] を $\mathbf{\Lambda}$ で偏微分すると、

$$\frac{\partial -2E \ln L(\mathbf{\Lambda}, \mathbf{\Psi})}{\partial \mathbf{\Lambda}} = \mathbf{\Psi}^{-1} \left(-2\mathbf{Y}' \mathbf{J} \mathbf{F}^* + 2\mathbf{\Lambda} (\mathbf{F}^*{}' \mathbf{J} \mathbf{F}^* + n\mathbf{V}^*) \right) \quad (271)$$

であるため、これを \mathbf{O} と置くことにより

$$\mathbf{\Lambda} = \mathbf{Y}' \mathbf{J} \mathbf{F}^* (\mathbf{F}^*{}' \mathbf{J} \mathbf{F}^* + n\mathbf{V}^*)^{-1} \quad [\text{Lambda-em}] \quad (272)$$

を得る。また、 $\mathbf{\Psi}$ に関しては、Eq.(270) [elcd12] の形から、 $\log |\mathbf{\Psi}|$ を $-\log |\mathbf{\Psi}^{-1}|$ と書き換え、 $\mathbf{\Psi}^{-1}$ で偏微分してゼロと置くことにより

$$\mathbf{\Psi} = \frac{1}{n} \text{diag} \left((\mathbf{Y} - \mathbf{F}^* \mathbf{\Lambda}')' \mathbf{J} (\mathbf{Y} - \mathbf{F}^* \mathbf{\Lambda}') + n\mathbf{\Lambda} \mathbf{V}^* \mathbf{\Lambda}' \right) \quad [\text{Psi-em}] \quad (273)$$

が最小値を与えることがわかる。

なお、上記の解は、 $n \times r$ の行列 \mathbf{F}^* を含んでいるが、以下の様に変形することにより \mathbf{S} を用いた表現を得ることが可能である。なぜなら、

$$\mathbf{F}^*{}' \mathbf{J} \mathbf{F}^* = n\mathbf{V}^* \mathbf{\Lambda}'_0 \mathbf{\Psi}_0^{-1} \mathbf{S} \mathbf{\Psi}_0^{-1} \mathbf{\Lambda}_0 \mathbf{V}^* \quad [\text{FF}] \quad (274)$$

$$\mathbf{Y}' \mathbf{J} \mathbf{F}^* = n\mathbf{S} \mathbf{\Psi}_0^{-1} \mathbf{\Lambda}_0 \mathbf{V}^* \quad [\text{YF}] \quad (275)$$

であるため、これらを Eq.(272) [Lambda-em] と Eq.(273) [Psi-em] に代入すると

$$\mathbf{\Lambda} = \mathbf{S} \mathbf{\Psi}_0^{-1} \mathbf{\Lambda}_0 \mathbf{V}^* (\mathbf{V}^* \mathbf{\Lambda}'_0 \mathbf{\Psi}_0^{-1} \mathbf{S} \mathbf{\Psi}_0^{-1} \mathbf{\Lambda}_0 \mathbf{V}^* + \mathbf{V}^*)^{-1} \quad [\text{Lambda-em2}] \quad (276)$$

$$= \mathbf{S} \mathbf{\Psi}_0^{-1} \mathbf{\Lambda}_0 (\mathbf{V}^* \mathbf{\Lambda}'_0 \mathbf{\Psi}_0^{-1} \mathbf{S} \mathbf{\Psi}_0^{-1} \mathbf{\Lambda}_0 + \mathbf{I})^{-1} \quad [\text{Lambda-em3}] \quad (277)$$

$$\mathbf{\Psi} = \frac{1}{n} \text{diag} \left(\mathbf{Y}' \mathbf{J} \mathbf{Y} - \mathbf{Y}' \mathbf{J} \mathbf{F}^* \mathbf{\Lambda}' - \mathbf{\Lambda} \mathbf{F}^* \mathbf{J} \mathbf{Y} + \mathbf{\Lambda} \mathbf{F}^* \mathbf{J} \mathbf{F}^* \mathbf{\Lambda}' + n\mathbf{\Lambda} \mathbf{V}^* \mathbf{\Lambda}' \right) \quad (278)$$

$$= \text{diag} \left(\mathbf{S} - \mathbf{S} \mathbf{\Psi}_0^{-1} \mathbf{\Lambda}_0 \mathbf{V}^* \mathbf{\Lambda}' - \mathbf{\Lambda} \mathbf{V}^* \mathbf{\Lambda}'_0 \mathbf{\Psi}_0^{-1} \mathbf{S} \right) \quad (279)$$

$$+ \mathbf{\Lambda} \mathbf{V}^* \mathbf{\Lambda}'_0 \mathbf{\Psi}_0^{-1} \mathbf{S} \mathbf{\Psi}_0^{-1} \mathbf{\Lambda}_0 \mathbf{V}^* \mathbf{\Lambda}' + \mathbf{\Lambda} \mathbf{V}^* \mathbf{\Lambda}' \quad [\text{Psi-em2}] \quad (280)$$

$$= \text{diag} \left(\mathbf{S} - 2\mathbf{S} \mathbf{\Psi}_0^{-1} \mathbf{\Lambda}_0 \mathbf{V}^* \mathbf{\Lambda}' + \mathbf{\Lambda} \mathbf{V}^* \mathbf{\Lambda}'_0 \mathbf{\Psi}_0^{-1} \mathbf{S} \mathbf{\Psi}_0^{-1} \mathbf{\Lambda}_0 \mathbf{V}^* \mathbf{\Lambda}' + \mathbf{\Lambda} \mathbf{V}^* \mathbf{\Lambda}' \right) \quad [\text{Psi-em22}] \quad (281)$$

を得る。

なお、上記の解は、付録に示す行列間の様々な関係を用いると以下の様に表現することもできるが、

$$\Lambda_0 = S \Sigma_0^{-1} \Lambda_0 \Phi (\Phi \Lambda_0' \Sigma_0^{-1} S \Sigma_0^{-1} \Lambda_0 \Phi + \Phi - \Phi \Lambda_0' \Sigma_0^{-1} \Lambda_0 \Phi)^{-1} \quad [\text{Lambda} - \text{RT10}] \quad (282)$$

$$= S \Sigma_0^{-1} \Lambda_0 \Phi (\Phi \Lambda_0' \Sigma_0^{-1} S \Sigma_0^{-1} \Lambda_0 \Phi + V^*)^{-1} \quad [\text{Lambda} - \text{RT10} - 1] \quad (283)$$

$$= S \Psi_0^{-1} \Lambda_0 (\Phi \Lambda_0' \Sigma_0^{-1} S \Psi^{-1} \Lambda_0 + I)^{-1} \quad [\text{Lambda} - \text{ZYZ14}] \quad (284)$$

$$\Psi = \text{diag} \left(S - S \Sigma_0^{-1} \Lambda_0 \Phi (\Phi \Lambda_0' \Sigma_0^{-1} S \Sigma_0^{-1} \Lambda_0 \Phi + V^*)^{-1} \Phi \Lambda_0' \Sigma_0^{-1} S \right) \quad [\text{Psi} - \text{RT10}] \quad (285)$$

$$= \text{diag} \left(S (I - \Sigma_0^{-1} \Lambda_0 \Phi \Lambda_0') \right) = \text{diag} \left(S \Sigma_0^{-1} \Psi_0 \right) \quad [\text{Psi} - \text{em4}] \quad (286)$$

$$= \text{diag} \left(S (I - \Psi^{-1} \Lambda_0 V^* \Lambda_0') \right) \quad [\text{Psi} - \text{em3}] \quad (287)$$

Eq.(282) [Lambda-RT10] と Eq.(285) [Psi-RT10] が Rubin & Thayer (1982) で用いられているものであり、Eq.(284) [Lambda-ZYZ14] と Eq.(286) [Psi-em4] の表現は Zhao. et al. (2008) に見られる。なお、Eq.(286) [Psi-em4] と Eq.(287) [Psi-em3] が成り立つのは Λ_0 が EM アルゴリズムの解であり場合に限ることに注意が必要である。

また、EM アルゴリズムにおいては、Lawley の方法で仮定した Eq.(206) [Delta] が対角であるという制約を用いていないが、このことは、最終的に得られる解がどのような性質を持つかは初期値に依存して決まることを意味している。したがって、通常は、繰り返し計算が収束した時点で、たとえば $\Lambda' \Lambda$ が対角行列となる様に、すなわち、因子の軸をその主軸方向に向くような回転を施して最終解とすることが多い。なお、この操作は、Lawley の方法を用いた得られた解にも施すことが出来る。

最後に、Eq.(273) [Psi-em] を他の方法による Ψ の推定式 (Eq.(158) [PsiLS], Eq.(168) [PsiLS2], Eq.(243) [PsiML]) と比較してみると、EM アルゴリズムを用いた場合に限り、その推定値が非負となることが保証されていることが分かる。なぜなら、Eq.(273) [Psi-em] は二つの2次形式の和の対角要素で構成されているからである。なお、次節に述べる制約条件を課した際の解も負にはならない。

2.2.4 誤差分散に構造を仮定した場合の解法

上記のランダムモデルの最小2乗解や最尤解には解析解が存在しなかった。しかし、全ての誤差分散が等しい、すなわち

$$\Psi = \psi I \quad (288)$$

を仮定する事により、解析解を得ることが出来る。

2.2.4.1 最小2乗解

最小2乗解に関しては、 S の固有値からなる対角行列を D_1 とすれば $S - cI$ の固有値が $D_1 - cI$ となり、固有ベクトルは同一であることに注意すれば、Eq.(153) [eigen] と Eq.(154) [LambdaLS] に現れる Γ は、 S の r 次元のスペクトル分解を

$$S \approx P_1 D_1 P_1' \quad [\text{eigenS}] \quad (289)$$

とすれば、

$$\Gamma = D_1 - \psi I \quad \text{and} \quad P = P_1 \quad (290)$$

となる。したがって、Eq.(154) [LambdaLS] は

$$\Lambda = P \Gamma^{1/2} = P_1 (D_1 - \psi I)^{1/2} \quad [\text{LambdaLS} - \text{hs}] \quad (291)$$

