THE ASYMPTOTIC DISTRIBUTION OF A NONITERATIVE ESTIMATOR IN EXPLORATORY FACTOR ANALYSIS

BY YUTAKA KANO

Osaka University

This paper presents the asymptotic distribution of Ihara and Kano’s noniterative estimator of the uniqueness in exploratory factor analysis. When the number of factors is overestimated, the estimator is not a continuous function of the sample covariance matrix and its asymptotic distribution is not normal, but the consistency holds. It is also shown that the first-order moment of the asymptotic distribution does not exist.

1. Introduction. Factor analysis is an important branch of statistical science designed to analyze the internal relationship among a set of observed variables. The definition of the factor analysis model is as follows: A family of probability distributions of a \( p \times 1 \) random vector \( \mathbf{x} \) is called a factor analysis model with \( k \) common factors if there exist a \( p \times k \) matrix \( \Lambda \) and a \( p \times p \) positive definite diagonal matrix \( \Psi \) such that the covariance matrix \( \Sigma \) of \( \mathbf{x} \) is represented in the form

\[
\text{Var}(\mathbf{x}) = \Sigma = \Lambda \Lambda' + \Psi,
\]

where \( \Lambda \) and \( \Psi \) consist of factor loadings and unique variances, respectively [see, e.g., Lawley and Maxwell (1971), page 6]. This paper deals with estimation of exploratory factor analysis in which there is no prior information about the number \( k \) of factors, values of \( \Lambda \) and \( \Psi \). Let \( \mathbf{x}_1, \ldots, \mathbf{x}_N \) be a random sample of size \( N \) drawn from the factor analysis model; the parameters \( \Lambda \) and \( \Psi \) are then estimated, after choosing an appropriate \( k \), using the sample covariance matrix \( S \) defined as

\[
S = \frac{1}{n} \sum_{k=1}^{N} (\mathbf{x}_k - \bar{x})(\mathbf{x}_k - \bar{x})',
\]

where \( n = N - 1 \) and \( \bar{x} = (1/N) \sum_{k=1}^{N} \mathbf{x}_k \).

Many methods for estimating these parameters have been developed; these include maximum likelihood (ML [Lawley (1940)]), the canonical factor analysis of Rao (1955) and the generalized least squares (GLS) method due to Jöreskog and Goldberger (1972). Although they are statistically efficient, these methods require iterative processes and may cause several difficulties, such as improper solutions, starting-value problems, nonconvergence and heavy computation [see, e.g., Driel (1978), Anderson and Gerbing (1984), Boomsma (1985) and Sato (1987)]. On the other hand, Ihara and Kano (1986) proposed a closed form estimator of the uniqueness \( \Psi \), and Kano (1989) showed that the

Received February 1989; revised October 1989.


Key words and phrases. Anderson and Rubin’s condition, Bartlett decomposition, exploratory factor analysis, multivariate \( t \)-distribution, Wishart distribution.
inverse matrix involving the Ihara–Kano (I–K) estimator can be replaced by a
generalized inverse matrix. It is a surprising result that the I–K estimator is
consistent even when the number \( k \) of factors is overestimated, which was
shown by Kano (1990a). He also showed that this property ensures the rare
occurrence of improper solutions. The traditional estimation methods (includ-
ing the ML and GLS methods) often encounter some serious difficulties when
\( k \) is overestimated: For example, the estimators are inconsistent and the
distributions of the estimators still do not have been obtained even with the
help of the asymptotic theory [see, Geweke and Singleton (1980), Section 2,
and Kano (1990a), Section 1]. These problems happen because the parameter
is not identified. These facts may also make it difficult to determine the
number of factors because the distribution theory of statistics for choosing
models, for example, Akaike’s Information Criterion (AIC), is based on the
asymptotic normality of the estimators [see Akaike (1987)].

After consistency, distributions of estimators are important because they
are used to construct confidence intervals and to test statistical hypotheses.
The case has already been treated in which the number \( k \) of factors is
correctly chosen. Ihara and Kano (1986) showed that the I–K estimator with
the true \( k \) is asymptotically normally distributed, and the asymptotic variance
was given by Kano (1990b). This paper investigates the asymptotic distribution
of the I–K estimator, when \( k \) is overestimated. In this case the analysis is not
straightforward because the estimator is not a continuous function of the
sample covariance matrix \( S \) and the usual technique using derivatives is not
available.

