Economic Review, Vol.49, No.1, 22-69,
Otaru University of Commerce, July 1998

Standard Errors of Several Indices for Unrotated and Rotated Factors

Haruhiko Ogasawara

Abstract

The asymptotic standard errors of several indices for unrotated and rotated factors are derived. These are obtained from the augmented information matrix for parameters with restrictions. For unrotated factors, the asymptotic standard errors of communalities are presented. In the case of orthogonal rotation, the asymptotic standard errors for contributions of rotated factors are obtained. For oblique rotation, the asymptotic standard errors of loadings, structures, factor correlations and four types of factor contributions are derived. Accuracies of these estimators are evaluated by the Monte Carlo simulation with true values of parameters.

Keywords: standard errors, varimax, quartimin, communalities, contributions, reference factors, augmented information matrix.

Request for reprints should be sent to Haruhiko Ogasawara, Otaru University of Commerce, 3-5-21 Midori, Otaru 047-8501 Japan. E-mail: hogasa@res.otaru-uc.ac.jp

Results of factor analysis can be described by various indices, i.e., factor loadings, factor structures, correlations among factors, communalities, factor contributions and so on. For evaluating estimates of these indices, information on their statistical behavior is required. Some standard errors of the indices in factor analysis are available. Lawley (1967) derived the asymptotic standard errors of the factor loadings of unrotated factors (canonical solution). (See also Jennrich & Thayer, 1973.) Jöreskog (1969) provided the asymptotic standard errors of the parameters in the confirmatory factor analysis model. The asymptotic standard errors of the loadings of orthogonally rotated factors were given by Archer and Jennrich (1973). Subsequently, Archer and Jennrich (1976) showed the accuracy of their results by simulation. The asymptotic standard errors for obliquely rotated loadings were derived by Jennrich (1973a). For communalities, Ichikawa (1992) gave their asymptotic results.

These results are mostly for the factor analysis model in covariance matrices, that is, the scale-dependent model for unstandardized manifest variables. (The model will simply be called "covariance model" hereafter.) For unrotated factor loadings (canonical solution) based on correlation matrices (standardized loadings), Lawley and Maxwell (1971) derived their asymptotic standard errors. (Models with standardized loadings will simply be called "correlation models" hereafter.) Jennrich (1974) obtained the asymptotic standard errors of orthogonally rotated factor-loadings for the correlation model by using the augmented information matrix.

However, these results for rotated loadings are not utilized in typical packaged programs such as SAS (SAS Institute Inc., 1990), BMDP (Dixon, 1992) and SPSS (SPSS Inc., 1988). On the other hand, factor analysis model can be seen as one of the models of covariance structure analysis. In the computer programs for these models (e.g., LISREL, Jöreskog & Sörbom, 1984; EQS, Bentler, 1989; CALIS, SAS Institute Inc., 1990), several methods for factor rotation in exploratory model are available, but the standard errors of their results are not provided.

This paper aims to show the simple derivation of restrictions for rotated parameters, where the restrictions indicate those for just identifying an exploratory factor-analysis model, and to present the asymptotic standard errors of several indices which are used in factor analysis.

Orthogonal Factor Model

The orthogonal exploratory-factor-analysis model for the covariance matrix Σ of p manifest variables (the covariance model) is

$$\Sigma = \Lambda\Lambda' + \Psi, \tag{1}$$

where $\Lambda(p \times k)$ is the loading matrix for k orthogonal common factors and $\Psi(p \times p)$ is the diagonal matrix with the diagonal elements of the variances of unique factors.

The correlation model for orthogonal factors is described as follows (Jennrich, 1974):

$$\Sigma = D(\Lambda\Lambda' + \mathrm{Diag}(I_p - \Lambda\Lambda'))D, \qquad (2)$$

where $D = \{\mathrm{Diag}(\Sigma)\}^{1/2}$; $\mathrm{Diag}(\cdot)$ indicates the diagonal matrix with the diagonal elements of the matrix in parentheses; and I_p is the $p \times p$ identity matrix. Without appropriate restrictions for Λ, the loading matrix Λ has rotational indeterminacy. Let the equation, $\underline{g}(\Lambda) = \underline{0}(m \times 1)$, represent the m restrictions for Λ and $\underline{\alpha}(p \times 1)$ be the non-fixed parameters in Ψ or in D. Assume that the manifest variables are normally distributed with covariance matrix Σ. Let L be the Wishart likelihood for the parameter vector $\underline{\theta}$, i.e.,

$$\ln L = -\frac{n}{2}\{\ln|\Sigma| + \mathrm{tr}(\Sigma^{-1}S)\} + \mathrm{const.}, \qquad (3)$$

where $n=N-1$; N is the number of observations; and S is a sample covariance matrix. Then the (i, j)th element of the information matrix is

$$I[\theta_i, \theta_j] = E(-\partial^2 \ln L / \partial\theta_i \partial\theta_j), \qquad (4)$$

where θ_i is the i-th element of $\underline{\theta}$. Further, the augmented information matrix (see e.g., Silvey, 1975) is described as

$$I_A = \begin{bmatrix} I(\Lambda;\underline{\alpha}), & \dfrac{\partial \underline{g}'}{\partial \mathrm{vec}\Lambda} \\ & O \\ \dfrac{\partial \underline{g}}{\partial(\mathrm{vec}\Lambda)'}, \; O, & O \end{bmatrix} = \begin{bmatrix} I[\Lambda,\Lambda], & I[\Lambda,\underline{\alpha}], & \dfrac{\partial \underline{g}'}{\partial \mathrm{vec}\Lambda} \\ I[\underline{\alpha},\Lambda], & I[\underline{\alpha},\underline{\alpha}], & O \\ \dfrac{\partial \underline{g}}{\partial(\mathrm{vec}\Lambda)'}, & O, & O \end{bmatrix} \qquad (5)$$

where $I(\Lambda;\underline{\alpha})$ stands for the information matrix of the parameters Λ and $\underline{\alpha}$; vec(·) indicates the column vector consisting of the columns of the parenthesized matrix in consecutive order ; $\partial \underline{g}'/\partial \underline{\theta}$ denotes the matrix whose (*i, j*) th element is the partial derivative of the *j*-th element of the vector $\underline{g}$ with respect to the *i*-th element of the vector $\underline{\theta}$; and $\partial \underline{g}/\partial \underline{\theta}' = (\partial \underline{g}'/\partial \underline{\theta})'$.