となる。

また、RSS の ψ に関する偏微分は、Eq.(150) [drssSdPsi] より

$$\frac{\partial \text{RSS}}{\partial \psi} = 2(p\psi + \text{tr}(\Lambda\Lambda' - S)) \quad (292)$$

となるためこれをゼロと置くと

$$\psi = \text{tr}(\Lambda\Lambda' - S)/p \quad (293)$$

であるが、そこに Eq.(291) [LambdaLS-hs] を用いると

$$\Lambda\Lambda' = P_1(D_1 - \psi I)P_1' \quad (294)$$

$$= P_1 D_1 P_1' - \psi P_1 P_1' \quad (295)$$

となる。したがって、

$$\psi = \frac{\text{tr}(S) - \text{tr}(D_1)}{p - r} \quad [\text{psi} - \text{hs}] \quad (296)$$

を得る。

すなわち、このモデルの解は、Eq.(289) [eigenS] で与えられる P_1 と D_1 を用いて Eq.(296) [psi-hs] により ψ を求めし、それを Eq.(291) [LambdaLS-hs] に代入して Λ を求めることにより得られる。

2.2.4.2 最尤解

最尤解の場合は、より一般的に、 Ψ が既知の対角行列 C に比例する、すなわち

$$\Psi = \psi C \quad [\text{Psi}C] \quad (297)$$

というモデルの解析解が存在する。いま、このモデルの下では Eq.(211) [eigenML] が

$$\frac{1}{\psi} C^{-\frac{1}{2}} S C^{-\frac{1}{2}} \approx \Omega \Theta \Omega' \quad [\text{eigenML2}] \quad (298)$$

と書けるため、Eq.(103) [sd2] と同様に、 $C^{-\frac{1}{2}}$ を用いてスケーリングされた分散共分散行列の r 次元のスペクトル分解を

$$S^\dagger = C^{-\frac{1}{2}} S C^{-\frac{1}{2}} \approx P^\dagger \Gamma^\dagger P'^\dagger \quad [\text{sd3}] \quad (299)$$

とすれば、

$$\Theta = (1/\psi)\Gamma^\dagger \quad \text{and} \quad \Omega = P^\dagger \quad [\text{T0}] \quad (300)$$

である。また、Eq.(217) [LambdaML] より、

$$\Lambda = \sqrt{\psi} C^{\frac{1}{2}} \Omega (\Theta - I)^{\frac{1}{2}} \Phi^{-\frac{1}{2}} \quad (301)$$

$$= C^{\frac{1}{2}} P^\dagger (\Gamma^\dagger - \psi I)^{\frac{1}{2}} \Phi^{-\frac{1}{2}} \quad [\text{LambdaML} - \text{hs}] \quad (302)$$

であるため、

$$\Lambda \Phi \Lambda' = C^{\frac{1}{2}} P^\dagger (\Gamma^\dagger - \psi I) P'^\dagger C^{\frac{1}{2}} \quad (303)$$

$$= C^{\frac{1}{2}} P^\dagger \Gamma^\dagger P'^\dagger C^{\frac{1}{2}} - \psi C^{\frac{1}{2}} P^\dagger P'^\dagger C^{\frac{1}{2}} \quad [\text{LL}] \quad (304)$$

となる。

次に, Eq.(297) [PsiC] を仮定した場合の対数尤度関数の ψ により微分は,

$$\frac{d\Psi}{d\psi} = C \quad (305)$$

であることから, Eq.(241) [Sigma2Psi] を用いると, Eq.(190) [dFdPsi] は

$$\frac{\partial F}{\partial \psi} = \text{tr} \left(\text{diag} \left(\Psi^{-1} (\Sigma - S) \Psi^{-1} \right) C \right) \quad (306)$$

$$= \frac{1}{\psi^2} \text{tr} \left(\text{diag} \left(C^{-1} (\Sigma - S) \right) \right) = \frac{1}{\psi^2} \text{tr} \left(C^{-1} (\Sigma - S) \right) \quad (307)$$

となるが, これを 0 と置くことにより, Eq.(304) [LL] を利用して

$$0 = \text{tr} \left(C^{-1} (\Sigma - S) \right) \quad (308)$$

$$= \text{tr} \left(C^{-1} (\Lambda \Phi \Lambda' + \psi C - S) \right) \quad (309)$$

$$= \text{tr} \left(C^{-1} \left(C^{\frac{1}{2}} P^\dagger \Gamma^\dagger P'^{\dagger'} C^{\frac{1}{2}} - \psi C^{\frac{1}{2}} P^\dagger P'^{\dagger'} C^{\frac{1}{2}} + \psi C - S \right) \right) \quad (310)$$

$$= \text{tr} \left(C^{-\frac{1}{2}} P^\dagger \Gamma^\dagger P'^{\dagger'} C^{\frac{1}{2}} - \psi C^{-\frac{1}{2}} P^\dagger P'^{\dagger'} C^{\frac{1}{2}} + \psi C^{-1} C - C^{-1} S \right) \quad (311)$$

$$= \text{tr} \left(P^\dagger \Gamma^\dagger P'^{\dagger'} C^{\frac{1}{2}} C^{-\frac{1}{2}} \right) - \psi \text{tr} \left(P^\dagger P'^{\dagger'} C^{\frac{1}{2}} C^{-\frac{1}{2}} \right) + \psi \text{tr} (I_p) - \text{tr} \left(C^{-\frac{1}{2}} C^{-\frac{1}{2}} S \right)$$

$$= \text{tr} \left(\Gamma^\dagger \right) - \text{tr} \left(C^{-\frac{1}{2}} S C^{-\frac{1}{2}} \right) - \psi \times r + \psi \times p \quad (312)$$

を得る。したがって,

$$\psi = \frac{\text{tr} \left(S^\dagger \right) - \text{tr} \left(\Gamma^\dagger \right)}{p - r} \quad [\text{psi} - \text{hs2}] \quad (313)$$

である

すなわち, このモデルの解は, Eq.(299) [sd3] で与えられる Γ^\dagger を用いて Eq.(313) [psi-hs2] により ψ を求め, それを用いて Eq.(300) [T0] の Θ と Ω を計算し, それらを Eq.(302) [LambdaML-hs] に代入して Λ を求めることにより得られる。

なお, $C = I$ の場合には, この解は Eq.(291) [LambdaLS-hs] と Eq.(296) [psi-hs] に示す最小2乗解の場合と同一となる。。

2.3 因子分析の解析解のまとめ

前節までに, 誤差分散に以下の構造

$$\Psi = \psi C \quad (314)$$

を仮定した場合には, 因子分析モデルのパラメタ (Λ, ψ) は解析的に求まることが示された。本節では, それらを再度掲載してその相互比較を行う。なお, データから計算される分散共分散行列 S を上記の C の平方根の逆数を用いてスケールしたものの r 次元のスペクトル分解を

$$\frac{1}{n} C^{-\frac{1}{2}} Y' J Y C^{-\frac{1}{2}} = C^{-\frac{1}{2}} S C^{-\frac{1}{2}} = S^\dagger \approx P^\dagger \Gamma^\dagger P'^{\dagger'} \quad (315)$$

とする。

- 母数モデルの最尤解

Eq.(105) [MLf-Beta]、Eq.(106) [MLf-X]、Eq.(107) [MLf-psi]

$$\boldsymbol{\beta} = \mathbf{C}^{\frac{1}{2}} \mathbf{P}^{\dagger} \boldsymbol{\Gamma}^{\dagger \frac{1}{2}} \quad (316)$$

$$\mathbf{X} = \mathbf{J} \mathbf{Y} \mathbf{C}^{-\frac{1}{2}} \mathbf{P}^{\dagger} \boldsymbol{\Gamma}^{\dagger -\frac{1}{2}} \quad (317)$$

$$\psi = \frac{\text{tr}(\mathbf{S}^{\dagger}) - \text{tr}(\boldsymbol{\Gamma}^{\dagger})}{p} \quad (318)$$

- ランダムモデルの最尤解

Eq.(313) [psi-hs2]、Eq.(302) [LambdaML-hs]

$$\psi = \frac{\text{tr}(\mathbf{S}^{\dagger}) - \text{tr}(\boldsymbol{\Gamma}^{\dagger})}{p - r} \quad (319)$$

$$\boldsymbol{\Lambda} = \mathbf{C}^{\frac{1}{2}} \mathbf{P}^{\dagger} (\boldsymbol{\Gamma}^{\dagger} - \psi \mathbf{I})^{\frac{1}{2}} \boldsymbol{\Phi}^{-\frac{1}{2}} \quad (320)$$

これらの解の特徴は、因子パタン行列 $\boldsymbol{\Lambda}$ は、行列 \mathbf{S}^{\dagger} の固有ベクトルに定数を掛けることにより、また、誤差分散 ψ は、 \mathbf{S}^{\dagger} 行列の対角要素の和からその固有値の最大 r 個の和を引いたものを何かで割った形で与えられていることである。したがって、 ψ は負にはならない。

なお、 $\mathbf{C} = \mathbf{I}$ を仮定した場合の直交解、すなわち

$$\boldsymbol{\Psi} = \psi \mathbf{I}, \quad \text{and} \quad \boldsymbol{\Phi} = \mathbf{I} \quad (321)$$

の場合は、 \mathbf{S} の r 次元のスペクトル分解を

$$\frac{1}{n} \mathbf{Y}' \mathbf{J} \mathbf{Y} = \mathbf{S} \approx \mathbf{P} \boldsymbol{\Gamma} \mathbf{P} \quad (322)$$

