

The Library of the Department of Statistics  
North Carolina State University

**THE ASYMPTOTIC COVARIANCE MATRIX FOR  
COVARIANCE ESTIMATORS WITH STANDARDIZATION  
AND RAW-VARIMAX ROTATION IN FACTOR ANALYSIS**

by

**Kentaro Hayashi**

and

**Pranab K. Sen**

Department of Biostatistics  
University of North Carolina

Institute of Statistics  
Mimeo Series No. 2173

November 1996

**Abstract.** This work focuses on obtaining the analytical formulas for covariance matrix for covariance estimators of MLEs of standardized factor loadings with raw-varimax rotation in factor analysis. This is an extension of our previous work for unrotated case and we employed the same approach; as a first step, we expressed all the coordinatewise formulas that were previously derived by other researchers in terms of matrices. Then we obtained the necessary derivatives which map the differential of the sample covariance (or correlation) matrix (in vector form) into the differential of covariance estimators (in vector form). In obtaining the derivatives, we made extensive use of vec operator, Kronecker (direct) product, and the relevant properties.

*Key words and phrases:* asymptotic normality, covariance estimator, asymptotic covariance matrix, factor analysis, factor loadings, Kronecker (direct) product, maximum likelihood estimator, vec operator, standardization, raw varimax rotation (transformation).

## 1. Introduction

Let  $x_i$  ( $i = 1, \dots, n$ ) be a  $p \times 1$  random vector of observations with  $E(x_i) = 0$  and covariance matrix  $\Sigma$ ,  $\Lambda$  be a  $p \times m$  matrix of factor loadings;  $f_i$  be a  $m \times 1$  vector of factor scores, and  $\varepsilon_i$  be a  $p \times 1$  vector of unique factors. Then the factor analysis model is expressed as

$$x_i = \Lambda f_i + \varepsilon_i,$$

for  $i = 1, \dots, n$ , with the assumptions

$$(1.1) \quad E(f_i) = 0; \quad E(\varepsilon_i) = 0; \quad \text{Cov}(f_i, \varepsilon_i) = E(f_i \varepsilon_i') = 0; \quad \text{and} \quad \text{Cov}(\varepsilon_i) = E(\varepsilon_i \varepsilon_i') = \Psi,$$

where  $\Psi$  is a positive definite diagonal matrix. That is, the assumptions in (1.1) state that the factor scores  $f_i$  and the unique factors  $\varepsilon_i$  are centered;  $f_i$  and  $\varepsilon_i$  are uncorrelated; the elements of  $\varepsilon_i$  corresponding to different variables in  $x_i$  are uncorrelated, and the variances of  $\varepsilon_i$  are positive.

Furthermore, we assume the orthogonal model in which the factors are uncorrelated. Thus the covariance matrix of  $f$  is diagonal, that is,

$$\text{Cov}(f_i) = E(f_i f_i') = I_m.$$

and it follows that the covariance (or correlation) matrix of  $x_i$  is expressed as

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

Thus, in factor analysis, the covariance structure of  $x_i$  is expressed in terms of the factor loading matrix  $\Lambda$  and the unique variances  $\Psi$ .

In maximum likelihood estimation in factor analysis, we further assume that the vectors of observation  $x_i$ 's are random samples from a normal population

with mean vector 0 and covariance matrix  $\Sigma$ . Maximization of the log likelihood function is equivalent to minimization of the discrepancy function  $F(\Sigma, S)$ :

$$F(\Sigma, S) = \text{tr}(S\Sigma^{-1}) - \log|\Sigma^{-1}| - p,$$

where  $S$  is an unbiased estimate of  $\Sigma$ . By differentiating  $F(\Sigma, S_n)$  with respect to  $\Lambda$  and  $\Psi$ , and setting them to a null matrix, with some further algebra, we obtain the following two equations:

$$(1.2) \quad (S - \Sigma)\Psi^{-1}\Lambda = 0,$$

$$(1.3) \quad \Psi = \text{Diag}(S - \Lambda\Lambda'),$$

where  $\text{Diag}(A)$  denotes the diagonal matrix whose elements are the diagonal elements of the square matrix  $A$ . The two equations (1.2) and (1.3) cannot be solved analytically, and we must use some numerical approximation to solve them. Several numerical methods for maximum-likelihood factor analysis have been developed based on such algorithms as the Fletcher-Powell, the Newton-Raphson, and the Gauss-Newton algorithms (See e.g., Clarke (1970), Jennrich and Robinson (1969), Jöreskog (1967, 1977), and Lee and Jennrich (1979)).

The formulas for the asymptotic covariance matrix of MLEs of (unrotated) factor loadings were first obtained by Lawley (1967), and they were systematically presented in Chapter 5 of Lawley and Maxwell (1971), including the case with standardized factor loadings. There are some mistakes in the formulas presented by Lawley and Maxwell (1971), and they were found and corrected by Jennrich and Thayer (1973). The formulas for the asymptotic covariance matrix of estimators of orthogonally rotated factor loadings (with the orthomax criterion, including quartimax, varimax, and equimax rotations) were obtained by Archer and Jennrich (1973). In the case of the asymptotic covariance matrix for obliquely rotated factor loadings (i.e., direct oblimin and

direct Crawford-Ferguson rotations), Jennrich (1973) gave the necessary formulas. Jennrich (1974) introduced a simplified method of obtaining the asymptotic covariance matrix (including for the standardized and rotated case) by use of the inverse of an augmented information matrix.

This work focuses on obtaining the analytical formulas for the asymptotic covariance matrix for covariance estimators of MLEs of standardized factor loadings with raw-varimax rotation in factor analysis. Hayashi and Sen (1996) showed the asymptotic normality of covariance estimators of MLEs of factor loadings in factor analysis and obtained the explicit formulas for the asymptotic covariance matrix for covariance estimators of MLEs of factor loadings.

We should note that the formulas presented in Hayashi and Sen (1996) are for the unstandardized and unrotated case. In actual practice of factor analysis, however, the estimates of factor loadings  $\hat{\Lambda} = (\hat{\lambda}_{ir})$  and unique variances  $\hat{\Psi} = (\hat{\psi}_{ij})$  are often obtained from a sample correlation matrix rather than an unstandardized sample covariance matrix. In fact, most of the examples in standard factor analysis texts (e.g., Basilevsky (1994), Lawley and Maxwell (1971)) obtain solutions from a sample correlation matrix.

In addition, a rotation (i.e., a transformation) is normally performed on the factor loading matrix after the initial solutions of MLEs have been obtained. To obtain the MLEs of factor loadings, we introduce the constraints that

$$(1.4) \quad \Gamma = \Lambda \Psi^{-1} \Lambda$$

is a diagonal matrix. The constraints in (1.4) are to remove the indeterminacy regarding orthogonal rotations and for convenience from the point of view of computing the solutions of MLEs, not because the estimates of factor loadings obtained under such constraints have an intrinsic meaning. Therefore, almost

always we prefer the rotated solution which is obtained under the criterion that the interpretation of factors is easy and clear.

Thus considering the importance of standardization and rotation, the present study extends our previous work (Hayashi and Sen (1996)) to the case of standardized and rotated (estimators of) factor loadings.

