HIROSHIMA MATH. J.
22 (1992), 479–524

A study of an identification problem and substitute use of principal component analysis in factor analysis

Manabu SATO

(Received May 20, 1991)

Contents

0. Introduction

Factor Analysis (FA) is a branch of multivariate statistical analysis which is concerned with the internal relationships of a set of variables. Since Spearman [28] originated FA, it was developed by psychometricians. From 1940's, statisticians have been concerned with FA (see e.g. Lawley [15], Rao [20], Anderson and Rubin [3], Lawley and Maxwell [16], [17]). Factor analysis has been used in many fields of sciences in addition to psychology. Recently program packages applying FA have been developed. However, it may be noted that FA still involves some fundamental problems, and hence an investigation of it is very important.

In an FA model, we assume that an observed p-dimensional vector $\boldsymbol{x}$ follows

$$\boldsymbol{x} = \boldsymbol{\mu} + \Lambda \boldsymbol{f} + \boldsymbol{u}, \tag{0.1}$$

where $\boldsymbol{\mu}$ is a mean vector, Λ is a $p \times k$ $(p > k)$ factor loading matrix of rank k, $\boldsymbol{f}$ is a common factor vector and $\boldsymbol{u}$ is a unique factor vector. Further, suppose that $\mathrm{E}\{\boldsymbol{f}\} = \boldsymbol{0}$, $\mathrm{E}\{\boldsymbol{u}\} = \boldsymbol{0}$, $\mathrm{E}\{\boldsymbol{u}\boldsymbol{u}'\}$ is a diagonal matrix with positive diagonal elements, say Ψ, $\mathrm{E}\{\boldsymbol{f}\boldsymbol{u}'\} = O$ and $\mathrm{E}\{\boldsymbol{f}\boldsymbol{f}'\} = I$ (a unit matrix). Then, a variance-covariance matrix Σ of $\boldsymbol{x}$ can be decomposed as

$$\Sigma = \Lambda\Lambda' + \Psi. \tag{0.2}$$

Since the righthand side of (0.2) is a sum of a positive semidefinite matrix and a positive definite matrix, Σ is positive definite. Formula (0.2) is called *a fundamental equation of factor analysis.*

If a column of ΛG contains only one nonzero element for some nonsingular matrix G, a factor corresponding to this column is called *a specific factor.* If ΛG contains more than one nonzero element in every column for any nonsingular matrix G, Λ is called *a common factor matrix.*

When $k = 1$, it is called a *monofactor* case. This model is quite simple, however, it is useful in practice. In fact, in the analysis of empirical data, researchers often assume that the data have *a complete simple structure*; each row of Λ has only one nonzero element. This structure can be reduced to some sets of monofactor structure. For example, consider the case where Λ is of the following form after changing the order of rows suitably;

$$\Lambda = \begin{bmatrix} \lambda_{11} & \lambda_{21} & \lambda_{31} & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & \lambda_{42} & \lambda_{52} & \lambda_{62} & \lambda_{72} \end{bmatrix}'.$$

If we set

$$\boldsymbol{x}_1 = (x_1\ x_2\ x_3)',\ \boldsymbol{x}_2 = (x_4\ x_5\ x_6\ x_7)',\ \boldsymbol{\mu}_1 = (\mu_1\ \mu_2\ \mu_3)',$$

$$\boldsymbol{\mu}_2 = (\mu_4\ \mu_5\ \mu_6\ \mu_7)',\ \Lambda = \begin{bmatrix} \boldsymbol{\lambda}_1 & \boldsymbol{0} \\ \boldsymbol{0} & \boldsymbol{\lambda}_2 \end{bmatrix},\ \boldsymbol{f} = (f_1\ f_2)',$$

$$\boldsymbol{u}_1 = (u_1\ u_2\ u_3)' \text{ and } \boldsymbol{u}_2 = (u_4\ u_5\ u_6\ u_7)',$$

then,

$$\boldsymbol{x}_j = \boldsymbol{\mu}_j + \boldsymbol{\lambda}_j f_j + \boldsymbol{u}_j, \qquad j = 1, 2.$$

Namely, this structure is reduced to two sets of monofactor structure.

The present paper treats an identification problem of the FA model (Part I) and an adequacy problem of Principal Component Analysis as substitute use for FA (Part II).

Main inferential problems in the FA model are to estimate a number k of factors and matrices Λ and Ψ, based on samples of $\boldsymbol{x}$. However, before getting these estimates, we need to clear the identification problem which is divided into two parts:

(P1) the existence of a decomposition,

(P2) uniqueness of the decompositions.

In fact, if the existence of a decomposition is not guaranteed, the object of estimation is vague. Further, even if the decompositions exist, it is not clear which common (or unique) factors are estimated on the condition that the decompositions are not unique.

Problem (P1) has been recognized insufficiently. First we review the results which have been obtained hitherto. Next we study the region where the decompositions of Σ exist. For the case $p=3$ and $k=1$, its area is calculated. As for (P2), main sufficient conditions for uniqueness which have been obtained up to now are due to Anderson and Rubin [3] and Tumura and Sato [34]. For a review on (P2), see Shapiro [26] and some comments on his paper due to Sato [25]. In the present paper, we give necessary and/or sufficient conditions for their sufficient conditions, in the forms commonly met in practice. Using the results, it is seen that we can examine uniqueness easily. Further we propose the loading matrix whose *most elements are unique*. For such a loading matrix Λ, even if Λ is not unique, the estimates corresponding to the unique part is meaningful.

It is well known that Principal Component Analysis (PCA) and FA resemble each other but have rather different aims (Chap. 7 of Jolliffe [9]; Chap. 14 of Anderson [2]). However, PCA is very often used for the same purpose as FA without careful consideration. In fact, when PCA is applied, researchers calculate not only principal components but correlations between principal components and original variables (see e.g. §4.3.7 of Chatfield and Collins [6]). The correlations are called factor loadings. Using the (rotated) factor loadings, it is quite common to try to discern a latent structure. This is what is called substitute use of PCA for FA (see e.g. Chap. 3 of Okuno, Kume, Haga and Yoshizawa [19]).

One of the reasons why substitute use is often applied is that there exists a serious difficulty in estimating parameters in FA, that is, we quite often encounter an improper solution (Jöreskog [10]; Tumura, Fukutomi and Asoo [33]). Several ideas for overcoming the difficulty have been proposed (Jöreskog [10]; Koopman [14]; Martin and McDonald [18]; Akaike [1]). Some causes of improper solutions have been investigated (van Driel [38]; Tumura and Sato [35], [36], [37]). A method (Sato [23]) of overcoming them, which works well for many sets of empirical data, has been proposed. However, the difficulty in the estimation problem has not been solved completely. As a result, PCA is quite often used for the same purpose as FA. Of course there are several advantages of FA as compared with PCA. First, FA admits a formal statistical model, and hence factor loadings are estimated, considering the effects of error variances. In contrast, PCA does not have such a structural model. Second, the FA model has a property of *scale invariance*. Consequently, if we use an estimation method with scale invariance (for example, the maximum likelihood method, and the generalized least-squares method with a weight matrix S^{-1} or $\{\text{diag } S\}^{-1}$ where S is a sample variance-covariance matrix) and its solution is determined uniquely, then the estimates are *scale equivariant* (see e.g. Chap. 14 of Anderson [2]). This means the following: if we obtain an estimate $\hat{\Lambda}$ based on $\boldsymbol{x}$, then an estimate based on $C\boldsymbol{x}$ is $C\hat{\Lambda}$ where C is any nonsingular diagonal matrix. As a result, we can ignore measurement units of observations. On the other hand, the loadings calculated with PCA do not have such a property. From these viewpoints, it is important to examine whether PCA as substitute use for FA is adequate or not.

Part I consists of Sections 1 to 3. In Section 1, the identification problem is described in detail. In Section 2, the existence of a decomposition is discussed. In Section 3, uniqueness of the decompositions is discussed. Part II consists of Sections 4 to 6. In Section 4, an approach of investigating PCA as a substitute for FA is introduced. In Section 5, monofactor cases ($k = 1$) are treated. Finally, in Section 6, multifactor cases ($k \geq 2$) are treated.

Part I. Identification problem

1. Preliminary

The identification problems (P1) and (P2) may be stated as follows:
(P1) For any p-order positive definite symmetric matrix Σ, can it be decomposed as

$$\Sigma = \Lambda_k \Lambda_k' + \Psi_k,$$

where Λ_k is a $p \times k$ real matrix of rank k and Ψ_k is a diagonal matrix with positive diagonal elements, for assumed $k(< p)$?
(P2) If a decomposition exists, is it unique?

The subscript k of Λ_k indicates the number of columns of Λ_k and the subscript k of Ψ_k means that Ψ_k depends on Λ_k; for the sake of simplicity, either or both of the subscripts are sometimes omitted in the following text.

Before we discuss the problems in detail, we take two notes. First, the decomposition may be discussed in the term of a correlation matrix $P = (\text{diag } \Sigma)^{-1/2} \Sigma (\text{diag } \Sigma)^{-1/2} = (\rho_{ij})$ instead of $\Sigma = (\sigma_{ij})$; because structure (0.2) is equivalent to

$$\begin{aligned} P = &\{(\text{diag } \Sigma)^{-1/2} \Lambda\}\{(\text{diag } \Sigma)^{-1/2} \Lambda\}' \\ &+ (\text{diag } \Sigma)^{-1/2} \Psi (\text{diag } \Sigma)^{-1/2}. \end{aligned}$$

Therefore, we may deal with the decomposition of either Σ or P. Second, there exists an indeterminacy of a rotation of a factor loading matrix; multiplication on the right side of Λ by an orthogonal matrix, since

$$\Lambda\Lambda' + \Psi = (\Lambda T)(\Lambda T)' + \Psi,$$

where T is an arbitrary k-order orthogonal matrix. We ignore this indeterminacy in the following.

2. Existence of a decomposition

The following proposition treats the existence problem of a decomposition when factor size is increased.

PROPOSITION 2.1. (Reirsøl [22]) *If there exists a decomposition for factor size k, then there exist infinitely many decompositions for $k+1$.* □

PROOF. A loading matrix Λ_{k+1} for factor size $k+1$ can be constructed as follows; Suppose

$$\Lambda_{k+1} = [\Lambda_k\ \boldsymbol{\gamma}],$$

where $\boldsymbol{\gamma} = (0 \cdots 0\ \ \gamma\ \ 0 \cdots 0)'$, γ is the ith component of $\boldsymbol{\gamma}$ and $0 < \gamma^2 < \psi_i$. Without loss of generality, we may assume that rank $\Lambda_{k+1} = k+1$. Then, we have

$$\begin{aligned}\Sigma &= \Lambda_k\Lambda_k' + \operatorname{diag}\{\psi_1 \cdots \psi_p\}\\ &= \Lambda_{k+1}\Lambda_{k+1}' + \operatorname{diag}\{\psi_1 \cdots \psi_{i-1}\quad \psi_i - \gamma^2\quad \psi_{i+1} \cdots \psi_p\}.\end{aligned}$$

Consequently, there exist infinitely many decompositions for $k+1$ since we can take any γ such that $0 < \gamma^2 < \psi_i$. □

REMARK. In the proof of Reirsøl [22], the form of Λ_{k+1} is not apparent, however, the above proof shows it explicitly.

PROPOSITION 2.2. *When $k = p-1$, there exist infinitely many decompositions for any Σ.* □

PROOF. Let θ_p be the smallest eigenvalue of Σ. Set

$$\Sigma^* = \Sigma - \operatorname{diag}\{\varepsilon_1 \cdots \varepsilon_p\},$$

where $0 < \varepsilon_i < \theta_p$ $(i = 1,\ldots,p)$, then, Σ^* is a positive definite symmetric matrix since

$$\Sigma^* = \Sigma - \theta_p I + \operatorname{diag}\{\theta_p - \varepsilon_1\quad \cdots\quad \theta_p - \varepsilon_p\}.$$

Let L be a $p \times p$ lower triangular matrix (Cholesky decomposition) such that

$$\Sigma^* = LL',$$

and let L partition as

$$L = \begin{bmatrix} & 0 \\ \Lambda_{p-1} & \vdots \\ & 0 \\ & d \end{bmatrix}.$$

Then, rank $\Lambda_{p-1} = p - 1$, $d \neq 0$ and

$$\Sigma^* = \Lambda_{p-1}\Lambda'_{p-1} + \operatorname{diag}\{0 \cdots 0 \ d^2\},$$

hence

$$\Sigma = \Lambda_{p-1}\Lambda'_{p-1} + \operatorname{diag}\{\varepsilon_1 \ \cdots \ \varepsilon_{p-1} \ \varepsilon_p + d^2\}.$$

The matrix Λ_{p-1} depends on ε_i and we can choose ε_i arbitrarily under $0 < \varepsilon_i < \theta_p$ $(i = 1,\dots,p)$. Consequently, there exist infinitely many decompositions for $k = p - 1$. □

REMARK. Guttman [7] has given *one* decomposition for $k = p - 1$ under the assumption that the smallest eigenvalue of Σ is simple. As a special case of Proposition 2.2 we obtain that for the case $p = 2$ and $k = 1$, there exist infinitely many decompositions.

PROPOSITION 2.3. (Theorem 5 of Bekker and Leeuw [5]) *There exists no decomposition for $k < p - 1$ if and only if all elements of Σ^{-1} are positive, possibly after sign changes of rows and corresponding columns.* □

REMARK. Guttman [7] has shown that a tridiagonal matrix with nonzero subdiagonal elements *has no decomposition for $k < p - 1$.*

PROPOSITION 2.4. (Theorem 1 of Bekker and Leeuw [5]) *For $p \geq 4$ and $k = 1$, a decomposition exists if and only if, after sign changes of rows and corresponding columns, all elements of $\Sigma = (\sigma_{ab})$ are positive and*

$$\sigma_{ih}\sigma_{jl} - \sigma_{il}\sigma_{jh} = 0 \ \textit{and}$$

$$\sigma_{ih}\sigma_{ji} - \sigma_{ii}\sigma_{jh} < 0 \ (i \neq j,\ h,\ l;\ j \neq h,\ l;\ h \neq l). \quad \square$$

PROPOSITION 2.5. *For the case $p = 3$ and $k = 1$, the following* (1)–(3) *hold*:

(1) *If the following four inequalities*

$$\rho_{21}\rho_{31}\rho_{32} > 0,\ \rho_{21}\rho_{31}/\rho_{32} < 1,\ \rho_{21}\rho_{32}/\rho_{31} < 1 \ \textit{and} \ \rho_{31}\rho_{32}/\rho_{21} < 1$$

are satisfied, there exists a unique decomposition with

$$\boldsymbol{\lambda} = (\ (\operatorname{sgn} \rho_{32})\sqrt{(\rho_{21}\rho_{31}/\rho_{32})} \quad (\operatorname{sgn} \rho_{31})\sqrt{(\rho_{21}\rho_{32}/\rho_{31})} \quad (\operatorname{sgn} \rho_{21})\sqrt{(\rho_{31}\rho_{32}/\rho_{21})}\ )'.$$

(2) *If two or three of ρ_{ij}'s $(i > j)$ equal* 0, *there exist infinitely many decompositions.*
(3) *Otherwise, there is no decomposition.* □

PROOF. From the identity

$$P = \boldsymbol{\lambda}\boldsymbol{\lambda}' + \Psi,$$

we obtain

$$\rho_{21} = \lambda_2\lambda_1, \ \rho_{31} = \lambda_3\lambda_1 \text{ and } \rho_{32} = \lambda_3\lambda_2,$$

where $P = (\rho_{ij})$ and $\boldsymbol{\lambda} = (\lambda_1\ \lambda_2\ \lambda_3)'$. Therefore, using these equations, we have the following:
(i) If one of the elements ρ_{ij}'s $(i > j)$ equals 0, there is no decomposition.
(ii) If two or three of ρ_{ij}'s $(i > j)$ equal 0, there exist infinitely many decompositions.
(iii) If $\rho_{21}\rho_{31}\rho_{32} < 0$, there is no decomposition.
(iv) If $\rho_{21}\rho_{31}\rho_{32} > 0$, the above equations yield

$$\boldsymbol{\lambda} = \pm (\ (\operatorname{sgn} \rho_{32})\sqrt{(\rho_{21}\rho_{31}/\rho_{32})} \quad (\operatorname{sgn} \rho_{31})\sqrt{(\rho_{21}\rho_{32}/\rho_{31})} \quad (\operatorname{sgn} \rho_{21})\sqrt{(\rho_{31}\rho_{32}/\rho_{21})}\ )'.$$

If the following three conditions

$$\rho_{21}\rho_{31}/\rho_{32} < 1 \tag{2.1}$$

$$\rho_{21}\rho_{32}/\rho_{31} < 1 \tag{2.2}$$

$$\rho_{31}\rho_{32}/\rho_{21} < 1 \tag{2.3}$$

are satisfied, $\Psi = \operatorname{diag}(I - \boldsymbol{\lambda}\boldsymbol{\lambda}')$ is positive definite; consequently, there exists a unique decomposition. Otherwise, Ψ is not positive definite and consequently, there is no decomposition.

Summarizing above results (i)–(iv), we obtain results (1)–(3). □

Now we investigate more precisely the case where there exists a unique decomposition in the case $p = 3$ and $k = 1$. This case is very simple, however, its investigation is useful in practice. Because it is fundamental for a complete simple structure. First, we will consider the region where P is positive definite. Since $\rho_{11} = 1 > 0$, positive definiteness of P is equivalent to

$$\begin{aligned} &-1 < \rho_{21} < 1 \text{ and} \\ &\det P > 0. \end{aligned} \tag{2.4}$$

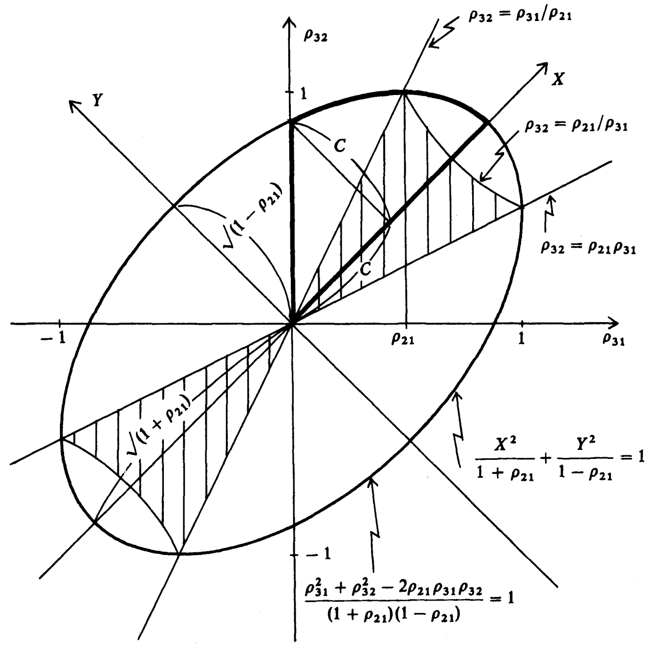

Fig. 2.1. Region where the decomposition exists uniquely (shaded portion) and region where P is positive definite (inside the ellipse)

Inequality (2.4) becomes

$$\frac{(\rho_{31}+\rho_{32})^2}{2(1+\rho_{21})} + \frac{(\rho_{31}-\rho_{32})^2}{2(1-\rho_{21})} < 1 \text{ or}$$

$$(\rho_{31}^2 + \rho_{32}^2 - 2\rho_{21}\rho_{31}\rho_{32})/\{(1+\rho_{21})(1-\rho_{21})\} < 1. \tag{2.5}$$

Since there are three variables in (2.5), we fix ρ_{21} and regard the lefthand side of (2.5) as a function of two variables ρ_{31} and ρ_{32}. Let the coordinate axes rotate and let the current axes denote X (which is direction of the major axis) and Y (which is direction of the minor axis). Then, the region where P is positive definite is given by

$$X^2/(1+\rho_{21}) + Y^2/(1-\rho_{21}) < 1,$$

(see Fig. 2.1). From (2.1)–(2.3), the region where the decomposition exists uniquely in the first quadrant is given by

$$\rho_{32} > \rho_{21}\rho_{31},\ \rho_{32} < \rho_{31}/\rho_{21} \text{ and } \rho_{32} < \rho_{21}/\rho_{31}.$$

Similar inequalities hold in the third quadrant. The shaded portion of Fig. 2.1

shows the region where the decomposition exists uniquely.