2. Ihara and Kano estimator of the uniqueness. Let \( \Sigma = \Lambda \Lambda' + \Psi \),
with \( \Lambda \) being \( p \times k \), and suppose that the parameter \( (\Lambda, \Psi) \) satisfies Anderson
and Rubin’s sufficient condition for identifiability [see Theorem 5.1 in Anderson
and Rubin (1955)]: If any row vector of \( \Lambda \) is deleted, there remain two
disjoint nonsingular submatrices of order \( k \). The condition will be abbreviated
to the A–R condition hereafter. Note that the Anderson–Rubin (A–R) condition
requires that \( p \geq 2k + 1 \).

Let \( m \) be the number of assumed factors. Since the (true) number \( k \) of
factors is generally unknown in exploratory factor analysis, \( m \) is not necessar-
ily equal to \( k \). Partition \( \Lambda, \Psi, \Sigma \) and \( S \) as follows:

\[
\Lambda = \begin{bmatrix}
\lambda_1 \\
\lambda_2 \\
\lambda_3 \\
\lambda_4 \\
\end{bmatrix}
\begin{bmatrix}
1 \\
m \\
m \\
p - 2m - 1 \\
\end{bmatrix}, \quad \Psi = \begin{bmatrix}
\psi_1 \\
\psi_2 \\
\psi_3 \\
\psi_4 \\
\end{bmatrix}
\]

\[
\Sigma = \begin{bmatrix}
\sigma_{11} & \text{sym} \\
\sigma_{21} & \Sigma_{22} \\
\sigma_{31} & \Sigma_{32} \\
\sigma_{41} & \Sigma_{42} \\
\end{bmatrix}
\quad \text{and} \quad S = \begin{bmatrix}
s_{11} & \text{sym} \\
s_{21} & S_{22} \\
s_{31} & S_{32} \\
s_{41} & S_{42} \\
\end{bmatrix}
\begin{bmatrix}
s_{11} & \text{sym} \\
s_{21} & S_{22} \\
s_{31} & S_{32} \\
s_{41} & S_{42} \\
\end{bmatrix}
\]
We assume that \( m \) is not greater than \( (p - 1)/2 \) to enable the preceding partition. Ihara and Kano (1986) then proposed the following simple estimator of \( \psi_1 \):

\[
\hat{\psi}_1^{(m)} = s_{11} - s_{12}S_{32}^{-1}s_{31}.
\]

We will call \( \hat{\psi}_1^{(m)} \) the I–K estimator for \( \psi_1 \). We can calculate estimates \( \hat{\psi}_i^{(m)} \) of \( \psi_i \), \( i = 2, \ldots, p \), in a similar manner after interchanging the variates \( X_1, \ldots, X_p \), and an estimator of \( \Psi \) is then defined as \( \hat{\Psi}^{(m)} = \text{diag}(\hat{\psi}_1^{(m)}, \ldots, \hat{\psi}_p^{(m)}) \), and inference of \( \Lambda \) is based on \( S - \hat{\Psi}^{(m)} \). The basic idea of the estimator (2.1) is the direct application of the moment method based on the relation

\[
\psi_1 = \sigma_{11} - \sigma_{12}\Sigma_{32}^{-1}\sigma_{31},
\]

which holds when \( m = k \) and \( \Sigma_{32} \) (\( = \Lambda_3\Lambda_2 \)) is nonsingular (this is ensured by the A–R condition). Thus, when \( m = k \), consistency of \( \hat{\psi}_1^{(k)} \) holds and \( \psi_1^{(k)} \) is totally differentiable at \( S = \Sigma \), which guarantees its asymptotic normality [see Ihara and Kano (1986)].

Assume that the observed vector \( \mathbf{x} \) is normally distributed with the covariance matrix \( \Sigma \) in (1.1), and \( nS \) follows a Wishart distribution \( W_p(n, \Sigma) \). Kano (1990b) then presented the asymptotic variance of \( \hat{\psi}_1^{(k)} \) as follows:

\[
\text{Var}(\hat{\psi}_1^{(k)}) = \psi_1^2 + \left( \sigma_{12}\Sigma_{32}^{-1}\Sigma_3\Sigma_{23}^{-1}\sigma_{21} + \psi_1 \right) \left( \sigma_{13}\Sigma_{23}^{-1}\Psi_2\Sigma_{32}^{-1}\sigma_{31} + \psi_1 \right)
\]

and

\[
\text{Var}(\hat{\psi}_1^{(k)}/s_{11}) = \{\text{Var}(\hat{\psi}_1^{(k)}) - c\}/\sigma_{11}^2
\]

for a standardized case, where \( c = 2\psi_1^2(2\psi_1^2/\sigma_{11} - 1) \).