## 2. Effects of standardization

When we use the correlation matrix, the analysis results in the (estimates of) standardized factor loadings  $\hat{\lambda}_{ir}^{\#}$  and the (estimates of) standardized unique variances  $\hat{\psi}_{ii}^{\#}$ , which are related to the (estimates of) unstandardized factor loadings  $\hat{\lambda}_{ir}$  and the (estimates of) unstandardized unique variances  $\hat{\psi}_{ii}$  with the following equations:

$$(2.1) \quad \hat{\lambda}_{ir}^{\#} = \frac{\hat{\lambda}_{ir}}{s_{ii}^{1/2}},$$

$$(2.2) \quad \hat{\psi}_{ii}^{\#} = \frac{\hat{\psi}_{ii}}{s_{ii}},$$

where  $s_{ii}$  is the sample variance for the  $i$ -th variable. By use of an expansion, (2.1) and (2.2) can also be expressed as

$$(2.3) \quad \hat{\lambda}_{ir}^{\#} = \hat{\lambda}_{ir} - \left(\frac{1}{2}\right)\lambda_{ir}(s_{ii} - 1) + o_p(n^{-1/2}),$$

$$(2.4) \quad \hat{\psi}_{ii}^{\#} = \hat{\psi}_{ii} - \psi_{ii}(s_{ii} - 1) + o_p(n^{-1/2}),$$

as are given in Lawley and Maxwell (1971). The matrix expressions for (2.3) and (2.4) are

$$(2.5) \quad \hat{\Lambda}^{\#} = \hat{\Lambda} - \left(\frac{1}{2}\right)(\text{Diag}(S) - I_p)\Lambda + o_p, p \times m(n^{-1/2}),$$

$$(2.6) \quad \hat{\Psi}^{\#} = \hat{\Psi} - \Psi(\text{Diag}(S) - I_p) + o_p, p \times p(n^{-1/2}),$$

where  $S$  is the sample correlation matrix;  $\text{Diag}(S)$  is a diagonal matrix whose elements are the diagonal elements of  $S$ ;  $I_p$  is the  $p \times p$  identity matrix; and the order of the matrix of the remainder term  $o_p, p \times m(n^{-1/2})$  is  $p \times m$ .

Let  $\lambda = \text{vec}(\Lambda)$  denote the  $pm \times 1$  vector listing  $m$  columns of the  $p \times m$  matrix  $\Lambda$  starting from the first column. Likewise, denote  $\psi = \text{vec}(\Psi)$ . Also, let  $S_1 = \text{vech}(S)$  denote the column vector consisting of elements on and below the diagonal of the square matrix  $S$ , starting with the first column (c.f., e.g., Searle (1982)). Then using (2.5) and (2.6), the derivatives of  $\lambda$  and  $\psi$  with respect to  $S_1$  (evaluated at  $S = \Sigma$ ) are expressed as

$$(2.7) \quad \frac{\partial \lambda^\#}{\partial S_1'} = \frac{\partial \lambda}{\partial S_1'} - \left(\frac{1}{2}\right)(\Lambda' \otimes I_p) K_{p^{**}} G_p,$$

$$(2.8) \quad \frac{\partial \psi^\#}{\partial S_1'} = \frac{\partial \psi}{\partial S_1'} - (I_p \otimes \Psi) K_{p^{**}} G_p,$$

where  $K_{p^{**}} = \text{diag}(\text{vec}(I_p))$ , i.e., the diagonal matrix whose diagonal elements are  $\text{vec}(I_p)$ ;  $G_p$  is a  $p^2 \times (1/2)p(p+1)$  unit matrix defined such that  $\text{vec}(S) = G_p \text{vech}(S)$  for any  $p \times p$  matrix  $S$ ; and  $\otimes$  is the Kronecker (direct) product operator.

Alternatively, instead of taking into account the standardization effect at the level of the derivatives of  $\lambda$  and  $\psi$  as in (2.7) and (2.8), we can make the adjustment for standardization at the level of the derivatives of the covariance estimators of the MLEs of factor loadings. Let  $V^\# = (n) \text{Cov}(\hat{\lambda}^\#)$  be the sample size  $n$  times the estimated covariance matrix for the MLEs of the standardized factor loadings  $\hat{\lambda}^\#$  arranged as a vector, and let  $\hat{V}^\# = (n) \hat{\text{Cov}}(\hat{\lambda}^\#)$  be the estimator of  $V^\#$ . Note that the order of  $V^\#$  is  $pm \times pm$ . Let  $v^\# = \text{vech}(V^\#)$  and let  $\hat{v}^\# = (n) \text{vech}(\hat{\text{Cov}}(\hat{\lambda}^\#))$  be the estimator of  $v^\#$ . The following theorem states the asymptotic normality of  $\sqrt{n}(\hat{v}^\# - v^\#)$  when we have the asymptotic normality of  $\sqrt{n}(S - \Sigma)$ .

Theorem 1. Let  $v^\# = (n)\text{vech}(\text{Cov}(\hat{\lambda}^\#))$  and let  $\hat{v}^\# = (n)\text{vech}(\hat{\text{Cov}}(\hat{\lambda}^\#))$  be the estimator of  $v^\#$ . Then  $\hat{v}^\# \rightarrow v^\#$  a.s. as  $n \rightarrow \infty$ . Furthermore, let the assumptions (i) - (iii) in Theorem 12.1 of Anderson and Rubin (1956) hold, i.e.,

(i)  $\Phi\#\Phi$  is nonsingular, i.e., the determinant  $|\Phi\#\Phi| \neq 0$ , where  $\Phi$  is defined as

$$(2.9) \quad \Phi = \Psi^{-1} - \Psi^{-1}\Lambda\Gamma^{-1}\Lambda'\Psi^{-1}.$$

(ii)  $\Lambda$  and  $\Psi$  are identified by the condition that  $\Gamma$  in (1.4) is diagonal and the diagonal elements are different and ordered.

(iii) The sample correlation matrix  $S$  converges to  $\Sigma = \Lambda\Lambda' + \Psi$ , in prob, and  $\sqrt{n}(S - \Sigma)$  has a limiting normal distribution.

Then  $\sqrt{n}(\hat{v}^\# - v^\#)$  has a limiting normal distribution.

*Proof.* Let  $V$  be the sample size  $n$  times the estimated covariance matrix for the MLEs of the unstandardized factor loadings  $\hat{\lambda}$  arranged as a vector, and let  $\hat{V}$  be the estimator of  $V$ . Let  $v = \text{vech}(V)$  and  $\hat{v}$  be the estimator of  $v$ . First note that just as  $\hat{v}$  is a function of the sample covariance matrix  $S^*$  and  $v$  is the function of the population covariance matrix  $\Sigma^*$  (See Theorems 1 and 2 of Hayashi and Sen (1996)),  $\hat{v}^\#$  is a function of the sample correlation matrix  $S$  and  $v^\#$  is the function of the population correlation matrix  $\Sigma$ . Also,  $\hat{v}^\#$  is continuous at  $S = \Sigma$ . Since  $S \rightarrow \Sigma$  a.s. as  $n \rightarrow \infty$ , we instantly obtain the strong consistency of  $\hat{v}^\#$  by use of Theorem 2.3.4 of Sen and Singer (1993). For the asymptotic normality, we express  $\sqrt{n}(\hat{v}^\# - v^\#)$  as

$$(2.10) \quad \sqrt{n}(\hat{v}^\# - v^\#) = \left(\frac{\partial v^\#}{\partial S_1'}\right) \text{vech}(\sqrt{n}(S - \Sigma)) + o_{p,q \times 1}(1),$$

where  $\frac{\partial v^\#}{\partial S_1'}$  is the matrix of partial derivatives of  $v^\#$  with respect to  $S_1 = \text{vech}(S)$  evaluated at  $S = \Sigma$ , which is a nonstochastic matrix depending only on the population parameters, and  $q = (1/2)pm(pm + 1)$ . Conditions (i) and (ii) imply that the elements of  $v^\#$  are identifiable. Thus the result follows directly from the asymptotic normality of  $\sqrt{n}(S - \Sigma)$  in condition (iii) and equation (2.10). Q.E.D.