The vector $\underline{g}$ varies with models. A simple case for just identified unrotated factors is that the appropriate $(k^2 - k)/2$ elements of Λ (e.g., the upper right triangle of Λ) are set to zeros and remaining elements are free parameters. In this case $\underline{g}$ in (3) vanishes and I_A is simply $I(\Lambda;\underline{\alpha})$. Suppose that we have some restriction $\underline{g} = \underline{0}$. Then the inverse of I_A is

$$I_A^{-1} = \begin{bmatrix} \multicolumn{2}{c}{I^*(\Lambda;\underline{\alpha})} & \# \\ & & \# \\ \#, & \#, & \# \end{bmatrix} = \begin{bmatrix} I^*[\Lambda,\Lambda], & I^*[\Lambda,\underline{\alpha}], & \# \\ I^*[\underline{\alpha},\Lambda], & I^*[\underline{\alpha},\underline{\alpha}], & \# \\ \#, & \#, & \# \end{bmatrix}, \quad (6)$$

where $I^*[\cdot,\cdot]$ denotes a submatrix in I_A^{-1} corresponding to the parameters in brackets and the #'s indicate the matrices which will not be used here. Note that $I^*[\cdot,\cdot]$ does not indicate that it is the function of the parameters in brackets alone. The matrix $I^*[\Lambda,\Lambda]$ is the asymptotic variance-covariance matrix of the loadings. Hence its estimate is obtained by replacing Λ by the maximum likelihood estimate $\hat{\Lambda}$. Let $\underline{\tilde{\theta}}$ be the function of Λ and $\underline{\alpha}$. Then the asymptotic covariance matrix for $\underline{\tilde{\theta}}$, $I^*[\underline{\tilde{\theta}},\underline{\tilde{\theta}}]$, is obtained by the chain rule of derivatives as follows (Efron & Tibshirani, 1993):

$$I^*[\underline{\tilde{\theta}},\underline{\tilde{\theta}}] = \frac{\partial\underline{\tilde{\theta}}}{\partial((\mathrm{vec}\Lambda)',\underline{\alpha}')} I^*(\Lambda;\underline{\alpha}) \frac{\partial\underline{\tilde{\theta}}'}{\partial((\mathrm{vec}\Lambda)',\underline{\alpha}')} \quad (7)$$

In the following sections, this basic result will be applied to several indices.

The Standard Errors for Communalities

Ichikawa (1992) has derived the asymptotic standard errors of communalities in the covariance model using the relationship among parameter estimates, $\mathrm{Diag}(\hat{\Lambda}\hat{\Lambda}') = \mathrm{Diag}(\hat{\Sigma} - \hat{\Psi})$, with the variances and covariances of the estimates $\hat{\Sigma}$ and $\hat{\Psi}$. Since in the case of the correlation model, a communality and its corresponding uniqueness sum to one, the standard errors of communalities are equal to those for corresponding uniquenesses, which are available from the results of Lawley and Maxwell (1971).

Though the asymptotic standard errors of the communalities both for the covariance and correlation models are already available as described above, they are for the exploratory factor analysis model. The formula cannot be used when restrictions beyond rotation (e.g., equalities among loadings) are imposed. Thus, the following simple method is presented using the standard errors of factor loadings. When we are interested in standard errors of communalities, those for factor loadings are also often of interest. Let the following Λ represent the Λ in (1) or (2). Let h_i^2 be the communality of the i-th variable. Then

$$h_i^2 = (\Lambda\Lambda')_{ii}, \qquad (8)$$

where $(\cdot)_{ij}$ indicates the (i, j)th element of the matrix in parentheses. Let $\underline{\lambda}_{(i)}$ be the column vector consisting of the non-fixed parameters in the i-th row of Λ, then

$$I^*[h_i^2, h_j^2] = \frac{\partial h_i^2}{\partial \underline{\lambda}'_{(i)}} I^*[\underline{\lambda}_{(i)}, \underline{\lambda}_{(j)}] \frac{\partial h_j^2}{\partial \underline{\lambda}_{(j)}} \quad (9)$$
$$= 4 \underline{\lambda}'_{(i)} I^*[\underline{\lambda}_{(i)}, \underline{\lambda}_{(j)}] \underline{\lambda}_{(j)}$$

is obtained. Substituting $\underline{\lambda}_{(i)}$ and $\underline{\lambda}_{(j)}$ by their estimates, we have the estimates of the asymptotic covariance of $\hat{h}_i^2$ and $\hat{h}_j^2$.