とすると以下の通りである。

- 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

Eq.(105) [MLf-Beta]、Eq.(106) [MLf-X]、Eq.(107) [MLf-psi]

$$\boldsymbol{\beta} = \mathbf{P} \boldsymbol{\Gamma}^{\frac{1}{2}} \quad (323)$$

$$\psi = \frac{\text{tr}(\mathbf{S}) - \text{tr}(\boldsymbol{\Gamma})}{p} \quad (324)$$

$$\mathbf{X} = \mathbf{J} \mathbf{Y} \mathbf{P} \boldsymbol{\Gamma}^{-\frac{1}{2}} \quad (325)$$

- ランダムモデルの最尤解 = ランダムモデルの最小2乗解

Eq.(296) [psi-hs]、Eq.(302) [LambdaML-hs]

$$\boldsymbol{\Lambda} = \mathbf{P} (\boldsymbol{\Gamma} - \psi \mathbf{I})^{\frac{1}{2}} \quad (326)$$

$$\psi = \frac{\text{tr}(\mathbf{S}) - \text{tr}(\boldsymbol{\Gamma})}{p - r} \quad (327)$$

3 確認的因子分析

本節では、因子パターン行列や因子間相関行列の一部の要素の値が所与である場合のパラメタの推定法に関して述べる。特に、因子パタンの特定の要素 λ_{jq} が 0 であるということは、因子 q が変数 j には全く影響を与えていないことを示している。逆に言うと、因子 q は、 Λ 行列の第 q 列の非負要素に対応する変数のみに影響を与えているということになるが、そのような意味で、因子パターン行列の 0 の要素を指定するということは、データの分析に先立って、求めるべき因子の意味づけを行っていると考えられる。したがって、このような方法は、仮説検証的因子分析、または、確認的因子分析と呼ばれている。また、パラメタ推定の観点からは、一部の要素の値を指定するということは、推定されるべきパラメタに制約が加わる訳であるから、尤度関数の値は、制約無しの場合よりも小さくなることが予想される。

3.1 EM アルゴリズムによる方法

通常、確認的因子分析は、尤度比検定等を行う場合があるため、最尤法によりランダムモデルのパラメタを推定する場合が多い。前節で述べたように、最尤解には Lawley の方法と EM アルゴリズムを用いる方法があるが、ここでは、より柔軟な EM アルゴリズムを用いる方法に手を入れた形で、パラメタの一部が既知であるという制約を入れた解を導く。

いま、改めて因子分析モデルを想起してみると、因子分析モデルとは、Eq.(28) [col1] に示すように、第 j 番目の変数 $\mathbf{y}_{(j)}$ を潜在変数からなる説明変数行列 \mathbf{X} または \mathbf{F} へ回帰するモデルであり、因子パターン行列の第 j 行、すなわち β'_j または λ'_j はその回帰モデルの回帰係数であった。

いま、第 j 番目の変数の因子パターン λ_j の第 I_1, I_2, \dots, I_{m_j} 要素が既知であり、それ以外の要素が未知であるとす。そして、既知の要素を λ_{j2} という $m_j \times 1$ のベクトルで、また、それに対応する \mathbf{F}^* 行列の列をまとめたものを \mathbf{F}_{j2}^* と記す。同様に、制約の無い λ_j の要素を並べた $(r - m_j) \times 1$ のベクトルを λ_{j1} 、それに対応する \mathbf{F}^* の要素を \mathbf{F}_{j1}^* と記すと、期待対数完全データ尤度は

$$E \ln L(\Lambda, \Psi) = -\frac{1}{2} \left(n \log |\Psi| + \sum_{j=1}^p \frac{1}{\psi_j} \left((\mathbf{y}_{(j)} - \mathbf{F}^* \lambda_j)' \mathbf{J} (\mathbf{y}_{(j)} - \mathbf{F}^* \lambda_j) + n \lambda_j' \mathbf{V}^* \lambda_j \right) \right) \quad (328)$$

[elcd13]

$$= -\frac{1}{2} \left(n \log |\Psi| + \sum_{j=1}^p \frac{1}{\psi_j} (Q_j + n \lambda_j' \mathbf{V}^* \lambda_j) \right) \quad (329)$$

と書けるが、ここに現れる λ_j の 2 次形式の部分を Q_j とすると

$$Q_j = (\mathbf{y}_{(j)} - \mathbf{F}^* \lambda_j)' \mathbf{J} (\mathbf{y}_{(j)} - \mathbf{F}^* \lambda_j) \quad (330)$$

$$= (\mathbf{y}_{(j)} - \mathbf{F}_{j2}^* \lambda_{j2} - \mathbf{F}_{j1}^* \lambda_{j1})' \mathbf{J} (\mathbf{y}_{(j)} - \mathbf{F}_{j2}^* \lambda_{j2} - \mathbf{F}_{j1}^* \lambda_{j1}) \quad (331)$$

であり、 \mathbf{V}^* 行列を λ_j の分割に対応して $\mathbf{V}_{j11}^*, \mathbf{V}_{j12}^* = \mathbf{V}_{j21}^*, \mathbf{V}_{j22}^*$ と分割したものをを用いると

$$\lambda_j' \mathbf{V}^* \lambda_j = \lambda_{j1}' \mathbf{V}_{j11}^* \lambda_{j1} + 2 \lambda_{j1}' \mathbf{V}_{j12}^* \lambda_{j2} + \lambda_{j2}' \mathbf{V}_{j22}^* \lambda_{j2} \quad (332)$$

である。したがって、

$$\frac{\partial Q_j + n \lambda_j' \mathbf{V}^* \lambda_j}{\partial \lambda_{j1}} = -2\mathbf{F}_{j1}^{*'} \mathbf{J}(\mathbf{y}_{(j)} - \mathbf{F}_{j2}^* \lambda_{j2}) + 2\mathbf{F}_{j1}^{*'} \mathbf{J} \mathbf{F}_{j1}^* \lambda_{j1} + 2n \mathbf{V}_{j12}^* \lambda_{j2} + 2n \mathbf{V}_{j11}^* \lambda_{j1} \quad (333)$$

であるため、 λ_{j1} は

$$\lambda_{j1} = (\mathbf{F}_{j1}^{*'} \mathbf{J} \mathbf{F}_{j1}^* + n \mathbf{V}_{j11}^*)^{-1} \left(\mathbf{J} \mathbf{F}_{j1}^{*'} \mathbf{y}_{(j)} - (\mathbf{F}_{j1}^{*'} \mathbf{J} \mathbf{F}_{j2}^* + n \mathbf{V}_{j12}^*) \lambda_{j2} \right) \quad [\text{Lambda-em41}] \quad (334)$$

となる。

なお、 Ψ の推定は、通常の EM アルゴリズムと同じ形、すなわち Eq.(281) [Psi-em22] や Eq.(285) [Psi-RT10] を用いれば良い。また、 Λ の全ての要素が既知の場合と同様に、Eq.(286) [Psi-em4] や Eq.(287) [Psi-em3] を用いてはいけなない。

3.2 因子の平均と相関行列の推定

確認的因子分析においては、制約条件として、因子パタンのいくつかの要素が既知であるという制約を課すが、それ故、探索的因子分析に付随している回転に関する不定性は解消する。しかし、逆に、因子間相関をモデルのパラメタとして取り扱うことが可能となる。Eq.(186) [F] に含まれる Σ が Φ の関数であることに注意すれば、

$$\frac{\partial F}{\partial \Phi} = \Lambda' \Sigma^{-1} \Lambda - \Lambda' \Sigma^{-1} \mathbf{S} \Sigma^{-1} \Lambda \quad (335)$$

$$= \Lambda' \Sigma^{-1} (\mathbf{I} - \mathbf{S} \Sigma^{-1}) \Lambda \quad (336)$$

であるが、このままの形では取り扱いが難しい。しかし、この、因子間相関行列 Φ の推定は、前節同様、EM アルゴリズムを用いることにより、簡単に行うことが出来る。

EM アルゴリズムの中に現れる対数完全データ尤度は Eq.(252) [1cd1] に示されているが、そこでは、パラメタ Λ と Ψ を含まない部分を定数 C として省略した形で示していた。しかし、この部分は 因子平均ベクトル μ_0 と因子相関行列 Φ を含む部分であり、以下の様な形をしていた。

$$C = -\frac{1}{2} \left(n \log |\Phi| + \sum_{i=1}^n (\mathbf{f}_i - \mu_0)' \Phi^{-1} (\mathbf{f}_i - \mu_0) \right) \quad (337)$$

したがって、これらを推定すべき母数と見なす際には、Eq.(270) [e1cd12] に示した期待対数完全データ尤度にこの部分の期待値

$$\mathcal{E}(C) = -\frac{1}{2} \left(n \log |\Phi| + \sum_{i=1}^n (\mathbf{f}_i^* - \mu_0)' \Phi^{-1} (\mathbf{f}_i^* - \mu_0) + n \text{tr} (\Phi^{-1} \mathbf{V}^*) \right) \quad (338)$$

$$= -\frac{1}{2} \left(n \log |\Phi| + \text{tr} (\Phi^{-1} ((\mathbf{F}^* - \mathbf{1} \mu_0')' (\mathbf{F}^* - \mathbf{1} \mu_0') + n \mathbf{V}^*)) \right) \quad [\text{EC}] \quad (339)$$

$$= -\frac{1}{2} \left(n \log |\Phi| + \text{tr} (\Phi^{-1} (n (\frac{1}{n} \mathbf{1}' \mathbf{F}^* - \mu_0') (\frac{1}{n} \mathbf{1}' \mathbf{F}^* - \mu_0')' + \mathbf{F}^{*'} \mathbf{J} \mathbf{F}^* + n \mathbf{V}^*)) \right) \quad (340)$$

[EC1]