Now, we will give the explicit expression for the covariance matrix for  $\hat{v}^\#$ :

$$(2.11) \quad \text{Cov}(\hat{v}^\#) = \left(\frac{\partial v^\#}{\partial S_1'}\right) (\text{Cov}(S_1)) \left(\frac{\partial v^\#}{\partial S_1'}\right)',$$

where  $\frac{\partial v^\#}{\partial S_1'}$  is the matrix of partial derivatives of  $v^\#$  with respect to  $S_1 = \text{vech}(S)$  evaluated at  $S = \Sigma$ . (Here,  $S$  is the sample correlation matrix and  $\Sigma$  is the population correlation matrix, as before.) The formula for  $\text{Cov}(S_1)$  is given by Neudecker and Wesselman (1990) and is

$$(2.12) \quad \text{Cov}(S_1) = L_p(\text{Cov}(\text{vec}(S^*)))L_p',$$

where

$$L_p = H_p \left\{ I_{p^2} - (1/2)(I_{p^2} + K_{pp})(I_p \otimes \Sigma)(H_p' H_p K_{pp} H_p' H_p) \right\} \\ \left\{ (dg(\Sigma^*))^{-1/2} \otimes (dg(\Sigma^*))^{-1/2} \right\};$$

$dg(U) = \text{vec}(\text{Diag}(U))$  for any square matrix  $U$ ;  $K_{pp}$  is defined such that  $K_{pp}\text{vec}(S) = \text{vec}(S')$  for any  $p \times p$  matrix  $S$ ;  $H_p = (G_p' G_p)^{-1} G_p'$ , and  $\text{Cov}(\text{vec}(S^*))$ , i.e., the formula for covariance matrix for the vector form of a sample covariance matrix  $S^*$  when the vectors of observation  $x_i$ 's are from a normal distribution, is given in, e.g., Muirhead (1982) and is

$$(2.13) \quad \text{Cov}(\text{vec}(S^*)) = \left(\frac{1}{n-1}\right)(I_{p^2} + K_{pp})(\Sigma^* \otimes \Sigma^*),$$

where  $\Sigma^*$  is the population covariance matrix. Thus our only remaining task is to obtain the explicit expression for the matrix of partial derivatives  $\frac{\partial v^\#}{\partial S_1'}$ . We employed the same approach as in Hayashi and Sen (1996) in obtaining  $\frac{\partial v^\#}{\partial S_1'}$ ; as a first step, we expressed all the coordinatewise formulas derived by Archer and Jennrich (1973) in terms of matrices. Then we obtained the necessary derivatives which map the differential of the sample correlation matrix (in vector form) into the differential of covariance estimators (in vector form). In obtaining the derivatives, we made extensive use of vec operator, Kronecker (direct) product, and the relevant properties discussed in Magnus and Neudecker (1988). The generic steps are given in section 5 of Hayashi and Sen (1996) and are replicated in the appendix. The actual derivation is, however, too involved to show here. The details of proofs are given in Hayashi and Sen (in preparation).

Let  $V = (n) \text{Cov}(\hat{\lambda}) = (n) \text{Cov}(\hat{\lambda}_{ir}, \hat{\lambda}_{js}) = (\hat{v}_{ir,js})$  be  $n$  times the covariance matrix for the MLEs of the unstandardized factor loadings, as before; let  $\text{vdg}(S)$  denote the column vector whose elements are the diagonal elements of the square matrix  $S$ ; let  $E^\# = (n/2) \text{Cov}(\text{vdg}(\hat{\psi}^\#)) = (e_{ij}^\#)$  be  $n/2$  times the covariance matrix for the MLEs of the standardized unique variances; likewise let  $E = (n/2) \text{Cov}(\text{vdg}(\hat{\psi})) = (e_{ij})$  be  $n/2$  times the covariance matrix for the MLEs of the unstandardized unique variances. The formula for  $E$  is given by Lawley and Maxwell (1971), which is

$$(2.14) \quad E = (\Phi^\# \Phi)^{-1},$$

where # is the elementwise product operator, and  $\Phi$  has already been defined in (2.9). Furthermore, let  $W = (n)(\text{Cov}(\hat{\lambda}_{ir}, s_{jj})) = (w_{ir,j})$  be  $n$  times the covariance matrix for the MLEs of the unstandardized factor loadings and the sample variances  $s_{jj}$ . Note that the order of  $V$  is  $pm \times pm$ , the order of  $W$  is  $pm \times p$ , and the orders of  $E^\#$  and  $E$  are  $p \times p$ .

The formula that connects ( $n$  times) the covariance of the standardized and the unstandardized factor loadings as well as the formula that connects ( $n$  times) the covariance of the standardized and the unstandardized unique variances are given in coordinate form in Lawley and Maxwell (1971), which are

$$(2.15) \quad v_{ir,js}^\# = v_{ir,js} - \left(\frac{1}{2}\right)\lambda_{js}w_{ir,j} - \left(\frac{1}{2}\right)w_{js,i} + \left(\frac{1}{2}\right)\lambda_{ir}\lambda_{js}\sigma_{ij}^2,$$

$$(2.16) \quad 2e_{ij}^\# = 2e_{ij} - 4\delta_{ij}\psi_{ii}^3 + 2\psi_{ii}\psi_{jj}\sigma_{ij}^2,$$

where  $\sigma_{ij}$  is the  $(i, j)$ -th element of  $\Sigma$ ;  $\delta_{ij} = 1$  if  $i = j$ , and  $\delta_{ij} = 0$  if  $i \neq j$ ; and

$$w_{ir,j} = 2 \sum_{s=1}^m (\lambda_{js}a_{ir,js}) + 2\psi_{jj}^2b_{j,ir},$$

with the expressions for the coefficients  $a_{ir,js}$  and  $b_{j,ir}$  are given in equations (16), (17), and (18) of Hayashi and Sen (1996).

Next, we express the above coordinatewise formulas in terms of matrices. The matrix expressions for the equations (2.15) and (2.16) are

$$(2.17) \quad V^\# = V - (1/2)(1_m' \otimes W)(\text{diag}(\lambda)) - (1/2)(\text{diag}(\lambda))(1_m \otimes W') \\ + (1/2)(\text{diag}(\lambda))(1_m 1_m' \otimes (\Sigma \# \Sigma))(\text{diag}(\lambda)),$$

$$(2.18) \quad 2E^\# = 2E - 4\Psi^3 + 2\Psi(\Sigma \# \Sigma)\Psi,$$

where

$$W = 2A(\text{diag}(\lambda))(1_m \otimes I_p) + 2B'\Psi^2,$$

with the matrix expressions for A and B being given in equations (21) and (28) of Hayashi and Sen (1996); and E being defined in (2.14) above. Thus the matrices of derivatives of  $v^\#$  and of  $e^\#$  with respect to  $S_1$  are obtained from (2.16) and (2.17), and are