Standard Errors for Contributions of Orthogonally Rotated Factors

The loading matrix of orthogonally rotated factors, $B(p \times k)$, is

$$B = \Lambda T, (TT' = T'T = I_k), \quad (10)$$

where B optimizes a criterion of orthogonal rotation, $h(B)$. This optimization gives rise to $(k^2 - k)/2$ restrictions for rotated factors, which were provided by Archer and Jennrich (1973). On the other hand, the restrictions, which must be satisfied by estimates, have been utilized in methods of estimating rotated loadings. In particular, the several expressions of the restrictions with respect to rotated loadings by the varimax method have been presented (Horst, 1965, Chapters 18 and 19; Sherin, 1966, p.537, (18); Magnus & Neudecker, 1988, p.375, (12)). Furthermore, general condition for obtaining rotated loadings including the varimax rotation was given by Takeuchi and Yanai (1972, p.230, (7.78)), which is closely related to the results of Archer and Jennrich (1973).

In the following corollary, the result is presented with the minor extension to the case where several loading matrices are simultaneously rotated by the same matrix T. The optimizing function for rotation is assumed to be the sum of the values of individual functions for each loading matrix. (For actual examples see e.g., Hakstian, 1976 and Ogasawara, 1997.)

Corollary 1. *Let an orthonormal matrix* $T(k \times k)$ *transform s matrices* $\Lambda_1(p_1 \times k),\ldots,\Lambda_s(p_s \times k)$ *into* $B_1 = \Lambda_1 T,\ldots,B_s = \Lambda_s T$. *Suppose that the criterion of rotation is* $t(B_1,\ldots,B_s) = \sum_{i=1}^{s} t_i(B_i)$, *where* $t_i(B_i)$ *is an individual criterion for* B_i. *Then the restriction which should be satisfied by* $B_1,\ldots,B_s$ *is*

$$\sum_{i=1}^{s}\left(B'_i \frac{\partial t_i(B_i)}{\partial B_i} - \frac{\partial t_i(B_i)}{\partial B'_i} B_i\right) = O. \tag{11}$$

The proof will be given in Appendix 1.

When s=1, by forming (11) in a vector with appropriate order and rewriting B_1 as Λ, we have $\underline{g}$ in (3). For actual forms, Archer and Jennrich (1973) provided the case of the raw orthomax rotation and Ogasawara (1996) gave the result of the orthomax rotation with Kaiser's normalization. Part of Ogasawara's (1996) results will be shown in Appendix 2.

The asymptotic standard errors for the contributions of orthogonally rotated factors $(c_j = \sum_{i=1}^{p} \beta_{ij}^2,\ \beta_{ij} = (B)_{ij}, j = 1,\ldots,k)$ follow in a similar manner to the standard errors for communalities. For simplicity, we deal with the case of s=1. The asymptotic covariance of $\hat{c}_i$ and $\hat{c}_j$ is obtained as

$$
\begin{aligned}
I^*[c_i, c_j] &= \frac{\partial c_i}{\partial \underline{\beta}'_{i:}} I^*[\underline{\beta}_{i:}, \underline{\beta}_{j:}] \frac{\partial c_j}{\partial \underline{\beta}_{j:}} \\
&= 4\underline{\beta}'_{i:} I^*[\underline{\beta}_{i:}, \underline{\beta}_{j:}] \underline{\beta}_{j:},
\end{aligned}
\qquad (12)
$$

where $\underline{\beta}_{i:}$ denotes the i-th column of B. Replacing $\underline{\beta}_{i:}$ by their estimates, we have the estimate of the asymptotic covariance. For the case of contributions of unrotated factors, B is equivalent to Λ and actual forms of $\underline{g}$ vary with models. For instance, in the case of usual canonical solution, the vector $\underline{g}(\Lambda, \Psi)$ consists of the unique off-diagonal elements of $\Lambda' \Psi^{-1} \Lambda$.

Oblique Factor Model

Let Λ be the loading matrix of unrotated orthogonal factors and T be the transformation matrix for oblique rotation. Let B and Φ denote the loading and correlation matrices of obliquely rotated factors, respectively. Then

$$B = \Lambda T'^{-1},\ \Phi = T'T \quad \text{and} \quad \text{Diag}\,\Phi = I_k. \tag{13}$$

The covariance and correlation models for the oblique factors are

$$\Sigma = B\Phi B' + \Psi \quad \text{and} \tag{14a}$$

$$\Sigma = D(B\Phi B' + \text{Diag}(I_p - B\Phi B'))D, \tag{14b}$$

respectively, where Ψ and D are the same as those for orthogonal rotation. Let $\Gamma(p \times k)$ denote the covariances or correlations between common factors and manifest variables, then we have $\Gamma = B\Phi = \Lambda T$.

We have two types of optimization for oblique rotation, that is, those for B and Γ. Historically, the optimization for Γ preceded that of B. The restriction for B with direct optimization for B was provided by Jennrich (1973a), though the derivation with the result of the standard errors of the loadings was complicated. Here, the same restriction is derived by using Lagrange multipliers with the similar minor extension in Corollary 1.

Corollary 2. *Let the matrix for oblique rotation* $T(k \times k)$ *transform* s *matrices* $\Lambda_1(p_1 \times k),\ldots,\Lambda_s(p_s \times k)$ *into* $B_1 = \Lambda_1 T'^{-1},\ldots,B_s = \Lambda_s T'^{-1}$. *Suppose that the criterion of rotation is* $u(B_1,\ldots,B_s) = \sum_{i=1}^{s} u_i(B_i)$, *where* $u_i(B_i)$ *is an* *individual criterion for* B_i. *Then the restriction which should be satisfied by* $B_1, \ldots, B_s$ *is*

$$\sum_{i=1}^{s} B'_i \frac{\partial u_i(B_i)}{\partial B_i} \Phi^{-1} = L_d, \tag{15}$$

where L_d *is a diagonal matrix.*

The proof will be given in Appendix 3.