The area S_E of the ellipse in Fig. 2.1, that is, the region where P is positive definite, is

$$S_E = \pi\sqrt{(1-\rho_{21}^2)}.$$

On the other hand, the area S_P of the shaded portion in Fig. 2.1, that is, the region where the decomposition exists uniquely, is

$$S_P = -2|\rho_{21}|\ln|\rho_{21}|.$$

Because, if $\rho_{21} > 0$, the area in the first quadrant is given by

$$\begin{aligned} S_P/2 &= \int_0^{\rho_{21}} (\rho_{31}/\rho_{21} - \rho_{21}\rho_{31})\,\mathrm{d}\rho_{31} + \int_{\rho_{21}}^1 (\rho_{21}/\rho_{31} - \rho_{21}\rho_{31})\,\mathrm{d}\rho_{31} \\ &= -\rho_{21}\ln|\rho_{21}|. \end{aligned}$$

If $\rho_{21} < 0$, we obtain similarly

$$S_P/2 = \rho_{21}\ln|\rho_{21}|.$$

Next we consider the area S_R of the ellipse where the elements of $\boldsymbol{\lambda}$ are real numbers, under ρ_{21} being fixed. The area of the shape surrounded by the bold line in Fig. 2.1, which is equal to $S_R/4$, is

$$c^2/2 + \frac{b}{a}\int_c^a \sqrt{(a^2 - X^2)}\,\mathrm{d}X,$$

where $a^2 = 1 + \rho_{21}^2$, $b^2 = 1 - \rho_{21}^2$ and $c = ab/\sqrt{(a^2+b^2)}$. Using the formula

$$\int \sqrt{(d^2 - X^2)}\,\mathrm{d}X = (X\sqrt{(d^2 - X^2)} + d^2 \arcsin(X/d))/2 \qquad \text{for } d > 0,$$

we have

$$S_R = \sqrt{(1-\rho_{21}^2)}(\pi - 2\arcsin\sqrt{((1-\rho_{21})/2)}).$$

Table 2.1 presents S_P, S_R, S_E, S_P/S_R and S_P/S_E for $|\rho_{21}| = .05(.05).95$. We note that the ratio S_P/S_E is not large, at most .26.

Finally, we obtain the area of the region where P is positive definite and the region where the decomposition exists uniquely. These area are obtained by integrating S_P and S_E with respect to ρ_{21} from -1 to 1:

$$\int_{-1}^1 S_P\,\mathrm{d}\rho_{21} = 1 \quad \text{and} \quad \int_{-1}^1 S_E\,\mathrm{d}\rho_{21} = \frac{\pi^2}{2}.$$

Therefore, the ratio is given by

Table 2.1. Existence of the unique decomposition

$\lvert\rho_{21}\rvert$	S_P	S_R	S_E	S_P/S_R	S_P/S_E
.05	.300	1.619	3.138	.185	.095
.10	.461	1.663	3.126	.277	.147
.15	.569	1.702	3.106	.334	.183
.20	.644	1.736	3.078	.371	.209
.25	.693	1.766	3.042	.393	.228
.30	.722	1.789	2.997	.404	.241
.35	.735	1.806	2.943	.407	.250
.40	.733	1.817	2.879	.403	.255
.45	.719	1.820	2.806	.395	.256
.50	.693	1.814	2.721	.382	.255
.55	.658	1.798	2.624	.366	.251
.60	.613	1.771	2.513	.346	.244
.65	.560	1.731	2.387	.323	.235
.70	.499	1.676	2.244	.298	.223
.75	.432	1.600	2.078	.270	.208
.80	.357	1.499	1.885	.238	.189
.85	.276	1.363	1.655	.203	.167
.90	.190	1.173	1.369	.162	.138
.95	.097	.882	.981	.111	.099

S_P: area of the region where the decomposition exists uniquely
S_R: area of the region where the elements of λ are real numbers
S_E: area of the region where P is positive definite

$$\int_{-1}^{1} S_P \, d\rho_{21} \Big/ \int_{-1}^{1} S_E \, d\rho_{21} = 2/\pi^2$$
$$\fallingdotseq .203.$$

Summarizing this result, we obtain the following Proposition.

PROPOSITION 2.6. *For the case $p=3$ and $k=1$, if ρ_{ij} $(i>j)$ are independently uniformly distributed, the probability that the decomposition exists uniquely is $2/\pi^2$.* $\square$

For a sample case, we consider the estimate $\hat{\lambda}$ obtained by replacing P by R where $R=(r_{ij})$ is a sample correlation matrix. If $r_{21}r_{31}r_{32}>0$, the elements of $\hat{\lambda}$ are real numbers;

$$\hat{\lambda} = \pm\,(\ (\operatorname{sgn} r_{32})\sqrt{(r_{21}r_{31}/r_{32})}\quad (\operatorname{sgn} r_{31})\sqrt{(r_{21}r_{32}/r_{31})}\quad (\operatorname{sgn} r_{21})\sqrt{(r_{31}r_{32}/r_{21})}\ )'.$$

Now we will obtain $\Pr(r_{21}r_{31}r_{32}>0)$ based on samples for given P. Konishi ([12], [13]) has obtained an asymptotic expansion for the distribution of an analytic function of r_{ij}, based on a sample of size n from a multivariate normal distribution. Using his result (Theorem 6.2 of [13]), we see

$$\begin{aligned}\Pr(\sqrt{n}(r_{21}r_{31}r_{32} - \rho_{21}\rho_{31}\rho_{32})/\tau < u)\\ = \Phi(u) - \frac{1}{\sqrt{n}}(a_1\Phi^{(1)}(u)/(2\tau) + a_3\Phi^{(3)}(u)/(2\tau)^3)\\ + \frac{1}{n}\sum_{j=1}^{3} b_{2j}\Phi^{(2j)}(u)/(2\tau)^{2j} + O(n^{-3/2}),\end{aligned}$$

where $\tau^2 = \rho_{21}^2\rho_{31}^2 + \rho_{31}^2\rho_{32}^2 + \rho_{32}^2\rho_{21}^2$
$+ 2\rho_{21}\rho_{31}\rho_{32}(\rho_{21}^2(1 - 2\rho_{31}^2) + \rho_{31}^2(1 - 2\rho_{32}^2) + \rho_{32}^2(1 - 2\rho_{21}^2))$
$+ \rho_{21}^2\rho_{31}^2\rho_{32}^2(4\rho_{21}^2 + 4\rho_{31}^2 + 4\rho_{32}^2 - 9)$,

$\Phi^{(h)}(u)$ denotes the hth derivative of the standard normal distribution function of $\Phi(u)$. The coefficients are

$$\begin{aligned}a_1 &= 2\{\rho_{21}\rho_{31}\rho_{32}(2\rho_{21}^2 + 2\rho_{31}^2 + 2\rho_{32}^2 - 3)\\ &\quad + \rho_{21}^2(1 - 2\rho_{31}^2) + \rho_{31}^2(1 - 2\rho_{32}^2) + \rho_{32}^2(1 - 2\rho_{21}^2)\},\\ a_3 &= \textstyle\sum_{i\neq j}\{\rho_{ij}(3d_{ii} + d_{jj}) - 4d_{ij}\}d_{ii}f_{ij}\\ &\quad + \textstyle\sum_{i\neq j}\sum_{k\neq \ell}(d_{ij} - \rho_{ij}d_{ii})(d_{k\ell} - \rho_{k\ell}d_{kk})f_{ijk\ell}\\ &\quad + \frac{4}{3}\textstyle\sum_{i\neq j}\sum_{k\neq \ell}\sum_{q\neq r}\rho_{ir\cdot q}\rho_{jk\cdot i}\rho_{\ell q\cdot k}f_{ij}f_{k\ell}f_{qr},\end{aligned}$$

$$\begin{aligned}b_2 &= a_1^2/2 + \textstyle\sum_{i\neq j}(1 - 3\rho_{ij}^2)(\rho_{ij}d_{ii} - d_{ij})f_{ij}\\ &\quad + \frac{1}{4}\textstyle\sum_{i\neq j}\sum_{k\neq \ell}\rho_{ik\cdot j}\{\rho_{j\ell}(\rho_{ik}^2 + 3\rho_{j\ell}^2 + 12\rho_{jk}^2)\\ &\qquad - \rho_{jk}\rho_{k\ell}(\rho_{i\ell}^2 + 6\rho_{j\ell}^2 + 9\rho_{jk}^2)\}f_{ij}f_{k\ell}\\ &\quad + \textstyle\sum_{i\neq j}\sum_{k\neq \ell}\Bigg(\left\{\frac{1}{2}\rho_{ij}\rho_{k\ell}(1 - \rho_{k\ell}^2) + 3\rho_{ij}\rho_{ik}\rho_{i\ell\cdot k} - \rho_{j\ell\cdot k}(2\rho_{ik} - \rho_{ij}\rho_{jk})\right\}d_{ii}\\ &\qquad - \left\{\frac{1}{2}\rho_{k\ell}(1-\rho_{k\ell}^2) + 2\rho_{ik}\rho_{i\ell\cdot k}\right\}d_{ij} + 2(\rho_{j\ell\cdot i} - \rho_{ij}\rho_{i\ell\cdot k})d_{ik}\Bigg)f_{ijk\ell}\\ &\quad + \frac{1}{2}\textstyle\sum_{i\neq j}\sum_{k\neq \ell}\sum_{q\neq r}(\rho_{kq}(\rho_{\ell r\cdot k} - \rho_{k\ell}\rho_{kr\cdot q})(d_{ij} - \rho_{ij}d_{ii})f_{ijk\ell qr}\\ &\qquad + \rho_{kq\cdot r}\{\rho_{iq}\rho_{kr}\rho_{k\ell}\rho_{jq\cdot i} - \rho_{ir}\rho_{\ell r}\rho_{jr\cdot i}\\ &\qquad\quad + \rho_{\ell r\cdot k}(3\rho_{ij}\rho_{ir}^2 - \rho_{iq}\rho_{jq} - 2\rho_{ir}\rho_{jr})\}f_{ijk\ell}f_{qr})\\ &\quad + \frac{1}{4}\textstyle\sum_{i\neq j}\sum_{k\neq \ell}\sum_{q\neq r}\sum_{s\neq t}\rho_{ir\cdot q}\rho_{jq\cdot i}\rho_{kt\cdot s}\rho_{s\ell\cdot k}f_{ijk\ell}f_{qrst},\end{aligned}$$

$$
\begin{aligned}
b_4 = {} & a_1 a_3 + \\
& \textstyle\sum_{i \neq j} \{2(3d_{ii} + d_{jj})(2\rho_{ij} d_{ii}^* + d_{ij} d_{ii}) - 8(d_{ij} d_{ii}^* + d_{ii} d_{ij}^*) - \rho_{ij}(5d_{ii} + 3d_{jj}) d_{ii}^2\} f_{ij} \\
& \quad + 2\textstyle\sum_{i \neq j} \sum_{k \neq \ell} \Bigg(\Bigg\{ \rho_{j\ell} \rho_{ik \cdot j} + \rho_{jk} \rho_{i\ell \cdot j} - 3\rho_{ij}(\rho_{i\ell} \rho_{ik \cdot \ell} + \rho_{ik} \rho_{i\ell \cdot k}) \\
& \qquad\qquad + \frac{1}{2} \rho_{ij} \rho_{k\ell} (\rho_{j\ell}^2 + 3\rho_{ik}^2) \Bigg\} d_{ii} d_{kk} + 2\rho_{ik}^2 d_{ij} d_{k\ell} \\
& \qquad\qquad - 2\{\rho_{ij}(\rho_{jk}^2 + \rho_{ik}^2) - 2\rho_{ik} \rho_{jk \cdot i}\} d_{ii} d_{k\ell} \Bigg) f_{ij} f_{k\ell} \\
& \quad + \textstyle\sum_{i \neq j} \sum_{k \neq \ell} (d_{ij} - \rho_{ij} d_{ii}) \{d_{kk}(\rho_{k\ell} d_{\ell\ell} + 3\rho_{k\ell} d_{kk} - 4d_{k\ell}) \\
& \qquad\qquad + 4(d_{k\ell}^* - \rho_{k\ell} d_{kk}^*)\} f_{ijk\ell} \\
& \quad + \textstyle\sum_{i \neq j} \sum_{k \neq \ell} \sum_{q \neq r} \Bigg(2\rho_{q\ell \cdot k}(d_{ij} - \rho_{ij} d_{ii}) \{\rho_{qr} \rho_{qk}(3d_{qq} + d_{rr}) \\
& \qquad\qquad - 2(\rho_{qk} d_{qr} + \rho_{kr} d_{rr})\} f_{ijk\ell} f_{qr} \\
& \qquad + \frac{1}{3} \{(d_{qr} - 3\rho_{qr} d_{qq}) d_{ij} d_{k\ell} + \rho_{ij} \rho_{k\ell} (3d_{qr} - \rho_{qr} d_{qq}) d_{ii} d_{kk}\} f_{ijk\ell qr} \Bigg) \\
& \quad + \textstyle\sum_{i \neq j} \sum_{k \neq \ell} \sum_{q \neq r} \sum_{s \neq t} (2\rho_{it \cdot s} \rho_{jk \cdot i} \rho_{\ell q \cdot k} \rho_{rs \cdot q} f_{ij} f_{k\ell} f_{qr} f_{st} \\
& \qquad\qquad + \{\rho_{ks}(\rho_{\ell t \cdot s} - \rho_{st} \rho_{s\ell \cdot k})(d_{ij} d_{qr} + \rho_{ij} \rho_{qr} d_{ii} d_{qq}) \\
& \qquad\qquad - \rho_{qr} \rho_{kt \cdot s} \rho_{s\ell \cdot k} d_{ij} d_{qq}\} f_{ijk\ell} f_{qrst}) \quad \text{and}
\end{aligned}
$$

$$b_6 = a_3^2/2,$$

where

$$
\begin{aligned}
& \rho_{jk \cdot i} = \rho_{jk} - \rho_{ij} \rho_{ik}, \ d_{ij} = \textstyle\sum_{\alpha \neq \beta} \rho_{i\alpha}(\rho_{j\beta} - \rho_{j\alpha} \rho_{\alpha\beta}) f_{\alpha\beta}, \\
& d_{ij}^* = \textstyle\sum_{\alpha \neq \beta} d_{i\alpha}(\rho_{j\beta} - \rho_{j\alpha} \rho_{\alpha\beta}) f_{\alpha\beta}, \\
& f_{12} = f_{21} = \rho_{13} \rho_{23}, \ f_{13} = f_{31} = \rho_{12} \rho_{23}, \ f_{23} = f_{32} = \rho_{12} \rho_{13}, \\
& f_{1212} = f_{1221} = f_{2112} = f_{2121} = f_{1313} = f_{1331} = f_{3113} = f_{3131} \\
& \quad = f_{2323} = f_{2332} = f_{3223} = f_{3232} = 0, \\
& f_{1213} = f_{1231} = f_{2113} = f_{2131} = \rho_{23}, \\
& f_{1223} = f_{1232} = f_{2312} = f_{2321} = \rho_{13}, \\
& f_{1323} = f_{1332} = f_{3123} = f_{3132} = \rho_{12}, \\
& f_{121323} = f_{121332} = f_{122313} = f_{122331} = f_{131223} = f_{131232} \\
& \quad = f_{132312} = f_{132321} = f_{231213} = f_{231231} = f_{231312} = f_{231321} = 1,
\end{aligned}
$$

$f_{ijk\ell qr} = 0$ for other $1 \leq i, j, k, \ell, q, r \leq 3$.

Here, the summation $\sum_{a \neq b}$ stands for $\sum_{a,b=1,a \neq b}^{3}$. Putting $u = -\sqrt{n}\rho_{21}\rho_{31}\rho_{32}/\tau$, we can obtain approximations for $\Pr(r_{21}r_{31}r_{32} < 0)$, and consequently, $\Pr(r_{21}r_{31}r_{32} > 0)$.

Table 2.2. Probability that the elements of $\hat{\lambda}$ are real numbers

λ'	$\rho_{21}\rho_{31}\rho_{32}$	n	(1)	(2)	(3)	(4)
(.4 .4 .4)	.004096	50	.720	.653	.614	.762
				−.067	−.039	
		100	.796	.796	.799	.859
				−.000	.003	
		150	.845	.872	.888	.929
				.027	.016	
(.5 .5 .5)	.015625	50	.815	.835	.857	.923
				.020	.022	
		100	.899	.945	.967	.989
				.046	.022	
		150	.941	.986	.999	.995
				.045	.013	
(.4 .5 .6)	.0144	50	.802	.816	.835	.900
				.014	.019	
		100	.886	.929	.951	.973
				.043	.022	
		150	.930	.975	.990	.992
				.045	.015	
($\sqrt{.24}$ .5 $\sqrt{.24}$)	.0144	50	.809	.825	.845	.909
				.016	.020	
		100	.893	.938	.961	.986
				.045	.023	
		150	.936	.982	.997	.993
				.045	.015	
(.4 .9 .4)	.020736	50	.799	.820	.846	.884
				.021	.026	
		100	.884	.921	.938	.950
				.037	.018	
		150	.928	.965	.975	.977
				.037	.010	
(.6 .6 .4)	.020736	50	.828	.858	.885	.940
				.029	.027	
		100	.911	.957	.976	.993
				.045	.020	
		150	.951	.991	1.001	.995
				.040	.010	

(1) the limiting term
(2) upper: up to the term of $1/\sqrt{n}$
lower: the term of $1/\sqrt{n}$
(3) upper: up to the term of $1/n$
lower: the term of $1/n$
(4) values obtained by simulation (1000 replications)

Numerical examples are presented in Table 2.2. Here we assume $\boldsymbol{x}$ is distributed as a multivariate normal distribution with mean $\mathbf{0}$ and a variance-covariance matrix $\boldsymbol{\lambda\lambda}' + \mathrm{diag}\,(I - \boldsymbol{\lambda\lambda}')$, and we use the asymptotic expansion up to the term of $1/n$. Table 2.2 shows the probability that the elements of $\hat{\boldsymbol{\lambda}}$ are real numbers for some cases of $\boldsymbol{\lambda}$. It is seen that (i) when $\rho_{21}\rho_{31}\rho_{32}$ is large or n is large, $\Pr(r_{21}r_{31}r_{32} > 0)$ is large, and (ii) for the same $\rho_{21}\rho_{31}\rho_{32}$, $\Pr(r_{21}r_{31}r_{32} > 0)$ is smaller when smaller loading exists.