(2.19)

$$\begin{aligned} \frac{\partial v^\#}{\partial S_1'} &= \frac{\partial v}{\partial S_1'} - (1/2)H_{pm}(I_{p^2m^2} + K_{pm,pm})(\text{diag}(\lambda) \otimes I_{pm}) \\ &\quad (I_m \otimes K_{p1} \otimes I_{pm})(1_m \otimes I_{p^2m})\left(\frac{\partial w}{\partial S_1'}\right) \\ &\quad + H_{pm}\{\text{diag}(\lambda) \otimes \text{diag}(\lambda)\}(I_m \otimes K_{pm} \otimes I_p)(1_{m^2} \otimes I_{p^2})G_p(\text{diag}(\sigma)) \\ &\quad + (1/2)H_{pm}\{(\text{diag}(\lambda))(1_m 1_m' \otimes (\Sigma \# \Sigma)) \otimes I_{pm} - 1_m' \otimes W \otimes I_{pm} \\ &\quad + I_{pm} \otimes (\text{diag}(\lambda))(1_m 1_m' \otimes (\Sigma \# \Sigma)) - I_{pm} \otimes 1_m' \otimes W\}K_{pm}^*\left(\frac{\partial \lambda}{\partial S_1'}\right), \end{aligned}$$

(2.20)

$$\begin{aligned} \frac{\partial e^\#}{\partial S_1'} &= \frac{\partial e}{\partial S_1'} + H_p\{\Psi(\Sigma \# \Sigma) \otimes I_p + I_p \otimes \Psi(\Sigma \# \Sigma) - 6(I_p \otimes \Psi^2)\}\left(\frac{\partial \psi_1}{\partial S_1'}\right) \\ &\quad + 2H_p(\Psi \otimes \Psi)G_p(\sigma 1_q'), \end{aligned}$$

where

$$\begin{aligned} \frac{\partial w}{\partial S_1'} &= 2\{(1_m' \otimes I_p)(\text{diag}(\lambda)) \otimes I_{pm}\}G_{pm}\left(\frac{\partial a}{\partial S_1'}\right) + 2(\Psi^2 \otimes I_{pm})K_{p,pm}\left(\frac{\partial b}{\partial S_1'}\right) \\ &\quad + 2(1_m' \otimes I_p \otimes A)K_{pm}^*\left(\frac{\partial \lambda}{\partial S_1'}\right) + 4(I_p \otimes B' \Psi)\left(\frac{\partial \psi_1}{\partial S_1'}\right), \end{aligned}$$

and  $H_{pm}$  and  $G_{pm}$  are of order  $(1/2)pm(pm + 1) \times p^2m^2$  and  $p^2m^2 \times (1/2)pm(pm + 1)$ , and defined such that  $\text{vec}(V^*) = G_{pm}\text{vech}(V^*)$  and  $H_{pm} = (G_{pm}'G_{pm})^{-1}G_{pm}'$ ;  $K_{a,b}$  is of order  $ab \times ab$  and is defined such that  $\text{vec}(Y') = K_{a,b}\text{vec}(Y)$  for any  $a \times b$  matrix  $Y$ ;  $K_{pm}^*$  is of order  $p^2m^2 \times pm$  and is defined such that  $K_{pm}^*\lambda = \text{vec}(\text{diag}(\lambda))$ ;  $q^{**} = (1/2)p(p + 1)$  and  $\sigma = \text{vech}(\Sigma)$ .

Finally, if we wish to obtain the asymptotic covariance matrix for  $\hat{e}^\# = (n/2)\text{vech}(\hat{\text{Cov}}(\text{vdg}(\hat{\psi}^\#)))$ , i.e.  $\text{Cov}(\hat{e}^\#)$ , replace  $\frac{\partial v^\#}{\partial S_1'}$  in RHS of (2.11) by  $\frac{\partial e^\#}{\partial S_1'}$  in (2.19). The strong consistency of  $\hat{e}^\#$  and the asymptotic normality of  $\sqrt{n}(\hat{e}^\# - e^\#)$  are proven in exactly the same way as in the proof for Theorem 1.

### 3. Raw varimax rotation

Next, we discuss the effects of a rotation. There are several different criteria for rotation. In this study, we employ the raw varimax rotation (Kaiser (1958)), which is a special case of a class called the orthomax criteria. (See e.g., Harman (1976) for the orthomax criteria and other criteria for rotation.)

Let  $\Lambda^* = (\lambda_{ir}^*)$  be the factor loading matrix after rotation;  $\Lambda = (\lambda_{ir})$  be the factor loading matrix before rotation, and  $T = (t_{sr})$  is the  $m \times m$  transformation matrix. Then the raw varimax rotation is an orthogonal rotation

$$(3.1) \quad \Lambda^* = \Lambda T \quad \text{and} \quad T'T = I_m$$

such that the variance of squared factor loadings

$$\left(\frac{1}{p}\right) \sum_{j=1}^m \left\{ \sum_{i=1}^p \lambda_{ij}^4 - \left(\frac{1}{p}\right) \left(\sum_{i=1}^p \lambda_{ij}^2\right)^2 \right\}$$

is maximized.

The necessary formulas for the covariance matrix of estimators of factor loadings with raw varimax rotation are given in Archer and Jennrich (1973), Jennrich (1974), and Jennrich and Clarkson (1980). Let  $\hat{\Lambda}^*$  be the raw-varimax rotated estimator of factor loading matrix and let  $\hat{\lambda}^* = \text{vec}(\hat{\Lambda}^*)$  be its vector form. Let  $V^* = (n)\text{Cov}(\hat{\lambda}^*) = (n \text{Cov}(\hat{\lambda}_{ir}^*, \hat{\lambda}_{js}^*))$  be the sample size  $n$  times covariance

matrix for the estimators of raw Varimax rotated factor loadings arranged as a vector, and  $V = (n)\text{Cov}(\hat{\lambda}) = (n \text{Cov}(\hat{\lambda}_{ir}, \hat{\lambda}_{js}))$  be  $n$  times covariance matrix for the estimators of unrotated factor loadings arranged as a vector. Note that both  $V^*$  and  $V$  are  $pm \times pm$  matrices. The coordinatewise expression given in Archer and Jennrich (1973) is

$$(3.2) \quad \text{Cov}(\hat{\lambda}_{ir}^*, \hat{\lambda}_{js}^*) = \sum_{mnuv} \frac{\partial h_{ir}}{\partial \lambda_{mu}} \text{Cov}(\hat{\lambda}_{mu}, \hat{\lambda}_{nv}) \frac{\partial h_{js}}{\partial \lambda_{nv}},$$

where

$$(3.3) \quad \frac{\partial h_{ir}}{\partial \lambda_{js}} = \delta_{ijtsr} - \sum_{u=1}^{m-1} \sum_{v=u+1}^m \sum_{t=1}^m (e_{iruv} \left( \frac{\partial \xi_{uv}}{\partial \lambda_{it}^*} \right) t_{st}).$$