To date the restriction for Γ which optimizes some criterion has not been provided. Though Takeuchi and Yanai (1972, p.233, Theorem 7.8) showed an equation which is essentially the same as the following Theorem 3, the viewpoint of restriction among estimators was not employed, nor the derivation was given. Accordingly, we present Theorem 3.

Theorem 3. *Let the matrix for oblique rotation* $T(k \times k)$ *transform* s *matrices* $\Lambda_1(p_1 \times k), \ldots, \Lambda_s(p_s \times k)$ *into the structures* $\Gamma_1 = \Lambda_1 T, \ldots, \Gamma_s = \Lambda_s T$. *Suppose that the criterion of rotation is* $v(\Gamma_1, \ldots, \Gamma_s) = \sum_{i=1}^{s} v_i(\Gamma_i)$, *where* $v_i(\Gamma_i)$ *is an individual criterion for* Γ_i. *Then the restriction which should be satisfied by* $\Gamma_1, \ldots, \Gamma_s$ *is*

$$\sum_{i=1}^{s} \frac{\partial v_i(\Gamma_i)}{\partial \Gamma'_i} \Gamma_i \Phi^{-1} = L_d, \tag{16}$$

where L_d *is a diagonal matrix.*

The proof will be given in Appendix 4. Among applications of Corollary 2 and Theorem 3 we consider the case of s=1 for simplicity. Let $\underline{g} = \underline{0}$ be the vector of restrictions for B and $\underline{\alpha}$ be the vector of the parameters in Ψ or D. Then the inverse of the augmented information matrix becomes

$$I_A^{-1} = \begin{bmatrix} I(B;\Phi;\underline{\alpha}), & \dfrac{\partial \underline{g}'}{\partial \mathrm{vec} B} \\ & O \\ & O \\ \dfrac{\partial \underline{g}}{\partial (\mathrm{vec} B)'}, O, O & O \end{bmatrix}^{-1} = \begin{bmatrix} & & & \# \\ I^*(B;\Phi;\underline{\alpha}), & & & \# \\ & & & \# \\ \#, & \#, & \#, & \# \end{bmatrix}, \quad (17)$$

where $I^*(\cdot\,;\cdot\,;\cdot)$, the submatrix of the inverse corresponding to the three matrices in parentheses, is the asymptotic variance-covariance matrix of the estimates of the parameters. In the case when $\underline{\alpha}$ consists of the diagonal elements of Ψ (i.e., the covariance model, (14a)), the information matrix $I(B;\Phi;\Psi)$ is described in various literatures (e.g., Jöreskog, 1969). For the case in which $\underline{\alpha}$ consists of the diagonal elements of D (i.e., the correlation model, (14b)), the information matrix is given in Appendix 5.

When restrictions are imposed on Γ, (14a) is reparameterized in the following way. For the covariance model,

$$\Sigma = B\Phi B' + \Psi = \Gamma\Phi^{-1}\Gamma' + \Psi,\ \ \mathrm{Diag}\Phi = I_k. \qquad (18)$$

For the correlation model,

$$\Sigma = D(\Gamma\Phi^{-1}\Gamma' + \mathrm{Diag}(I_p - \Gamma\Phi^{-1}\Gamma'))D,\ \ \mathrm{Diag}\Phi = I_k. \quad (19)$$

Let $\underline{g}$ and $\underline{\alpha}$ be defined similarly to those in (15). Then the inverse of the augmented information matrix becomes

$$I_A^{-1} = \begin{bmatrix} & & & \frac{\partial \underline{g}'}{\partial \mathrm{vec}\Gamma} \\ & I(\Gamma;\Phi;\underline{\alpha}), & & O \\ & & & O \\ \frac{\partial \underline{g}}{\partial(\mathrm{vec}\Gamma)'}, & O, & O, & O \end{bmatrix}^{-1} = \begin{bmatrix} & & & \# \\ & I^*(\Gamma;\Phi;\underline{\alpha}), & & \# \\ & & & \# \\ \#, & \#, & \#, & \# \end{bmatrix}. \quad (20)$$

The information matrices $I(\Gamma;\Phi;\Psi)$ and $I(\Gamma;\Phi;D)$ will be given in Appendix 6.

For the actual expression of $\underline{g}$ for oblique rotation, Jennrich (1973a) provided the case of the generalized Crawford-Ferguson family which includes the direct oblimin criterion as a special case.

Standard Errors for Loadings, Structures and Correlations of Oblique Factors

In this section we deal with the indices for oblique factors which are obtained as functions of initial parameters. (Factor contributions will be treated in next section.) The method of simplifying Γ has been used as an indirect method for simplifying B, where the structures of reference factors (factors in reference axes) are initially obtained and then transformed into the loadings of corresponding primary factors (see e.g., Harman, 1976). Though this may be seen as a historical method when the direct solution was not available, it still has the theoretical meaning as constructing rescaled anti-image factors.

We begin with the direct method.