When the number of factors is overestimated, i.e., \( m > k \), the matrix \( \Sigma_{32} \) is singular and hence the I–K estimator cannot be defined at \( S = \Sigma \). Kano (1990a), however, proved that \( \hat{\psi}_1^{(m)} \) converges to \( \psi_1 \) in probability. It follows from the property that \( S - \hat{\Psi}^{(m)} \) converges to \( \Lambda\Lambda' \) in probability, where \( m = [(p - 1)/2] \), denoting the maximum integer not greater than \( (p - 1)/2 \) (the Gauss symbol). He has used this to propose a new method for determining the number of factors.

It is important to investigate which estimator is the best among the class of consistent estimators \( \hat{\psi}_1^{(m)} \) with \( k \leq m \leq (p - 1)/2 \). The aim of this paper is to present the asymptotic distribution of the I–K estimator when \( m > k \); this will also provide useful information about the choice of the best I–K estimator.

3. Main results. We shall first consider the following lemma, which also gives notations used in this paper.

**Lemma 1.** Assume that \( m \geq k \) and that \( \Lambda_2 \) and \( \Lambda_3 \) are of full column rank. Decompose the \((2m + 1) \times (2m + 1)\) submatrix

\[
\begin{bmatrix}
\sigma_{11} & \text{sym} \\
\sigma_{21} & \Sigma_{22} \\
\sigma_{31} & \Sigma_{32} \\
\end{bmatrix}
\]
of $\Sigma$ in (1.1) into $PP'$, with

$$P = \begin{bmatrix}
    p_{11} & 0 & 0 \\
    p_{21} & p_{22} & 0 \\
    p_{31} & p_{32} & p_{33}
\end{bmatrix}. $$

Then there exist $m \times k$ matrices $B$ and $C$ of rank $k$ and $k$-vectors $d$ and $e$ such that $P_{32}^{-1}P_{32} = CB'$, $P_{22}^{-1}p_{21} = Bd$ and $P_{33}^{-1}p_{31} = Ce$. Furthermore, $d'e + 1 = \sigma_{11}/\psi_1$.

**Proof.** Since $\Sigma$ is nonsingular, so are $P_{22}$ and $P_{32}$. From the definition of $P$, we see that $P_{32}P_{22}' = \Lambda_3(I_k - \lambda_1'p_{11}^{-2}\lambda_1')\Lambda_2'$, which implies that rank($P_{32}$) = $k$. Hence, there exist $m \times k$ matrices $B$ and $C$ with $P_{33}^{-1}P_{32} = CB'$, and we may then define

$$d = C'P_{33}\Lambda_3(\Lambda_2'\Lambda_3)^{-1}\lambda_1'p_{11}\psi_1^{-1} \quad \text{and} \quad e = B'P_{22}'\Lambda_2(\Lambda_2'\Lambda_2)^{-1}\lambda_1'p_{11}\psi_1^{-1}.$$  

We have easily $d'e = p_{11}^{-2}\lambda_1(I_k - \lambda_1'p_{11}^{-2}\lambda_1')\lambda_1'p_{11}\psi_1^{-1} - 1$. □

We use here the symbols $\rightarrow_L$, $\rightarrow_p$, $\Rightarrow_a$ and $\Rightarrow_d$ to mean convergence in law and in probability, asymptotic equivalence in probability and equality in distribution, respectively. The following theorem will be established.

**Theorem.** Assume that the observed vector $x$ is distributed as $N_p(0, \Sigma)$ and that $\Sigma = \Lambda\Lambda' + \Psi$, with $\Lambda$ being $p \times k$, which satisfies Anderson and Rubin's sufficient condition for identifiability. Let $m$, $k \leq m \leq (p - 1)/2$, be the number of assumed factors and let the $I-K$ estimator $\hat{\psi}_1^{(m)}$ be defined in (2.1). The constants $B$, $C$, $d$ and $e$ are defined in Lemma 1. Then the following holds:

$$\sqrt{n}(\hat{\psi}_1^{(m)} - \psi_1) \rightarrow_L \psi_1Z_1 + (\psi_2^2/\sigma_{11})c_1c_2(Z_2 + z_1z_2/\chi_1),$$

where $Z_1$ and $Z_2$ are distributed as $N(0, 1)$, $z_1$ and $z_2$ also follow $N_{m-k}(0, I_{m-k})$ and $\chi_1^2$ conforms to a chi-square distribution with one degree of freedom, which are all mutually independent, and where

$$c_1^2 = d'(B'B)d + d'(C'C)^{-1}d + 1 \quad \text{and} \quad c_2^2 = e'(B'B)^{-1}e + 1.$$ 

