

因子間の相関関係を保持させながら因子得点を推定する方法

モチベーション：

因子得点の推定は Thurstone(1947)の回帰法が一般的である。この最良の推定量には、因子軸回転の因子間相関とは一致しないという欠点がある。因子得点の推定方法には、独自性を最小化する方法など因子分析のモデルとその取り扱いの違いで、いくつかの流れがある（芝, 1969, 1972）。Anderson & Rubin(1956)による因子軸間の直交性を推定間にも保持させることを条件とした方法は、独自性の最小化に条件を挿入したものであった。回帰法に直交の条件をもたせた方法は、Shiba(1969)によって、提案されている。

延長因子分析のような応用的な利用では、構成する尺度の方向性が重要な尺度構成の成否のチェックポイントとなる。芝(1972)の数理的展開をベースとして、清水(1981)は、斜交の因子間相関を拘束条件とする推定方法を提案した。Amos などの SEM ソフトでは、回帰法の重み行列しか提供されず、応用利用においては、不満が残るからである。そして、同じ推定方法がオランダの研究者からも提案されていることを紹介するためでもある。

清水(1981)の概略：

使用する行列の定義

N ：被験者数

n ：観測変数の数

m ：因子の数

Z ：観測変数の標準得点 ($N \times n$) 次

R ：観測変数間の相関行列 ($n \times n$) 次

V ：直交因子解 ($n \times m$) 次

V_{fp} ：因子パターン行列 ($n \times m$) 次

V_{fs} ：因子構造行列 ($n \times m$) 次

T_f ：因子軸変換行列 ($m \times m$) 次、 $V_{fs} = VT_f$ そして $V_{fp} = V(T_f')^{-1}$

F ：因子得点行列 ($N \times m$) 次

C_f ：因子間相関行列 ($m \times m$) 次、 $C_f = N^{-1}Z'Z$ あるいは $C_f = T_f' T_f$

\hat{F} あるいは F_x ：因子得点の推定値行列 ($N \times m$) 次、添字の x は芝(1972)に従う。

W ：因子得点推定値を求めるための重み行列 ($n \times m$) 次、 $\hat{F} = ZW$

回帰法： $F_{13} = ZR^{-1}V_{fs}$

直交条件： $F_{14} = ZR^{-1}V(V'R^{-1}V)^{-\frac{1}{2}}$ 芝(1972, p.301)

斜交条件： $F_{24} = ZR^{-1}V_{fs}T_f'(T_fV_{fs}'R^{-1}V_{fs}T_f')^{-\frac{1}{2}}T_f$ 清水(1981)で新に 24 とした。

T_f は直交から斜交への因子軸変換行列であり、因子軸の回転から得られる。因子間の相

関行列を固有分解して $C_f = QAQ'$ とし、 $B = A^{\frac{1}{2}}Q'$ とすると、この行列 B を因子軸変換行列と置き換えても同じ推定値を得ることができる (清水(1981, p.119))。

本稿では新たに $C_f^{\frac{1}{2}} = QA^{\frac{1}{2}}Q'$ とおき、そして $V_{fs} = V_{fp}C_f$ との関係から因子パターン行列で上の推定式を表現し、簡単に展開してみることにする。

まず、因子得点の推定値行列を $F_{24} = ZWC_f^{\frac{1}{2}}$ と定義する。この推定値行列間の相関行列

は、 $W'RW = I$ の条件を満たせば、 $N^{-1}F'_{24}F_{24} = C_f^{\frac{1}{2}}W'RWC_f^{\frac{1}{2}} = C_f$ となる。次に、因子

得点と推定値との差の行列を $E = F - ZWC_f^{\frac{1}{2}}$ とし、ラグランジュの未定乗数からなる行列を M ($m \times m$) 次として、上の条件を挿入した W についての関数を次のように定義する。

$$f(W) = tr(N^{-1}F'F) - 2tr\left(N^{-1}F'ZWC_f^{\frac{1}{2}}\right) + tr\left(C_f^{\frac{1}{2}}W'RWC_f^{\frac{1}{2}}\right) - tr\{M(W'RW - I)\}$$