1. $\Gamma(= B\Phi)$ When B is Simplified

It is not easy to imagine the configuration of observed variables in the common factor space when B and Φ are simultaneously considered, especially if the number of factors is more than two. For these cases, structures Γ may be of help to grasp the configuration. The asymptotic covariance matrix for Γ is easily obtained as

$$I^*(\Gamma) = \frac{\partial \text{vec}\Gamma}{\partial((\text{vec}B)', (\text{ve}\Phi)')} I^*(B;\Phi) \frac{\partial(\text{vec}\Gamma)'}{\partial((\text{vec}B)', (\text{ve}\Phi)')'} \quad (21)$$

where $\text{ve}(\cdot)$ denotes the vector consisting of the unique off-diagonal elements of the matrix in parentheses, which are arranged in an appropriate order such as

$$(\text{ve}\,\Phi)' = (\phi_{21}, \phi_{31}, \phi_{32}, \ldots, \phi_{k,k-1}).$$

2. $B(=\Gamma\,\Phi^{-1})$ When Γ is Simplified

This is the case for loadings when structures are simplified. Using the similar notation in previous subsection, we have

$$I^*(B) = \frac{\partial \text{vec}B}{\partial((\text{vec}\Gamma)', (\text{ve}\Phi)')} I^*(\Gamma;\Phi) \frac{\partial(\text{vec}B)'}{\partial((\text{vec}\Gamma)', (\text{ve}\Phi)')'}. \quad (22)$$

3. Loadings of Primary Factors When Γ is Simplified

This is the case for simple loadings obtained by the indirect method. From the identity

$$\begin{aligned}\Lambda\Lambda' = \Gamma\,\Phi^{-1}\Gamma' &= \Gamma(\text{Diag}\,\Phi^{-1})^{1/2}(\text{Diag}\,\Phi^{-1})^{-1/2}\,\Phi^{-1} \\ &\quad\times(\text{Diag}\,\Phi^{-1})^{-1/2}(\text{Diag}\,\Phi^{-1})^{1/2}\Gamma' \\ &= E(\text{Diag}\,\Phi^{-1})^{-1/2}\,\Phi^{-1}(\text{Diag}\,\Phi^{-1})^{-1/2}E'\end{aligned}$$

with the definition of $E = \Gamma(\text{Diag}\,\Phi^{-1})^{1/2}$, we have

$$I^*(E) = \frac{\partial \text{vec}E}{\partial((\text{vec}\Gamma)', (\text{ve}\Phi)')} I^*(\Gamma;\Phi) \frac{\partial(\text{vec}E)'}{\partial((\text{vec}\Gamma)', (\text{ve}\Phi)')'}. \quad (23)$$

4. Correlations among Primary Factors When Γ is Simplified

Using the notation in previous subsection, we have the correlation matrix of the primary factors $P = (\text{Diag}\,\Phi^{-1})^{-1/2}\,\Phi^{-1}(\text{Diag}\,\Phi^{-1})^{-1/2}$. When k=2, $\rho_{21} = -\phi_{21}$. Since P is the function of Φ, we obtain

$$I^*(P) = \frac{\partial \text{ve}P}{\partial(\text{ve}\,\Phi)'} I^*(\Phi) \frac{\partial(\text{ve}P)'}{\partial \text{ve}\,\Phi}. \tag{24}$$

The partial derivatives for the above four types of parameters are shown in Appendix 7.

Standard Errors for Contributions of Oblique Factors

Contributions of oblique factors are defined in several ways. Here we consider the following four types of contributions (see e.g., Harman, 1976; Yanai et al., 1990). Though the results are restricted to the case when B is simplified, they apply both to the covariance and correlation models.

1. Direct Contributions

$$c_{di} = \underline{\beta}'_{i:}\underline{\beta}_{i:},\ (i = 1,\ldots,k). \tag{25}$$

2. Contributions Neglecting Other Factors

$$c_{ni} = \underline{\gamma}'_{i:}\underline{\gamma}_{i:} = \sum_{j=1}^{p}(\sum_{l=1}^{k}\beta_{jl}\phi_{li})^2,\ (i = 1,\ldots,k), \tag{26}$$

where $\underline{\gamma}_{i:}$ is the i-th column of Γ.

3. Contributions Eliminating the Effects of Other Factors

$$\begin{aligned} c_{ei} &= \{\underline{\beta}'_{i:}(\phi^{ii})^{-1/2}\}\{(\phi^{ii})^{-1/2}\underline{\beta}_{i:}\} \\ &= \underline{\beta}'_{i:}\underline{\beta}_{i:}/\phi^{ii},\ (i = 1,\ldots,k), \end{aligned} \tag{27}$$

where $\phi^{ii} = (\Phi^{-1})_{ii}$. This is the sum of the squares of the covariances between the oblique factors and normalized anti-image factors.

4. Joint Contribution

$$c_{ij} = 2\phi_{ij}\underline{\beta}'_{i:}\underline{\beta}_{j:},\quad (i, j = 1, \ldots, k; i > j). \tag{28}$$

From the definition of direct contribution, we have

$$\sum_{i=1}^{k} c_{di} + \sum_{i>j}^{k} c_{ij} = \sum_{i=1}^{p} h_i^2. \tag{29}$$

The asymptotic variances of these contributions are obtained by the similar method to that in previous section. That is

$$I^*(c^*) = \frac{\partial c^*}{\partial((\mathrm{vec}B)', (\mathrm{ve}\Phi)')} I^*(B;\Phi) \frac{\partial c^*}{\partial((\mathrm{vec}B)', (\mathrm{ve}\Phi)')'}, \tag{30}$$

where c^* indicates one of the four types of contributions. The partial derivatives in (30) are in Appendix 8.

Numerical Examples

Data and Methods

The proposed methods are applied to two sets of data. One is the correlation matrix of Harman's (1976) Twelve Psychological Tests (N=355). The other is the correlation matrix of Lawley and Maxwell's (1971) Six School Subjects (N=220). The methods of rotation are the normal varimax, raw direct oblimin (quartimin) and raw indirect oblimin (quartimin) methods. The factor patterns including the usual canonical solution are shown in Table 1. We observe fairly simple patterns in the rotated loadings, though the loadings of Factor Ⅲ in the Twelve Psychological Tests are not so clear-cut as those of Factors Ⅰ and Ⅱ.