In particular, if $\boldsymbol{\lambda} = (\lambda\ \lambda\ \lambda)'$ $(\lambda > 0)$ and $\Psi = \mathrm{diag}\,(I - \boldsymbol{\lambda\lambda}')$, then

$$\tau^2 = \{\sqrt{3}\lambda^4(\lambda^2 - 1)(2\lambda^2 + 1)\}^2 \text{ and } u = \sqrt{n}\lambda^2/\{\sqrt{3}(\lambda^2 - 1)(2\lambda^2 + 1)\}.$$

If n is large, we can approximate $\Pr(r_{21}r_{31}r_{32} < 0)$ by $\Phi(u)$. As λ tends to 1 from 0, u is monotone decreasing, because

$$\frac{\mathrm{d}u}{\mathrm{d}\lambda} = \frac{\sqrt{n}\cdot 2(-2\lambda^4 - 1)\lambda}{3(\lambda^2 - 1)^2(2\lambda^2 + 1)^2} < 0.$$

Therefore, $\Pr(r_{21}r_{31}r_{32} > 0)$ tends to 1 from $1/2$ monotonously as λ tends to 1 from 0.

3. Uniqueness of the decompositions

Throughout this section, we assume that Σ has a decomposition

$$\Sigma = \Lambda_k\Lambda_k' + \Psi_k,$$

where Λ_k is a $p \times k$ $(p > k)$ real matrix of rank k and Ψ_k is a diagonal matrix with positive diagonal elements. The uniqueness problem for factor size m is as follows: Does there exist $\Delta \neq \Psi$ such that

$$\Sigma = FF' + \Delta$$

where F is a $p \times m$ $(p > m)$ real matrix of rank m and Δ is a diagonal matrix with positive diagonal elements, for given m?

First, we will discuss sufficient conditions for uniqueness. For factor size $m = k$, the main result which has been obtained hitherto is as follows:

THEOREM 3.1. (Theorem 5.1 of Anderson and Rubin [3]) *A sufficient condition for uniqueness is that if any one row of Λ is deleted then there remain two disjoint submatrices of rank k.* □

PROPOSITION 3.1. (p. 211 of Takeuchi and Yanai [30]) *If a decomposition is unique for factor size k, then k is the smallest number of all k satisfying* (0.2).

We will consider the situation $m > k$. In general, researchers often try to extract more factors. In fact, factor size is usually unknown in practice, and thus, then we try to estimate Λ, the hypothetical factor size is increased step by step. Further a statistical test almost always indicates more factors than the factors researchers postulated in advance ([31]).

Of course Λ_{k+1} is not unique from Proposition 2.1. Note that Λ_{k+1} does not always have specific factor loading. For example, when $p = 3$ and $k = 1$, suppose that

$$P = \Lambda_1 \Lambda_1' + \Psi_1,$$

where $\Lambda_1 = (\lambda\ \lambda\ \lambda)'$, $0 < \lambda^2 < 1/4$ and $\Psi_1 = \operatorname{diag}\{1-\lambda^2 \quad 1-\lambda^2 \quad 1-\lambda^2\}$. If we set

$$\Lambda_2 = \begin{bmatrix} 2\lambda & \lambda/2 & \lambda/2 \\ 0 & \lambda & 3\lambda/4 \end{bmatrix}' \text{ and}$$
$$\Psi_2 = \operatorname{diag}\{1-4\lambda^2 \quad 1-5\lambda^2/4 \quad 1-13\lambda^2/16\},$$

then,

$$P = \Lambda_2 \Lambda_2' + \Psi_2.$$

Now we can observe that Λ_2 does not have specific factor loading. This follows from the following proposition:

PROPOSITION 3.2. (Theorem 2.1 of Tumura and Fukutomi [32]) *A necessary and sufficient condition that Λ does not have specific factor loading is that the rank of Λ remains invariant after deletion of any one row of Λ.*

The aim of FA is to extract common factors. Next theorem gives a sufficient condition for the following property: If factor size is increased up to $k + s$, s specific factor loadings are added, and, the common factor loading matrix Λ_k remains invariant. This property is called *the extended uniqueness.*

THEOREM 3.2. (Theorem 1 of Tumura and Sato [34]) *If there remain two disjoint submatrices of rank k in Λ_k after deletion of any $(r + 1)$ rows of Λ_k $(0 \leq r \leq p - 2k - 1)$. Then, for other decompositions such that*

$$\Sigma = \Lambda_{k+s} \Lambda_{k+s}' + \Psi_{k+s},$$

where Λ_{k+s}: $p \times (k + s)$, rank $\Lambda_{k+s} = k + s$, $0 \leq s \leq r$,

Ψ_{k+s}: *a diagonal matrix,*

Λ_{k+s} *is a following form*

$$\Lambda_{k+s} T_{k+s} = [\Lambda_k \quad \Gamma_s],$$

where T_{k+s} *is some orthogonal matrix of order* $k+s$ *and* off-diag $\Gamma_s \Gamma_s' = O$. □

This theorem is an improvement of Lemma 2.1 of Tumura and Fukutomi [32] (see Sato [25]). The juxtaposed matrix Γ_s contains s specific factor loadings, not common factor. In the case $r = 0$, this theorem is reduced to Theorem 3.1.

Next, we will discuss necessary conditions for uniqueness. For $k = 1$ and 2, the condition of Theorem 3.1 is also necessary one ([3]). For $k = 3$, the condition is necessary for the cases $p \geq 7$ ([34]), but is never satisfied for the case $p = 6$, because $p < 2k + 1$. However, for the latter case $k = 3$ and $p = 6$, there exist unique loading matrices ([34]).

PROPOSITION 3.3. (Theorem 5.6 of Anderson and Rubin [3]) *A necessary condition for uniqueness is that each column of* ΛG *has at least three nonzero elements for every nonsingular* G. □

The following theorem is an extension of Proposition 3.3.

THEOREM 3.3. *A necessary condition for satisfying the condition of Theorem* 3.1 *is that the submatrices which consist of each* q *columns of* ΛG *have at least* $(2q + 1)$ *nonzero rows for every nonsingular* G $(q = 1, 2, \ldots, k)$. □

Consider the cases where the condition of Theorem 3.1 is a necessary and sufficient condition (that is, the cases $k = 3$ for $p \geq 7$ and $k = 1$ and 2). For these cases, the condition of Theorem 3.3 is a necessary condition for uniqueness. In particular, when $q = 1$, Theorem 3.3 agrees with Proposition 3.3. For other cases, that is, the cases $k = 3$ & $p \leq 6$ and $k \geq 4$, if the condition of Theorem 3.3 is not satisfied, we must examine the uniqueness by other ways not based on Theorem 3.1.

THEOREM 3.4. *A necessary condition for satisfying the condition of Theorem* 3.2 *is that the submatrices which consist of each* q *columns of* ΛG *have at least* $(2q + r + 1)$ *nonzero rows for every nonsingular* G $(q = 1, 2, \ldots, k)$. □

For the cases where the rank of a submatrix of Λ is not full, we will propose methods to examine whether the condition of Theorems 3.1 or 3.2 is satisfied or not. Let the rank of a submatrix which consists of p_2 rows of Λ be $k_1 (< k)$ and suppose the submatrix is the last p_2 rows of Λ. Then, by a suitable orthogonal matrix T, we can obtain

$$\Lambda T = \begin{bmatrix} \Lambda_{11} & \Lambda_{12} \\ \Lambda_{21} & O \end{bmatrix}, \qquad \begin{array}{l} \Lambda_{11} : p_1 \times k_1, \ \Lambda_{12} : p_1 \times k_2, \\ \Lambda_{21} : p_2 \times k_1, \ \ O \ : p_2 \times k_2, \\ p = p_1 + p_2, \ k = k_1 + k_2. \end{array}$$

THEOREM 3.5. *A necessary condition for satisfying the condition of Theorem* 3.1 *is that* Λ_{12} *satisfies the condition of Theorem* 3.1. □

PROOF. If Λ_{12} does not satisfy the condition of Theorem 3.1, the submatrix which consists of last p_2 columns of Λ does not satisfy the condition of Theorem 3.1. Then, Λ can not satisfy the condition of Theorem 3.1. □

In the same manner, Theorems 3.6 to 3.8 can be proved.

THEOREM 3.6. *Suppose* Λ_{21} *satisfies the condition of Theorem* 3.1. *Then, a necessary and sufficient condition that* Λ *satisfies the condition of Theorem* 3.1 *is that* Λ_{12} *satisfies the condition of Theorem* 3.1. □

THEOREM 3.7. *A necessary condition for satisfying the condition of Theorem* 3.2 *is that* Λ_{12} *satisfies the condition of Theorem* 3.2. □

THEOREM 3.8. *Suppose* Λ_{21} *satisfies the condition of Theorem* 3.2. *Then, a necessary and sufficient condition that* Λ *satisfies the condition of Theorem* 3.2 *is that* Λ_{12} *satisfies the condition of Theorem* 3.2. □

Next, we consider a loading matrix whose *most elements are unique.*

THEOREM 3.9. *Suppose that* Λ *has the following form*

$$\Lambda = \begin{bmatrix} \Lambda_{11} & \Lambda_{12} \\ \Lambda_{21} & O \end{bmatrix}, \qquad \begin{array}{l} \Lambda_{11}: p_1 \times k_1, \; \Lambda_{12}: p_1 \times k_2 \\ \Lambda_{21}: p_2 \times k_1, \quad O \;: p_2 \times k_2 \\ p = p_1 + p_2, \; k = k_1 + k_2, \end{array}$$

and Λ_{21} *is unique. Then, a loading matrix* F *for factor size* k *can be expressed as*

$$FT = \begin{bmatrix} \Lambda_{11} & F_{12} \\ \Lambda_{21} & O \end{bmatrix},$$

where T *is some orthogonal matrix.* □

PROOF. Let us partition F as $F = [F_1' \; F_2']'$, where F_2 is last p_2 rows of F. Since Λ_{21} is unique, there exists an orthogonal matrix T such that

$$F_2 T = [\Lambda_{21} \; O].$$

Letting

$$FT = \begin{bmatrix} F_{11} & F_{12} \\ \Lambda_{21} & O \end{bmatrix},$$

we obtain

$$\Lambda_{11}\Lambda_{21}' = F_{11}\Lambda_{21}'.$$

Since Λ_{21} is unique, rank $\Lambda_{21} = k_1 < p_2$, and hence $F_{11} = \Lambda_{11}$. $\square$

This theorem means that submatrices $\Lambda_{11}, \Lambda_{21}$ and O are determined uniquely though Λ is not unique. When Λ is not unique, the object of estimation is not clear. However, if there exist unique submatrices, the estimates for the unique parts are meaningful.

We give an example showing usefulness of this theorem. The hypothetical factor loading matrix $\Lambda_H(17 \times 6)$ of the data treated by Bechtoldt [4] was the following form:

$$\Lambda_H = \begin{bmatrix} \Lambda_{11} & \Lambda_{12} \\ \Lambda_{21} & \mathbf{0} \end{bmatrix}, \qquad \begin{matrix} \Lambda_{11}:\ 2 \times 5, & \Lambda_{12}:\ 2 \times 1 \\ \Lambda_{21}: 15 \times 5, & \mathbf{0}\ : 15 \times 1, \end{matrix}$$

where

$$\Lambda_{21} = \begin{bmatrix} * * * & & & & \\ & * * * & & O & \\ & & * * * & & \\ & O & & * * * & \\ & & & & * * * \end{bmatrix}',$$

$$\Lambda_{12} = (*\ *)'$$

and $*$ denotes a nonzero element. From Proposition 3.3, we observe that Λ_H is not unique. As a matter of course, when researchers estimate Λ and Ψ, several difficulties arise. Estimates largely depend on methods of estimation (e.g. the maximum likelihood method, the least squares method), samples which are divided randomly to two sets, and initial approximations for iterative schemes (e.g. the value recommended by Jöreskog [10], the highest correlation, both of which are used widely). Further improper solutions are met, and the structure of the data cannot be recognized. Now we reexamine the estimates of Λ precisely. We find that the submatrices corresponding to $\Lambda_{11}, \Lambda_{21}$ and $\mathbf{0}$ are stable and only the elements corresponding to Λ_{12} are fluctuate (see Table 3.1. The loading matrices are rotated by the incomplete Procrustean method [23; §4.1].); this can be interpreted from this theorem, and we can expect to obtain information from such data.

Table 3.1. Example which indicates usefulness of Theorem 3.9: Bechtoldt's data

(1) ML, Sample I

$$\begin{bmatrix}
.3 & .2 & .0 & .2 & .2 & \mathbf{.0} \\
.1 & .3 & .1 & .1 & .1 & \mathbf{7.3} \\
\mathbf{.8} & .2 & -.0 & .1 & .3 & -.0 \\
\mathbf{.8} & .3 & -.0 & .2 & .2 & -.0 \\
\mathbf{.8} & .2 & .3 & .1 & .2 & .0 \\
.2 & \mathbf{.7} & .1 & .2 & .2 & -.0 \\
.1 & \mathbf{.8} & .1 & .2 & .1 & .0 \\
.3 & \mathbf{.6} & -.0 & .1 & .1 & -.0 \\
.0 & .1 & \mathbf{.7} & .2 & .2 & -.0 \\
.0 & .1 & \mathbf{.9} & .0 & .0 & -.0 \\
.1 & .0 & \mathbf{.8} & .0 & .1 & .0 \\
.1 & .2 & .1 & \mathbf{.6} & .2 & .0 \\
.1 & .2 & -.0 & \mathbf{.9} & .2 & .0 \\
.2 & .2 & .2 & \mathbf{.5} & .3 & -.0 \\
.3 & .2 & .1 & .3 & \mathbf{.7} & .0 \\
.4 & .2 & .1 & .1 & \mathbf{.7} & .0 \\
.2 & .2 & .2 & .3 & \mathbf{.6} & -.0
\end{bmatrix}$$

(2) ML, Sample II

$$\begin{bmatrix}
.1 & .2 & .0 & .1 & .2 & \mathbf{1.0} \\
.2 & .1 & .1 & .2 & .1 & \mathbf{.4} \\
\mathbf{.8} & .2 & .0 & .1 & .2 & .1 \\
\mathbf{.8} & .3 & .0 & .2 & .2 & .1 \\
\mathbf{.8} & .3 & .1 & .1 & .2 & .1 \\
.2 & \mathbf{.8} & .1 & .2 & .2 & .1 \\
.2 & \mathbf{.7} & .2 & .1 & .2 & .1 \\
.3 & \mathbf{.6} & .0 & .1 & .1 & .2 \\
.0 & .1 & \mathbf{.7} & .1 & .1 & .1 \\
.0 & .0 & \mathbf{.8} & .0 & .1 & .0 \\
.1 & .1 & \mathbf{.8} & .1 & .1 & .0 \\
.1 & .1 & .1 & \mathbf{.9} & .1 & .1 \\
.1 & .2 & .1 & \mathbf{.7} & .2 & .1 \\
.2 & .1 & .3 & \mathbf{.5} & .3 & .1 \\
.2 & .2 & .2 & .3 & \mathbf{.7} & .1 \\
.4 & .1 & .0 & .1 & \mathbf{.5} & .3 \\
.2 & .3 & .2 & .2 & \mathbf{.6} & .1
\end{bmatrix}$$

(3) LS, Sample I

$$\begin{bmatrix}
.2 & .2 & .0 & .2 & .2 & \mathbf{.4} \\
.1 & .1 & .1 & .0 & .1 & \mathbf{1.0} \\
\mathbf{.8} & .2 & -.0 & .1 & .3 & .1 \\
\mathbf{.8} & .3 & -.0 & .2 & .2 & .1 \\
\mathbf{.8} & .2 & .3 & .1 & .2 & .1 \\
.2 & \mathbf{.7} & .1 & .2 & .2 & .1 \\
.1 & \mathbf{.8} & .1 & .2 & .1 & .1 \\
.3 & \mathbf{.5} & -.0 & .1 & .1 & .1 \\
.1 & .1 & \mathbf{.7} & .2 & .2 & -.0 \\
.0 & .1 & \mathbf{.9} & .0 & .0 & .0 \\
.1 & -.0 & \mathbf{.8} & .0 & .1 & .1 \\
.2 & .2 & .1 & \mathbf{.6} & .2 & .1 \\
.1 & .2 & -.0 & \mathbf{.9} & .2 & .1 \\
.2 & .2 & .2 & \mathbf{.5} & .3 & -.0 \\
.3 & .2 & .1 & .3 & \mathbf{.7} & .1 \\
.4 & .2 & .1 & .1 & \mathbf{.7} & .1 \\
.2 & .2 & .2 & .3 & \mathbf{.5} & .1
\end{bmatrix}$$

(4) LS, Sample II

$$\begin{bmatrix}
.1 & .2 & .0 & .1 & .2 & \mathbf{.8} \\
.2 & .1 & .1 & .2 & .1 & \mathbf{.5} \\
\mathbf{.8} & .2 & .0 & .2 & .2 & .1 \\
\mathbf{.8} & .3 & .0 & .2 & .2 & .2 \\
\mathbf{.8} & .3 & .1 & .1 & .2 & .1 \\
.2 & \mathbf{.8} & .1 & .2 & .2 & .1 \\
.2 & \mathbf{.7} & .2 & .1 & .3 & .1 \\
.3 & \mathbf{.6} & .0 & .1 & .1 & .2 \\
.0 & .1 & \mathbf{.7} & .1 & .1 & .1 \\
.0 & .0 & \mathbf{.8} & .0 & .1 & -.0 \\
.1 & .1 & \mathbf{.8} & .1 & .1 & .1 \\
.1 & .1 & .1 & \mathbf{.9} & .1 & .1 \\
.1 & .2 & .1 & \mathbf{.7} & .2 & .1 \\
.2 & .1 & .3 & \mathbf{.5} & .3 & .1 \\
.2 & .2 & .2 & .3 & \mathbf{.7} & .1 \\
.4 & .1 & .0 & .1 & \mathbf{.6} & .3 \\
.2 & .3 & .2 & .2 & \mathbf{.6} & .1
\end{bmatrix}$$