この関数を W に関して偏微分し、その結果を $\mathbf{0}$ とおくわけであるが、その前に、ここでは清水(1981)を変更して $N^{-1}F'Z = V_{fs} = V_{fp}C_f$ 因子パターン行列から表現しておくことにする。偏微分の結果の整理については、清水(1981, Pp. 117-118)を参照してもらうことにして、ここでは、推定式のみを示す。

$$\begin{aligned} F_{24} &= ZWC_f^{\frac{1}{2}} \\ &= ZR^{-1}V_{fp}C_fC_f^{\frac{1}{2}}\left(C_f^{\frac{1}{2}}C_fV_{fp}'R^{-1}V_{fp}C_fC_f^{\frac{1}{2}}\right)^{-\frac{1}{2}}C_f^{\frac{1}{2}} \\ &= ZR^{-1}V_{fp}C_f^{\frac{3}{2}}\left(C_f^{\frac{3}{2}}V_{fp}'R^{-1}V_{fp}C_f^{\frac{3}{2}}\right)^{-\frac{1}{2}}C_f^{\frac{1}{2}} \end{aligned}$$

ここで、上で因子間相関行列の逆行列から因子軸変換行列に相当する部分をおいたところを応用すると次のように展開することができるからである。

$$\begin{aligned} C_f^{\frac{3}{2}} &= C_f^{\frac{1}{2}} C_f^{\frac{1}{2}} C_f^{\frac{1}{2}} \\ &= Q A^{\frac{1}{2}} Q' Q A^{\frac{1}{2}} Q' Q A^{\frac{1}{2}} Q' \\ &= Q A^{\frac{1}{2}} A^{\frac{1}{2}} A^{\frac{1}{2}} Q' \\ &= Q A^{\frac{1}{2}} Q' \\ &= C_f^{\frac{1}{2}} \end{aligned}$$

この結果は F24 は次のようにも表すことができる。

$$F_{24} = Z R^{-1} V_{fp} C_f^{\frac{1}{2}} \left(C_f^{\frac{1}{2}} V_{fp}' R^{-1} V_{fp} C_f^{\frac{1}{2}} \right)^{-\frac{1}{2}} C_f^{\frac{1}{2}}$$

Beauducel(2007)などは、この式を使っている。

同じ結果を ten Berge, Krijinen, Wansbeek, & Shapiro (1999)は、Anderson & Rubin(1956)の流れの下で、Green(1969)、McDonald(1981)、ten Berge (1983)および Krijinen, Wansbeek, & ten Berge (1996)から導いている。彼らの式は、共分散行列を対象とするものであるが、ここでは、これまでの表記にあわせて簡単に紹介してみることにする。

彼らの重み行列は $W_{ten} = R^{-\frac{1}{2}} C C_f^{\frac{1}{2}}$ (ten Berge et al.(1999)式(3)より) である。この C ($n \times m$) 次の正規直交の行列を Green(1969)の結果から

$$C = R^{-\frac{1}{2}} V_{fp} C_f^{\frac{3}{2}} \left(C_f^{\frac{3}{2}} V_{fp}' R^{-1} V_{fp} C_f^{\frac{3}{2}} \right)^{-\frac{1}{2}}$$

としている(p.317)。

因子得点の比較研究をおこなっている Grice(2001)は、ten Berge et al.(1999)を紹介しているが、式(8)のカッコ内の相関行列を逆行列とする指示をミスしている。

引用文献

- Anderson, R. D., & Rubin, H. (1956). Statistical inference in factor analysis. *Proceedings of the Third Berkeley Symposium of Mathematical Statistics and Probability*, 5, 111-150.
- Beauducel, A. (2007). In spite of indeterminacy many common factor score estimates yield and identical reproduced covariance matrix. *Psychometrika*, 72, 437-441.
- Green, B. F. (1969). Best linear composites with a specific structure. *Psychometrika*, 34,

301-318.

Grice, J. W. (2001). Computing and evaluating factor scores. *Psychological Methods*, **6**, 430-450.

Krijnen, W. P., Wansbeek, T. J., & ten Berge, T. M. F. (1996). Best linear predictors for factor scores. *Communications in Statistics: Theory and Methods*, **25**, 3013-3025.

McDonald, R. P. (1981). Constrained least squares estimators of oblique common factors. *Psychometrika*, **46**, 337-341.

Shiba, S. (1969). New estimates of factor scores. *Japanese Psychological Research*, **11**, 129-133.

芝 祐順 (1972) 因子分析法 東京大学出版会 (第2版 1979)

清水和秋 (1981) 因子間相関を固定した斜交因子得点 関西大学社会学部紀要、**12**(2)、113-128

ten Berge, J. M. F. (1983). On Green's best linear composites with a specific structure, and oblique estimates of factor scores. *Psychometrika*, **48**, 371-375.

ten Berge, T. M. F., Krijnen, W. P., Wansbeek, T. J., & Shapiro, A. (1999). Some new results on correlation-preserving factor scores prediction methods. *Linear Algebra and its Applications*, **289**, 311-318.