Here,  $t_{sr}$  is the  $(s, r)$  element of the transformation matrix  $T$ , and  $E^* = (e_{iruv})$  has elements

$$(3.4) \quad e_{iruv} = \sum_{t=1}^{r-1} \lambda_{it}^* L^{l(t,r),l(u,v)} - \sum_{t=r+1}^m \lambda_{it}^* L^{l(r,t),l(u,v)},$$

where  $L^{ij}$  in equation (3.4) is the  $(i,j)$  element of the inverse of the matrix  $L$ , and the  $(l(r,s), l(u,v))$  element of  $L$  is

$$(3.5) \quad L_{l(r,s),l(u,v)} = \sum_{i=1}^p \left( \lambda_{iu}^* \frac{\partial \xi_{rs}}{\partial \lambda_{iv}^*} - \lambda_{iv}^* \frac{\partial \xi_{rs}}{\partial \lambda_{iu}^*} \right),$$

with  $l(r, s) = (1/2)(r - 1)(2m - r) + (s - r)$ ,  $1 \leq r < s \leq m$ , and  $l(u, v) = (1/2)(u - 1)(2m - u) + (v - u)$ ,  $1 \leq u < v \leq m$ . For  $1 \leq r < s \leq m$ ,

$$(3.6) \quad \frac{\partial \xi_{rs}}{\partial \lambda_{ir}^*} = 3\lambda_{ir}^* 2\lambda_{is}^* - \lambda_{is}^* 3 - \left( \frac{1}{p} \right) \left\{ \lambda_{is}^* \sum_{j=1}^p (\lambda_{jr}^* 2 - \lambda_{js}^* 2) + 2 \lambda_{ir}^* \sum_{j=1}^p (\lambda_{jr}^* \lambda_{js}^*) \right\},$$

$$\frac{\partial \xi_{rs}}{\partial \lambda_{is}^*} = - \frac{\partial \xi_{sr}}{\partial \lambda_{is}^*},$$

$$\frac{\partial \xi_{rs}}{\partial \lambda_{it}^*} = 0, \text{ for all } t \neq s.$$

Now, we give the matrix expressions for all the coordinatewise formulas presented above. First, n time the equation (3.2) is expressed in matrix form as

$$(3.7) \quad V^* = \left( \frac{\partial h}{\partial \lambda'} \right) V \left( \frac{\partial h}{\partial \lambda'} \right)',$$

where  $\frac{\partial h}{\partial \lambda'}$  is the matrix expression for  $\frac{\partial h_{ir}}{\partial \lambda_{js}}$  in (3.3) and is of the form

$$(3.8) \quad \frac{\partial h}{\partial \lambda'} = (I_{pm} - E^* \frac{\partial \xi}{\partial \lambda^{*'}})(T' \otimes I_p),$$

with the matrix expressions corresponding to equations (3.4) - (3.6) given by the following three equations:

$$(3.9) \quad E^* = (I_m \otimes \Lambda^*) K_m L^{-1},$$

$$(3.10) \quad L = \left( \frac{\partial \xi}{\partial \lambda^{*'}} \right) (I_m \otimes \Lambda^*) (J_m^{(2)} - J_m^{(1)}),$$

$$(3.11) \quad \frac{\partial \xi}{\partial \lambda^{*'}} = J_m^{(2)'} (K_{mm} - I_{m2}) \sum_{i=1}^m \left\{ D_{m2,i} K_{mm} \left( \frac{\partial \xi_{(+)} }{\partial \lambda^{*'}} \right) D_{pm,i} \right\},$$

where

$$(3.12) \quad \begin{aligned} \frac{\partial \xi_{(+)} }{\partial \lambda^{*'}} &= 3(1_m 1_m' \otimes (\Lambda^* \# \Lambda^*)) \# (1_m' \otimes \Lambda^{*'} \otimes 1_m) - 1_m' \otimes (\Lambda^* \# \Lambda^* \# \Lambda^*)' \otimes 1_m \\ &- \left( \frac{1}{p} \right) \left[ (1_m' \otimes \Lambda^{*'} \otimes 1_m) \# \left\{ 1_m \otimes J_m^{(3)} (\text{vec}(\Lambda^* \Lambda^*)) 1_{pm}' \right. \right. \\ &\left. \left. - J_m^{(3)} (\text{vec}(\Lambda^* \Lambda^*)) \otimes 1_m 1_{pm}' \right\} + 2(1_m 1_m' \otimes \Lambda^*) \# (\text{vec}(\Lambda^* \Lambda^*)) 1_{pm}' \right]; \end{aligned}$$

$K_m = (K_{(1)'}, K_{(2)'}, \dots, K_{(m)'})'$  is a  $m^2 \times (1/2)m(m-1)$  matrix with a  $m \times (1/2)m(m-1)$  unit matrix  $K_{(r)}$  which has 1's in the  $(i, l(i,r))$  elements,  $1 \leq i \leq r-1$ ; -1's in the  $(r+i, l(r, r+i))$  elements,  $1 \leq i \leq m-r$ ; 0's in the rest of the elements; and  $l(i, r)$  is

defined such that  $l(i, r) = (1/2)(i - 1)(2m - i) + (r - i)$ ,  $1 \leq i < r \leq m$ . (The components of  $K_{(r)}$  are arranged as in Table 1.)  $D_{m^2, i}$  is a  $m^2 \times m^2$  block diagonal matrix with  $m$  blocks (of order  $m \times m$ ) whose  $i$ -th block is an identity matrix and the rest of  $m - 1$  blocks are null matrices;  $D_{pm, i}$  is a  $pm \times pm$  block diagonal matrix with  $m$  blocks (of order  $p \times p$ ) whose  $i$ -th block is an identity matrix and the rest of  $m - 1$  blocks are null matrices. That is,  $D_{m^2, i} = (J_{m, i} J_{m, i}') \otimes I_m$  and  $D_{pm, i} = (J_{m, i} J_{m, i}') \otimes I_p$  where  $J_{m, i} = (0, \dots, 0, 1, 0, \dots, 0)'$  is a  $m$ -dimensional unit vector whose  $i$ -th element is 1 and the rest are 0's.  $J_m^{(1)}$  and  $J_m^{(2)}$  in (3.10) and (3.11) are  $m^2 \times (1/2)m(m - 1)$  unit matrices whose components are presented in Tables 2 and Table 3, respectively. Finally,  $J_m^{(3)}$  in (3.12) is a  $m \times m^2$  unit matrix defined as

$$\begin{array}{c}
 \text{(Column numbers of 1's)} \\
 \begin{array}{cccccc}
 & 1 & m+2 & 2m+3 & (m-2)m+(m-1) & m^2 \\
 \text{(row 1)} & 1 & 0 & \dots & \dots & 0 \\
 \text{(row 2)} & 0 & 0 & \dots & 0 & 1 & 0 & \dots & \dots & 0 \\
 \text{(row 3)} & 0 & 0 & \dots & \dots & 0 & 1 & 0 & \dots & 0 \\
 \dots & \dots \\
 \text{(row } m-1) & 0 & 0 & \dots & \dots & 0 & 1 & 0 & \dots & 0 \\
 \text{(row } m) & 0 & 0 & \dots & \dots & \dots & \dots & 0 & 1 & 0
 \end{array}
 \end{array}
 \left[ \begin{array}{c} \\ \\ \\ \\ \\ \\ \\ \\ \\ \end{array} \right]$$

Let  $\hat{V}^* = (n)\hat{Cov}(\hat{\lambda}^*)$  be the estimator of  $V^*$ . Let  $v^* = \text{vech}(V^*)$  and  $\hat{v}^* = \text{vech}(\hat{V}^*)$ . The following theorem states the strong consistency of  $\hat{v}^*$  and the asymptotic normality of  $\sqrt{n}(\hat{v}^* - v^*)$  when we have the asymptotic normality of  $\sqrt{n}(S - \Sigma)$ .