(5) ML, Sample I

$$\begin{bmatrix}
.2 & .3 & .0 & .2 & .2 & \mathbf{8.1} \\
.1 & .2 & .1 & .1 & .1 & \mathbf{.0} \\
\mathbf{.8} & .2 & -.0 & .1 & .3 & -.0 \\
\mathbf{.8} & .3 & -.0 & .2 & .2 & .0 \\
\mathbf{.8} & .2 & .3 & .1 & .2 & -.0 \\
.2 & \mathbf{.7} & .1 & .2 & .2 & -.0 \\
.1 & \mathbf{.8} & .1 & .2 & .1 & -.0 \\
.3 & \mathbf{.6} & -.0 & .1 & .1 & .0 \\
.0 & .1 & \mathbf{.7} & .2 & .2 & .0 \\
.0 & .1 & \mathbf{.9} & .0 & .0 & .0 \\
.1 & -.0 & \mathbf{.8} & .0 & .1 & -.0 \\
.1 & .2 & .1 & \mathbf{.6} & .2 & .0 \\
.1 & .2 & -.0 & \mathbf{.9} & .1 & -.0 \\
.2 & .2 & .2 & \mathbf{.5} & .3 & .0 \\
.3 & .2 & .1 & .3 & \mathbf{.7} & -.0 \\
.4 & .2 & .1 & .1 & \mathbf{.6} & .0 \\
.2 & .2 & .2 & .3 & \mathbf{.6} & .0
\end{bmatrix}$$

estimation methods
- ML: the Maximum Likelihood method
- LS : the Least Squares method

sets of sample
- sample I : Size equals 212.
- sample II: Size equals 213.

initial approximation for an iterative process
- (1)–(4): the value recommended by Jöreskog
- (5): the highest correlation

convergence/not convergence
- (2)–(4): convergence
- (1), (5): not convergence (after 100 iterative counts)

Part II. Examination of adequacy of substitute use of principal component analysis

4. Preliminary

First, we describe precisely the use of PCA as a substitute for FA. In FA, an observed vector $\boldsymbol{x}$ is assumed to follow model (0.1), and consequently Σ or P has a decomposition

$$\Lambda\Lambda' + \Psi.$$

Then, to discern a latent structure, we estimate not only Λ but an error variance matrix Ψ. In contrast, PCA does not require such a structural model. In substitute use for FA, a sample correlation matrix R is decomposed as

$$\begin{aligned} R &= (QD^{1/2})(QD^{1/2})' \\ &= \tilde{\Lambda}\tilde{\Lambda}' + E, \text{ say,} \end{aligned}$$

where D is a p-order diagonal matrix with the ith largest eigenvalue of R as the ith diagonal element, Q is an orthogonal matrix such that $Q'RQ = D$, $\tilde{\Lambda}$ is the first k columns of $QD^{1/2}$ and $E = R - \tilde{\Lambda}\tilde{\Lambda}'$ (see e.g. §4.3.7. of Chatfield and Collins [6]); $\tilde{\Lambda}$ is named "a factor loading matrix" after FA. A sample variance-covariance matrix S instead of R may be used. If k is unknown, it is often used to determine k as the number of eigenvalues of R which are greater than one.

The present study attempts to answer the following question: Can substitute use be justified? More precisely, we examine the following points:
(1) Is it justifiable to use a rule where factor size is taken as the number of eigenvalues of R which are greater than one?
(2) Is it justifiable to use the first some columns of $QD^{1/2}$ for factor loadings?
(3) In what situation and to what extent does the result using PCA differ from the one using FA?

We will study the above problems (1)–(3) under the following setup. First we assume that Σ or P has a decomposition

$$\Lambda\Lambda' + \Psi.$$

This will be natural, because, when researchers want to interpret the loadings $\tilde{\Lambda}$ calculated with PCA, it is assumed implicitly that an FA model holds or at least approximately. Next, we will restrict our discussion to the population case, because it is difficult to express the loadings $\hat{\Lambda}$ estimated with FA explicitly. If substitute use is not justified in the population, it cannot expect to work well in a sample. Finally, we assume that Λ satisfies the condition of Theorem

3.1. Then, Ψ can be expressed as a function of Σ explicitly (Ihara and Kano [8]), and hence an estimated loading matrix is determined the true value Λ uniquely (up to multiplication on the right by an orthogonal matrix) from Σ or P. Consequently, our problems are reduced to compare a factor loading matrix $\tilde{\Lambda}$ calculated with PCA to the true value Λ.

5. Monofactor case

In this section, we consider the monofactor case, i.e., $\Lambda = (\lambda_1, \ldots, \lambda_p)'$, say $\boldsymbol{\lambda}$. We can assume $\lambda_1 \geq \cdots \geq \lambda_p \geq 0$ without loss of generality. Because, if the sign of the ith variable of an observation vector is changed, the sign of λ_i is inverted, and, even if the order of variables is changed, the model is invariant. When $p \geq 3$ and $\lambda_p \neq 0$, the assumption of Theorem 3.1 is satisfied, so that $\boldsymbol{\lambda}$ is determined uniquely.

First, we give a property of eigenvalues.

THEOREM 5.1. *Suppose that a population correlation matrix P has a structure*

$$P = \boldsymbol{\lambda}\boldsymbol{\lambda}' + \Psi,$$

where $\boldsymbol{\lambda} = (\lambda_1, \ldots, \lambda_p)'$ is a factor loading vector and $\Psi = \mathrm{diag}\,(\psi_1, \ldots, \psi_p)$ is an error variance matrix. Assume that $1 > \lambda_1 \geq \cdots \geq \lambda_p > 0$ and $p \geq 3$. Then the following inequalities for the eigenvalues $\theta_1 \geq \cdots \geq \theta_p$ of P are established;

$$\boldsymbol{\lambda}'\boldsymbol{\lambda} + \psi_p \geq \theta_1 \geq \boldsymbol{\lambda}'\boldsymbol{\lambda} + \psi_1 > 1 > \psi_p \geq \theta_2 \geq \psi_{p-1} \geq \cdots \geq \theta_p \geq \psi_1. \tag{5.1}$$

The equalities $\boldsymbol{\lambda}'\boldsymbol{\lambda} + \psi_p = \theta_1 = \boldsymbol{\lambda}'\boldsymbol{\lambda} + \psi_1$ hold if and only if $\psi_1 = \cdots = \psi_p$. The equalities $\psi_{p+2-i} = \theta_i = \psi_{p+1-i}$ hold if and only if $\psi_{p+2-i} = \psi_{p+1-i}$ $(i = 2, \ldots, p)$. □

PROOF. An eigenvalue of P is a zeropoint of an eigenpolynomial $|P - \theta I|$. We will examine signs of the eigenpolynomial at the upper and lower bounds of θ_i in (5.1). First, consider the sign of $|P - (\alpha + \psi_1)I|$, where $\alpha = \boldsymbol{\lambda}'\boldsymbol{\lambda}$.

Adding $\sum_{i=2}^{p}$(ith row) $\times\ \lambda_i/\lambda_1$ to the first row in the matrix $P - (\alpha + \psi_1)I$, we have

$$|P - (\alpha + \psi_1)I|$$
$$= \begin{vmatrix} 0 & (\psi_2 - \psi_1)\lambda_2/\lambda_1 & (\psi_3 - \psi_1)\lambda_3/\lambda_1 & \cdots & (\psi_p - \psi_1)\lambda_p/\lambda_1 \\ \lambda_2\lambda_1 & \lambda_2^2 - \alpha + \psi_2 - \psi_1 & \lambda_2\lambda_3 & \cdots & \lambda_2\lambda_p \\ \lambda_3\lambda_1 & \lambda_3\lambda_2 & \lambda_3^2 - \alpha + \psi_3 - \psi_1 & \cdots & \lambda_3\lambda_p \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ \lambda_p\lambda_1 & \lambda_p\lambda_2 & \lambda_p\lambda_3 & \cdots & \lambda_p^2 - \alpha + \psi_p - \psi_1 \end{vmatrix}.$$

Let us divide the first column by λ_1, multiply the first row by λ_1, and subtract (the first column) $\times\ \lambda_j$ from the jth column ($j=2,\dots,p$). Then, the determinant is equal to

$$\begin{vmatrix} 0 & (\psi_2-\psi_1)\lambda_2 & (\psi_3-\psi_1)\lambda_3 & \cdots & (\psi_p-\psi_1)\lambda_p \\ \lambda_2 & -(\alpha+\psi_1-\psi_2) & 0 & \cdots & 0 \\ \lambda_3 & 0 & -(\alpha+\psi_1-\psi_3) & \cdots & 0 \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ \lambda_p & 0 & 0 & \cdots & -(\alpha+\psi_1-\psi_p) \end{vmatrix}$$
$$=\prod_{j=2}^{p}(-(\alpha+\psi_1-\psi_j))\sum_{i=2}^{p}(\psi_i-\psi_1)\lambda_i^2/(\alpha+\psi_1-\psi_i).$$

The last reduction is obtained by using the formula

$$\begin{vmatrix} a & \boldsymbol{b}' \\ \boldsymbol{c} & D \end{vmatrix} = |D|\cdot|a-\boldsymbol{b}'D^{-1}\boldsymbol{c}|.$$

Thus, if all ψ_i's are not equal, the sign of $|P-(\alpha+\psi_1)I|$ is $(-1)^{p-1}$. Similarly, we get

$$|P-(\alpha+\psi_p)I|$$
$$=\prod_{j=1}^{p-1}(-(\alpha+\psi_p-\psi_j))\sum_{i=1}^{p-1}(\psi_i-\psi_p)\lambda_i^2/(\alpha+\psi_p-\psi_i).$$

Thus, if all ψ_i's are not equal, the sign of $|P-(\alpha+\psi_p)I|$ is $(-1)^p$. Next, consider the sign of $|P-\psi_\ell I|$. In the matrix $P-\psi_\ell I$, subtracting (the ℓth column) $\times\ \lambda_j/\lambda_\ell$ from the jth column ($j=1,\dots,p,\ j\neq\ell$), we have

$$|P-\psi_\ell I|$$
$$=\begin{vmatrix} \psi_1-\psi_\ell & 0 & \cdots & \lambda_1\lambda_\ell & \cdots & 0 & 0 \\ 0 & \psi_2-\psi_\ell & \cdots & \lambda_2\lambda_\ell & \cdots & 0 & 0 \\ 0 & 0 & & \lambda_3\lambda_\ell & \cdots & 0 & 0 \\ \cdots & \cdots & & \cdots & & \cdots & \cdots \\ 0 & 0 & & \lambda_\ell^2 & & 0 & 0 \\ \cdots & \cdots & \cdots & \cdots & & \cdots & \cdots \\ 0 & 0 & \cdots & \lambda_{p-1}\lambda_\ell & & \psi_{p-1}-\psi_\ell & 0 \\ 0 & 0 & \cdots & \lambda_p\lambda_\ell & \cdots & 0 & \psi_p-\psi_\ell \end{vmatrix}$$
$$=\lambda_\ell^2\prod_{j\neq\ell}(\psi_j-\psi_\ell).$$

Thus, when all ψ_ℓ's ($\ell=1,\dots,p$) are distinct, the sign of $|P-\psi_\ell I|$ is $(-1)^{\ell-1}$. Therefore, noting that the eigenpolynomial is a continuous function, we obtain the following inequalities;

$$\boldsymbol{\lambda}'\boldsymbol{\lambda} + \psi_p > \theta_1 > \boldsymbol{\lambda}'\boldsymbol{\lambda} + \psi_1 > 1 > \psi_p > \theta_2 > \psi_{p-1} > \cdots > \theta_p > \psi_1.$$

Consider the case where some ψ_i's, say ψ_i^*'s, are equal. Separate ψ_i^*'s temporarily, and then approach them the original values. This leads to the required inequalities (5.1). The equality conditions can be obtained easily. □

This theorem makes the following remarks;
(1) The rule taking the number of $\theta_i > 1$ as the number of "factors" is justified. (Kendall (p. 27 in [11]) stated that this rule is a very rough-and-ready procedure for which it is difficult to advance a convincing theoretical justification.)
(2) An addition of variables or an increase of $|\lambda_2|, |\lambda_3|, \ldots, |\lambda_p|$ makes the lower bound of θ_1 larger.
(3) A decrease of ψ_p, or equivalently an increase of $|\lambda_p|$, makes the upper bound of θ_2 smaller.

Next we will examine behavior for factor loadings $\tilde{\boldsymbol{\lambda}} = (\tilde{\lambda}_1 \tilde{\lambda}_2, \ldots, \tilde{\lambda}_p)'$ calculated with PCA. Some relationships between $\tilde{\boldsymbol{\lambda}}$ and $\boldsymbol{\lambda}$ are given in the following theorem.

THEOREM 5.2. *Suppose that the same assumptions as Theorem* 5.1 *hold and* $\tilde{\lambda}_1 \geq 0$. *Then the following properties can be proved.*
(1) $1 \geq \tilde{\lambda}_1 \geq \cdots \geq \tilde{\lambda}_p > 0$. *The equality* $\tilde{\lambda}_i = \tilde{\lambda}_j$ *holds if and only if* $\lambda_i = \lambda_j$.
(2) *If* $\lim_{\boldsymbol{\lambda}'\boldsymbol{\lambda} \to \infty} p/(\boldsymbol{\lambda}'\boldsymbol{\lambda})$ is finite, $\tilde{\boldsymbol{\lambda}} = (1 + O(1/(\boldsymbol{\lambda}'\boldsymbol{\lambda})))\boldsymbol{\lambda}$.
(3) $\dfrac{\lambda_i}{\lambda_j} \cdot \dfrac{\boldsymbol{\lambda}'\boldsymbol{\lambda} - \lambda_1^2 + \lambda_j^2}{\boldsymbol{\lambda}'\boldsymbol{\lambda} - \lambda_1^2 + \lambda_i^2} \leq \dfrac{\tilde{\lambda}_i}{\tilde{\lambda}_j} \leq \dfrac{\lambda_i}{\lambda_j} \cdot \dfrac{\boldsymbol{\lambda}'\boldsymbol{\lambda} - \lambda_p^2 + \lambda_j^2}{\boldsymbol{\lambda}'\boldsymbol{\lambda} - \lambda_p^2 + \lambda_i^2}$ for $i < j$. *The equalities hold if and only if* $\psi_1 = \cdots = \psi_p$.
(4) $\boldsymbol{\lambda}'\boldsymbol{\lambda} + \psi_p \geq \tilde{\boldsymbol{\lambda}}'\tilde{\boldsymbol{\lambda}} \geq \boldsymbol{\lambda}'\boldsymbol{\lambda} + \psi_1$. *The equalities hold if and only if* $\psi_1 = \cdots = \psi_p$.
□

PROOF.

(1) This property has been proved by Sato [24; (1) of Theorem 1].
(2) From Theorem 5.1, we can express θ_1 as

$$\theta_1 = \boldsymbol{\lambda}'\boldsymbol{\lambda} + \psi_0,$$

where $0 < \psi_1 \leq \psi_0 \leq \psi_p < 1$. Since θ_1 is a simple root of the characteristic equation,

$$\operatorname{rank}(P - \theta_1 I) = p - 1.$$

Let us permute rows and columns of $P - \theta_1 I$ in order to have a form

$$\begin{bmatrix} 1 - \theta_1 & \boldsymbol{\rho}' \\ \boldsymbol{\rho} & P^* - \theta_1 I^* \end{bmatrix},$$

such that $|P^* - \theta_1 I^*| \neq 0$ where P^* is a $(p-1)$-order matrix and I^* is a $(p-1)$-order unit matrix. The same permutation is done for $\boldsymbol{\lambda}$ and Ψ. We will denote the permuted results by P, $\boldsymbol{\lambda}$ and Ψ again. (Now, the relation $\lambda_1 \geq \cdots \geq \lambda_p$ does not hold.) Since an eigenvector has indefiniteness of its length, the eigenvector corresponding to θ_1 can be put $\boldsymbol{\lambda} + \boldsymbol{\delta}$, where $\boldsymbol{\delta} = (0\ \boldsymbol{\delta}^{*\prime})'$. Then

$$(P - \theta_1 I)(\boldsymbol{\lambda} + \boldsymbol{\delta}) = \mathbf{0},$$

which leads to

$$(\Psi - \psi_0 I)\boldsymbol{\lambda} + (P - \theta_1 I)\boldsymbol{\delta} = \mathbf{0},$$

and

$$(\Psi^* - \psi_0 I^*)\boldsymbol{\lambda}^* + (P^* - \theta_1 I^*)\boldsymbol{\delta}^* = \mathbf{0},$$

where $\Psi = \begin{bmatrix} \psi_1 & \mathbf{0}' \\ \mathbf{0} & \Psi^* \end{bmatrix}$ and $\boldsymbol{\lambda} = (\lambda_1\ \boldsymbol{\lambda}^{*\prime})'$.