**Comment.** The constants $c_1$ and $c_2$ are free from the choice of $m \times k$ matrices $B$ and $C$ in Lemma 1. Note that the term $z_1^2/z_2/\chi_1^2$ in (3.1) vanishes when $m = k$, which implies the asymptotic normality of $\hat{\psi}_1^{(k)}$, and then the asymptotic variance due to (3.1) is exactly equal to that in (2.2). This will be shown in Appendix A. When the number of factors is overestimated, the asymptotic distribution is not normal, and the expectation of the asymptotic distribution does not exist because that of $1/\sqrt{\chi_1^2}$ does not. Therefore, it could be said that $\hat{\psi}_1^{(k)}$ is the best among the set of consistent estimators $\hat{\psi}_1^{(m)}$, $k \leq m \leq (p - 1)/2$. Hence, we can recommend $\hat{\psi}_1^{(k)}$ as an estimator of $\psi_1$ when $k$ is known.
Proof of the Theorem. It follows from the Bartlett decomposition [see, e.g., Anderson (1984), page 251] that

\[ n \begin{bmatrix} s_{11} & \text{sym} \\ s_{21} & S_{22} \\ s_{31} & S_{32} \\ s_{31} & S_{33} \end{bmatrix} = PXX'P' \quad \text{with} \quad X = \begin{bmatrix} x_{11} & 0 & 0 \\ x_{21} & X_{22} & 0 \\ x_{31} & X_{32} & X_{33} \end{bmatrix} \]

and that

\[ x_{11}/\sqrt{n} \to_p 1 \]

\[ X_{22}/\sqrt{n} \to_p I_m \]

\[ \sqrt{n}(x_{11}^2/n - 1) \to_L Y \sim N(0,2) \]

\[ \sqrt{n}(X_{22}X_{22}/n - I_m)a \to_L y \sim N_m(0, \sum a'a \cdot I_m + aa') \]

where \( a = B'd. \) Note that \( Y, y, x_{21}, x_{31} \) and \( X_{32} \) are mutually independent. Put

\[
\begin{bmatrix}
t_{11} & 0 & 0 \\
t_{21} & T_{22} & 0 \\
t_{31} & T_{32} & T_{33}
\end{bmatrix}
\]

\[ = PX = \begin{bmatrix} p_{11}x_{11} & 0 & 0 \\ p_{21}x_{11} + p_{22}x_{21} & p_{22}X_{22} & 0 \\ p_{31}x_{11} + p_{32}x_{21} + p_{33}x_{31} & p_{32}X_{22} + p_{33}X_{32} & p_{33}X_{33} \end{bmatrix}. \]

Then the I–K estimator can be written as

\[ \hat{\psi}_1^{(m)} = s_{11} - s_{12}s_{32}^{-1}s_{31} \]

\[ = (t_{11}/n)(1 + t_{21}'T_{22}^{-1}T_{32}^{-1}t_{31})^{-1} \]

\[ = \sigma_{11}(x_{11}^2/n)G^{-1}, \quad \text{say,} \]

and the matrix \( G \) is represented in the form

\[ G = 1 + t_{21}'T_{22}^{-1}T_{32}^{-1}t_{31} \]

\[ = 1 + \left( p_{21}x_{11} + p_{22}x_{21} \right) \left( p_{22}X_{22} \right)^{-1} \left( p_{32}X_{22} + p_{33}X_{32} \right)^{-1} \]

\[ \times \left( p_{31}x_{11} + p_{32}x_{21} + p_{33}x_{31} \right) \]

\[ = 1 + \left( B'dx_{11} + x_{21} \right)'H^{-1} \left( Cex_{11} + CB'x_{21} + x_{31} \right) \]

in view of Lemma 1, where \( H^{-1} = (X_{22})^{-1}(CB'X_{22} + X_{32})^{-1} \).

The following lemma is needed to prove the theorem.

Lemma 2. Decompose \( B = PB_1B_D \) and \( C = PC_1D_C \), where \( D_B \) and \( D_C \) are \( k \times k \) nonsingular matrices and \( P_B = [P_{B1} : P_{B2}] \) and \( P_C = [P_{C1} : P_{C2}] \) are orthogonal matrices of order \( m \). Define

\[
\begin{bmatrix}
y_{11} \\
y_{12} \\
y_{21} \\
y_{22}
\end{bmatrix} = P'C_X_{32}P_B, \quad \text{where} \quad Y_{11} \quad \text{is} \quad k \times k.
\]

Then the elements of \( P'C_X_{32}P_B \) are independently identically distributed.
as $N(0,1)$ and

(i) \[ \sqrt{n} H^{-1} \rightarrow_L P_B \begin{bmatrix} 0 & 0 \\ 0 & Y_{22}^{-1} \end{bmatrix} \tilde{P}', \]