をも加えた形の期待対数完全データ尤度を M ステップで最大化する必要がある。そして、

$$\frac{\partial \mathcal{E}(C)}{\partial \Phi} = n\Phi^{-1} - \Phi^{-1}((\mathbf{F}^* - \mathbf{1}\mu'_0)'(\mathbf{F}^* - \mathbf{1}\mu'_0) + n\mathbf{V}^*)\Phi^{-1} \quad (341)$$

$$= \Phi^{-1} \left(n\mathbf{I} - ((\mathbf{F}^* - \mathbf{1}\mu'_0)'(\mathbf{F}^* - \mathbf{1}\mu'_0) + n\mathbf{V}^*)\Phi^{-1} \right) \quad (342)$$

$$= \Phi^{-1} (n\Phi - ((\mathbf{F}^* - \mathbf{1}\mu'_0)'(\mathbf{F}^* - \mathbf{1}\mu'_0) + n\mathbf{V}^*))\Phi^{-1} \quad [\text{dCdPhi}] \quad (343)$$

$$= \Phi^{-1} \left(n\Phi - \left(n\left(\frac{1}{n}\mathbf{1}'\mathbf{F}^* - \mu'_0\right)\left(\frac{1}{n}\mathbf{1}'\mathbf{F}^* - \mu'_0\right)' + \mathbf{F}^{*'}\mathbf{J}\mathbf{F}^* + n\mathbf{V}^* \right) \right)\Phi^{-1} \quad (344)$$

[dCdPhi1]

であり、 Φ は正則であると仮定されているため、

$$\mu_0 = \frac{1}{n}\mathbf{F}^{*'}\mathbf{1} \quad [\text{mu0}] \quad (345)$$

$$\Phi = \frac{1}{n} \left(\mathbf{F}^{*'}\mathbf{J}\mathbf{F}^* + n\mathbf{V}^* \right) \quad [\text{Phi}]$$

が、Eq.(339) [EC] を最大とするものであることが分かる。ただし、因子の平均と分散に関する不定性を除去するために、通常の単一母集団からのデータを用いる分析の場合には

$$\mu_0 = \mathbf{0}, \quad \text{and} \quad \text{diag}(\Phi) = \mathbf{I} \quad (346)$$

という制約を掛けるため、M ステップで推定されるものは Eq.(346) [Phi] で表される Φ 行列の対角成分が 1 となる様に基準化したものとなる。

また、 m 個の因子が、互いに無相関な複数のブロックに分かれる形で因子間相関行列に 0 要素の指定がある場合は、Eq.(346) [Phi] で得られる Φ の推定値のうち、0 と指定されている部分を 0 としたものが最尤解となる。

それ以外のパタンの場合には、Eq.(339) [EC] を最大とするような、すなわち、Eq.(343) [dCdPhi] を 0 とするような制約付きの Φ の値を数値的に求める必要が生じる。

4 共分散構造分析

本節では、以下の公式を用いて、モデルごとに観測される変数の分散共分散行列の構造を導く。

公式

x と e をそれぞれ平均が $\mathbf{0}$ である互いに独立な確率変数とし、 μ と A を定数とする。また、 x と e の分散共分散行列をそれぞれ

$$\mathcal{D}(x) = \Sigma_x, \quad \mathcal{D}(e) = \Sigma_e \quad (347)$$

とする。このとき、

$$y = \mu + Ax + e \quad (348)$$

で y が定義されているならば、

$$\mathcal{D}(y) = A\Sigma_x A' + \Sigma_e \quad (349)$$

である。

なお、以下の記述では、 x の平均が μ 、分散共分散行列が Σ であることを正規分布の表現

$$x \sim N(\mu, \Sigma) \quad (350)$$

と表現しているが、正規分布の仮定は必要ではない。

4.1 モデルと共分散構造

4.1.1 因子分析（多因子）

$$y = \mu + \Lambda f + \varepsilon \quad (351)$$

$$f \sim N(0, \Phi) \quad (352)$$

$$\varepsilon \sim N(0, \Psi) \quad (353)$$

$$\text{Cov}(f, \varepsilon) = \mathbf{0} \quad (354)$$

$$y \sim N(\mu, \Lambda\Phi\Lambda' + \Psi) \quad (355)$$

4.1.2 因子分析（1因子）

$$y = \mu + \lambda f + \varepsilon \quad (356)$$

$$f \sim N(0, \phi) \quad (f \text{ はスカラー}) \quad (357)$$

$$\varepsilon \sim N(0, \Psi) \quad (358)$$

$$\text{Cov}(f, \varepsilon) = \mathbf{0}' \quad (359)$$

$$y \sim N(\mu, \phi \times \lambda\lambda' + \Psi) \quad (360)$$

4.1.3 2次因子分析 (多因子)

$$\mathbf{y} = \boldsymbol{\mu} + \boldsymbol{\Lambda}\mathbf{f} + \boldsymbol{\varepsilon} \quad (361)$$

$$\mathbf{f} = \boldsymbol{\Gamma}\boldsymbol{\xi} + \boldsymbol{\zeta} \quad (362)$$

$$\boldsymbol{\varepsilon} \sim N(\mathbf{0}, \boldsymbol{\Psi}) \quad (363)$$

$$\boldsymbol{\xi} \sim N(\mathbf{0}, \boldsymbol{\Sigma}_\xi) \quad (364)$$

$$\boldsymbol{\zeta} \sim N(\mathbf{0}, \boldsymbol{\Sigma}_\zeta) \quad (365)$$

$$\text{Cov}(\boldsymbol{\xi}, \boldsymbol{\zeta}) = \mathbf{O}, \text{Cov}(\boldsymbol{\xi}, \boldsymbol{\varepsilon}) = \mathbf{O}, \text{Cov}(\boldsymbol{\zeta}, \boldsymbol{\varepsilon}) = \mathbf{O} \quad (366)$$

$$\mathbf{y} \sim N(\boldsymbol{\mu}, \boldsymbol{\Lambda}(\boldsymbol{\Gamma}\boldsymbol{\Sigma}_\xi\boldsymbol{\Gamma}' + \boldsymbol{\Sigma}_\zeta)\boldsymbol{\Lambda}' + \boldsymbol{\Psi}) \quad (367)$$

4.1.4 多重指標モデル

$$\mathbf{y}_1 = \boldsymbol{\Lambda}_1\mathbf{f}_1 + \boldsymbol{\varepsilon}_1 \quad (368)$$

$$\mathbf{y}_2 = \boldsymbol{\Lambda}_2\mathbf{f}_2 + \boldsymbol{\varepsilon}_2 \quad (369)$$

$$\boldsymbol{\varepsilon}_1 \sim N(\mathbf{0}, \boldsymbol{\Psi}_1) \quad (370)$$

$$\boldsymbol{\varepsilon}_2 \sim N(\mathbf{0}, \boldsymbol{\Psi}_2) \quad (371)$$

$$\mathbf{f}_1 \sim N(\mathbf{0}, \boldsymbol{\Phi}_1) \quad (372)$$

$$\mathbf{f}_2 = \boldsymbol{\Gamma}\mathbf{f}_1 + \boldsymbol{\zeta} \quad (373)$$

$$\boldsymbol{\zeta} \sim N(\mathbf{0}, \boldsymbol{\Sigma}_\zeta) \quad (374)$$

$$\text{Cov}(\mathbf{f}_1, \boldsymbol{\varepsilon}_1) = \mathbf{O}, \text{Cov}(\mathbf{f}_2, \boldsymbol{\varepsilon}_2) = \mathbf{O}, \text{Cov}(\mathbf{f}_2, \boldsymbol{\zeta}) = \mathbf{O} \quad (375)$$

$$\mathbf{y}_1 \sim N(\mathbf{0}, \boldsymbol{\Lambda}_1\boldsymbol{\Phi}_1\boldsymbol{\Lambda}_1' + \boldsymbol{\Psi}_1) \quad (376)$$

$$\mathbf{y}_2 \sim N(\mathbf{0}, \boldsymbol{\Lambda}_2(\boldsymbol{\Gamma}\boldsymbol{\Phi}_1\boldsymbol{\Gamma}' + \boldsymbol{\Sigma}_\zeta)\boldsymbol{\Lambda}_2' + \boldsymbol{\Psi}_2) \quad (377)$$

$$\text{cov}(\mathbf{y}_1, \mathbf{y}_2) = \boldsymbol{\Lambda}_1\boldsymbol{\Phi}_1\boldsymbol{\Gamma}'\boldsymbol{\Lambda}_2' \quad (378)$$

なお、このモデルは以下の様にまとめて記すことが出来る。

$$\mathbf{y} = \boldsymbol{\Lambda}\mathbf{u} \quad (379)$$

ただし

$$\mathbf{y} = \begin{pmatrix} \mathbf{y}_1 \\ \mathbf{y}_2 \end{pmatrix} \quad \text{and} \quad \boldsymbol{\Lambda} = \begin{bmatrix} \boldsymbol{\Lambda}_1 & \mathbf{O} & \mathbf{I} & \mathbf{O} \\ \boldsymbol{\Lambda}_2\boldsymbol{\Gamma} & \boldsymbol{\Lambda}_2 & \mathbf{O} & \mathbf{I} \end{bmatrix} \quad \text{and} \quad \mathbf{u} = \begin{pmatrix} \mathbf{f}_1 \\ \boldsymbol{\zeta} \\ \boldsymbol{\varepsilon}_1 \\ \boldsymbol{\varepsilon}_2 \end{pmatrix} \quad (380)$$

$$(381)$$

である。このとき、 \mathbf{u} の分散共分散行列を

$$\mathcal{D}(\mathbf{u}) = \begin{bmatrix} \boldsymbol{\Phi}_1 & \mathbf{O} & \mathbf{O} & \mathbf{O} \\ \mathbf{O} & \boldsymbol{\Sigma}_\zeta & \mathbf{O} & \mathbf{O} \\ \mathbf{O} & \mathbf{O} & \boldsymbol{\Psi}_1 & \mathbf{O} \\ \mathbf{O} & \mathbf{O} & \mathbf{O} & \boldsymbol{\Psi}_2 \end{bmatrix} \quad (382)$$