**Theorem 2.** *Let  $v^* = (n)\text{vech}(\text{Cov}(\hat{\lambda}^*))$  and let  $\hat{v}^* = (n)\text{vech}(\hat{Cov}(\hat{\lambda}^*))$  be the estimator of  $v^*$ . Then  $\hat{v}^* \rightarrow v^*$  a.s. as  $n \rightarrow \infty$ . Furthermore, let the assumptions (i) - (iii) in Theorem 1 hold. Then  $\sqrt{n}(\hat{v}^* - v^*)$  has a limiting normal distribution.*

*Proof.* Exactly as in the proof of Theorem 1, replacing  $\hat{v}^\#$  and  $v^\#$  with  $\hat{v}^*$  and  $v^*$ , respectively. Q.E.D.

Now, we will give the explicit expression for the covariance matrix for  $\hat{v}^*$ ; it is too involved to include the actual derivation steps. Details of proofs are, again, given in Hayashi and Sen (in preparation). The covariance matrix  $\text{Cov}(\hat{v}^*)$  is given by

$$(3.13) \quad \text{Cov}(\hat{v}^*) = \left( \frac{\partial v^*}{\partial S_1'} \right) (\text{Cov}(S_1)) \left( \frac{\partial v^*}{\partial S_1'} \right)',$$

where  $\frac{\partial v^*}{\partial S_1'}$  is the matrix of partial derivatives of  $v^*$  with respect to  $S_1 = \text{vech}(S)$  evaluated at  $S = \Sigma$ . The formula for  $\text{Cov}(S_1)$  is given by (2.12) for the standardized case, or by (2.13) (which need to be premultiplied by  $H_p$  and postmultiplied by  $H_p'$ ) for the unstandardized case. What remains to be given is the explicit expression for the derivative of  $V^*$  with respect to  $S_1$ . The matrix of derivatives of  $v^*$  can be obtained exactly in the same way as in the standardized case discussed in section 2. Let  $h^* = \text{vec}\left(\frac{\partial h}{\partial \lambda'}\right)$ . Then by the chain rule,

$$\frac{\partial v^*}{\partial S_1'} = \left( \frac{\partial v^*}{\partial v'} \right) \left( \frac{\partial v}{\partial S_1'} \right) + \left( \frac{\partial v^*}{\partial h^{*'}} \right) \left( \frac{\partial h^*}{\partial S_1'} \right),$$

where the expression for  $\frac{\partial v}{\partial S_1'}$  is given in equation (40) of Hayashi and Sen (1996) for the unstandardized case, or, for the standardized case, it is given by  $\frac{\partial v^\#}{\partial S_1'}$  in (2.19) in section 2; and

$$(3.14) \quad \begin{aligned} \frac{\partial v^*}{\partial v'} &= H_{pm} \left( \frac{\partial h}{\partial \lambda'} \otimes \frac{\partial h}{\partial \lambda'} \right) G_{pm}, \\ \frac{\partial v^*}{\partial h^{*'}} &= H_{pm} (I_{p^2 m^2} + K_{pm, pm}) \left( \left( \frac{\partial h}{\partial \lambda'} \right) V \otimes I_{pm} \right), \end{aligned}$$

$$(3.15) \quad \frac{\partial h^*}{\partial S_{1'}} = - (T \otimes I_{p^2 m}) \left\{ \left( \left( \frac{\partial \xi}{\partial \lambda^{*'}} \right)' \otimes I_{pm} \right) \left( \frac{\partial e^*}{\partial S_{1'}} \right) + (I_{pm} \otimes E^*) \left( \frac{\partial \xi^*}{\partial S_{1'}} \right) \right\} \\ + \left\{ I_{pm} \otimes (I_{pm} - E^* \left( \frac{\partial \xi}{\partial \lambda^{*'}} \right)) \right\} (I_m \otimes K_{pm} \otimes I_p) (I_{m^2} \otimes \text{vec}(I_p)) \\ K_{mm} \left( \frac{\partial t}{\partial S_{1'}} \right),$$

where  $e^* = \text{vec}(E^*)$ ;  $\xi^* = \text{vec}\left(\frac{\partial \xi}{\partial \lambda^{*'}}\right)$ ;  $\frac{\partial \xi}{\partial \lambda^{*}}$  and  $\frac{\partial h}{\partial \lambda'}$  in (3.15) and (3.14) are given in (3.11) and (3.8), respectively. Furthermore, note that  $\frac{\partial h^*}{\partial S_{1'}}$  in (3.15) is expressed in terms of the three matrices of derivatives:  $\frac{\partial e^*}{\partial S_{1'}}$ ,  $\frac{\partial \xi^*}{\partial S_{1'}}$ , and  $\frac{\partial t}{\partial S_{1'}}$ . First,  $\frac{\partial e^*}{\partial S_{1'}}$  is of the form

$$(3.16) \quad \frac{\partial e^*}{\partial S_{1'}} = (L^{-1} \otimes I_{pm}) (K_m' \otimes I_{pm}) (I_m \otimes K_{mm} \otimes I_p) (\text{vec}(I_m) \otimes I_{pm}) \left( \frac{\partial \lambda^*}{\partial S_{1'}} \right) \\ - \left\{ I_{q^*} \otimes ((I_m \otimes \Lambda^*) K_m) \right\} (L^{-1} \otimes L^{-1}) G_{q^*} \left( \frac{\partial \ell}{\partial S_{1'}} \right),$$

where  $\ell = \text{vech}(L)$ ,  $q^* = (1/2)m(m-1)$ ; and  $G_{pm}$ ,  $H_{pm}$ , and  $G_{q^*}$  are defined such that  $\text{vec}(V) = G_{pm}v$ ,  $H_{pm} = (G_{pm}'G_{pm})^{-1}G_{pm}'$ , and  $\text{vec}(L) = G_{q^*}\text{vech}(L) = G_{q^*}\ell$ , respectively. Second,  $\frac{\partial \xi^*}{\partial S_{1'}}$  is of the form

$$(3.17) \quad \frac{\partial \xi^*}{\partial S_{1'}} = \left\{ I_{pm} \otimes J_m^{(2)'} (K_{mm} - I_{m^2}) \right\} \left\{ \sum_{i=1}^m (D_{pm,i} \otimes D_{m^2,i} K_{mm}) \right\} \left( \frac{\partial \xi_{(+)}^*}{\partial S_{1'}} \right),$$

where  $\xi_{(+)}^* = \text{vec}\left(\frac{\partial \xi_{(+)}^*}{\partial \lambda^{*'}}\right)$ , and finally the expression for  $\frac{\partial t}{\partial S_{1'}}$  is

$$\frac{\partial t}{\partial S_{1'}} = G_{(m)} \left( \frac{\partial \xi}{\partial \lambda^{*'}} \right) (T' \otimes I_p) \left( \frac{\partial \lambda}{\partial S_{1'}} \right),$$

where the element of  $G_{(m)}$  (of order  $m^2 \times (1/2)m(m - 1)$ ) is obtained by noting the similarity to  $e_{iruv}$  in (3.4) and is of the form

$$g_{iruv} = \sum_{t=1}^{r-1} t_{it} L^{l(t,r),l(u,v)} - \sum_{t=r+1}^m t_{it} L^{l(r,t),l(u,v)},$$

with  $l(r,s) = (1/2)(r - 1)(2m - r) + (s - r)$ ;  $1 \leq i, r, s \leq m$ ,  $1 \leq j \leq p$ ,  $1 \leq u < v \leq m$ , or alternatively,  $G_{(m)}$  has the matrix form

$$G_{(m)} = (I_m \otimes T) K_m L^{-1},$$

where  $K_m = (K_{(1)'}, K_{(2)'}, \dots, K_{(m)'})'$  is, as before, a  $m^2 \times (1/2)m(m - 1)$  matrix with a  $m \times (1/2)m(m - 1)$  unit matrix  $K_{(r)}$  defined in Table 1.