(ii) \[ nH^{-1}C \rightarrow_L P_B \begin{bmatrix} I_k \\ -Y_{22}^{-1}Y_{21} \end{bmatrix} D_B^{-1}, \]

(iii) \[ n \bar{d}'B'H^{-1} \rightarrow_L - \bar{y}' P_B \begin{bmatrix} 0 & 0 \\ 0 & Y_{22}^{-1} \end{bmatrix} P_C' + \bar{d}'D_C^{-1} \left[ I_k : -Y_{12}Y_{22}^{-1} \right] P_C', \]

(iv) \[ \sqrt{n} \bar{d}'(nB'H^{-1}C - I_k) \]

\[ \rightarrow_L - \left( \bar{y}' P_B \begin{bmatrix} I_k \\ -Y_{22}^{-1}Y_{21} \end{bmatrix} D_B^{-1} + \bar{d}'D_C^{-1} \left( Y_{11} - Y_{12}Y_{22}^{-1}Y_{21} \right) D_B^{-1} \right), \]

where $\bar{y}$ is defined by (3.3).

A proof of Lemma 2 will be given in Appendix B. Note that (iv) in Lemma 2 implies \( n \bar{d}'B'H^{-1}C \rightarrow_p \bar{d}' \). From (3.5) and Lemma 2, we have \( G = a_1 + x_{11}'d'B'H^{-1}C\bar{e} = a_1 + \bar{d}'\bar{e} \), which means, in view of Lemma 1, that

\[ G \rightarrow_p \frac{\sigma_{11}}{\psi_1}. \]

It follows from (3.3), (3.4) and (3.6) that $\hat{\psi}_1^{(m)}$ is a (weakly) consistent estimator of $\psi_1$, and this is an alternative proof of consistency [cf. Kano (1990a)].

Since we have, from (3.3), (3.4) and (3.6),

\[ \sqrt{n} \left( \hat{\psi}_1^{(m)} - \psi_1 \right) = \sigma_{11}G^{-1} \sqrt{n} \left( \frac{x_{11}^2}{n} - 1 \right) - \psi_1G^{-1} \sqrt{n} \left( G - \frac{\sigma_{11}}{\psi_1} \right) \]

\[ \rightarrow_L \psi_1 \psi_1 - \frac{\sigma_{11}^2}{\psi_1} \sqrt{n} \left( G - \frac{\sigma_{11}}{\psi_1} \right), \]

we may investigate the distribution of $\sqrt{n} \left( G - \frac{\sigma_{11}}{\psi_1} \right)$. This can be represented from (3.5) and Lemma 1 as

\[ \sqrt{n} \left( G - \frac{\sigma_{11}}{\psi_1} \right) = \sqrt{n} \left( (Bd x_{11} + x_{21})'H^{-1}(C e x_{11} + CB'x_{21} + x_{31}) - \bar{d}'\bar{e} \right) \]

\[ = \sqrt{n} \left( \bar{d}'(x_{11}^2B'H^{-1}C - I_k)e + \sqrt{n} x_{11} \bar{d}'B'H^{-1}(CB'x_{21} + x_{31}) \right) \]

\[ + \sqrt{n} x_{11} x_{21}'H^{-1}C e + \sqrt{n} x_{21}'H^{-1}(CB'x_{21} + x_{31}) \]

\[ = I_1 + I_2 + I_3 + I_4, \quad \text{say}. \]
Each term of (3.8) can be evaluated using (3.3) and Lemma 2 as follows:

\[ I_1 = \mathbf{d}' \left( \sqrt{n} \left( \frac{x_{11}^2}{n} - 1 \right) nB'H^{-1}C + \sqrt{n} \left( nB'H^{-1}C - I_k \right) \right) \mathbf{e} \]

\[ \rightarrow_L \mathbf{d}' \mathbf{e} Y - \mathbf{y}' P_B \begin{bmatrix} I_k \\ -Y_{22}^{-1}Y_{21} \end{bmatrix} D_B^{-1} \mathbf{e} \]

\[ - \mathbf{d}' D_C^{-1} (Y_{11} - Y_{12} Y_{22}^{-1} Y_{21}) D_B^{-1} \mathbf{e}, \]

\[ I_2 = \mathbf{d}' B' \mathbf{x}_{21} + \sqrt{n} x_{11} \mathbf{d}' B'H^{-1} \mathbf{x}_{31} \]

\[ \rightarrow_L \mathbf{d}' B' \mathbf{x}_{21} - \mathbf{y}' P_B \begin{bmatrix} 0 \\ 0 \end{bmatrix} Y_{22}^{-1} Y_{21}^{-1} P_C' \mathbf{x}_{31} + \mathbf{d}' D_C^{-1} [I_k : -Y_{12} Y_{22}^{-1}] P_C' \mathbf{x}_{31}, \]

\[ I_3 \rightarrow_L \mathbf{x}_{21}' P_B \begin{bmatrix} I_k \\ -Y_{22}^{-1}Y_{21} \end{bmatrix} D_B^{-1} \mathbf{e} \]

and

\[ I_4 \rightarrow_L \mathbf{x}_{21}' P_B \begin{bmatrix} 0 \\ 0 \end{bmatrix} Y_{22}^{-1} P_C' \mathbf{x}_{31}. \]