To evaluate the accuracies of the estimators for the standard errors proposed in this paper, the Monte Carlo simulation was carried out, where a true factor structure (three and two common factors for the first and second data sets, respectively) was assumed. Based on the correlation matrix which was reconstructed from the results of exploratory factor analysis applied to the two data sets, the same numbers of random observations (N=355 or N=220) were generated with the assumption of multivariate normal distribution, where the reconstructed matrix was regarded as a population covariance matrix. From the generated data, the parameters in factor analysis model were estimated. This procedure was repeated 1,000 times and 1,000 estimates were obtained for each parameter. The standard error obtained by the simulation was defined as the standard deviation of the 1,000 estimates.

The simulation was carried out both for the covariance and correlation models. For the unrotated Λ, the just identified confirmatory model was used, where $(k^2 - k)/2$ elements in Λ were set to zeros. Table 2 shows the numbers of irregular cases. The Heywood cases (denoted by H in Table 2) had occurred until 1,000 regular cases were obtained. As for the results of factor rotation, the case without convergence beyond 50 iterations (denoted by ND in Table 2) in the direct quartimin method was deleted.

After regular samples are obtained, there remains indeterminacy of the signs and permutations of common factors. The initial rotated results were transformed to the patterns which are as close as possible to those in Table 1 by Clarkson's (1979) method. The number of cases which required any transformation in 1,000 cases is denoted by P in Table 2.

Table 1
Unrotated and Rotated Factor Loadings

Variable	No.	Unrotated (Canonical) I	II	III	Normal Varimax I	II	III	Direct Quartimin I	II	III	Indirect Quartimin I	II	III
	1	.68	.38	-.07	.27	.73	.14	.11	.74	-.04	.10	.75	-.06
	2	.72	.45	-.00	.22	.80	.20	.01	.83	.04	-.02	.84	.03
	3	.64	.45	-.05	.19	.76	.13	-.00	.80	-.03	-.01	.82	-.04
Twelve	4	.54	.43	.04	.10	.66	.17	-.11	.72	.07	-.14	.73	.07
Psycho.	5	.73	-.26	-.22	.75	.27	.15	.84	.07	-.13	.89	.06	-.18
Tests [a]	6	.64	-.31	-.01	.62	.15	.31	.65	-.05	.13	.66	-.08	.13
	7	.75	-.38	-.06	.76	.17	.32	.82	-.06	.09	.84	-.09	.07
	8	.78	-.30	-.10	.75	.26	.29	.79	.04	.04	.82	.02	.02
	9	.58	-.12	.24	.36	.24	.47	.24	.12	.40	.18	.08	.45
	10	.54	-.04	.24	.29	.28	.44	.15	.19	.38	.08	.16	.43
	11	.51	-.14	.46	.23	.15	.65	.04	.04	.66	-.08	-.01	.76
	12	.34	-.12	.38	.14	.08	.50	-.01	-.01	.54	-.11	-.05	.62
	1	.55	.43		.23	.66		.04	.68		.01	.69	
	2	.57	.29		.32	.55		.17	.53		.16	.54	
Six	3	.39	.45		.09	.59		-.11	.65		-.13	.66	
School	4	.74	-.27		.77	.17		.81	-.03		.82	-.05	
Subjects [b]	5	.72	-.21		.72	.22		.74	.03		.75	.01	
	6	.60	-.13		.57	.21		.57	.07		.57	.06	

Note: a. Harman (1976, p.401) based on Holtzinger's data, *N*=355.
b. Lawley and Maxwell (1971, p.66), *N*=220.

Table 2
Numbers of Irregular Cases

Rotation	Normal Varimax		Direct Quartimin			Indirect Quartimin	
Data	H	P	H	P	ND	H	P
Twelve Psychological Tests	1	1	1	1	1	1	384 (352)
Six School Subjects	12	4	12	7	0	12	90 (78)

Note: H=the numbers of Heywood cases until 1,000 regular cases were obtained,
P=the numbers of cases which required permutation of factors,
ND=the numbers of cases of no convergence in direct oblimin rotation until 1,000 regular cases were obtained.
The numbers in parentheses indicate the frequencies for the correlation model, when they differ from those for the covariance model.

Table 3
Unrotated and Orthogonally Rotated Results
for Twelve Psychological Tests

Variable	No.	Unrotated(Cov) (canonical)	(Standard Errors) (Cov) (N)	(Standard Errors) (Cor) (N)	Rotated (varimax)	(Standard Errors) (Cov. Model) (N S)	(Standard Errors) (Cor. Model) (N S)
Loadings	1	.685	(.054)	(.036)	.270	(.042 .042)	(.039 .040)
for	5	.733	(.051)	(.031)	.751	(.047 .047)	(.031 .032)
Factor Ⅰ	9	.582	(.051)	(.039)	.364	(.058 .059)	(.054 .055)
Loadings	1	.381	(.046)	(.046)	.725	(.046 .046)	(.030 .030)
for	5	-.263	(.050)	(.050)	.267	(.041 .042)	(.038 .039)
Factor Ⅱ	9	-.120	(.057)	(.057)	.237	(.049 .049)	(.047 .047)
Loadings	1	-.069	(.044)	(.044)	.140	(.047 .047)	(.046 .046)
for	5	-.224	(.044)	(.044)	.145	(.048 .046)	(.047 .046)
Factor Ⅲ	9	.239	(.062)	(.061)	.472	(.067 .070)	(.062 .066)
	1	.381				(.038 .036)	The same as
Ψ	5	.344				(.039 .040)	those for
	9	.589				(.052 .051)	communalities.
Communa-	1	.619				(.074 .073)	(.040 .039)
lities	5	.659				(.077 .076)	(.042 .043)
	9	.411				(.067 .066)	(.049 .048)
Contribu-	Ⅰ	4.813	(.394)	(.235)	2.540	(.230 .234)	(.148 .154)
tions	Ⅱ	1.189	(.120)	(.114)	2.520	(.211 .208)	(.106 .105)
	Ⅲ	.548	(.086)	(.082)	1.491	(.189 .191)	(.160 .159)

Note: Cov=covariance model, Cor=correlation model,
N=normal theory, S=simulation with true values.