Next,  $\frac{\partial e^*}{\partial S_{1'}}$  in (3.16) is expressed in terms of the two matrices of derivatives  $\frac{\partial \lambda^*}{\partial S_{1'}}$  and  $\frac{\partial \ell}{\partial S_{1'}}$ , and we need the expressions for these. First, the expression for  $\frac{\partial \lambda^*}{\partial S_{1'}}$  is obtained from (3.1) and is of the form

$$\frac{\partial \lambda^*}{\partial S_{1'}} = (T' \otimes I_p) \left( \frac{\partial \lambda}{\partial S_{1'}} \right) + (I_m \otimes \Lambda) \left( \frac{\partial t}{\partial S_{1'}} \right),$$

and  $\frac{\partial \ell}{\partial S_{1'}}$ , which is another matrix of derivatives inside the expression of

$\frac{\partial e^*}{\partial S_{1'}}$  in (3.16), has the form

$$\begin{aligned} \frac{\partial \ell}{\partial S_{1'}} &= H_{q^*} \left\{ (J_m^{(2)'} - J_m^{(1)'}) (I_m \otimes \Lambda^{*'}) \otimes I_{q^*} \right\} \left( \frac{\partial \xi^*}{\partial S_{1'}} \right) \\ &+ H_{q^*} \left\{ (J_m^{(2)'} - J_m^{(1)'}) \otimes \left( \frac{\partial \xi}{\partial \lambda^{*'}} \right) \right\} (I_m \otimes K_{mm} \otimes I_p) (\text{vec}(I_m) \otimes I_{pm}) \left( \frac{\partial \lambda^*}{\partial S_{1'}} \right), \end{aligned}$$

where  $H_{q^*}$  is a matrix with dimension  $(1/2)q^*(q^* + 1) \times q^{*2} = (1/8)m(m - 1)(m^2 - m + 2) \times (1/4)m^2(m - 1)^2$  such that  $H_{q^*}\text{vec}(L) = \text{vech}(L) = \ell$ . Finally, we need one more matrix of derivatives  $\frac{\partial \xi_{(+)}^*}{\partial S_{1'}}$  to obtain  $\frac{\partial \xi^*}{\partial S_{1'}}$  in (3.17). The formula for  $\frac{\partial \xi_{(+)}^*}{\partial S_{1'}}$  is given by

$$\begin{aligned} \frac{\partial \xi_{(+)}^*}{\partial S_{1'}} &= [6E_1(\lambda^* 1_{q^{**}} \# (\frac{\partial \lambda^*}{\partial S_{1'}}))] \# [E_2(K_{pm} \lambda^* \otimes 1_m) 1_{q^{**}}] \\ &+ [3E_1(\lambda^* \# \lambda^*) 1_{q^{**}}] \# [E_2 E_3 (\frac{\partial \lambda^*}{\partial S_{1'}})] \\ &- 3E_4 E_3 (\lambda^* 1_{q^{**}} \# \lambda^* 1_{q^{**}} \# (\frac{\partial \lambda^*}{\partial S_{1'}})) \\ &- (\frac{1}{p}) \{ [E_4 E_3 (\frac{\partial \lambda^*}{\partial S_{1'}})] \# [E_5 (I_m \otimes \Lambda^*) \lambda^* 1_{q^{**}}] \\ &+ [E_4 (K_{pm} \lambda^* \otimes 1_m) 1_{q^{**}}] \# [E_5 \{ (\Lambda^* \otimes I_m) K_{pm} + (I_m \otimes \Lambda^*) \} (\frac{\partial \lambda^*}{\partial S_{1'}})] \\ &+ [2E_1 (\frac{\partial \lambda^*}{\partial S_{1'}})] \# [E_6 (I_m \otimes \Lambda^*) \lambda^* 1_{q^{**}}] \\ &+ [E_7 (1_{m^2} \otimes K_{pm} \lambda^*) 1_{q^{**}}] \# [E_6 \{ (\Lambda^* \otimes I_m) K_{pm} + (I_m \otimes \Lambda^*) \} (\frac{\partial \lambda^*}{\partial S_{1'}})] \}, \end{aligned}$$

where  $q^{**} = (1/2)p(p + 1)$ , and

$$\begin{aligned} E_1 &= (I_m \otimes K_{pm} \otimes I_m) (1_{m^2} \otimes I_{pm}) K_{pm}, \\ E_2 &= (I_m \otimes K_{p1} \otimes I_{m^2}) (1_m \otimes I_{pm^2}) (I_p \otimes K_{1m} \otimes I_m), \\ E_3 &= (I_{pm} \otimes 1_m) K_{pm}, \\ E_4 &= (I_m \otimes K_{p1} \otimes I_{m^2}) (1_m \otimes I_{pm^2}) (I_p \otimes K_{1m} \otimes I_m), \\ E_5 &= (K_{pm,m} \otimes I_m) \{ (1_m \otimes I_{pm^2}) (1_{pm} \otimes J_m^{(3)}) - (I_m \otimes 1_{pm^2}) J_m^{(3)} \}, \\ E_6 &= 1_{pm} \otimes I_{m^2}, \\ E_7 &= 2(I_m \otimes K_{pm} \otimes I_m). \end{aligned}$$

In summary, we have given the formulas for the matrices of partial derivatives, which are necessary for the expression of the asymptotic covariance matrix for

covariance estimators  $\hat{\nu}^*$  for estimates of factor loadings with raw-varimax rotation given in equation (3.13). The matrices of derivatives are defined in a nested fashion and may look more complicated than those necessary for computing the asymptotic covariance matrix for  $\hat{\nu}^\#$  for the standardized case defined in equation (2.11), however, both of them are straightforward matrix expressions.

## REFERENCES

- Anderson, T. W., and Rubin, H. (1956). Statistical inference in factor analysis. *Proc. Third Berkeley Symp. Math. Statist. Probab.*, **5**, 111-150. Univ. California Press.
- Archer, C. O., and Jennrich, R. I. (1973). Standard errors for orthogonally rotated factor loadings. *Psychometrika*, **38**, 581-592.
- Basilevsky, A. (1994). *Statistical factor analysis and related methods: Theory and applications*. Wiley, New York.
- Clarke, M. R. B. (1970). A rapidly convergent method for maximum likelihood factor analysis. *British J. Math. Statist. Psychol.*, **23**, 43-52.
- Harman, H. H. (1976). *Modern factor analysis (3rd ed.)*. Univ. of Chicago Press, Chicago.
- Hayashi, K., and Sen, P. K. (1996). The asymptotic covariance matrix for covariance estimators in factor analysis. *Institute of Statistics Mimeo Series*, No. 2172, The University of North Carolina.
- Hayashi, K., and Sen, P. K. (in preparation). *Matrix computations involved in factor analysis: The standardized, raw-varimax rotated case*.
- Jennrich, R. I. (1973). Standard errors for obliquely rotated factor loadings. *Psychometrika*, **38**, 593-604.
- Jennrich, R. I. (1974). Simplified formulae for standard errors in maximum-likelihood factor analysis. *British J. Math. Statist. Psychol.*, **27**, 122-131.
- Jennrich, R. I., and Clarkson, D. B. (1980). A feasible method for standard errors of estimate in maximum likelihood factor analysis. *Psychometrika*, **45**, 237-247.
- Jennrich, R. I., and Robinson, S. M. (1969). A Newton-Raphson algorithm for maximum likelihood factor analysis. *Psychometrika* **34**, 111-123.