Let

\[ \begin{bmatrix} y_{11} \\ y_{12} \end{bmatrix} = \begin{bmatrix} P_{B1}' \mathbf{y} \\ P_{B2}' \mathbf{y} \end{bmatrix} = P_B' \mathbf{y}, \quad \begin{bmatrix} y_{21} \\ y_{22} \end{bmatrix} = \begin{bmatrix} P_{B1}' \mathbf{x}_{21} \\ P_{B2}' \mathbf{x}_{21} \end{bmatrix} = P_B' \mathbf{x}_{21} \]

and

\[ \begin{bmatrix} y_{31} \\ y_{32} \end{bmatrix} = \begin{bmatrix} P_{C1}' \mathbf{x}_{31} \\ P_{C2}' \mathbf{x}_{31} \end{bmatrix} = P_C' \mathbf{x}_{31}. \]

The random vectors \[ \begin{bmatrix} y_{21} \\ y_{22} \end{bmatrix} \] and \[ \begin{bmatrix} y_{31} \\ y_{32} \end{bmatrix} \] are independently distributed as \[ N_m(0, I_m) \], and \[ \begin{bmatrix} y_{11} \\ y_{12} \end{bmatrix} \] is also normally distributed with covariance matrix

\[ \begin{bmatrix} D_B \mathbf{d}' \\ 0 \end{bmatrix} \begin{bmatrix} D_B \mathbf{d}' \\ 0 \end{bmatrix}' + \left( B \mathbf{d}' \right) I_m. \]

\[ \sqrt{n} \left( G - \frac{C_{11}}{\psi_{11}} \right) \rightarrow_L \mathbf{d}' \mathbf{e} Y - \mathbf{e}' D_B^{-1} y_{11} + (\mathbf{d}' D_B' + \mathbf{e}' D_B^{-1}) y_{21} + \mathbf{d}' D_C^{-1} y_{31} \]

\[ - \mathbf{d}' D_C^{-1} Y_{11} D_B^{-1} \mathbf{e} + (\mathbf{d}' D_C^{-1} Y_{12} + y_{12}' - y_{22}') \times Y_{22}^{-1} (y_{21}' D_B^{-1} \mathbf{e} - y_{32}). \]

Thus, we can see from (3.7) and (3.14) that the distribution of \( \sqrt{n} (\hat{\psi}_1^{(m)} - \psi_1) \)
converges to that of
\[
\frac{\psi_1^2}{\sigma_{11}} \left( Y + \mathbf{e}'D_B^{-1}\mathbf{y}_{11} - (\mathbf{d}'D_B^{-1} + \mathbf{e}'D_B^{-1})\mathbf{y}_{21} - \mathbf{d}'D_C^{-1}\mathbf{y}_{31} \\
+ \mathbf{d}'D_C^{-1}Y_{11}D_B^{-1}\mathbf{e} - (\mathbf{d}'D_C^{-1}Y_{12} + \mathbf{y}'_{12} - \mathbf{y}'_{22})Y_{22}^{-1}(Y_{21}D_B^{-1}\mathbf{e} - \mathbf{y}_{32}) \right)
\]
\[= \frac{\psi_1^2}{\sigma_{11}} (W - \mathbf{w}'_{1}Y_{22}^{-1}\mathbf{w}_{2}), \text{ say.} \]

Note that \( W, \mathbf{w}_{1}, \mathbf{w}_{2} \) and \( Y_{22} \) are mutually independent. Since \( D_B D_B' = B'B \) and \( D_C D_C' = C'C \), we have from (3.13),
\[\text{Var}(\mathbf{w}_{1}) = \{\mathbf{d}'(B'B)\mathbf{d} + \mathbf{d}'(C'C)^{-1}\mathbf{d} + 1\}I_{m-k} = c_1^2 I_{m-k}, \text{ say} \]
and
\[\text{Var}(\mathbf{w}_{2}) = \{\mathbf{e}'(B'B)^{-1}\mathbf{e} + 1\}I_{m-k} = c_2^2 I_{m-k}, \text{ say.} \]