と置くと y の分散共分散行列は、公式より。

$$\mathcal{D}(y) = \begin{bmatrix} \mathcal{D}(y_1) & cov(y_1, y_2) \\ cov(y_2, y_1) & \mathcal{D}(y_2) \end{bmatrix} \quad (383)$$

$$= \Lambda \mathcal{D}(u) \Lambda' \quad (384)$$

$$= \begin{bmatrix} \Lambda_1 \Phi_1 \Lambda_1' + \Psi_1 & \Lambda_1 \Phi_1 (\Lambda_2 \Gamma)' \\ \Lambda_2 \Gamma \Phi_1 \Lambda_1' & \Lambda_2 \Sigma_\zeta \Lambda_2' + \Lambda_2 \Gamma \Phi_1 (\Lambda_2 \Gamma)' + \Psi_2 \end{bmatrix} \quad (385)$$

となる。すなわち、

$$\mathcal{D}(y_1) = \Lambda_1 \Phi_1 \Lambda_1' + \Psi_1 \quad (386)$$

$$\mathcal{D}(y_2) = \Lambda_2 \Sigma_\zeta \Lambda_2' + \Lambda_2 \Gamma \Phi_1 (\Lambda_2 \Gamma)' + \Psi_2 \quad (387)$$

$$cov(y_1, y_2) = \Lambda_1 \Phi_1 \Gamma' \Lambda_2' \quad (388)$$

である。

4.1.5 一般形 (RAM)

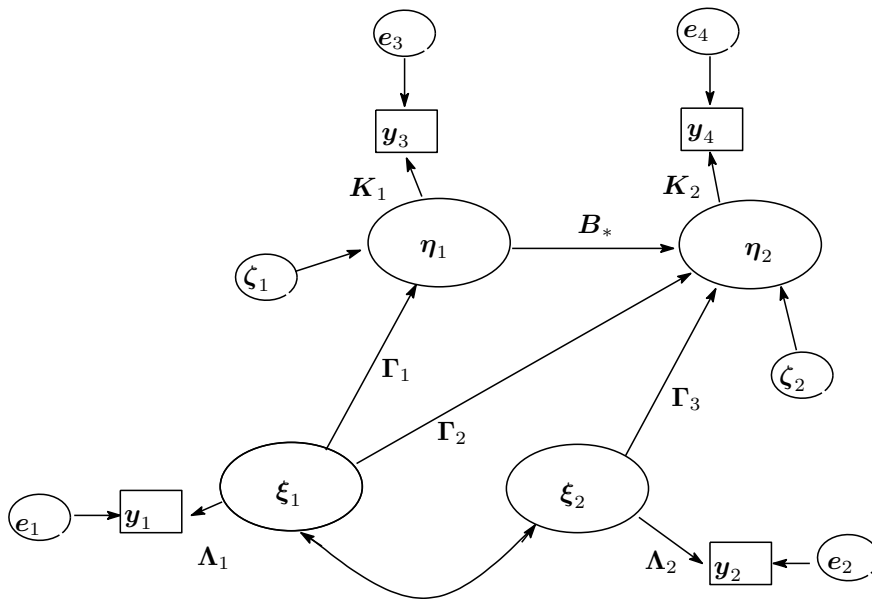


Figure1 Example of a SEM [Example1]

Figure 1 [Example1] の様な複雑なモデルを定式化することを考える。ただし、 y_1, y_2, \dots, y_4 は観測変数、 e_1, e_2, \dots, e_4 は観測変数に加わる誤差変数、 ξ_1, ξ_2 は外性的潜在変数、 η_1, η_2 は内性的潜在変数、また、 ζ_1, ζ_2 は内性的潜在変数に加わる誤差変数である。図中、各変数は 1 次元の変数の様に表現しているが、一般的にはベクトルであり、それぞれの変数のサイズを $size(\xi_1)$ 等で示す。したがって、係数行列のサイズ

は以下の様になる。

$$\Lambda_i \text{ size}(\mathbf{y}_i) \times \text{size}(\boldsymbol{\xi}_i), \quad i = 1, 2 \quad (389)$$

$$\mathbf{K}_i \text{ size}(\mathbf{y}_{i+2}) \times \text{size}(\boldsymbol{\eta}_i), \quad i = 1, 2 \quad (390)$$

$$\Gamma_i \text{ size}(\boldsymbol{\eta}_i) \times \text{size}(\boldsymbol{\xi}_1), \quad i = 1, 2 \quad (391)$$

$$\Gamma_3 \quad \text{size}(\boldsymbol{\eta}_2) \times \text{size}(\boldsymbol{\xi}_2) \quad (392)$$

$$\mathbf{B}_* \quad \text{size}(\boldsymbol{\eta}_2) \times \text{size}(\boldsymbol{\eta}_1) \quad (393)$$

$$(394)$$

なお、この図は以下の方程式を表している。

$$\mathbf{y}_1 = \Lambda_1 \boldsymbol{\xi}_1 + \mathbf{e}_1 \quad [\mathbf{y}1] \quad (395)$$

$$\mathbf{y}_2 = \Lambda_2 \boldsymbol{\xi}_2 + \mathbf{e}_2 \quad [\mathbf{y}2] \quad (396)$$

$$\mathbf{y}_3 = \mathbf{K}_1 \boldsymbol{\eta}_1 + \mathbf{e}_3 \quad [\mathbf{y}3] \quad (397)$$

$$\mathbf{y}_4 = \mathbf{K}_2 \boldsymbol{\eta}_2 + \mathbf{e}_4 \quad [\mathbf{y}4] \quad (398)$$

$$\boldsymbol{\eta}_1 = \Gamma_1 \boldsymbol{\xi}_1 + \boldsymbol{\zeta}_1 \quad [\text{eta}1] \quad (399)$$

$$\boldsymbol{\eta}_2 = \Gamma_2 \boldsymbol{\xi}_1 + \Gamma_3 \boldsymbol{\xi}_2 + \mathbf{B}_* \boldsymbol{\eta}_1 + \boldsymbol{\zeta}_2 \quad [\text{eta}2] \quad (400)$$

なお、上記の \mathbf{y} に関する方程式を測定方程式（観測変数と潜在変数の間の関係を表す方程式）、 $\boldsymbol{\eta}$ に関するものを構造方程式（潜在変数間の関係を表す方程式）と呼ぶ。（なお、構造方程式という名称は、より一般的にモデル全体を指し示す場合もあるので注意が必要である。）また、外性的潜在変数間の曲がった矢印に関しては、それが表す共分散を、後に示す Φ 行列の要素として指定することになる。

いま、それぞれの変数を縦に並べたベクトルを以下の様に定義する。

$$\mathbf{y} = \begin{pmatrix} \mathbf{y}_1 \\ \mathbf{y}_2 \\ \mathbf{y}_3 \\ \mathbf{y}_4 \end{pmatrix}, \quad \mathbf{e} = \begin{pmatrix} \mathbf{e}_1 \\ \mathbf{e}_2 \\ \mathbf{e}_3 \\ \mathbf{e}_4 \end{pmatrix}, \quad \boldsymbol{\xi} = \begin{pmatrix} \boldsymbol{\xi}_1 \\ \boldsymbol{\xi}_2 \end{pmatrix}, \quad \boldsymbol{\eta} = \begin{pmatrix} \boldsymbol{\eta}_1 \\ \boldsymbol{\eta}_2 \end{pmatrix}, \quad \boldsymbol{\zeta} = \begin{pmatrix} \boldsymbol{\zeta}_1 \\ \boldsymbol{\zeta}_2 \end{pmatrix} \quad (401)$$

また、 \mathbf{v} を観測変数と誤差以外の潜在変数を縦に並べたベクトル、 \mathbf{u} を外性変数を縦に並べたベクトルとする。

$$\mathbf{v} = \begin{pmatrix} \boldsymbol{\xi} \\ \boldsymbol{\eta} \\ \mathbf{y} \end{pmatrix} \quad \text{and} \quad \mathbf{u} = \begin{pmatrix} \boldsymbol{\xi} \\ \boldsymbol{\zeta} \\ \boldsymbol{\varepsilon} \end{pmatrix} \quad [\text{lse1vv}] \quad (402)$$

なお、これらの変数の順序は必ずしもこの順でなくてもかまわないが、ここでは \mathbf{v} をその因果の順に並べている。そうすることにより、以下に定義する \mathbf{B}^* 行列が下三角ブロック行列となるという利点がある。

この \mathbf{v} と \mathbf{u} を用いると、前述の全ての方程式は（ $\boldsymbol{\xi}$ を含めた形で）

$$\mathbf{v} = \mathbf{B}^* \mathbf{v} + \mathbf{u} \quad [\text{Model1}] \quad (403)$$

と書き表される。ただし、 \mathbf{B}^* 行列は

$$\mathbf{B}^* = \begin{bmatrix} \mathbf{O} & \mathbf{O} & \mathbf{O} \\ \Gamma & \mathbf{B}_{**} & \mathbf{O} \\ \Lambda & \mathbf{K} & \mathbf{O} \end{bmatrix} \quad (404)$$

とであり、 Λ , Γ , \mathbf{K} は先に示した添え字付きの係数行列を以下の様に配置したものであり、

$$\Lambda = \begin{bmatrix} \Lambda_1 & \mathbf{O} \\ \mathbf{O} & \Lambda_2 \\ \mathbf{O} & \mathbf{O} \\ \mathbf{O} & \mathbf{O} \end{bmatrix}, \quad \mathbf{K} = \begin{bmatrix} \mathbf{O} & \mathbf{O} \\ \mathbf{O} & \mathbf{O} \\ \mathbf{K}_1 & \mathbf{O} \\ \mathbf{O} & \mathbf{K}_2 \end{bmatrix}, \quad \Gamma = \begin{bmatrix} \Gamma_1 & \mathbf{O} \\ \Gamma_2 & \Gamma_3 \end{bmatrix} \quad (405)$$