Table 4

Unrotated and Orthogonally Rotated Results for Six School Subjects

Variable	No.	Unrotated (canonical)	(Standard Errors) (Cov) (N)	(Standard Errors) (Cor) (N)	Rotated (varimax)	(Standard Errors) (Cov. Model) N	(Cov. Model) S	(Cor. Model) N	(Cor. Model) S
Factor I	1	.553	(.086)	(.073)	.232	(.063	.066)	(.061	.063)
	4	.740	(.077)	(.058)	.770	(.076	.079)	(.059	.063)
Factor II	1	.429	(.085)	(.084)	.660	(.091	.099)	(.079	.087)
	4	-.273	(.066)	(.067)	.173	(.060	.060)	(.058	.059)
Ψ	1	.510				(.097	.106)	The same as those	
	4	.377				(.083	.085)	for communalities	
Communalities	1	.490				(.118	.130)	(.097	.107)
	4	.623				(.116	.120)	(.085	.090)
Contributions	I	2.209	(.262)	(.174)	1.606	(.196	.194)	(.119	.124)
	II	.606	(.115)	(.108)	1.209	(.170	.165)	(.124	.121)

Note: Cov=covariance model, Cor=correlation model.
N=normal theory, S=simulation with true values.

Table 5
Rotated Results by Direct Quartimin Method for Twelve Psychological Tests

Variable	No.	Estimates	(Standard Errors) (Cov. Model) N	(Cov. Model) S	(Cor. Model) N	(Cor. Model) S
Loadings	1	.107	(.053	.052)	(.052	.052)
for	5	.837	(.056	.057)	(.041	.042)
Factor I	9	.244	(.082	.083)	(.081	.082)
Structures	1	.479	(.052	.051)	(.043	.044)
for	5	.802	(.047	.047)	(.026	.026)
Factor I	9	.537	(.052	.052)	(.042	.042)
ϕ_{21}		.531	(.046	.045)	(.045	.045)
ϕ_{31}		.583	(.067	.065)	(.066	.063)
ϕ_{32}		.404	(.068	.069)	(.067	.067)
Direct	I	2.520	(.238	.245)	(.128	.140)
Contri.	II	2.452	(.217	.217)	(.103	.105)
(c_{di})	III	1.084	(.177	.178)	(.149	.149)
SS of	I	4.046	(.379	.380)	(.236	.237)
Structures	II	3.656	(.361	.362)	(.238	.240)
(c_{ni})	III	2.786	(.438	.433)	(.376	.374)
Anti-	I	1.401	(.187	.192)	(.168	.167)
Image	II	1.727	(.168	.170)	(.127	.129)
(c_{ei})	III	.702	(.110	.109)	(.100	.098)
Joint	(II, I)	.090	(.042	.047)	(.041	.044)
Contri.	(III, I)	.283	(.076	.069)	(.073	.065)
(c_{ij})	(III, II)	.122	(.038	.040)	(.037	.038)

Note: N=normal theory, S=simulation with true values.

Table 6
Rotated Results by Direct Quartimin Method
for Six School Subjects

Variable	No.	Estimates	Standard Errors (Cov. Model) N	Standard Errors (Cov. Model) S	Standard Errors (Cor. Model) N	Standard Errors (Cor. Model) S
Loadings	1	.036	(.088	.084)	(.087	.084)
Factor I	4	.806	(.083	.087)	(.066	.071)
Structures	1	.393	(.072	.071)	(.066	.065)
Factor I	4	.789	(.073	.076)	(.053	.057)
ϕ_{21}		.525	(.088	.083)	(.086	.081)
Direct	I	1.563	(.205	.205)	(.127	.141)
Contri.	II	1.167	(.186	.175)	(.141	.135)
SS of	I	1.969	(.253	.242)	(.167	,164)
Structures	II	1.682	(.262	.247)	(.204	.194)
Anti-	I	1.133	(.201	.209)	(.172	.177)
Image	II	.846	(.155	.150)	(.134	.128)
Joint (II, I)		.085	(.065	.058)	(.063	.054)

Note: N=normal theory, S=simulation with true values.

Table 7
Rotated Results by Indirect Quartimin Method
for Twelve Psychological Tests

Variable No.		Estimates	(Standard Errors) (Cov. Model) N	S	(Cor. Model) N	S
Structures	1	.068	(.038	.039)	(.038	.039)
Reference	5	.587	(.050	.050)	(.045	.045)
Axis I	9	.119	(.067	.072)	(.067	.072)
Loadings	1	.103	(.059	.061)	(.059	.061)
Primary	5	.895	(.070	.073)	(.056	.061)
Factor I	9	.181	(.097	.104)	(.097	.103)
Cor.among	ϕ_{21}	-.346	(.090	.101)	(.089	.099)
Reference	ϕ_{31}	-.597	(.078	.081)	(.076	.079)
Axis	ϕ_{32}	-.181	(.098	.105)	(.097	.103)
Cor.among	ρ_{21}	.575	(.046	.047)	(.046	.046)
Primary	ρ_{31}	.715	(.060	.065)	(.058	.063)
Factors	ρ_{32}	.515	(.066	.073)	(.065	.071)

Note: N=normal theory, S=simulation with true values,
ϕ_{ij}=correlation between the *i*-th and *j*-th reference factors,
ρ_{ij}=correlation between the *i*-th and *j*-th primary factors.