After some calculations we get
\[\text{Var}(W) = (\mathbf{d}'\mathbf{e} + 1)^2 + c_1^2 c_2^2 = \left(\frac{\sigma_{11}}{\psi_1}\right)^2 + c_1^2 c_2^2 \]
in view of Lemma 1. Let \( Z_1 \) and \( Z_2 \) be independently distributed as \( N(0,1) \). Then
\[\left(\frac{\psi_1^2}{\sigma_{11}}\right)W = d_1 \psi_1 Z_1 + \left(\frac{\psi_1^2}{\sigma_{11}}\right)c_1 c_2 Z_2. \]

Define \( \mathbf{z}_1 = -c_1^{-1}\mathbf{w}_1 \) and \( \mathbf{z}_2 = c_2^{-1}\mathbf{w}_2 \). Then \( \mathbf{z}_1 \) and \( \mathbf{z}_2 \) follow \( N_{m-k}(0, I_{m-k}) \) independently. We thus obtain
\[\sqrt{n} \left( \psi_1^{(m)} - \psi_1 \right) \rightarrow_L \psi_1 Z_1 + \left(\frac{\psi_1^2}{\sigma_{11}}\right)c_1 c_2 (Z_2 + \mathbf{z}'_1 Y_{22}^{-1}\mathbf{z}_2). \]

The following lemma [see, e.g., Johnson and Kotz (1972), page 144] is useful in completing the proof.

**Lemma 3.** Let \( \mathbf{x} \) be distributed as \( N_q(0, I_q) \) and \( J \) be a \( q \times q \) random matrix, independent of \( \mathbf{x} \), such that \( J'J \) follows \( W_q(n, R^{-1}) \). Then \( \sqrt{\nu} J^{-1} \mathbf{x} \) has a multivariate \( t \)-distribution with parameter matrix \( R \) and \( \nu \) degrees of freedom, where \( \nu = n - q + 1 \), that is,
\[\sqrt{\nu} J^{-1} \mathbf{x} = d_\nu \frac{\mathbf{y}}{\sqrt{\chi^2/\nu}}, \]
where \( \mathbf{y} \) and \( \chi^2 \) are independently distributed as \( N_q(0, R) \) and \( \chi^2(\nu) \), respectively.

Since \( Y_{22}^{-1} \mathbf{z}_2 \) is distributed according to \( W_{m-k}(m-k, I_{m-k}) \), it follows from Lemma 3 that \( Y_{22}^{-1} \mathbf{z}_2 \) has a multivariate \( t \)-distribution with parameter matrix
$I_{m-k}$ and one degree of freedom. This fact and (3.15) complete the proof of the theorem. □

Kano (1990b) dealt with the asymptotic distribution of the $I$–$K$ estimator with $m = k$ based on the sample correlation matrix as well. We can also obtain, in the same way, the distribution based on the sample correlation matrix in a case when $m > k$.

APPENDIX A

We shall here show that when $m = k$, the asymptotic variance based on (3.1) is the same as in (2.2). As stated just after the theorem, the asymptotic distribution of $\sqrt{n} (\hat{\psi}_1^{(k)} - \psi_1)$ is given by

\begin{equation}
\psi_1 Z_1 + \left( \frac{\psi_1^2}{\sigma_{11}} \right) c_1 c_2 Z_2, \tag{A1}
\end{equation}

and its variance is

\begin{equation}
\psi_1^2 + \left( \frac{\psi_1^2}{\sigma_{11}} \right)^2 c_1^2 c_2^2. \tag{A2}
\end{equation}

Note that $B$ and $C$ are nonsingular matrices of $k \times k$. Then $c_1^2$ and $c_2^2$ in (3.2) are written as

\begin{equation}
c_1^2 = \mathbf{p}_{21}' (P_{32}P_{22})^{-1} \Sigma_{33} (P_{22}P_{32})^{-1} \mathbf{p}_{21} - \left( \mathbf{p}_{21}' (P_{32}P_{22})^{-1} \mathbf{p}_{31} \right)^2 + 1 \tag{A3}
\end{equation}

and

\begin{equation}
c_2^2 = \mathbf{p}_{31}' (P_{22}P_{32})^{-1} \Sigma_{22} (P_{32}P_{22})^{-1} \mathbf{p}_{31} - \left( \mathbf{p}_{21}' (P_{32}P_{22})^{-1} \mathbf{p}_{31} \right)^2 + 1, \tag{A4}
\end{equation}

in view of the definitions of $B$, $C$, $d$ and $e$. We can easily check

\begin{equation}
\mathbf{p}_{21}' (P_{32}P_{22})^{-1} = \left( \frac{p_{11}}{\psi_1} \right) \sigma_{12} \Sigma_{32}^{-1}, \quad \mathbf{p}_{31}' (P_{22}P_{32})^{-1} = \left( \frac{p_{11}}{\psi_1} \right) \sigma_{13} \Sigma_{23}^{-1} \tag{A5}
\end{equation}

and

\begin{equation}
\mathbf{p}_{21}' (P_{32}P_{22})^{-1} \mathbf{p}_{31} = \psi_1^{-1} \sigma_{12} \Sigma_{32}^{-1} \sigma_{31} = \psi_1^{-1} \lambda_{1} \lambda_{1}'. \tag{A6}
\end{equation}

Substituting (A4) and (A5) for (A2) and (A3) and using (A1) lead to the asymptotic variance in (2.2).