それぞれのサイズは $size(\mathbf{y}) \times size(\boldsymbol{\xi})$, $size(\mathbf{y}) \times size(\boldsymbol{\eta})$, $size(\boldsymbol{\eta}) \times size(\boldsymbol{\xi})$ である。なお、内生的変数変数 $\boldsymbol{\eta}$ は、 $\boldsymbol{\xi}$ （ならびに付随する誤差）のみから影響を受ける部分 $\boldsymbol{\eta}_1$ と、 $\boldsymbol{\xi}$ ならびに $\boldsymbol{\eta}_1$ からの影響を受ける部分 $\boldsymbol{\eta}_2$ に分割されているが、 $\boldsymbol{\eta}_2$ に関する、

$$\boldsymbol{\eta}_2 = \boldsymbol{\Gamma}_2 \boldsymbol{\xi}_1 + \boldsymbol{\Gamma}_3 \boldsymbol{\xi}_2 + \mathbf{B}_* \boldsymbol{\eta}_1 + \boldsymbol{\zeta}_2 \quad (406)$$

という構造方程式は、 $size(\boldsymbol{\eta}_1) + size(\boldsymbol{\eta}_2)$ の正方行列 \mathbf{B}_{**} を

$$\mathbf{B}_{**} = \begin{bmatrix} \mathbf{O} & \mathbf{O} \\ \mathbf{B}_* & \mathbf{O} \end{bmatrix} \quad (407)$$

と置くことにより $\boldsymbol{\eta}$ 間の依存関係を

$$\mathbf{B}_{**} \boldsymbol{\eta} = \begin{pmatrix} \mathbf{0} \\ \mathbf{B}_* \boldsymbol{\eta}_1 \end{pmatrix} \quad (408)$$

と表現している。

これらの表現を使うと、Eq.(403) [Model11]

$$\begin{pmatrix} \boldsymbol{\xi} \\ \boldsymbol{\eta} \\ \mathbf{y} \end{pmatrix} = \begin{bmatrix} \mathbf{O} & \mathbf{O} & \mathbf{O} \\ \boldsymbol{\Gamma} & \mathbf{B}_{**} & \mathbf{O} \\ \boldsymbol{\Lambda} & \mathbf{K} & \mathbf{O} \end{bmatrix} \begin{pmatrix} \boldsymbol{\xi} \\ \boldsymbol{\eta} \\ \mathbf{y} \end{pmatrix} + \begin{pmatrix} \boldsymbol{\xi} \\ \boldsymbol{\zeta} \\ \boldsymbol{\varepsilon} \end{pmatrix} \quad (409)$$

が表す各方程式、すなわち

$$\boldsymbol{\xi} = \boldsymbol{\xi} \quad (410)$$

$$\boldsymbol{\eta} = \boldsymbol{\Gamma} \boldsymbol{\xi} + \mathbf{B}_{**} \boldsymbol{\eta} + \boldsymbol{\zeta} \quad (411)$$

$$\mathbf{y} = \boldsymbol{\Lambda} \boldsymbol{\xi} + \mathbf{K} \boldsymbol{\eta} + \boldsymbol{\varepsilon} \quad (412)$$

が、Eq.(395) [y1]-Eq.(400) [eta2] と同一のものであることが確かめられる。

(なお、 $\boldsymbol{\xi}$ に関しては、不自然ではあるが、これは以下の操作のためのトリックであると理解されたい。)

さて、これまでの準備の元に Eq.(403) [Model11] より $\mathbf{I} - \mathbf{B}^*$ 行列がフルランクである場合には

$$\mathbf{v} = (\mathbf{I} - \mathbf{B}^*)^{-1} \mathbf{u} \quad (413)$$

を得るため、 \mathbf{u} の要素である外性変数の誤差分散共分散行列を

$$\mathcal{D}(\mathbf{u}) = \begin{bmatrix} \boldsymbol{\Phi} & \mathbf{O} & \mathbf{O} \\ \mathbf{O} & \boldsymbol{\Delta} & \mathbf{O} \\ \mathbf{O} & \mathbf{O} & \boldsymbol{\Psi} \end{bmatrix} \quad (414)$$

と置くと、 \mathbf{v} の分散共分散行列、 $\boldsymbol{\Sigma}$ は、公式より、

$$\boldsymbol{\Sigma} = (\mathbf{I} - \mathbf{B}^*)^{-1} \mathcal{D}(\mathbf{u}) (\mathbf{I} - \mathbf{B}^*)'^{-1} \quad [\text{Disp1}] \quad (415)$$

となる。

より具体的には

$$\mathbf{I} - \mathbf{B}^* = \begin{pmatrix} (\mathbf{I}) & (\mathbf{O}) & (\mathbf{O}) \\ (-\boldsymbol{\Gamma}) & (\mathbf{I} - \mathbf{B}_{**}) & (\mathbf{O}) \\ (-\boldsymbol{\Lambda}) & (-\mathbf{K}) & (\mathbf{I}) \end{pmatrix} \quad (416)$$

ただし

$$\mathbf{I} - \mathbf{B}_{**} = \begin{bmatrix} \mathbf{I} & \mathbf{O} \\ -\mathbf{B}_* & \mathbf{I} \end{bmatrix} \quad (417)$$

であるが、ここで下三角ブロック行列の逆行列の公式

$$\begin{bmatrix} \mathbf{A}_{11} & \mathbf{O} & \mathbf{O} \\ \mathbf{A}_{21} & \mathbf{A}_{22} & \mathbf{O} \\ \mathbf{A}_{31} & \mathbf{A}_{32} & \mathbf{A}_{3,3} \end{bmatrix}^{-1} = \begin{bmatrix} \mathbf{A}_{11}^{-1} & \mathbf{O} & \mathbf{O} \\ -\mathbf{A}_{22}^{-1} \mathbf{A}_{21} \mathbf{A}_{11}^{-1} & \mathbf{A}_{22}^{-1} & \mathbf{O} \\ \mathbf{A}_{33}^{-1} \mathbf{A}_{32} \mathbf{A}_{22}^{-1} \mathbf{A}_{21} \mathbf{A}_{11}^{-1} - \mathbf{A}_{33}^{-1} \mathbf{A}_{31} \mathbf{A}_{11}^{-1} & -\mathbf{A}_{33}^{-1} \mathbf{A}_{32} \mathbf{A}_{22}^{-1} & \mathbf{A}_{33}^{-1} \end{bmatrix} \quad (418)$$

を用いると

$$(\mathbf{I} - \mathbf{B}^*)^{-1} = \begin{pmatrix} (\mathbf{I}) & (\mathbf{O}) & (\mathbf{O}) \\ ((\mathbf{I} - \mathbf{B}_{**})^{-1} \cdot \mathbf{\Gamma}) & ((\mathbf{I} - \mathbf{B}_{**})^{-1}) & (\mathbf{O}) \\ (\mathbf{\Lambda} + \mathbf{K} \cdot (\mathbf{I} - \mathbf{B}_{**})^{-1} \cdot \mathbf{\Gamma}) & (\mathbf{K} \cdot (\mathbf{I} - \mathbf{B}_{**})^{-1}) & (\mathbf{I}) \end{pmatrix} \quad (419)$$

を得る。したがって、 Σ 行列の各要素は

$$\Sigma_{11} : \mathcal{D}(\xi) = \Phi \quad (420)$$

$$\Sigma_{21} : cov(\eta, \xi) = (\mathbf{I} - \mathbf{B}_{**})^{-1} \cdot \mathbf{\Gamma} \cdot \Phi \quad (421)$$

$$\Sigma_{22} : \mathcal{D}(\eta) = (\mathbf{I} - \mathbf{B}_{**})^{-1} \cdot (\mathbf{\Delta} + \mathbf{\Gamma} \cdot \Phi \cdot \mathbf{\Gamma}^T) \cdot (\mathbf{I} - \mathbf{B}_{**})^{-1} \quad (422)$$

$$\Sigma_{31} : cov(\mathbf{y}, \xi) = (\mathbf{\Lambda} + \mathbf{K} \cdot (\mathbf{I} - \mathbf{B}_{**})^{-1} \cdot \mathbf{\Gamma}) \cdot \Phi \quad (423)$$

$$\begin{aligned} \Sigma_{32} : cov(\mathbf{y}, \eta) &= \mathbf{K} \cdot (\mathbf{I} - \mathbf{B}_{**})^{-1} \cdot \mathbf{\Delta} \cdot (\mathbf{I} - \mathbf{B}_{**})^{-1} \\ &+ (\mathbf{\Lambda} + \mathbf{K} \cdot (\mathbf{I} - \mathbf{B}_{**})^{-1} \cdot \mathbf{\Gamma}) \cdot \Phi \cdot \mathbf{\Gamma}^T \cdot (\mathbf{I} - \mathbf{B}_{**})^{-1} \end{aligned} \quad (424)$$

$$\begin{aligned} \Sigma_{33} : \mathcal{D}(\mathbf{y}) &= \Psi + (\mathbf{\Lambda} + \mathbf{K} \cdot (\mathbf{I} - \mathbf{B}_{**})^{-1} \cdot \mathbf{\Gamma}) \cdot \Phi \cdot (\mathbf{\Gamma}^T \cdot (\mathbf{I} - \mathbf{B}_{**})^{-1} \cdot \mathbf{K}^T + \mathbf{\Lambda}^T) \\ &+ \mathbf{K} \cdot (\mathbf{I} - \mathbf{B}_{**})^{-1} \cdot \mathbf{\Delta} \cdot (\mathbf{I} - \mathbf{B}_{**})^{-1} \cdot \mathbf{K}^T \end{aligned} \quad (425)$$