Hence

$$\boldsymbol{\delta}^* = (I^* - P^*/\theta_1)^{-1}(\Psi^* - \psi_0 I^*)\boldsymbol{\lambda}^*/\theta_1.$$

Thus each element δ_i^* of $\boldsymbol{\delta}^*$ is $O(1/(\boldsymbol{\lambda}'\boldsymbol{\lambda}))$, and hence

$$\delta_i^* = h_i/(\boldsymbol{\lambda}'\boldsymbol{\lambda}) + o(1/(\boldsymbol{\lambda}'\boldsymbol{\lambda})).$$

Letting $\max_i h_i = h$, we obtain

$$\begin{aligned} &\boldsymbol{\delta}^{*\prime}\boldsymbol{\delta}^* \leq (p-1)h^2/(\boldsymbol{\lambda}'\boldsymbol{\lambda})^2 \text{ and} \\ &(\boldsymbol{\lambda}^{*\prime}\boldsymbol{\delta}^*)^2 \leq (\boldsymbol{\lambda}^{*\prime}\boldsymbol{\lambda}^*)(\boldsymbol{\delta}^{*\prime}\boldsymbol{\delta}^*) \\ &\qquad\qquad \leq (p-1)h/(\boldsymbol{\lambda}'\boldsymbol{\lambda}). \end{aligned}$$

These inequalities imply

$$(\boldsymbol{\lambda} + \boldsymbol{\delta})'(\boldsymbol{\lambda} + \boldsymbol{\delta}) \leq (\boldsymbol{\lambda}'\boldsymbol{\lambda})(1 + 2\sqrt{(p-1)h}/(\boldsymbol{\lambda}'\boldsymbol{\lambda})^{3/2} + (p-1)h^2/(\boldsymbol{\lambda}'\boldsymbol{\lambda})^3).$$

Therefore, if $\boldsymbol{\lambda}'\boldsymbol{\lambda}$ is large enough, the loading vector calculated with PCA is

$$\begin{aligned} &\sqrt{((\boldsymbol{\lambda}'\boldsymbol{\lambda} + \psi_0)/((\boldsymbol{\lambda} + \boldsymbol{\delta})'(\boldsymbol{\lambda} + \boldsymbol{\delta})))} \cdot (\boldsymbol{\lambda} + \boldsymbol{\delta}) \\ &= \sqrt{((\boldsymbol{\lambda}'\boldsymbol{\lambda} + \psi_0)/((\boldsymbol{\lambda}'\boldsymbol{\lambda})(1 + O(1))))} \cdot (\boldsymbol{\lambda} + O(1/(\boldsymbol{\lambda}'\boldsymbol{\lambda}))) \\ &= (1 + O(1/(\boldsymbol{\lambda}'\boldsymbol{\lambda})))\boldsymbol{\lambda}. \end{aligned}$$

(3) Let $\boldsymbol{q} = (q_i)$ be the eigenvector of P corresponding to θ_1 with $\boldsymbol{q}'\boldsymbol{q} = 1$. Since $P\boldsymbol{q} = \theta_1 \boldsymbol{q}$, we obtain

$$(1-\theta_1)q_i + \lambda_i\lambda_j q_j + \lambda_i \textstyle\sum_{h\neq i,j}\lambda_h q_h = 0 \text{ and}$$
$$\lambda_i\lambda_j q_i + (1-\theta_1)q_j + \lambda_j \textstyle\sum_{h\neq i,j}\lambda_h q_h = 0 \text{ for } i\neq j,$$

where $P=(\rho_{ij})$, $\rho_{ii}=1$ and $\rho_{ij}=\lambda_i\lambda_j$ $(i\neq j)$.
Thus

$$((1-\theta_1)/\lambda_i - \lambda_i)q_i = ((1-\theta_1)/\lambda_j - \lambda_j)q_j.$$

Then, using (1) and $\tilde{\boldsymbol{\lambda}} = \sqrt{\theta_1}\boldsymbol{q}$, we see $q_j \neq 0$ $(j=1,\dots,p)$ and

$$q_i/q_j = (\lambda_i/\lambda_j)\cdot((\theta_1 - 1 + \lambda_j^2)/(\theta_1 - 1 + \lambda_i^2)).$$

By using an inequalities for θ_1 such that $\theta_- \leq \theta_1 \leq \theta_+$, we see

$$\frac{\lambda_i}{\lambda_j}\cdot\frac{\theta_- - 1 + \lambda_j^2}{\theta_- - 1 + \lambda_i^2} \leq \frac{q_i}{q_j} \leq \frac{\lambda_i}{\lambda_j}\cdot\frac{\theta_+ - 1 + \lambda_j^2}{\theta_+ - 1 + \lambda_i^2} \qquad \text{for } i<j.$$

From Theorem 5.1 and $q_i/q_j = \tilde{\lambda}_i/\tilde{\lambda}_j$, we obtain

$$\frac{\lambda_i}{\lambda_j}\cdot\frac{\boldsymbol{\lambda}'\boldsymbol{\lambda} - \lambda_1^2 + \lambda_j^2}{\boldsymbol{\lambda}'\boldsymbol{\lambda} - \lambda_1^2 + \lambda_i^2} \leq \frac{\tilde{\lambda}_i}{\tilde{\lambda}_j} \leq \frac{\lambda_i}{\lambda_j}\cdot\frac{\boldsymbol{\lambda}'\boldsymbol{\lambda} - \lambda_p^2 + \lambda_j^2}{\boldsymbol{\lambda}'\boldsymbol{\lambda} - \lambda_p^2 + \lambda_i^2} \qquad \text{for } i<j.$$

The equality condition can be obtained easily.
(4) Since $\theta_1 = \tilde{\boldsymbol{\lambda}}'\tilde{\boldsymbol{\lambda}}$, the result can be obtained from Theorem 5.1. □

Each of the results (1)–(4) in this theorem states the following properties;
(1) The order and the signs of $\tilde{\lambda}_i$'s coincide with those of λ_i's respectively.
(2) If $\boldsymbol{\lambda}'\boldsymbol{\lambda}$ is large, $\tilde{\boldsymbol{\lambda}}$ is good approximation of $\boldsymbol{\lambda}$. Note that $\tilde{\boldsymbol{\lambda}}$ depends on both the largest eigenvalue and its corresponding eigenvector. In multifactor case, this property does not always hold (see Section 6.1).
(3) Ratio $\tilde{\lambda}_i/\tilde{\lambda}_j$ underestimates λ_i/λ_j $(i<j)$.
(4) Usually $\tilde{\lambda}_i > \lambda_i$, and rarely $\tilde{\lambda}_i < \lambda_i$ ([24]); however, $\tilde{\boldsymbol{\lambda}}'\tilde{\boldsymbol{\lambda}}$ satisfies the above inequalities.

Some properties of $\tilde{\boldsymbol{\lambda}}$ in a case that $\boldsymbol{\lambda}$ has a special form have been described in Sato [24; Theorem 2, Corollaries 2.1–2.4 and §3]. For a ratio between two loadings, inequalities (3) of Theorem 5.2 assert that $\tilde{\lambda}_i/\tilde{\lambda}_j$ underestimates λ_i/λ_j. These inequalities are generalization and improvement on (2) of Theorem 1 in Sato [24]; formerly only $\tilde{\lambda}_1/\tilde{\lambda}_p$ was treated and the upper bound was λ_1/λ_p.

Now we examine the efficiency of the bounds by numerical examples. Eight cases of loading vectors are treated; they contain frequently encountered magnitude of loadings (.7–.85), especially large one (.9) and very small one

(.4) in practice. Table 5.1 shows the cases and results. For loading vectors commonly met, the intervals between the upper bounds and the lower bounds are very short. Further, even if very small loading exists, the interval does not quite widen. Thus the inequalities are effective.

Careful attention should be paid to the ratio of a large loading to a small

Table 5.1. Upper and lower bounds of proposed inequalities for $\tilde{\lambda}_i/\tilde{\lambda}_j$

case	λ_i	λ_j	$\tilde{\lambda}_i$	$\tilde{\lambda}_j$	λ_i/λ_j	lower bound	$\tilde{\lambda}_i/\tilde{\lambda}_j$	upper bound	interval
	.90	.80	.905	.875	1.125	1.026	1.033	1.040	.014
1	.90	.70	.905	.831	1.286	1.074	1.088	1.104	.030
	.80	.70	.875	.831	1.143	1.046	1.053	1.061	.015
	.80	.75	.859	.843	1.067	1.018	1.020	1.022	.004
2	.80	.70	.859	.822	1.143	1.042	1.046	1.050	.008
	.75	.70	.843	.822	1.071	1.023	1.025	1.027	.004
	.80	.70	.847	.817	1.143	1.028	1.037	1.046	.018
3	.80	.60	.847	.768	1.333	1.083	1.102	1.122	.040
	.70	.60	.817	.768	1.167	1.053	1.063	1.073	.020
	.70	.60	.794	.763	1.167	1.029	1.041	1.053	.025
4	.70	.50	.794	.708	1.400	1.095	1.121	1.149	.055
	.60	.50	.763	.708	1.200	1.064	1.077	1.091	.027
	.90	.85	.910	.886	1.059	1.025	1.026	1.028	.003
	.90	.80	.910	.860	1.125	1.055	1.058	1.061	.006
	.90	.75	.910	.830	1.200	1.091	1.096	1.100	.009
5	.85	.80	.886	.860	1.062	1.029	1.031	1.032	.003
	.85	.75	.886	.830	1.133	1.065	1.068	1.071	.006
	.80	.75	.860	.830	1.067	1.034	1.036	1.037	.003
	.85	.80	.875	.852	1.062	1.026	1.028	1.029	.003
	.85	.75	.875	.825	1.133	1.058	1.062	1.065	.007
	.85	.70	.875	.794	1.214	1.097	1.102	1.108	.010
6	.80	.75	.852	.825	1.067	1.031	1.033	1.034	.003
	.80	.70	.852	.794	1.143	1.069	1.073	1.076	.007
	.75	.70	.825	.794	1.071	1.037	1.039	1.040	.003
	.80	.75	.851	.842	1.067	1.006	1.010	1.022	.016
7	.80	.40	.851	.631	2.000	1.295	1.347	1.479	.184
	.75	.40	.842	.631	1.875	1.288	1.334	1.447	.160
	.80	.75	.842	.821	1.067	1.022	1.025	1.031	.009
	.80	.70	.842	.797	1.143	1.050	1.056	1.069	.019
	.80	.40	.842	.557	2.000	1.482	1.512	1.588	.107
8	.75	.70	.821	.797	1.071	1.028	1.030	1.037	.009
	.75	.40	.821	.557	1.875	1.450	1.476	1.540	.091
	.70	.40	.797	.557	1.750	1.411	1.432	1.485	.075

treated factor loadings $\boldsymbol{\lambda}$

case 1 (.90 .80 .70)′; case 5 (.90 .85 .80 .75)′
case 2 (.80 .75 .70)′; case 6 (.85 .80 .75 .70)′
case 3 (.80 .70 .60)′; case 7 (.80 .75 .40)′
case 4 (.70 .60 .50)′; case 8 (.80 .75 .70 .40)′

one. For instance, in case 7 in Table 5.1,

$$(\tilde{.8}/\tilde{.4})/(.8/.4) = 1.347/2.000 = .67.$$

Here, a figure with a symbol $\sim$ denotes a value calculated with PCA. In case 8,

$$(\tilde{.8}/\tilde{.4})/(.8/.4) = 1.512/2.000 = .76 (> .67 \text{ in case 7}).$$

The length $\boldsymbol{\lambda}'\boldsymbol{\lambda}$ of case 8 is larger than that of case 7, so that $(\tilde{.8}/\tilde{.4})/(.8/.4)$ of case 8 is nearer to one.

Table 5.2. Ratio $(\tilde{\lambda}_1/\tilde{\lambda}_2)/(\lambda_1/\lambda_2)$ when $\boldsymbol{\lambda} = (\underbrace{\lambda_1,\ldots,\lambda_1}_{p'}, \underbrace{\lambda_2,\ldots,\lambda_2}_{p'})'$

	λ_1/λ_2			
p'	1.5	2	3	4
2	.820	.722	.623	.577
3	.877	.809	.742	.713
4	.907	.855	.805	.783
5	.925	.883	.843	.826
10	.962	.941	.920	.912
15	.975	.960	.947	.941
20	.981	.970	.960	.956

Let us consider the simple case;

$$\boldsymbol{\lambda} = (\underbrace{\lambda_1,\ldots,\lambda_1}_{p_1}, \underbrace{\lambda_2,\ldots,\lambda_2}_{p_2})',$$

where $\lambda_1 > \lambda_2 > 0$ and $p_1 + p_2 \geq 3$. In this structure, λ_1 and λ_2 do not effect t/τ individually, where $t = \tilde{\lambda}_1/\tilde{\lambda}_2$ and $\tau = \lambda_1/\lambda_2$ (Theorem 2 in [24]). In particular, when $p_1 = p_2$, say p',

$$t/\tau = 2/((1 - 1/p')(1 - \tau^2) + \{(1 + \tau^4)(p' - 1)^2/p'^2 + 2\tau^2(p'^2 + 2p' - 1)/p'^2\}^{1/2}).$$

Table 5.2 provides the values of t/τ for various values of p' and τ.

Now we will study sample behavior of $\tilde{\lambda}_i/\tilde{\lambda}_j$. Let $\hat{\boldsymbol{\lambda}} = (\hat{\lambda}_1,\ldots,\hat{\lambda}_p)'$ be factor loadings estimated with FA based on a sample variance-covariance matrix S of sample size $n+1$ from a p-variate normal distribution $(n \geq p)$. Then

$$\hat{\lambda} = \sqrt{\hat{\theta}_1} f,$$

where $\hat{\theta}_1$ is the largest eigenvalue of S, $f = (f_1, \dots, f_p)'$ is the eigenvector corresponding to $\hat{\theta}_1$, $f_1 > 0$ and $f'f = 1$.

THEOREM 5.3. *Expectation and variance of* $\hat{\lambda}_i/\hat{\lambda}_j$ *are given as follows*;

$$\text{(5.2)} \quad \mathrm{E}\{\hat{\lambda}_i/\hat{\lambda}_j\} = \tilde{\lambda}_i/\tilde{\lambda}_j + O(n^{-1}).$$

$$\begin{aligned}\text{(5.3)} \quad \mathrm{V}\{\hat{\lambda}_i/\hat{\lambda}_j\} = \frac{1}{n}\Bigg(&\frac{1}{h_{j1}^2}\Bigg(\sum_{\alpha=2}^{p}\frac{h_{i\alpha}^2}{\theta_{1\alpha}^2}(\sigma_{\alpha\alpha}\sigma_{11} + \sigma_{\alpha 1}^2) \\ &+ 2\sum_{2=\alpha<\beta}^{p}\frac{h_{i\alpha}h_{i\beta}}{\theta_{1\alpha}\theta_{1\beta}}(\sigma_{\alpha\beta}\sigma_{11} + \sigma_{\alpha 1}\sigma_{\beta 1})\Bigg) \\ &- 2\frac{h_{i1}}{h_{j1}^3}\Bigg(\sum_{\alpha=2}^{p}\frac{h_{i\alpha}h_{j\alpha}}{\theta_{1\alpha}^2}(\sigma_{\alpha\alpha}\sigma_{11} + \sigma_{\alpha 1}^2) + \sum_{\substack{\alpha\neq\beta \\ \alpha,\beta=2}}^{p}\frac{h_{i\alpha}h_{i\beta}}{\theta_{1\alpha}\theta_{1\beta}}(\sigma_{\alpha\beta}\sigma_{11} + \sigma_{\alpha 1}\sigma_{\beta 1})\Bigg) \\ &+ \frac{h_{i1}^2}{h_{j1}^4}\Bigg(\sum_{\alpha=2}^{p}\frac{h_{j\alpha}^2}{\theta_{1\alpha}^2}(\sigma_{\alpha\alpha}\sigma_{11} + \sigma_{\alpha 1}^2) + 2\sum_{2=\alpha<\beta}^{p}\frac{h_{j\alpha}h_{j\beta}}{\theta_{1\alpha}\theta_{1\beta}}(\sigma_{\alpha\beta}\sigma_{11} + \sigma_{\alpha 1}\sigma_{\beta 1})\Bigg)\Bigg) \\ &+ O(n^{-2}),\end{aligned}$$

where $H = (h_{ij})$ *is the orthogonal matrix such that*

$$H'\Sigma H = \operatorname{diag}\{\theta_1 \cdots \theta_p\},$$

$\theta_1 > \theta_2 \geq \theta_3 \geq \cdots \geq \theta_p$ *are the eigenvalues of* Σ *and* $\theta_{\alpha\beta} = \theta_\alpha - \theta_\beta$. □

PROOF. Let

$$H'SH = \Gamma + V/\sqrt{n},$$

where $\Gamma = \operatorname{diag}\{\theta_1 \cdots \theta_p\}$, and let $c = (c_1, \dots, c_p)'$ be the eigenvector corresponding to the largest eigenvalue g of $H'SH$ where $c_1 > 0$ and $c'c = 1$.

Since all the elements of Σ are positive, owing to Perron's theorem (see e.g. §1c(xi) of Rao [21]), θ_1 is a simple root. Therefore, an asymptotic expansion of the eigenvector is given as follows (see e.g. (6.1) of Sugiura [29]):

$$\text{(5.4)} \quad c_1 = 1 + \frac{1}{n}\left(-\frac{1}{2}\sum_{\alpha=2}^{p}\frac{V_{1\alpha}^2}{\theta_{1\alpha}^2}\right) + O_p(n^{-3/2}),$$

$$\text{(5.5)} \quad c_j = \frac{1}{\theta_{1j}}\left(\frac{V_{j1}}{\sqrt{n}} + \frac{1}{n}\left\{\sum_{\alpha=2}^{p}\frac{V_{j\alpha}V_{\alpha 1}}{\theta_{1\alpha}} - \frac{V_{j1}V_{11}}{\theta_{1j}}\right\}\right) + O_p(n^{-3/2}), \qquad \text{for } j \neq 1,$$

where $V = (V_{ij})$.

Multiplying H on the lefthand side to $H'SHc = gc$, we obtain $SHc = gHc$. This means that Hc is an eigenvector of S;

$$f = Hc. \tag{5.6}$$

Substituting (5.4) and (5.5) to (5.6), we obtain

$$\begin{aligned} f_i &= h_{i1}c_1 + \sum_{\alpha=2}^{p} h_{i\alpha}c_\alpha \\ &= h_{i1} + \frac{1}{\sqrt{n}}v_i^{(1)} + \frac{1}{n}\{\text{homogeneous polynomial of degree 2 in } V_{ij}\text{'s}\}, \end{aligned}$$

where $v_i^{(1)} = \sum_{\alpha=2}^{p} h_{i\alpha}V_{\alpha 1}/\theta_{1\alpha}$.

Since $\hat{\lambda}_i/\hat{\lambda}_j = f_i/f_j$ and $\tilde{\lambda}_i/\tilde{\lambda}_j = h_{i1}/h_{j1}$, we have

$$\begin{aligned} &\sqrt{n}(\hat{\lambda}_i/\hat{\lambda}_j - \tilde{\lambda}_i/\tilde{\lambda}_j) \\ &= \sqrt{n}(f_i/f_j - h_{i1}/h_{j1}) \\ &= \frac{v_i^{(1)}}{h_{j1}} - \frac{h_{i1}}{h_{j1}^2}v_j^{(1)} + \frac{1}{\sqrt{n}}\{\text{homogeneous polynomial of degree 2 in } V_{ij}\text{'s}\} \\ &\quad + O_p(1/n). \end{aligned} \tag{5.7}$$

Noting that $\mathrm{E}\{V_{ab}\} = 0$, we obtain

$$\mathrm{E}\{\sqrt{n}(\hat{\lambda}_i/\hat{\lambda}_j - \tilde{\lambda}_i/\tilde{\lambda}_j)\} = O(n^{-1/2}),$$

which proves (5.2).

Next, we calculate the variance. Noting that $\mathrm{E}\{V_{ab}V_{cd}V_{ef}\} = O(1/\sqrt{n})$ (see Siotani, Hayakawa and Fujikoshi [27; Problem 4.3.4]), from (5.7) we have

$$\begin{aligned} &\mathrm{V}\{\sqrt{n}(\hat{\lambda}_i/\hat{\lambda}_j - \tilde{\lambda}_i/\tilde{\lambda}_j)\} \\ &= \mathrm{V}\left\{\frac{1}{h_{j1}}\sum_{\alpha=2}^{p}\frac{h_{i\alpha}V_{\alpha 1}}{\theta_{1\alpha}} - \frac{h_{i1}}{h_{j1}^2}\sum_{\alpha=2}^{p}\frac{h_{j\alpha}V_{\alpha 1}}{\theta_{1\alpha}}\right\} + O(n^{-1}) \\ &= \frac{1}{h_{j1}^2}\mathrm{E}\left\{\left(\sum_{\alpha=2}^{p}\frac{h_{i\alpha}V_{\alpha 1}}{\theta_{1\alpha}}\right)^2\right\} - \frac{2h_{i1}}{h_{j1}^3}\mathrm{E}\left\{\left(\sum_{\alpha=2}^{p}\frac{h_{i\alpha}V_{\alpha 1}}{\theta_{1\alpha}}\right)\left(\sum_{\alpha=2}^{p}\frac{h_{j\alpha}V_{\alpha 1}}{\theta_{1\alpha}}\right)\right\} \\ &\quad + \frac{h_{i1}^2}{h_{j1}^4}\mathrm{E}\left\{\left(\sum_{\alpha=2}^{p}\frac{h_{j\alpha}V_{\alpha 1}}{\theta_{1\alpha}}\right)^2\right\} \\ &\quad + O(n^{-1}). \end{aligned}$$

Result (5.3) is obtained by using the well-known formula $\mathrm{E}\{V_{ab}V_{cd}\} = \sigma_{ac}\sigma_{bd} + \sigma_{ad}\sigma_{bc}$. □

6. Multifactor case

6.1. Examination of the method of deciding the number of factors

We examine the rule where factor size is taken as the number of eigenvalues of a correlation matrix which are greater than one.