APPENDIX B

Proof of Lemma 2. The elements of $X_{32}$ are independently distributed as $N(0, 1)$, and $P_B$ and $P_C$ are both orthogonal. The elements of $P_C' X_{32} P_B$, therefore, have the same distribution as those of $X_{32}$. By definition of $Y_{i,j}$ we
see easily that

\[(B'X_{32}^{-1}C)^{-1} = D_C^{-1}(Y^{11})^{-1}D_B^{-1} = D_C^{-1}(Y_{11} - Y_{12}Y_{22}^{-1}Y_{21})D_B^{-1},\]

\[(B'X_{32}^{-1}C)^{-1}B'X_{32}^{-1} = D_C^{-1}[I_k : -Y_{12}Y_{22}^{-1}]P_C',\]

\[X_{32}^{-1}C(B'X_{32}^{-1}C)^{-1} = P_B \begin{bmatrix} I_k \\ -Y_{22}^{-1}Y_{21} \end{bmatrix}D_B^{-1}\]

and

\[X_{32}^{-1} - X_{32}^{-1}C(B'X_{32}^{-1}C)^{-1}B'X_{32}^{-1} = P_B \begin{bmatrix} 0 \\ 0 \end{bmatrix}Y_{22}^{-1}P_C'.\]

Now we shall prove Lemma 2 by using (B1)–(B4). We get from (3.3),

\[\sqrt{n} H^{-1} = \sqrt{n} (X_{22}')^{-1}(CB'X_{22} + X_{32})^{-1}\]

\[= \sqrt{n} (X_{22}')^{-1}\left\{X_{32}^{-1} - X_{32}^{-1}C(I_k + B'X_{22}X_{32}^{-1}C)^{-1}B'X_{22}X_{32}^{-1}\right\}\]

\[= aX_{32}^{-1} - X_{32}^{-1}C(B'X_{32}^{-1}C)^{-1}B'X_{32}^{-1}.\]

This relation and (B4) imply (i). Similarly, we have

\[nH^{-1}C = aX_{32}^{-1}C(B'X_{32}^{-1}C)^{-1},\]

which, along with (B3), shows (ii). Since

\[n d'B'H^{-1} = d'B'(nI_m - X_{22}'X_{22})H^{-1} + d'B'X_{22}(CB'X_{22} + X_{32})^{-1}\]

\[= -d'\sqrt{n}\left(\frac{X_{22}'X_{22}}{n} - I_m\right)\sqrt{n} H^{-1}\]

\[+ d'(I_k + B'X_{22}X_{32}^{-1}C)^{-1}B'X_{22}X_{32}^{-1}\]

\[\rightarrow_L - y'\sqrt{n} H^{-1} + d'(B'X_{32}^{-1}C)^{-1}B'X_{32}^{-1}\]

in view of (3.3), we get (iii) from (i) and (B2). By similar calculation,

\[\sqrt{n} d'(nB'H^{-1}C - I_k) = \sqrt{n} d'\left(nB'X_{22}'X_{32}^{-1}C(I_k + B'X_{22}X_{32}^{-1}C)^{-1} - I_k\right)\]

\[= -d'\sqrt{n}\left(\frac{X_{22}'X_{22}}{n} - I_m\right)\left(\frac{X_{22}'}{\sqrt{n}}\right)^{-1}X_{32}^{-1}C + d'\]

\[\times \sqrt{n}(I_k + B'X_{22}X_{32}^{-1}C)^{-1}\]

\[\rightarrow_L - (y'X_{32}^{-1}C + d')(B'X_{32}^{-1}C)^{-1}.\]

This relation, (B1) and (B3) show (iv). The proof is complete. \(\square\)
Acknowledgment. The author thanks Julie Speckart of UCLA for production assistance.

REFERENCES


DEPARTMENT OF APPLIED MATHEMATICS
FACULTY OF ENGINEERING SCIENCE
OSAKA UNIVERSITY
TOKYONAKA, OSKA 560
JAPAN