付録

RによるF13とF24の計算のためのスクリプト

(標準形式での因子得点の推定)

メイン・スクリプト

```
#####
##因子得点の推定 (計算)
## メインの記述
## 1) データ入力
##     nv :観測変数の数
##     nf :因子の数 (/潜在変数の数)
##     X  :素点データ (SEMでは、標準得点ではなく素点で)
##     W  :因子得点の重み行列
##     n  :被験者数           <===
## 2) ファイル出力
##     F  : 因子得点
##     write.csv (outcF24, file="Fxx.csv", append = FALSE)
## step0 データ入力
##
# nv <- 15
# nf <- 5
# n  <- 585
in_data <- read.csv("data_for_R.csv",header=T)
X0      <- subset(in_data, select=c(N1:C3))
nv <- ncol(X0)
n0  <- nrow(X0)
```

```

print("number of participants")
print(n0)
X <- na.omit(X0)
print("nomissing participants")
n <- nrow(X)
print(n)
print("means of observed variables")
mean_v <- apply(X, 2, mean)
print(mean_v)
print("S.D.s of observed variables")
sd_v <- apply(X, 2, sd)
print(sd_v)
nnv <- n*nv
Z <- matrix(1:nnv,nrow=n, ncol=nv)
for (i in 1:n)
{
  for (j in 1:nv)
  {
    Z[i,j] <- ( X[i,j]-mean_v[j] )/ sd_v[j]
  }
}
apply(Z, 2, mean)
apply(Z, 2, sd)
R <- read.csv("correlation_matrix.csv", sep="," , header=T,row.names=1)
R <- as.matrix(R, nrow=nv, ncol=nv)
Vfp <- read.table("factor_pattern.csv", sep="," , header=T,row.names=1)
Vfp <- as.matrix(Vfp, nrow=nv, ncol=nf)
Cf <- read.csv("factor_correlation.csv",header=T,row.names=1)
Cf <- as.matrix(Cf, nrow=nf, ncol=nf)
nf <- ncol(Vfp)
print(nf)
print(Vfp)
print(Cf)
source("F13andF24.R")
F13andF24(nv, nf, R, Vfp, Cf, F24)
W <- read.csv("F24.csv",header=T, row.names=1,)
Z <- as.matrix(Z,nrow=n, ncol=nv)
W <- as.matrix(W,nrow=nv, ncol=nf)
## 因子得点行列の計算
F <- Z%%W
apply(F, 2, mean)
apply(F, 2, sd)
write.csv (F, file="factor_scores.csv", append = FALSE)
cor(F)

```

F13 と F24 を計算するための関数

```

F13andF24 <- function(nv, nf, R, Vfp, Cf, F24)
## weight matrix for factor scores
## F13 : regression method (Thurstone,1947)
## F13 = Z R-1 Vfs = Z R-1 Vfp Cf
##
## F24 : regression method
## with fixed correlations among estimates

```

```

##                               to correlations among factors (清水,1981)
##      F24 = Z R-1 Vfp Cf1/2 (Cf1/2 Vfp' R-1 Vfp Cf1/2)-1/2 Cf1/2
##
## input : correlation matrix among observed variables
##        : factor patter matrix
##        : factor correlation matrix
##
##        nv : number of variables
##        nf : number of factors
{
F13 <- solve(R)%*%Vfp)%*%Cf
print("F13:weight matrix by regression method (Thurstone,1947)")
print(F13)
print("correlation/covarinance matrix among estimates by F13")
outcF13 <- t(F13)%*%R)%*%F13
print(outcF13)
egv_Cf <- eigen(Cf)
D_v <- egv_Cf$values
Q <- egv_Cf$vectors
D_v_sqrt <- diag(sqrt(D_v))
Cf_Q_sqrt <- Q)%*%D_v_sqrt
Cf_half <- Cf_Q_sqrt)%*%t(Q)
#####因子間相関の C1/2 の行列 Cf_half
B <- Cf_half)%*%t(Vfp)%*%solve(R)%*%Vfp)%*%Cf_half
egv_B <- eigen(B)
B_v <- egv_B$values
B_Q <- egv_B$vectors
B_inv_sqrt <- diag(1/sqrt(B_v))
B_inv_sqrt <- B_Q)%*%B_inv_sqrt
B_inv_half <- B_inv_sqrt)%*%t(B_Q)
# B_inv_half)%*%B_inv_half
RVC <- solve(R)%*%Vfp)%*%Cf_half
F24 <- RVC)%*%B_inv_half)%*%Cf_half
print("F24:constrained weight matrix by regression method (清水,1981)")
print(F24)
print("correlation/covarinance matrix among estimates by F24")
outcF24 <- t(F24)%*%R)%*%F24
print(outcF24)
write.csv (F24, file="F24.csv", append = FALSE)
}

```