THEOREM 6.1. *Suppose that a population correlation matrix P has a structure*

$$P = \Lambda\Lambda' + \Psi,$$

where Λ is a $p \times k$ $(p > k)$ matrix of rank k and Ψ is a diagonal matrix with positive diagonal elements. Then, the number of eigenvalues θ_i of P greater than one is at most k. $\square$

Before a proof is described, a lemma is introduced.

LEMMA 6.1. (Wilkinson ([39, pp. 97–98]) *Suppose that A is a p-order symmetric matrix and let*

$$B = A + \mathrm{diag}\,\{d\ 0 \cdots 0\}.$$

Let $\tau_1 \geq \cdots \geq \tau_p$ and $t_1 \geq \cdots \geq t_p$ be the eigenvalues of A and B, respectively. Then

$$t_i = \tau_i + dw_i,$$

where $0 \leq w_i \leq 1$ and $\sum_{i=1}^{p} w_i = 1$. $\square$

PROOF of Theorem 6.1. Let $u_1 \geq \cdots \geq u_p$ be the eigenvalues of $P + I - \Psi$. Noting that

$$\begin{aligned} P + I - \Psi = P &+ \mathrm{diag}\,\{1-\psi_1\quad 0 \cdots 0\} + \mathrm{diag}\,\{0\quad 1-\psi_2\quad 0 \cdots 0\} \\ &+ \cdots + \mathrm{diag}\,\{0 \cdots 0\quad 1-\psi_p\}, \end{aligned}$$

and using Lemma 6.1 successively, we obtain

$$u_i \geq \theta_i \qquad \text{for } i = 1, \ldots, p.$$

Since

$$P + I - \Psi = \Lambda\Lambda' + I$$

and $\Lambda\Lambda'$ is a positive semidefinite matrix of rank k, the number of $u_i > 1$ is k. Therefore, the number of $\theta_i > 1$ is at most k. $\square$

Table 6.1. Substitute use of PCA: an inappropriate example

$$\Lambda = \begin{bmatrix} .3 & .0 \\ .3 & .0 \\ .9 & .4 \\ .9 & .4 \\ .9 & -.4 \\ .9 & -.4 \end{bmatrix}, \quad \tilde{\Lambda} = \begin{bmatrix} .40 & .67 & .62 & .0 \\ .40 & -.67 & .62 & .0 \\ .89 & .0 & -.14 & .41 \\ .89 & .0 & -.14 & .41 \\ .89 & .0 & -.14 & -.41 \\ .89 & .0 & -.14 & -.41 \end{bmatrix}.$$

$$\theta_i = \quad 3.15 \quad .91 \quad .85 \quad .67$$

$P = \Lambda\Lambda' + \text{diag}\,(I - \Lambda\Lambda')$. The other eigenvalues are .03 (multiple).

This theorem states that the number of $\theta_i > 1$ is *at most k*. Table 6.1 provides an example where this number is less than k. Further, we note that even if we know the true value k and take as largest k eigenvalues, the corresponding loadings may not be an appropriate approximate of Λ. Table 6.1 demonstrates an example; The loadings corresponding to the *fourth* eigenvalue are the appropriate values for the *second* column of Λ.

Further, if we make a sample correlation matrix, the $(k + 1)$-th eigenvalue will be sometimes more than one by sampling fluctuation.

6.2. Properties of the loadings calculated with PCA

First we treat *a complete simple structure.* This structure is reduced to a combination of monofactor cases by changing order of variables. Therefore, the remarks on a monofactor, which are given in Sato [24], are also valid. Further, it may be noted that even if some loadings of the FA model equal, the loadings corresponding to the same one calculated with PCA differ, because the latter ones depend on other loadings and the number of variables. Therefore, when we compare loadings among some factors, we must pay attention to this property. We provide some examples, relating to such a property.

EXAMPLE 6.1. In the following examples, $P = \Lambda\Lambda' + \text{diag}\,(I - \Lambda\Lambda')$.

$$\text{If } \Lambda = \begin{bmatrix} .7 & .35 & .05 & .0 & .0 & .0 & .0 & .0 & .0 & .0 & .0 & .0 & .0 \\ .0 & .0 & .0 & .7 & .35 & .9 & .9 & .9 & .9 & .9 & .9 & .9 & .9 \end{bmatrix}',$$

then, $\tilde{\Lambda} = \begin{bmatrix} \mathit{.785} & \mathit{.779} & .164 & .0 & .0 & .0 & \cdots & .0 \\ .0 & .0 & .0 & \mathit{.741} & \mathit{.391} & .909 & \cdots & .909 \end{bmatrix}'$. For the same value .35 in the FA model, the value calculated with PCA in the first column is about twice as that in the second column; $.779 \fallingdotseq 782 = .391 \times 2$.

$$\text{If } \Lambda = \begin{bmatrix} .5 & .5 & .5 & .0 & .0 & .0 & .0 & .0 \\ .0 & .0 & .0 & .6 & .6 & .6 & .6 & .6 \end{bmatrix}',$$

then, $.\tilde{5} = .707 > .699 = .\tilde{6}$. Here, a figure with a symbol $\sim$ denotes a value calculated with PCA.

$$\text{If } \Lambda = \begin{bmatrix} .4 & .4 & .4 & .0 & .0 & .0 & .0 & .0 & .0 & .0 & .0 \\ .0 & .0 & .0 & .6 & .6 & .6 & .6 & .6 & .6 & .6 & .6 \end{bmatrix}',$$

then, $.\tilde{4} = .663 = .\tilde{6}$.

Hereafter we investigate structures which are not complete simple.

6.2.1. A treated form and problems of rotation

Consider a structure which we encounter very often in the analysis of empirical data; many variables are affected by only one factor and few are by more than one. As a simple case, we investigate precisely the following structure:

$$\Lambda = \begin{bmatrix} \lambda_1 & \lambda_2 & \cdots & \lambda_{p_1} & \alpha & 0 & 0 & \cdots & 0 \\ 0 & 0 & \cdots & 0 & \beta & \nu_1 & \nu_2 & \cdots & \nu_{p_2} \end{bmatrix}', \tag{6.1}$$

where $p_1 \geq 2$ and $p_2 \geq 2$.

We are interested to know whether $\tilde{\Lambda}$ is near to Λ or not. Since Λ has so many parameters, we treat more simple case; suppose $\lambda_1 = \cdots = \lambda_{p_1}$, say λ, and $\nu_1 = \cdots = \nu_{p_2}$, say ν. To judge whether $\tilde{\Lambda}$ is near to Λ or not in the sense of configuration, we pay attention to the following indices, which researchers are interested in:

$$\text{(I1) } \beta/\alpha \quad \text{(I2) } \alpha/\lambda \text{ or } \beta/\nu \text{ and } \quad \text{(I3) } \lambda/\nu.$$

We compare $\tilde{\beta}/\tilde{\alpha}$ with β/α, $\tilde{\alpha}/\tilde{\lambda}$ with α/λ and so on.

Before starting an argument, it is necessary to determine which rotation should be adopted, since there exists indeterminacy of a rotation for a loading matrix in multifactor cases. A varimax or a quartimax rotation, which is widely used, is not suitable for structure (6.1); more precisely, the criteria of these rotations are not optimum for (6.1). Further, for a Procrustes rotation, whose criterion is minimizing the sum of squares of differences between the corresponding elements of a matrix ΛT and a predetermined target matrix where $TT' = I$, it is very difficult to specify a target matrix.

An appropriate rotation for the present study is proposed as follows:

ALGORITHM (varimax rotation for simple structure variables)

(1) Omit the row in which α and β exist from the loading matrix.

(2) Calculate the varimax rotation matrix for the current $(p-1)$-rowed loading matrix.

(3) Multiply this rotation matrix to the original p-rowed loading matrix.

In a practical situation, researchers have the following information on treated data: which variables are affected by only one factor. Hence, programming for this algorithm is easy. We applied this rotation, say, *varimax rotation for simple structure variables*, to numerical examples in the present paper.

EXAMPLE 6.2. We try to clarify validity of some rotations. Let

$$\Lambda = \begin{bmatrix} \lambda & \lambda & .4 & 0 & 0 \\ 0 & 0 & .8 & .7 & .7 \end{bmatrix}', \text{ where } \lambda = .1\,(.2)\,.9.$$

In order to discuss not substitute use but a rotation problem, we treat not $\tilde{\Lambda}$ but Λ. Four kinds of rotations are performed;

V: varimax rotation,

Q: quartimax rotation,

P_1: Procrustes rotation with a target matrix

$$\begin{bmatrix} \lambda & \lambda & .8 & 0 & 0 \\ 0 & 0 & .4 & .7 & .7 \end{bmatrix}',$$

P_2: Procrustes rotation with a target matrix

$$\begin{bmatrix} .7 & .7 & .8 & 0 & 0 \\ 0 & 0 & .4 & \lambda & \lambda \end{bmatrix}',$$

P_3: Procrustes rotation with a target matrix

$$\begin{bmatrix} \lambda & \lambda & .6 & 0 & 0 \\ 0 & 0 & .6 & .7 & .7 \end{bmatrix}',$$

Proposed: the proposed rotation.

In order to examine adequacy of these rotations, we calculate the following indices;

(I1) $(\lambda_{32}^*/\lambda_{31}^*)/(\lambda_{32}/\lambda_{31})$, (I2) $(\lambda_{31}^*/\lambda_{11}^*)/(\lambda_{31}/\lambda_{11})$, $(\lambda_{32}^*/\lambda_{52}^*)/(\lambda_{32}/\lambda_{52})$ and

(I3) $(\lambda_{11}^*/\lambda_{52}^*)/(\lambda_{11}/\lambda_{52})$,

where λ_{ij}^* is the (i, j) element of the rotated loading matrix. Desirable values are 1.000. Table 6.2 presents the results; None of the rotations except the proposed method (*the varimax rotation for simple structure variables*) are appropriate for (I1) and (I2).

6.2.2. Numerical Experiments

The aim of the following experiments is to compare Λ with $\tilde{\Lambda}$ from the viewpoint of the above indices.

EXPERIMENT 6.1. Suppose

Table 6.2. Validity of various rotations

λ	V	Q	P_1	P_2	P_3	Proposed
			$(\lambda_{32}^*/\lambda_{31}^*)/(\lambda_{32}/\lambda_{31})$			
.1	1.140	1.181	.538	.362	.727	1.000
.3	1.140	1.181	.565	.508	.746	1.000
.5	1.140	1.181	.610	.600	.776	1.000
.7	1.140	1.181	.663	.663	.810	1.000
.9	1.140	1.181	.714	.709	.842	1.000
			$(\lambda_{31}^*/\lambda_{11}^*)/(\lambda_{31}/\lambda_{11})$			
.1	.900	.874	1.585	2.093	1.279	1.000
.3	.900	.874	1.533	1.649	1.255	1.000
.5	.900	.874	1.453	1.471	1.218	1.000
.7	.900	.874	1.369	1.369	1.179	1.000
.9	.900	.874	1.296	1.304	1.145	1.000
			$(\lambda_{32}^*/\lambda_{52}^*)/(\lambda_{32}/\lambda_{52})$			
.1	1.025	1.032	.854	.739	.930	1.000
.3	1.025	1.032	.867	.838	.936	1.000
.5	1.025	1.032	.887	.882	.945	1.000
.7	1.025	1.032	.908	.908	.955	1.000
.9	1.025	1.032	.926	.924	.964	1.000
			$(\lambda_{11}^*/\lambda_{52}^*)/(\lambda_{11}/\lambda_{52})$			
.1	1.000	1.000	1.000	1.000	1.000	1.000
.3	1.000	1.000	1.000	1.000	1.000	1.000
.5	1.000	1.000	1.000	1.000	1.000	1.000
.7	1.000	1.000	1.000	1.000	1.000	1.000
.9	1.000	1.000	1.000	1.000	1.000	1.000

NOTE λ_{ij}^*: an element of the rotated matrix

$$\Lambda = \begin{bmatrix} \lambda & \lambda & \cdots & \lambda & \alpha & 0 & 0 & \cdots & 0 \\ \underbrace{0 \quad 0 \quad \cdots \quad 0}_{p_1} & & & & \beta & \underbrace{\nu \quad \nu \quad \cdots \quad \nu}_{p_2} & & & \end{bmatrix}',$$

$P = \Lambda\Lambda' + \mathrm{diag}\,(I - \Lambda\Lambda')$ and let

$$\tilde{\Lambda} = \begin{bmatrix} \tilde{\lambda} & \tilde{\lambda} & \cdots & \tilde{\lambda} & \tilde{\alpha} & \tilde{e} & \tilde{e} & \cdots & \tilde{e} \\ \tilde{e} & \tilde{e} & \cdots & \tilde{e} & \tilde{\beta} & \tilde{\nu} & \tilde{\nu} & \cdots & \tilde{\nu} \end{bmatrix}'.$$

(1) Set $p_1 = p_2 = 2, 3$ and 5; $\alpha = .5$, $\beta = .5$, $(\beta/\alpha = 1)$, $\alpha = .4$, $\beta = .8$ $(\beta/\alpha = 2)$, $\alpha = .3$, $\beta = .9$ $(\beta/\alpha = 3)$ and $\alpha = .24$, $\beta = .96$ $(\beta/\alpha = 4)$; $\lambda = \nu = .1\,(.2).9$.

(2) Set $p_1 = p_2 = 2, 3$ and 5; combinations of α and β are the same as (1); $\lambda = .7$ (fixed), $\nu = .1\,(.2).9$.

(3) Set $p_1 = 2$ (fixed), $p_2 = 3, 4$ and 6; combinations of α and β are the

Table 6.3. Validity of loadings calculated with PCA from the viewpoint of some indices

(1)	$p_1 = p_2 = 2$				$p_1 = p_2 = 3$				$p_1 = p_2 = 5$			
	β/α				β/α				β/α			
λ	1	2	3	4	1	2	3	4	1	2	3	4
	$(\tilde\beta/\tilde\alpha)/(\beta/\alpha)$				$(\tilde\beta/\tilde\alpha)/(\beta/\alpha)$				$(\tilde\beta/\tilde\alpha)/(\beta/\alpha)$			
.1	1.000	.979	.972	.970	1.000	.982	.976	.974	1.000	.986	.982	.980
.3	1.000	.993	.991	.991	1.000	.996	.995	.995	1.000	.998	.998	.998
.5	1.000	.998	.998	.998	1.000	.999	.999	.999	1.000	1.000	1.000	1.000
.7	1.000	.999	.999	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
.9	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
	$(\tilde\alpha/\tilde\lambda)/(\alpha/\lambda)$				$(\tilde\alpha/\tilde\lambda)/(\alpha/\lambda)$				$(\tilde\alpha/\tilde\lambda)/(\alpha/\lambda)$			
.1	.164	.124	.115	.110	.197	.150	.140	.134	.248	.191	.179	.172
.3	.473	.373	.352	.338	.539	.437	.414	.399	.626	.525	.501	.486
.5	.741	.608	.578	.559	.796	.676	.648	.630	.856	.756	.731	.751
.7	.963	.815	.781	.759	.979	.859	.831	.811	.989	.904	.882	.866
.9	1.143	.990	.954	.931	1.106	.996	.968	.949	1.069	.999	.980	.966
	$(\tilde\beta/\tilde\nu)/(\beta/\nu)$				$(\tilde\beta/\tilde\nu)/(\beta/\nu)$				$(\tilde\beta/\tilde\nu)/(\beta/\nu)$			
.1	.164	.142	.139	.135	.197	.171	.167	.162	.248	.215	.209	.203
.3	.473	.407	.393	.381	.539	.467	.451	.438	.626	.549	.531	.516
.5	.741	.638	.614	.595	.796	.607	.673	.655	.856	.768	.746	.728
.7	.963	.835	.803	.780	.979	.871	.843	.822	.989	.909	.887	.871
.9	1.143	1.000	.964	.937	1.106	1.001	.973	.952	1.069	1.001	.982	.967
	$(\tilde\lambda/\tilde\nu)/(\lambda/\nu)$				$(\tilde\lambda/\tilde\nu)/(\lambda/\nu)$				$(\tilde\lambda/\tilde\nu)/(\lambda/\nu)$			
.1	1.000	1.173	1.237	1.263	1.000	1.159	1.216	1.240	1.000	1.139	1.188	1.208
.3	1.000	1.098	1.128	1.139	1.000	1.074	1.096	1.103	1.000	1.048	1.062	1.065
.5	1.000	1.052	1.064	1.067	1.000	1.032	1.040	1.040	1.000	1.016	1.019	1.019
.7	1.000	1.025	1.029	1.028	1.000	1.014	1.015	1.014	1.000	1.006	1.006	1.005
.9	1.000	1.011	1.010	1.007	1.000	1.006	1.005	1.003	1.000	1.002	1.002	1.001

(2)	$p_1 = p_2 = 2$				$p_1 = p_2 = 3$				$p_1 = p_2 = 5$			
	β/α				β/α				β/α			
ν	1	2	3	4	1	2	3	4	1	2	3	4
	$(\tilde\beta/\tilde\alpha)/(\beta/\alpha)$				$(\tilde\beta/\tilde\alpha)/(\beta/\alpha)$				$(\tilde\beta/\tilde\alpha)/(\beta/\alpha)$			
.1	.280	.351	.361	.385	.387	.429	.481	.501	.531	.556	.595	.604
.3	.660	.638	.640	.637	.771	.726	.723	.716	.871	.805	.795	.786
.5	.868	.839	.833	.828	.922	.834	.877	.871	.963	.924	.915	.908
.7	1.000	.999	.999	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
.9	1.091	1.136	1.148	1.155	1.045	1.089	1.100	1.108	1.018	1.052	1.061	1.068
	$(\tilde\alpha/\tilde\lambda)/(\alpha/\lambda)$				$(\tilde\alpha/\tilde\lambda)/(\alpha/\lambda)$				$(\tilde\alpha/\tilde\lambda)/(\alpha/\lambda)$			
.1	1.202	1.267	1.320	1.330	1.130	1.146	1.159	1.157	1.070	1.064	1.065	1.063
.3	1.094	1.029	1.022	1.012	1.049	.991	.982	.973	1.020	.974	.966	.959
.5	1.015	.898	.873	.855	1.004	.910	.889	.874	.979	.930	.914	.902
.7	.963	.815	.781	.759	.979	.859	.831	.811	.989	.904	.882	.866
.9	.929	.758	.718	.693	.963	.825	.791	.768	.984	.888	.861	.843