となることが分かる。

なお、下三角行列の公式から

$$(\mathbf{I} - \mathbf{B}_{**})^{-1} = \begin{bmatrix} \mathbf{I} & \mathbf{O} \\ -\mathbf{B}_* & \mathbf{I} \end{bmatrix}^{-1} = \begin{bmatrix} \mathbf{I} & \mathbf{O} \\ \mathbf{B}_* & \mathbf{I} \end{bmatrix} \quad [\text{invBss}] \quad (426)$$

であるため、上記の表現は逆行列を含まない形で書き表すことが出来る。

上記の様に $\mathcal{D}(\mathbf{y})$ の構造が示されたため、データが与えられた場合には、最小2乗法もしくは最尤解を用いてパラメタを推定することが出来る。

4.1.6 モデルの拡張

- ξ_1 から出る矢印で y_1 以外に入るもの。
→ Λ と K のゼロである要素を変更する。
- y_i から y_j に出る矢印がある場合。
→ B^* 行列の右下隅の O の部分を η 間の相互依存を表す B_{**} と同様の行列を用いて表現する。
- 観測される変数 y_0 が外生変数の場合。
→ それらの y_0 を v の中の ξ の上、もしくは η の上に配置し、 u の中にも配置する。
- 誤差変数間の共分散の取り扱い。
→ 対応する分散共分散行列の要素の指定。
- 再帰的モデルの取り扱い
→ B_* 行列、したがって B_{**} 行列が下三角にならないため Eq.(426) [invBss] が成り立たないがパラメタの推定は可能。

A 付録 [appendix]

1.1 偶に使う公式集

1: 偶に使う公式 [formulae]

以下の様に Σ と V が定義されているとする。

$$\Sigma = \Lambda\Phi\Lambda' + \Psi \quad (427)$$

$$V = (\Phi^{-1} + \Lambda'\Psi^{-1}\Lambda)^{-1} \quad (428)$$

このとき

$$V = \Phi - \Phi\Lambda'\Sigma^{-1}\Lambda\Phi \quad (429)$$

$$\Psi^{-1}\Lambda V = \Sigma^{-1}\Lambda\Phi \quad (430)$$

$$\Psi^{-1}\Lambda V\Lambda' = \Sigma^{-1}\Lambda\Phi\Lambda' \quad (431)$$

$$\Lambda V\Lambda' = \Psi\Sigma^{-1}\Lambda\Phi\Lambda' = \Lambda\Phi\Lambda'\Sigma^{-1}\Psi \quad (\text{sym}) \quad (432)$$

$$\Lambda\Phi\Lambda' = \Sigma\Psi^{-1}\Lambda V\Lambda' = \Lambda V\Lambda'\Psi^{-1}\Sigma \quad (\text{sym}) \quad (433)$$

$$\Psi\Sigma^{-1} = I - \Lambda\Phi\Lambda'\Sigma^{-1} = I - \Lambda V\Lambda'\Psi^{-1} \quad (434)$$

$$\Sigma^{-1} = \Psi^{-1} - \Psi^{-1}\Lambda(\Phi^{-1} + \Lambda'\Psi^{-1}\Lambda)^{-1}\Lambda'\Psi^{-1} \quad (435)$$

$$= \Psi^{-1} - \Psi^{-1}\Lambda V\Lambda'\Psi^{-1} \quad (436)$$

$$\Sigma^{-1}\Lambda = \Psi^{-1}\Lambda(\Phi^{-1} + \Lambda'\Psi^{-1}\Lambda)^{-1}\Phi^{-1} = \Psi^{-1}\Lambda V\Phi^{-1} \quad (437)$$

$$\Lambda'\Sigma^{-1}\Lambda = \Phi^{-1} - \Phi^{-1}(\Phi^{-1} + \Lambda'\Psi^{-1}\Lambda)^{-1}\Phi^{-1} \quad (438)$$

$$= \Phi^{-1} - \Phi^{-1}V\Phi^{-1} \quad (439)$$

$$\Phi\Lambda'\Sigma^{-1}\Lambda\Phi = \Phi - V = \Phi\Lambda'\Psi^{-1}\Lambda V \quad (440)$$

[$\Sigma^{-1}\Lambda$ の導出]

Eq.(193)[bif] で与えられる Σ^{-1} の表現を用いると

$$\Sigma^{-1}\Lambda = \left(\Psi^{-1} - \Psi^{-1}\Lambda(\Phi^{-1} + \Lambda'\Psi^{-1}\Lambda)^{-1}\Lambda'\Psi^{-1} \right) \Lambda \quad (441)$$

ここで $D = \Lambda'\Psi^{-1}\Lambda$ と置くと

$$\Sigma^{-1}\Lambda = \Psi^{-1}\Lambda - \Psi^{-1}\Lambda(\Phi^{-1} + D)^{-1}D \quad (442)$$

であるが、これに公式

$$(A + B)^{-1} = A^{-1}(A^{-1} + B^{-1})^{-1}B^{-1} \quad [\text{mSI1}] \quad (443)$$

を用いると

$$\Sigma^{-1}\Lambda = \Psi^{-1}\Lambda - \Psi^{-1}\Lambda \left(\Phi(\Phi + D^{-1})^{-1}D^{-1}D \right) \quad (444)$$

$$= \Psi^{-1}\Lambda - \Psi^{-1}\Lambda\Phi(\Phi + D^{-1})^{-1} \quad (445)$$

$$(446)$$

を得る。続いて、公式

$$(A + B)^{-1} = A^{-1} - A^{-1}(A^{-1} + B^{-1})^{-1}A^{-1} \quad [\text{mSI3}] \quad (447)$$

を用いると

$$\Sigma^{-1}\Lambda = \Psi^{-1}\Lambda - \Psi^{-1}\Lambda\Phi\left(\Phi^{-1} - \Phi^{-1}(\Phi^{-1} + D)^{-1}\Phi^{-1}\right) \quad (448)$$

$$= \Psi^{-1}\Lambda - \Psi^{-1}\Lambda + \Psi^{-1}\Lambda(\Phi^{-1} + D)^{-1}\Phi^{-1} \quad (449)$$

$$= \Psi^{-1}\Lambda(\Phi^{-1} + D)^{-1}\Phi^{-1} \quad (450)$$

となる。

別解

Eq.(193)[bif] で与えられる Σ^{-1} の表現を用いると

$$\Sigma^{-1}\Lambda = \left(\Psi^{-1} - \Psi^{-1}\Lambda(\Phi^{-1} + \Lambda'\Psi^{-1}\Lambda)^{-1}\Lambda'\Psi^{-1}\right)\Lambda \quad (451)$$

ここで $D = \Lambda'\Psi^{-1}\Lambda$ と置くと

$$\Sigma^{-1}\Lambda = \Psi^{-1}\Lambda - \Psi^{-1}\Lambda(\Phi^{-1} + D)^{-1}D \quad (452)$$

であるが、これに公式 4 から導かれる

$$(A + B)^{-1}B = I - B^{-1}(A^{-1} + B^{-1})^{-1} \quad (453)$$

を用いて変形していくと

$$\Sigma^{-1}\Lambda = \Psi^{-1}\Lambda - \Psi^{-1}\Lambda(I - D^{-1}(\Phi + D^{-1})^{-1}) \quad (454)$$

$$= \Psi^{-1}\Lambda D^{-1}(\Phi + D^{-1})^{-1} \quad (455)$$

$$= \Psi^{-1}\Lambda D^{-1} D D^{-1} (\Phi + D^{-1})^{-1} \quad (456)$$

$$= \Psi^{-1}\Lambda D^{-1} D D^{-1} (\Phi + D^{-1})^{-1} \quad (457)$$

$$= \Psi^{-1}\Lambda ((\Phi + D^{-1})D)^{-1} \quad (458)$$

$$= \Psi^{-1}\Lambda (\Phi D + I)^{-1} \quad (459)$$

$$= \Psi^{-1}\Lambda (\Phi(D + \Phi^{-1}))^{-1} \quad (460)$$

$$= \Psi^{-1}\Lambda (D + \Phi^{-1})^{-1}\Phi^{-1} \quad (461)$$

$$= \Psi^{-1}\Lambda V\Phi^{-1} \quad (462)$$

を得る。

1.2 逆行列の公式 [matrixinversionformulae]

[基本公式] $(A^{-1})^{-1} = A, (A')^{-1} = (A^{-1})', (AB)^{-1} = B^{-1}A^{-1}$
 $A(B+C)^{-1}D = (D^{-1}(B+C)A^{-1})^{-1} = (D^{-1}BA^{-1} + D^{-1}CA^{-1})^{-1}$

1. $(A+B)^{-1} = A^{-1}(A^{-1} + B^{-1})^{-1}B^{-1}$
2. $(A+B)^{-1} = B^{-1}(A^{-1} + B^{-1})^{-1}A^{-1}$
3. $(A+B)^{-1} = A^{-1} - A^{-1}(A^{-1} + B^{-1})^{-1}A^{-1} = A^{-1} \left(I - (A^{-1} + B^{-1})^{-1}A^{-1} \right)$
 $= \left(I - A^{-1}(A^{-1} + B^{-1})^{-1} \right) A^{-1}$
4. $(A+B)^{-1} = B^{-1} - B^{-1}(A^{-1} + B^{-1})^{-1}B^{-1} = B^{-1} \left(I - (A^{-1} + B^{-1})^{-1}B^{-1} \right)$
 $= \left(I - B^{-1}(A^{-1} + B^{-1})^{-1} \right) B^{-1}$
5. $(A+B)^{-1} = A^{-1} - A^{-1}B(A+B)^{-1} = A^{-1} \left(I - B(A+B)^{-1} \right)$
6. $(A+B)^{-1} = A^{-1} - (A+B)^{-1}BA^{-1} = \left(I - (A+B)^{-1}B \right) A^{-1}$
7. $(A+B)^{-1} = B^{-1} - B^{-1}A(A+B)^{-1} = B^{-1} \left(I - A(A+B)^{-1} \right)$
8. $(A+B)^{-1} = B^{-1} - (A+B)^{-1}AB^{-1} = \left(I - (A+B)^{-1}A \right) B^{-1}$
9. $(A+BCD)^{-1} = A^{-1} - A^{-1}B(C^{-1} + DA^{-1}B)^{-1}DA^{-1}$
10. $(A+BCD)^{-1}B = A^{-1}B(C^{-1} + DA^{-1}B)^{-1}C^{-1}$
11. $D(A+BCD)^{-1} = C^{-1}(C^{-1} + DA^{-1}B)^{-1}DA^{-1}$
12. $D(A+BCD)^{-1}B = DA^{-1}B(C^{-1} + DA^{-1}B)^{-1}C^{-1}$
 $= C^{-1}(C^{-1} + DA^{-1}B)^{-1}DA^{-1}B$
 $= (C + B^{-1}AD^{-1})^{-1} = C^{-1} - C^{-1}(C^{-1} + DA^{-1}B)^{-1}C^{-1}$