- Jennrich, R. I., and Thayer, D. T. (1973). A note on Lawley's formulas for standard errors in maximum likelihood factor analysis. *Psychometrika* **38**, 571-580.
- Jöreskog, K. G. (1967). Some contributions to maximum likelihood factor analysis. *Psychometrika* **32**, 443-482.
- Jöreskog, K. G. (1977). Factor analysis by least-squares and maximum likelihood methods. In *Statistical methods for digital computers*, Vol. 3. Wiley, New York.
- Kaiser, H. F. (1958). The varimax criterion for analytic rotation in factor analysis. *Psychometrika*, **23**, 187-200.
- Lawley, D. N. (1967). Some new results in maximum likelihood factor analysis *Proc. Roy. Soc. Edinburgh Sect. A*, **67**, 256-264.
- Lawley, D. N., and Maxwell, A. E. (1971). *Factor analysis as a statistical method (2nd ed.)*. American Elsevier, New York.
- Lee, S.-Y., and Jennrich, R. I. (1979). A study of algorithms for covariance structure analysis with specific comparisons using factor analysis. *Psychometrika*, **44**, 99-113.
- Magnus, J. R., and Neudecker, H. (1988). *Matrix differential calculus with applications in statistics and econometrics*. Wiley, New York.
- Muirhead, R. J. (1982). *Aspects of multivariate statistical theory*. Wiley, New York.
- Neudecker, H., and Wesselman, A.M. (1990). The asymptotic variance matrix of the sample correlation matrix. *Linear Algebra Appl.*, **127**, 589-599.
- Searle, S. R. (1982). *Matrix algebra useful for statistics*. Wiley, New York.
- Sen, P. K., and Singer, J. M. (1990). *Large sample methods in statistics: An introduction with applications*. Chapman and Hall, New York.

Table 1. Components of unit matrix  $K_{(r)}$

<u>Column no. of <math>K_{(r)}</math></u>	<u>Row no. of <math>K_{(r)}</math></u>									
	1	2	3	r-2	r-1	r	r+1	r+2	r+3	m
1	0	.....	.....	.....	.....	.....	.....	.....	.....	0
r - 2	0	.....	.....	.....	.....	.....	.....	.....	.....	0
$l(1,r) = r - 1$	1	0	.....	.....	.....	.....	.....	.....	.....	0
r	0	.....	.....	.....	.....	.....	.....	.....	.....	0
	0	.....	.....	.....	.....	.....	.....	.....	.....	0
$l(2,r) = (r-1)+(m-2)$	0	1	0	.....	.....	.....	.....	.....	.....	0
	0	.....	.....	.....	.....	.....	.....	.....	.....	0
	0	.....	.....	.....	.....	.....	.....	.....	.....	0
$l(r-1,r) = (1/2)(r-2)(2m-r+1)+1$	0	.....	.....	.....	0	1	0	.....	.....	0
	0	.....	.....	.....	.....	.....	.....	.....	.....	0
	0	.....	.....	.....	.....	.....	.....	.....	.....	0
$l(r,r+1) = (1/2)(r-1)(2m-r)+1$	0	.....	.....	.....	.....	.....	0	-1	0	.....
	0	.....	.....	.....	.....	.....	.....	0	-1	0
	0	.....	.....	.....	.....	.....	.....	0	-1	0
	0	.....	.....	.....	.....	.....	.....	.....	.....	0
$l(r,m) = (1/2)(r-1)(2m-r)+(m-r)$	0	.....	.....	.....	.....	.....	.....	.....	.....	-1
	0	.....	.....	.....	.....	.....	.....	.....	.....	0
	0	.....	.....	.....	.....	.....	.....	.....	.....	-1
	0	.....	.....	.....	.....	.....	.....	.....	.....	0
$(1/2)m(m - 1)$	0	.....	.....	.....	.....	.....	.....	.....	.....	0

Table 2. Components of unit matrix  $J_m^{(1)}$

$0_{1 \times (m-1)}$	$0_{1 \times (m-2)}$	0	0	0
$I_{m-1}$	$0_{(m-1) \times (m-2)}$	0	0	0
$0_{2 \times (m-1)}$	$0_{2 \times (m-2)}$	0	0	0
$0_{(m-2) \times (m-1)}$	$I_{m-2}$	0	0	0
0	$0_{3 \times (m-2)}$	0	0	0
.....	.....	.....	.....	.....
0	0	0	$0_{(m-2) \times 2}$	$0_{(m-2) \times 1}$
0	0	0	$I_2$	$0_{2 \times 1}$
0	0	0	$0_{(m-1) \times 2}$	$0_{(m-1) \times 1}$
0	0	0	$0_{1 \times 2}$	$I_1$
0	0	0	$0_{m \times 2}$	$0_{m \times 1}$

Note: I's are identity (sub)matrices and 0's are null (sub)matrices. The subscripts indicate the order of the submatrix.

Table 3. Components of unit matrix  $J_m^{(2)}$

$0_{m \times 1}$	$0_{m \times 2}$	0	0	0
$I_1$	$0_{1 \times 2}$	0	0	0
$0_{(m-1) \times 1}$	$0_{(m-1) \times 2}$	0	0	0
$0_{2 \times 1}$	$I_2$	0	0	0
$0_{(m-2) \times 1}$	$0_{(m-2) \times 2}$	0	0	0
.....	.....	.....	.....	.....
0	0	0	$0_{3 \times (m-2)}$	0
0	0	0	$I_{m-2}$	$0_{(m-2) \times (m-1)}$
0	0	0	$0_{2 \times (m-2)}$	$0_{2 \times (m-1)}$
0	0	0	$0_{(m-1) \times (m-2)}$	$I_{m-1}$
0	0	0	$0_{1 \times (m-2)}$	$0_{1 \times (m-1)}$

Note: I's are identity (sub)matrices and 0's are null (sub)matrices. The subscripts indicate the order of the submatrix.

## APPENDIX

The matrices of partial derivatives in this paper were obtained in the following generic steps: Suppose we wish to obtain the matrix of partial derivatives of the product of two matrices in vector form, i.e.,  $\text{vec}(UW)$  with respect to  $S_1 = \text{vech}(S)$ . Let the orders of  $U$  and  $W$  be  $m \times n$  and  $n \times q$ , respectively.

(i) *Obtain the differential for the matrix. Use the product rule of the differential:*

$$d(UW) = (dU)W + U(dW) + o_{p, m \times q}(n^{-1/2}).$$

(ii) *Put vec operator to the differential, forming*

$$\text{vec}(d(UW)) = \text{vec}((dU)W) + \text{vec}(U(dW)) + o_{p, mq \times 1}(n^{-1/2}).$$

(iii) *Move the vec operator of the differential to the right-most side of the equation by use of formulas given in, e.g., Magnus and Neudecker (1988), i.e.,*

$$d(\text{vec}(UW)) = (W' \otimes I_m) \text{vec}(dU) + (I_q \otimes U) \text{vec}(dW) + o_{p, mq \times 1}(n^{-1/2}).$$

(iv) *Form a matrix of derivatives from the differential, i.e.,*

$$\frac{\partial}{\partial S_1'}(\text{vec}(UW)) = (W' \otimes I_m) \left( \frac{\partial u}{\partial S_1'} \right) + (I_q \otimes U) \left( \frac{\partial w}{\partial S_1'} \right),$$

where  $u = \text{vec}(U)$  and  $w = \text{vec}(W)$ .