Table 6.3. (Continued)

λ	1	2	3	4	1	2	3	4	1	2	3	4
	$(\tilde{\beta}/\tilde{v})/(\beta/v)$				$(\tilde{\beta}/\tilde{v})/(\beta/v)$				$(\tilde{\beta}/\tilde{v})/(\beta/v)$			
.1	.055	.069	.088	.100	.086	.103	.124	.135	.139	.154	.174	.181
.3	.359	.349	.361	.361	.457	.425	.428	.424	.576	.521	.515	.507
.5	.690	.614	.601	.588	.767	.682	.665	.649	.843	.760	.741	.725
.7	.963	.835	.803	.780	.979	.871	.843	.822	.989	.909	.887	.871
.9	1.176	1.017	.972	.942	1.119	1.010	.977	.955	1.073	1.005	.984	.969
	$(\tilde{\lambda}/\tilde{v})/(\lambda/v)$				$(\tilde{\lambda}/\tilde{v})/(\lambda/v)$				$(\tilde{\lambda}/\tilde{v})/(\lambda/v)$			
.1	.165	.173	.186	.196	.197	.209	.223	.232	.245	.261	.275	.283
.3	.498	.532	.551	.561	.565	.591	.603	.609	.648	.664	.671	.673
.5	.783	.815	.826	.829	.829	.847	.852	.853	.876	.884	.886	.886
.7	1.000	1.025	1.029	1.028	1.000	1.014	1.015	1.014	1.000	1.006	1.006	1.005
.9	1.160	1.181	1.180	1.176	1.111	1.124	1.123	1.121	1.070	1.076	1.076	1.075

(3)	$p_1 = 2, p_2 = 3$				$p_1 = 2, p_2 = 4$				$p_1 = 2, p_2 = 6$			
	β/α				β/α				β/α			
λ	1	2	3	4	1	2	3	4	1	2	3	4
	$(\tilde{\beta}/\tilde{\alpha})/(\beta/\alpha)$				$(\tilde{\beta}/\tilde{\alpha})/(\beta/\alpha)$				$(\tilde{\beta}/\tilde{\alpha})/(\beta/\alpha)$			
.1	1.131	1.133	1.133	1.134	1.226	1.256	1.263	1.268	1.461	1.600	1.634	1.653
.3	1.035	1.074	1.083	1.090	1.054	1.133	1.154	1.167	1.079	1.280	1.338	1.374
.5	.985	1.035	1.047	1.056	.974	1.061	1.084	1.100	.950	1.119	1.173	1.211
.7	.956	1.007	1.020	1.029	.932	1.012	1.035	1.051	.892	1.025	1.069	1.101
.9	.938	.985	.998	1.007	.906	.977	.997	1.013	.859	.963	.997	1.024
	$(\tilde{\alpha}/\tilde{\lambda})/(\alpha/\lambda)$				$(\tilde{\alpha}/\tilde{\lambda})/(\alpha/\lambda)$				$(\tilde{\alpha}/\tilde{\lambda})/(\alpha/\lambda)$			
.1	.150	.109	.100	.095	.141	.099	.090	.086	.120	.079	.071	.067
.3	.458	.349	.325	.311	.449	.333	.308	.293	.432	.300	.272	.256
.5	.738	.590	.556	.535	.736	.578	.541	.518	.734	.554	.510	.484
.7	.972	.807	.769	.743	.976	.803	.760	.733	.984	.794	.744	.711
.9	1.160	.994	.954	.926	1.168	.996	.953	.923	1.180	1.001	.952	.917
	$(\tilde{\beta}/\tilde{v})/(\beta/v)$				$(\tilde{\beta}/\tilde{v})/(\beta/v)$				$(\tilde{\beta}/\tilde{v})/(\beta/v)$			
.1	.213	.178	.170	.164	.253	.206	.195	.188	.364	.287	.267	.256
.3	.553	.475	.455	.441	.610	.525	.502	.485	.736	.647	.618	.599
.5	.799	.702	.676	.656	.835	.745	.719	.699	.903	.836	.813	.795
.7	.972	.873	.844	.823	.977	.896	.871	.852	.987	.940	.923	.909
.9	1.095	1.000	.973	.952	1.070	1.000	.978	.961	1.035	1.000	.988	.978
	$(\tilde{\lambda}/\tilde{v})/(\lambda/v)$				$(\tilde{\lambda}/\tilde{v})/(\lambda/v)$				$(\tilde{\lambda}/\tilde{v})/(\lambda/v)$			
.1	1.255	1.443	1.501	1.523	1.465	1.660	1.712	1.731	2.072	2.272	2.309	2.319
.3	1.167	1.267	1.293	1.300	1.289	1.390	1.413	1.418	1.579	1.683	1.699	1.702
.5	1.099	1.151	1.161	1.163	1.163	1.216	1.225	1.227	1.294	1.348	1.357	1.359
.7	1.046	1.073	1.077	1.077	1.074	1.102	1.107	1.106	1.125	1.155	1.160	1.161
.9	1.006	1.021	1.022	1.021	1.011	1.028	1.029	1.029	1.021	1.038	1.041	1.042

same as (1); $\lambda = \nu = .1\,(.2)\,.9$.

Table 6.3 shows the results:

(I1) β/α: If monofactor parts of a complete simple structure are identical, $\tilde{\beta}/\tilde{\alpha}$ approximates β/α for all β/α. Further, as $p_1(=p_2)$ increases, an approximation is closer. If monofactor parts of a complete simple structure are not identical, that is, $\lambda = \nu$ but $p_1 \neq p_2$, or $p_1 = p_2$ but $\lambda \neq \nu$, then $\tilde{\beta}/\tilde{\alpha}$ is far from β/α.

(I2) α/λ or β/ν: If λ or ν is small, $\tilde{\lambda} - \lambda$ and $\tilde{\nu} - \nu$ tend to positive. Values $\tilde{\alpha}/\tilde{\lambda}$ and $\tilde{\beta}/\tilde{\nu}$ are far from α/ν and β/ν, respectively.

Table 6.4. Calculated loadings with PCA for some typical cases

(1–1)

$$\Lambda = \begin{bmatrix} .7 & .7 & .7 & .0 & .0 \\ .0 & .0 & \nu & \nu & \nu \end{bmatrix}'.$$

$$\text{If } \nu = .1,\ \tilde{\Lambda} = \begin{bmatrix} .812 & .812 & .812 & -.002 & -.002 \\ -.002 & -.002 & .026 & .711 & .711 \end{bmatrix}'.$$

comm.	.660	.660	.660	.505	.505
model's comm.	.490	.490	.500	.010	.010

$$\text{If } \nu = .3,\ \tilde{\Lambda} = \begin{bmatrix} .817 & .817 & .798 & -.015 & -.015 \\ -.016 & -.016 & .205 & .733 & .733 \end{bmatrix}'.$$

comm.	.667	.667	.678	.538	.538
model's comm.	.490	.490	.580	.090	.090

$$\text{If } \nu = .5,\ \tilde{\Lambda} = \begin{bmatrix} .832 & .832 & .741 & -.021 & -.021 \\ -.021 & -.021 & .440 & .776 & .776 \end{bmatrix}'.$$

comm.	.692	.692	.743	.602	.602
model's comm.	.490	.490	.740	.250	.250

$$\text{If } \nu = .7,\ \tilde{\Lambda} = \begin{bmatrix} .850 & .850 & .653 & -.013 & -.013 \\ -.013 & -.013 & .653 & .850 & .850 \end{bmatrix}'.$$

comm.	.723	.723	.854	.723	.723
model's comm.	.490	.490	.980	.490	.490

ν	$\dfrac{\tilde{\lambda}_{32}/\tilde{\lambda}_{31}}{\lambda_{32}/\lambda_{31}}$	$\dfrac{\tilde{\lambda}_{31}/\tilde{\lambda}_{11}}{\lambda_{31}/\lambda_{11}}$	$\dfrac{\tilde{\lambda}_{32}/\tilde{\lambda}_{52}}{\lambda_{32}/\lambda_{52}}$	$\dfrac{\tilde{\lambda}_{11}/\tilde{\lambda}_{52}}{\lambda_{11}/\lambda_{52}}$
.1	.224	1.000	.037	.163
.3	.599	.976	.279	.477
.5	.832	.891	.568	.766
.7	1.000	.769	.769	1.000

Compare the above with monofactor cases; Let $\Lambda = (\nu\ \nu\ \nu)'$ and $\tilde{\Lambda} = (\tilde{\nu}\ \tilde{\nu}\ \tilde{\nu})'$.

ν	$\tilde{\nu}$
.1	.583
.3	.627
.5	.707
.7	.812

Table 6.4. (Continued)

(1–2)

$$\Lambda = \begin{bmatrix} .7 & .7 & .7 & .0 & .0 \\ .0 & .0 & .7 & v & v \end{bmatrix}'.$$

$$\text{If } v = .1,\ \tilde{\Lambda} = \begin{bmatrix} .816 & .816 & .801 & -.013 & -.013 \\ -.016 & -.016 & .177 & .706 & .706 \end{bmatrix}'.$$

comm.	.666	.666	.674	.499	.499
model's comm.	.490	.490	.980	.010	.010

$$\text{If } v = .3,\ \tilde{\Lambda} = \begin{bmatrix} .831 & .831 & .746 & -.022 & -.022 \\ -.026 & -.026 & .428 & .720 & .720 \end{bmatrix}'.$$

comm.	.691	.691	.740	.519	.519
model's comm.	.490	.490	.980	.090	.090

$$\text{If } v = .5,\ \tilde{\Lambda} = \begin{bmatrix} .842 & .842 & .694 & -.019 & -.019 \\ -.020 & -.020 & .567 & .772 & .772 \end{bmatrix}'.$$

comm.	.710	.710	.802	.597	.597
model's comm.	.490	.490	.980	.250	.250

$$\text{If } v = .7,\ \tilde{\Lambda} = \begin{bmatrix} .850 & .850 & .653 & -.013 & -.013 \\ -.013 & -.013 & .653 & .850 & .850 \end{bmatrix}'.$$

comm.	.723	.723	.854	.723	.723
model's comm.	.490	.490	.980	.490	.490

$$\text{If } v = .9,\ \tilde{\Lambda} = \begin{bmatrix} .855 & .855 & .623 & -.008 & -.008 \\ -.008 & -.008 & .714 & .943 & .943 \end{bmatrix}'.$$

comm.	.731	.731	.898	.890	.890
model's comm.	.490	.490	.980	.810	.810

v	$\dfrac{\tilde{\lambda}_{32}/\tilde{\lambda}_{31}}{\lambda_{32}/\lambda_{31}}$	$\dfrac{\tilde{\lambda}_{31}/\tilde{\lambda}_{11}}{\lambda_{31}/\lambda_{11}}$	$\dfrac{\tilde{\lambda}_{32}/\tilde{\lambda}_{52}}{\lambda_{32}/\lambda_{52}}$	$\dfrac{\tilde{\lambda}_{11}/\tilde{\lambda}_{52}}{\lambda_{11}/\lambda_{52}}$
.1	.221	.985	.036	.165
.3	.573	.899	.254	.494
.5	.817	.824	.524	.779
.7	1.000	.769	.769	1.000
.9	1.146	.729	.973	1.165

Compare the above with monofactor cases; Let $\Lambda = (.7\ v\ v)'$ and $\tilde{\Lambda} = (.\tilde{7}\ \tilde{v}\ \tilde{v})'$.

v	$.\tilde{7}$	$\tilde{v}$	$(.\tilde{7}/\tilde{v})/(.7/v)$
.1	.724	.538	.192
.3	.756	.622	.521
.5	.786	.713	.787
.7	.812	.812	1.000
.9	.836	.918	1.171

Table 6.4. (Continued)

(2–1)

$$\Lambda = \begin{bmatrix} \lambda & \lambda & \lambda & .0 & .0 \\ .0 & .0 & \lambda & \lambda & \lambda \end{bmatrix}'.$$

$$\text{If } \lambda = .1,\ \tilde{\Lambda} = \begin{bmatrix} .638 & .638 & .441 & -.073 & -.073 \\ -.073 & -.073 & .441 & .638 & .638 \end{bmatrix}'.$$

comm.	.412	.412	.388	.412	.412
model's comm.	.010	.010	.020	.010	.010

$$\text{If } \lambda = .3,\ \tilde{\Lambda} = \begin{bmatrix} .678 & .678 & .483 & -.060 & -.060 \\ -.060 & -.060 & .483 & .678 & .678 \end{bmatrix}'.$$

comm.	.464	.464	.466	.464	.464
model's comm.	.090	.090	.180	.090	.090

$$\text{If } \lambda = .5,\ \tilde{\Lambda} = \begin{bmatrix} .752 & .752 & .557 & -.038 & -.038 \\ -.038 & -.038 & .557 & .752 & .752 \end{bmatrix}'.$$

comm.	.567	.567	.621	.567	.567
model's comm.	.250	.250	.500	.250	.250

$$\text{If } \lambda = .7,\ \tilde{\Lambda} = \begin{bmatrix} .850 & .850 & .653 & -.013 & -.013 \\ -.013 & -.013 & .653 & .850 & .850 \end{bmatrix}'.$$

comm.	.723	.723	.854	.723	.723
model's comm.	.490	.490	.980	.490	.490

λ	$\dfrac{\tilde{\lambda}_{32}/\tilde{\lambda}_{31}}{\lambda_{32}/\lambda_{31}}$	$\dfrac{\tilde{\lambda}_{31}/\tilde{\lambda}_{11}}{\lambda_{31}/\lambda_{11}}$	$\dfrac{\tilde{\lambda}_{32}/\tilde{\lambda}_{52}}{\lambda_{32}/\lambda_{52}}$	$\dfrac{\tilde{\lambda}_{11}/\tilde{\lambda}_{52}}{\lambda_{11}/\lambda_{52}}$
.1	1.000	.691	.691	1.000
.3	1.000	.712	.712	1.000
.5	1.000	.741	.741	1.000
.7	1.000	.769	.769	1.000

Compare the above with monofactor cases; Let $\Lambda = (\lambda\ \lambda\ \lambda)'$ and $\tilde{\Lambda} = (\tilde{\lambda}\ \tilde{\lambda}\ \tilde{\lambda})'$.

λ	$\tilde{\lambda}$
.1	.583
.3	.627
.5	.707
.7	.812

(2–2)

$$\Lambda = \begin{bmatrix} \lambda & \lambda & \lambda/\sqrt{2} & .0 & .0 \\ .0 & .0 & \lambda/\sqrt{2} & \lambda & \lambda \end{bmatrix}'.$$

$$\text{If } \lambda = .1,\ (\lambda/\sqrt{2} = .071),\ \tilde{\Lambda} = \begin{bmatrix} .647 & .647 & .412 & -.064 & -.064 \\ -.064 & -.064 & .412 & .647 & .647 \end{bmatrix}'.$$

comm.	.422	.422	.340	.422	.422
model's comm.	.010	.010	.010	.010	.010

$$\text{If } \lambda = .3,\ (\lambda/\sqrt{2} = .212),\ \tilde{\Lambda} = \begin{bmatrix} .683 & .683 & .443 & -.056 & -.056 \\ -.056 & -.056 & .443 & .683 & .683 \end{bmatrix}'.$$

comm.	.469	.469	.393	.469	.469
model's comm.	.090	.090	.090	.090	.090

Table 6.4. (Continued)

If $\lambda = .5$, $(\lambda/\sqrt{2} = .354)$, $\tilde{\Lambda} = \begin{bmatrix} .749 & .749 & .500 & -.042 & -.042 \\ -.042 & -.042 & .500 & .749 & .749 \end{bmatrix}'$.

comm.	.563	.563	.500	.563	.563
model's comm.	.250	.250	.250	.250	.250

If $\lambda = .7$, $(\lambda/\sqrt{2} = .495)$, $\tilde{\Lambda} = \begin{bmatrix} .838 & .838 & .574 & -.025 & -.025 \\ -.025 & -.025 & .574 & .838 & .838 \end{bmatrix}'$.

comm.	.703	.703	.660	.703	.703
model's comm.	.490	.490	.490	.490	.490

If $\lambda = .9$, $(\lambda/\sqrt{2} = .636)$, $\tilde{\Lambda} = \begin{bmatrix} .943 & .943 & .661 & -.008 & -.008 \\ -.008 & -.008 & .661 & .943 & .943 \end{bmatrix}'$.

comm.	.889	.889	.873	.889	.889
model's comm.	.810	.810	.810	.810	.810

λ	$\dfrac{\tilde{\lambda}_{32}/\tilde{\lambda}_{31}}{\lambda_{32}/\lambda_{31}}$	$\dfrac{\tilde{\lambda}_{31}/\tilde{\lambda}_{11}}{\lambda_{31}/\lambda_{11}}$	$\dfrac{\tilde{\lambda}_{32}/\tilde{\lambda}_{52}}{\lambda_{32}/\lambda_{52}}$	$\dfrac{\tilde{\lambda}_{11}/\tilde{\lambda}_{52}}{\lambda_{11}/\lambda_{52}}$
.1	1.000	.901	.901	1.000
.3	1.000	.919	.919	1.000
.5	1.000	.944	.944	1.000
.7	1.000	.970	.970	1.000
.9	1.000	.991	.991	1.000

Compare the above with monofactor cases; Let $\Lambda = (\lambda \;\; \lambda \;\; \lambda/\sqrt{2})'$ and $\tilde{\Lambda} = (\tilde{\lambda}\;\tilde{\lambda}\;\tilde{\gamma})'$.