(註) A と B のみを含む公式においては、この両者は交換可能である。したがって、
 公式 1 と 2、公式 3 と 4、公式 5 と 7、そして公式 6 と 8 は同じことである。

公式 1-4 は A と B を A^{-1} と B^{-1} で置き換える。

公式 5-8 は $(A+B)^{-1}$ を残す。

公式 9-12 は A と C を A^{-1} と C^{-1} で置き換える。

公式 1-4 は、たとえば、

$$(A^{-1} + B^{-1})^{-1} = A(A+B)^{-1}B$$

という形で用いる場合が多い。

公式 3-8 は、 $A(A+B)^{-1}$ の様な形の式を単純化するとき利用される。

$$A(A+B)^{-1} = I - (A^{-1} + B^{-1})^{-1}A^{-1} \quad \text{公式 3}$$

$$A(A+B)^{-1} = I - B(A+B)^{-1} \quad \text{公式 5}$$

である。

[公式 12 の導出]

公式 9 の左から D 右から B を掛けると

$$D(A + BCD)^{-1}B = D(A^{-1} - A^{-1}B(C^{-1} + DA^{-1}B)^{-1}DA^{-1})B \quad (463)$$

$$= DA^{-1}B - DA^{-1}B(C^{-1} + DA^{-1}B)^{-1}DA^{-1}B \quad (464)$$

となるが、ここで改めて

$$A^* \leftarrow (DA^{-1}B)^{-1} \quad (465)$$

$$C^* \leftarrow C, \quad B^* \leftarrow I, \quad D^* \leftarrow I \quad (466)$$

と置いて Eq.(9) [mif9] を逆に使うと

$$D(A + BCD)^{-1}B = A^{*-1} - A^{*-1}(C^{*-1} + A^{*-1})^{-1}A^{*-1} \quad (467)$$

$$= (A^* + C^*)^{-1} \quad (468)$$

を得る。その後、Eq.(3) [mif3] を用いると

$$D(A + BCD)^{-1}B = C^{*-1} - C^{*-1}(C^{*-1} + A^{*-1})^{-1}C^{*-1} \quad (469)$$

$$= C^{-1} - C^{-1}(C^{-1} + DA^{-1}B)^{-1}C^{-1} \quad (470)$$

である。

別解

公式 9 の左から D 右から B を掛けたものに基本公式を使うと

$$D(A + BCD)^{-1}B = (B^{-1}AD^{-1} + B^{-1}BCDD^{-1})^{-1} \quad (471)$$

$$= (B^{-1}AD^{-1} + C)^{-1} \quad (472)$$

となるが、ここで公式 9 を使ったものに再度基本公式を使うと

$$(C + B^{-1}AD^{-1})^{-1} = C^{-1} - C^{-1}B^{-1}(A^{-1} + D^{-1}CB^{-1})^{-1}D^{-1}C^{-1} \quad (473)$$

$$= C^{-1} - C^{-1}(DA^{-1}B + C)^{-1}C^{-1} \quad (474)$$

を得る。

1.3 確率変数の線形結合 [linearcomb]

確率変数の線形結合

$$\mathbf{y} = \boldsymbol{\mu} + \mathbf{A}\mathbf{x} \quad (475)$$

$$\boldsymbol{\mu}, \mathbf{A} = \text{constant} \quad (476)$$

$$\mathbf{x} \sim N(\mathbf{m}_x, \boldsymbol{\Sigma}_x) \quad (477)$$

ならば

$$\mathbf{y} \sim N(\boldsymbol{\mu} + \mathbf{A}\mathbf{m}_x, \mathbf{A}\boldsymbol{\Sigma}_x\mathbf{A}') \quad (478)$$

系 1

$$\mathbf{x} = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}, \quad \text{and} \quad \boldsymbol{\Sigma} = \begin{bmatrix} \boldsymbol{\Sigma}_{11} & \boldsymbol{\Sigma}_{12} \\ \boldsymbol{\Sigma}_{21} & \boldsymbol{\Sigma}_{22} \end{bmatrix} \quad (479)$$

$$\mathbf{A} = \begin{bmatrix} \mathbf{A}_{11} & \mathbf{A}_{12} \\ \mathbf{A}_{21} & \mathbf{A}_{22} \end{bmatrix} \quad (480)$$

$$\mathbf{y} = \mathbf{A}\mathbf{x} = \begin{pmatrix} y_1 \\ y_2 \end{pmatrix} = \begin{bmatrix} \mathbf{A}_{11}x_1 + \mathbf{A}_{12}x_2 \\ \mathbf{A}_{21}x_1 + \mathbf{A}_{22}x_2 \end{bmatrix} \quad (481)$$

$$\text{Cov}(\mathbf{y}_1, \mathbf{y}_1) = \mathbf{A}_{11}\boldsymbol{\Sigma}_{11}\mathbf{A}'_{11} + \mathbf{A}_{12}\boldsymbol{\Sigma}_{21}\mathbf{A}'_{11} + \mathbf{A}_{11}\boldsymbol{\Sigma}_{12}\mathbf{A}'_{12} + \mathbf{A}_{12}\boldsymbol{\Sigma}_{22}\mathbf{A}'_{12} \quad (482)$$

$$\text{Cov}(\mathbf{y}_1, \mathbf{y}_2) = \mathbf{A}_{11}\boldsymbol{\Sigma}_{11}\mathbf{A}'_{21} + \mathbf{A}_{12}\boldsymbol{\Sigma}_{21}\mathbf{A}'_{21} + \mathbf{A}_{11}\boldsymbol{\Sigma}_{12}\mathbf{A}'_{22} + \mathbf{A}_{12}\boldsymbol{\Sigma}_{22}\mathbf{A}'_{22} \quad (483)$$

$$\text{Cov}(\mathbf{y}_2, \mathbf{y}_1) = \mathbf{A}_{21}\boldsymbol{\Sigma}_{11}\mathbf{A}'_{11} + \mathbf{A}_{22}\boldsymbol{\Sigma}_{21}\mathbf{A}'_{11} + \mathbf{A}_{21}\boldsymbol{\Sigma}_{12}\mathbf{A}'_{12} + \mathbf{A}_{22}\boldsymbol{\Sigma}_{22}\mathbf{A}'_{12} \quad (484)$$

$$\text{Cov}(\mathbf{y}_2, \mathbf{y}_2) = \mathbf{A}_{21}\boldsymbol{\Sigma}_{11}\mathbf{A}'_{21} + \mathbf{A}_{22}\boldsymbol{\Sigma}_{21}\mathbf{A}'_{21} + \mathbf{A}_{21}\boldsymbol{\Sigma}_{12}\mathbf{A}'_{22} + \mathbf{A}_{22}\boldsymbol{\Sigma}_{22}\mathbf{A}'_{22} \quad (485)$$

系 2

$$\mathbf{y} = \mathbf{x}_1 + \mathbf{x}_2 \quad (486)$$

$$\mathbf{x}_1 \sim N(\mathbf{m}_1, \boldsymbol{\Sigma}_1) \quad (487)$$

$$\mathbf{x}_2 \sim N(\mathbf{m}_2, \boldsymbol{\Sigma}_2) \quad (488)$$

$$\mathbf{x}_1 \text{ と } \mathbf{x}_2 \text{ は独立, } \text{Cov}(\mathbf{x}_1, \mathbf{x}_2) = \boldsymbol{\Sigma}_{12} = \boldsymbol{\Sigma}_{21} = \mathbf{O} \quad (489)$$

ならば

$$\mathbf{y} \sim N(\mathbf{m}_1 + \mathbf{m}_2, \boldsymbol{\Sigma}_1 + \boldsymbol{\Sigma}_2) \quad (490)$$

References

- Dempster, A.P., Laird, N.M., and Rubin, D.B. (1977) Maximum likelihood from incomplete data using the EM algorithm (with discussion). *J. R. Stat. Soc. B* 39(1), p1-38
- Lawley, D.N. (1940) The estimation of factor loadings by the method of maximum likelihood. *Proc. R. Soc. Edinb. A* 60, p64-82
- Lawley, D. N. and Maxwell, A. E. (1971) *Factor Analysis as a Statistical Method*. Second ed. London: Butterworths.
- Joereskog, Karl G. (1967) Some contributions to maximum likelihood factor analysis. *Psychometrika* vol. 32 p443-482
- Joereskog, Karl G. (1977) Factor analysis by least squares and maximum likelihood methods. In Enslein, K., Ralston, A., & Wilf, H.S. Eds. *Statistical Methods for Digital Computers*. Wiley, New York
- Rubin, D. B. and Thayer, D. T. (1982) EM algorithms for ML factor analysis. *Psychometrika* vol. 47 p69-76
- 柳井晴夫・繁榎算男・前川眞一・市川雅教 (1990) 因子分析 その理論と方法 朝倉書店
- Zhao, J.-H., Yu, P.L.H., & Jiang, Q. (2008) ML estimation for factor analysis:EM or non-EM? *Statistics and Computing*. vol. 18, p109-123.