λ	$\tilde{\lambda}$	$\tilde{\gamma}$	$(\tilde{\gamma}/\tilde{\lambda})\,/\,((\lambda/\sqrt{2})/\lambda)$
.1	.606	.530	1.236
.3	.644	.563	1.236
.5	.713	.623	1.236
.7	.805	.704	1.236
.9	.914	.799	1.236

(2–3)

$$\Lambda = \begin{bmatrix} .7 & .7 & .7 & .0 & .0 \\ .0 & .0 & \beta & .7 & .7 \end{bmatrix}'.$$

If $\beta = .1$, $\tilde{\Lambda} = \begin{bmatrix} .814 & .814 & .807 & -.007 & -.007 \\ -.007 & -.007 & .116 & .862 & .862 \end{bmatrix}'$.

comm.	.662	.662	.665	.743	.743
model's comm.	.490	.490	.500	.490	.490

If $\beta = .3$, $\tilde{\Lambda} = \begin{bmatrix} .824 & .824 & .771 & -.017 & -.017 \\ -.017 & -.017 & .331 & .856 & .856 \end{bmatrix}'$.

comm.	.679	.679	.704	.733	.733
model's comm.	.490	.490	.580	.490	.490

If $\beta = .5$, $\tilde{\Lambda} = \begin{bmatrix} .838 & .838 & .714 & -.018 & -.018 \\ -.018 & -.018 & .510 & .850 & .850 \end{bmatrix}'$.

comm.	.702	.702	.771	.723	.723
model's comm.	.490	.490	.740	.490	.490

Table 6.4. (Continued)

$$\text{If } \beta = .7,\ \tilde{\Lambda} = \begin{bmatrix} .850 & .850 & .653 & -.013 & -.013 \\ -.013 & -.013 & .653 & .850 & .850 \end{bmatrix}'.$$

comm.	.723	.723	.854	.723	.723
model's comm.	.490	.490	.980	.490	.490

β	$\dfrac{\tilde{\lambda}_{32}/\tilde{\lambda}_{31}}{\lambda_{32}/\lambda_{31}}$	$\dfrac{\tilde{\lambda}_{31}/\tilde{\lambda}_{11}}{\lambda_{31}/\lambda_{11}}$	$\dfrac{\tilde{\lambda}_{32}/\tilde{\lambda}_{52}}{\lambda_{32}/\lambda_{52}}$	$\dfrac{\tilde{\lambda}_{11}/\tilde{\lambda}_{52}}{\lambda_{11}/\lambda_{52}}$
.1	1.002	.992	.938	.944
.3	1.001	.936	.902	.963
.5	1.000	.853	.841	.985
.7	1.000	.769	.769	1.000

Compare the above with monofactor cases; Let $\Lambda = (\beta\ \ .7\ \ .7)'$ and $\tilde{\Lambda} = (\tilde{\beta}\ \ \tilde{.7}\ \ \tilde{.7})'$.

β	$\tilde{\beta}$	$\tilde{.7}$	$(\tilde{\beta}/\tilde{.7})\,/\,(\beta/.7)$
.1	.234	.853	1.924
.3	.544	.817	1.556
.5	.707	.805	1.229
.7	.812	.812	1.000

(I3) λ/ν: If monofactor parts of a complete simple structure resemble each other, $\tilde{\lambda}/\tilde{\nu}$ approximates λ/ν.

Experiment 6.2. The aim of this experiment is to investigate the case of a typical loading form more precisely. Let a general form of loading matrices be

$$\Lambda = \begin{bmatrix} \lambda & \lambda & \alpha & 0 & 0 \\ 0 & 0 & \beta & \nu & \nu \end{bmatrix}' \text{ and } P = \Lambda\Lambda' + \operatorname{diag}(I - \Lambda\Lambda').$$

Then, from Table 6.4 we can see the following properties:

(1) The cases where monofactor parts of a complete simple structure are not identical;

(1–1) Let $\lambda = \alpha = .7$ and $\beta = \nu = .1\,(.2).7$. In this form, the differences between two columns of Λ are larger with decreasing ν. The values of $\tilde{\beta} - \beta$ are negative, on the other hand, the ones of $\tilde{\nu} - \nu$ are positive; consequently, $\tilde{\beta}/\tilde{\nu}$ is far from β/ν. When ν is small, $|\tilde{\nu} - \nu|$ is large and $\tilde{\lambda}/\tilde{\nu}$ is far from λ/ν.

(1–2) Let $\lambda = \alpha = \beta = .7$ and $\nu = .1\,(.2).9$. In this form, the differences between two columns of Λ are smaller than the ones of (1–1). The values of $\tilde{\beta} - \beta$ are negative; $\tilde{\alpha}/\tilde{\beta}$ is far from α/β as ν is away from α. Note that $\tilde{\beta} \ll \tilde{\nu}$ even if $\beta \geq \nu$.

(2) The cases where monofactor parts of a complete simple structure are identical;

(2–1) Let $\lambda = \alpha = \beta = \nu = .1\,(.2).7$. In this case, both the values $\tilde{\alpha} - \alpha$ and $\tilde{\beta} - \beta$ are negative.

(2–2) Let $\alpha=\beta=\lambda/\sqrt{2}$ and $\nu=\lambda=.1(.2).7$. In this form, all the communalities are equal. The values of $\tilde{\alpha}-\alpha$ and $\tilde{\beta}-\beta$ are smaller than those of (2–1).
(2–3) Let $\lambda=\alpha=\nu=.7$ and $\beta=.1(.2).7$. This form is often assumed in practical situations, and researchers wish to know β. We note that $\tilde{\beta}$ is near to β for all β.

Experiments 6.1 and 6.2 deal with only the cases that $\lambda, \nu, \alpha, \beta>0$. However we can assume $\lambda, \nu>0$ without loss of generality. Further if $\alpha<0$ and/or $\beta<0$, the absolute values of the elements of $\tilde{\Lambda}$ are invariant. In fact, if $\alpha<0$ and/or $\beta<0$, then

$$\tilde{\Lambda}=\begin{bmatrix} \tilde{\lambda} & \tilde{\lambda} & \cdots & \tilde{\lambda} & u\tilde{\alpha} & uv\tilde{e} & uv\tilde{e} & \cdots & uv\tilde{e} \\ \underbrace{uv\tilde{e} \quad uv\tilde{e} \quad \cdots \quad uv\tilde{e}}_{p_1} & & & & v\tilde{\beta} & \underbrace{\tilde{\nu} \quad \tilde{\nu} \quad \cdots \quad \tilde{\nu}}_{p_2} \end{bmatrix}',$$

where $u=\operatorname{sgn}\alpha$ and $v=\operatorname{sgn}\beta$. Here $\tilde{\lambda}$, $\tilde{\nu}$, $\tilde{\alpha}$, $\tilde{\beta}$ and $\tilde{e}$ are the elements of $\tilde{\Lambda}$ for the case $\lambda, \nu, \alpha, \beta>0$.

6.2.3. Analytical Results

For some special cases, we can obtain $\tilde{\Lambda}$ explicitly, and, as a result, some properties are obtained.

THEOREM 6.2. *Let*

$$\Lambda=\begin{bmatrix} \lambda & \lambda & \cdots & \lambda & \alpha & 0 & 0 & \cdots & 0 \\ 0 & 0 & \cdots & 0 & \alpha & \lambda & \lambda & \cdots & \lambda \end{bmatrix}', \quad (\text{each block of length } q)$$

where $0<\lambda<1$, $0<\alpha<1/\sqrt{2}$, $p=2q+1$ *and* $q\geq 2$, *and*
$P=\Lambda\Lambda'+\operatorname{diag}(I-\Lambda\Lambda')$.

Then, after being rotated by the method proposed in Section 6.2.1, $\tilde{\Lambda}$ *can be expressed as a following form*:

$$\tilde{\Lambda}=\begin{bmatrix} \ell & \ell & \cdots & \ell & a & e & e & \cdots & e \\ e & e & \cdots & e & a & \ell & \ell & \cdots & \ell \end{bmatrix}',$$

where

$$\ell=\sqrt{\{\{\alpha^2(4+(p-3)\lambda^2+\lambda\sqrt{\{(p-3)^2\lambda^2+16(p-1)\alpha^2\}})\}/}$$
$$\{(p-3)^2\lambda^2+16(p-1)\alpha^2-(p-3)\lambda\sqrt{\{(p-3)^2\lambda^2+16(p-1)\alpha^2\}}\}\}$$
$$+\sqrt{\{(1+(p-3)\lambda^2/2)/(2(p-1))\}},$$
$$e=\sqrt{\{\{\alpha^2(4+(p-3)\lambda^2+\lambda\sqrt{\{(p-3)^2\lambda^2+16(p-1)\alpha^2\}})\}/}$$
$$\{(p-3)^2\lambda^2+16(p-1)\alpha^2-(p-3)\lambda\sqrt{\{(p-3)^2\lambda^2+16(p-1)\alpha^2\}}\}\}$$

$$- \sqrt{\{(1 + (p-3)\lambda^2/2)/(2(p-1))\}} \text{ and}$$

$$a = \sqrt{\{\{(p-1)\alpha^2(4 + (p-3)\lambda^2 + \lambda\sqrt{\{(p-3)^2\lambda^2 + 16(p-1)\alpha^2\}})\}/}$$
$$\{(p-3)^2\lambda^2 + 16(p-1)\alpha^2 + (p-3)\lambda\sqrt{\{(p-3)^2\lambda^2 + 16(p-1)\alpha^2\}}\}\}.$$

The largest and the second eigenvalues $\theta_1 > \theta_2$ of P are given by

$$\theta_1 = 1 + ((p-3)^2\lambda^2 + \lambda\sqrt{\{(p-3)^2\lambda^2 + 16(p-1)\alpha^2\}})/4 > 1,$$
$$\theta_2 = 1 + (p-3)\lambda^2/2 > 1. \quad \square$$

COROLLARY 6.1. *Under the same assumptions as Theorem* 6.2, *it follows that*

(1) $e < 0$.

(2) *The inequality $a > (<) \alpha$ holds according to*

$$\alpha > (<) \sqrt{\{2(p-1)(p-2)\lambda^2 + 4(p-1)^2 + 2(p-1)\lambda\sqrt{\{(p-2)^2\lambda^2} + 4(p-1)\}\}/(4(p-1)). \quad \square$$

Table 6.5 presents the boundary shown in (2) of Corollary 6.1 for $\lambda = .1(.2).9$ and $p = 5(2)21$. The numerical experiment for the cases of $p = 5$, $\lambda = .01(.01).99$ and $\alpha = .1(.1).9$ shows the following: $\tilde{\lambda} < \lambda$ when $\lambda = .97$ for $\alpha = .3-.4$; $\lambda = .98$ for $\alpha = .2-.5$; and $\lambda = .99$ for $\alpha = .2-.6$; otherwise, $\tilde{\lambda} > \lambda$.

Table 6.5. Boundary between $\tilde{\alpha} > \alpha$ and $\tilde{\alpha} < \alpha$, when $\Lambda = \begin{bmatrix} \lambda \cdots \lambda & \alpha & 0 \cdots 0 \\ 0 \cdots 0 & \alpha & \lambda \cdots \lambda \end{bmatrix}'$

	p								
λ	5	7	9	11	13	15	17	19	21
.1	.513	.511	.510	.509	.508	.508	.507	.507	.507
.3	.545	.540	.537	.535	.533	.532	.531	.530	.530
.5	.583	.577	.574	.572	.570	.569	.568	.567	.566
.7	.628	.624	.621	.619	.618	.617	.616	.616	.615
.9	.679	.678	.677	.676	.675	.675	.675	.675	.674

COROLLARY 6.2. *Let*

$$\Lambda = \begin{bmatrix} \lambda & \lambda & \cdots & \lambda & \lambda & 0 & 0 & \cdots & 0 \\ \underbrace{0 \quad 0 \quad \cdots \quad 0}_{q} & & & & \lambda & \underbrace{\lambda \quad \lambda \quad \cdots \quad \lambda}_{q} \end{bmatrix}',$$

where $0 < \lambda < 1/\sqrt{2}$, $p = 2q + 1$ *and* $q \geq 2$, *and* $P = \Lambda\Lambda' + \operatorname{diag}(I - \Lambda\Lambda')$.

Then, after being rotated by the method proposed in Section 6.2.1, $\tilde{\Lambda}$ *can be expressed as a following form*:

$$\tilde{\Lambda} = \begin{bmatrix} \ell & \ell & \cdots & \ell & a & e & e & \cdots & e \\ e & e & \cdots & e & a & \ell & \ell & \cdots & \ell \end{bmatrix}'.$$

Further it holds that

(1) $a > (<) \lambda$ *according to*

$$\lambda > (<) \sqrt{\{(p-1)(7p-5+\sqrt{\{p^2+10p-7\}})/\{2(3p^2-5p+2)\}\}}/2.$$

(2) $\ell > \lambda$. □

Table 6.6 presents the boundary shown in (1) of Corollary 6.2 for $p = 5(2)21$.

Table 6.6. Boundary between $\tilde{\lambda} > \lambda$ and $\tilde{\lambda} < \lambda$, when $\Lambda = \begin{bmatrix} \lambda \cdots \lambda & \lambda & 0 \cdots 0 \\ 0 \cdots 0 & \lambda & \lambda \cdots \lambda \end{bmatrix}'$

p	5	7	9	11	13	15	17	19	21
boundary	.606	.599	.595	.592	.590	.589	.587	.586	.586

6.2.4. Concluding Remarks

Consider the situation where researchers explore a latent structure in practice. They do not always examine loadings precisely; they are interested in signs of the loadings and see roughly whether absolute values of the loadings are large or small. Nevertheless, the loadings calculated with PCA or their ratios may be far from the ones in the FA model. Further, we note that the following difficulties (D1), (D2) and (D3) arise:

(D1) A varimax and a quartimax rotation, which are widely used without careful consideration, are not always appropriate for other cases except a complete simple structure.

(D2) Differences between the values of loadings calculated with PCA and the corresponding values of an FA model in multifactor cases tend to be larger than those in monofactor cases. Further, an order of calculated values may not coincide with an order of model's values; see (1–2) of Experiment 6.2 (On the other hand, in monofactor cases, the order of calculated values is guaranteed (Sato [24])).

(D3) When discrepancy between monofactor parts of a complete simple structure is large, substitute use is inappropriate.

Acknowledgments

The author would like to express his sincere gratitude to Professor Yasunori Fujikoshi for his instructive suggestions and discussions. He is also indebted to Drs. Yosiro Tumura and Kazuo Fukutomi for their continuous encouragement, and to the referees for careful reading. This study was carried out in part under the ISM Cooperative Research Program (90–ISM·CRP–9).

References

[1] Akaike, H., Factor analysis and AIC, Psychometrika, **52** (1987), 317–332.

[2] Anderson, T. W., An Introduction to Multivariate Statistical Analysis (second ed.), John Wiley & Sons, New York, 1984.

[3] Anderson, T. W. and Rubin, H., Statistical inference in factor analysis, Proceedings of the Third Berkeley Symposium on Mathematical Statistics and Probability, **5** (1956), 111–150.

[4] Bechtoldt, H. P., An empirical study of the factor analysis stability hypothesis, Psychometrika, **26** (1961), 405–432.

[5] Bekker, P. A. and de Leeuw, Jan, The rank of reduced dispersion matrices, Psychometrika, **52** (1987), 125–135.

[6] Chatfield, C. and Collins, A. J., Introduction to Multivariate Analysis, Chapman and Hall, London, 1980.

[7] Guttman, L., To what extent can communalities reduce rank? Psychometrika, **23** (1958), 297–308.

[8] Ihara, M. and Kano, Y., A new estimator of the uniqueness in factor analysis, Psychometrika, **51** (1986), 563–566.

[9] Jolliffe, I. T., Principal Component Analysis, Springer-Verlag, New York, 1986.

[10] Jöreskog, K. G., Some contributions to maximum likelihood factor analysis, Psychometrika, **32** (1967), 443–482.

[11] Kendall, M. G., Multivariate Analysis, Charles Griffin, London, 1975.

[12] Konishi, S., Asymptotic expansions for the distributions of functions of a correlation matrix, J. Multivariate Anal., **9** (1979), 259–266.

[13] Konishi, S., Asymptotic expansions for the distributions of statistics based on the sample correlation matrix in principal component analysis, Hiroshima Math. J., **9** (1979), 647–700.

[14] Koopman, R. P., On Bayesian estimation in unrestricted factor analysis, Psychometrika, **43** (1978), 109–110.

[15] Lawley, D. N., The estimation of factor loadings by the method of maximum likelihood, Proc. Roy. Soc. Edinburgh Sect. A, **60** (1940), 64–82.

[16] Lawley, D. N. and Maxwell, A. E., Factor Analysis as a Statistical Method, Butterworth, London, 1963.

[17] Lawley, D. N. and Maxwell, A. E., Factor Analysis as a Statistical Method (second ed.), Butterworth, London, 1971.

[18] Martin, J. K. and McDonald, R. P., Bayesian estimation in unrestricted factor analysis: a treatment for Heywood cases, Psychometrika, **40** (1975), 505–517.

[19] Okuno, T., Kume, H., Haga, T. and Yoshizawa, T., Multivariate Analysis (in Japanese), JUSE, Tokyo, 1971.

[20] Rao, C. R., Estimation and tests of significance in factor analysis, Psychometrika, **20** (1955), 93–111.

[21] Rao, C. R., Linear Statistical Inference and its Applications (second ed.), John Wiley & Sons, New York, 1973.

[22] Reirsøl, O., On the identifiability of parameters in Thurstone's multiple factor analysis, Psychometrika, **15** (1950), 121–149.

[23] Sato, M., Pragmatic treatment of improper solutions in factor analysis, Ann. Inst. Statist. Math., **39** (1987), 443–455.

[24] Sato, M., Some remarks on principal component analysis as a substitute for factor analysis in monofactor cases, J. Japan. Statist. Soc., **20** (1990), 23–31.

[25] Sato, M., Some comments on Shapiro's paper: Identifiability of factor analysis, Technical Report, **249**, Statistical Research Group, Hiroshima University, Hiroshima, 1989.

[26] Shapiro, A., Identifiability of factor analysis: some results and open problems, Linear Algebra Appl., **70** (1985), 1–7. Erratum, **125** (1989), 149.

[27] Siotani, M., Hayakawa, T. and Fujikoshi, Y., Modern Multivariate Statistical Analysis: A Graduate Course and Handbook, American Sciences Press, Ohio, 1985.

[28] Spearman, C., "General Intelligence," objectively determined and measured, American Journal of Psychology, **15** (1904), 201–293.

[29] Sugiura, N., Asymptotic expansions of the distributions of the latent roots and the latent vector of the Wishart and multivariate F matrices, J. Multivariate Anal., **6** (1976), 500–525.

[30] Takeuchi, K. and Yanai, H., The Foundations of Multivariate Analysis (in Japanese), Toyokeizai Shimposha, Tokyo, 1972.

[31] Tucker, L. R. and Lewis, C., A reliability coefficient for maximum likelihood factor analysis, Psychometrika, **38** (1973), 1–10.

[32] Tumura, Y. and Fukutomi, K., On the identification in factor analysis, Rep. Statist. Appl. Res. Un. Japan. Sci. Engrs., **15** (1968), 98–103.

[33] Tumura, Y., Fukutomi, K. and Asoo, Y., On the unique convergence of iterative procedures in factor analysis, TRU Math., **4** (1968), 52–59.

[34] Tumura, Y. and Sato, M., On the identification in factor analysis, TRU Math., **16**–2 (1980), 121–131.

[35] Tumura, Y. and Sato, M., On the convergence of iterative procedures in factor analysis, TRU Math., **17** (1981), 159–168.

[36] Tumura, Y. and Sato, M., On the convergence of iterative procedures in factor analysis (II), TRU Math., **17** (1981), 255–271.

[37] Tumura, Y. and Sato, M., On some notes in factor extraction, TRU Math., **18** (1982), 85–88.

[38] van Driel, O. P., On various causes of improper solutions in maximum likelihood factor analysis, Psychometrika, **43** (1978), 225–243.

[39] Wilkinson, J. H., The Algebraic Eigenvalue Problem, Oxford University Press, Oxford, 1965.

Department of Applied Mathematics,
Faculty of Engineering,
Hiroshima University