Table 8
Rotated Results by Indirect Quartimin Method
for Six School Subjects

Variable No.	Estimates	(Cov. Model) N	(Cov. Model) S	(Cor. Model) N	(Cor. Model) S
Structures 1	.009	(.065	.072)	(.064	.072)
Ref.Axi. I 4	.675	(.081	.086)	(.071	.074)
Loadings 1	.011	(.079	.084)	(.078	.083)
Pri. Fact. I 4	.817	(.085	.091)	(.069	.075)
$\rho_{21}(-\phi_{21})$	.564	(.080	.083)	(.078	.079)

Note: N=normal theory, S=simulation with true values,
ϕ_{21}=correlation between two reference factors,
ρ_{21}=correlation between two primary factors.

Results

In Tables 3 and 4 are shown the results of canonical solution and rotated ones by the normal varimax method. The canonical solutions are included for comparison with the rotated results. The patterns of the just identified confirmatory models which were used as unrotated patterns in simulation were different from the canonical solutions, but they can be rotated to them. The standard errors of communalities for the covariance and correlation models coincided with those by Ichikawa's (1992) and Lawley and Maxwell's (1971) methods, respectively, within reasonable accuracy of numerical computations. The rotated results in Tables 3 and 4 show that the theoretical values obtained by our methods (denoted by N) are close to those by simulation (denoted by S). In Tables 3 and 4, the standard errors in the correlation models seem to be smaller than those in the covariance models.

Tables 5 and 6 give the results for the raw direct quartimin rotation. The values for N in Tables 5 and 6 are close to those for S, which indicates the accuracy of theoretical values. Among the four kinds of contributions in Table 5, the values of the joint contribution are smallest and the ratios of the estimates to their standard errors are also smallest. From the fact that the estimates and [estimates / standard errors] of the three types of contributions (C_{di}, C_{ni} and C_{ei}) for Factor Ⅲ are smaller than those for Factors Ⅱ and Ⅰ, it is concluded that Factor Ⅲ has smaller effect than Factor Ⅱ and Factor Ⅰ. As for the standard errors for structures and those for corresponding loadings in Table 5, we see that some of them are similar, but that some values of structures are about half of those for corresponding loadings.

In Table 6 we see that the joint contributions are smaller than other types of contributions both in estimates and [estimates / standard errors]. The standard errors of structures tend to be smaller than corresponding loadings.

In Tables 5 and 6, the standard errors for the correlation models seem to be smaller than those for the covariance models, which is a similar tendency to the results of orthogonal rotation.

Tables 7 and 8 show the results of the raw indirect quartimin method. The theoretical values (N) are similar to those by simulation

(S). The loadings of primary factors are proportional to the structures of reference factors. The standard errors seem to reflect this relationship. As for the correlations among factors, the absolute values of the correlations among the reference factors in Table 7 are smaller than those for the primary factors. However, the standard errors of the correlations among the reference factors are greater than those for the primary factors..

It is of some interest to compare the results of the direct and indirect quartimin methods with respect to the standard errors of corresponding estimates. In the results of Twelve Psychological Tests (Tables 5 and 7), the standard errors of loadings by the direct method seem to be smaller. But, with respect to the standard errors of the factor correlations, they are similar. In the results of Six School Subjects (Tables 6 and 8) the standard errors of the loadings and factor correlations by the two quartimin methods are similar.

Some Concluding Remarks

The numerical examples indicate various tendencies, though they are based on only two correlation matrices.

1. From the theoretical and simulated results , the accuracy of our method of estimating standard errors is indicated. In addition to the two data sets, the data of Eight Physical Variables (Harman, 1976, p.22, N=305) were analyzed and the results, which are not included in this paper, showed the similar tendency. If a factor analysis model fits data well and the distribution of the manifest variables is

approximately normal, it is likely that the standard errors estimated by our method approximate actual ones.

2. The case in which the standard errors of factor loadings are reduced by rotation is known as Wexler phenomenon (Jennrich, 1973b). In the two numerical examples, neither the results of orthogonal rotation nor those for oblique rotation show the phenomenon. The standard errors for unrotated and rotated results are similar. However, it is observed that the standard errors of the contributions of rotated factors are equalized as well as the values of the contributions. It can be called leveling effect of the standard errors for the contributions of rotated factors.

3. Though the contributions of oblique factors are defined in several ways, they are of use for understanding the relative importance of a factor together with their estimated standard errors.

4. It cannot be said that the statistical behavior (stability) of the indirect quartimin method is inferior to that of the direct one.

5. The standard errors of the parameters in the correlation models tend to be smaller than those in the covariance models with unit population variances for manifest variables.

Appendix 1. The Proof of Corollary 1

Proof 1. Let $L(k \times k)$ be the symmetric matrix consisting of Lagrange multipliers. The orthogonal matrix T is obtained by optimizing the following function:

$$f_t = t(B_1, \ldots, B_s) - \frac{1}{2}\mathrm{tr}(L(T'T - I_k)).$$

Taking the derivative of f_t with respect to T, we have

$$df_t = \mathrm{tr}\sum_{i=1}^{s}\left(\frac{\partial t_i}{\partial B'_i} dB_i\right) - \mathrm{tr}(LT'dT)$$

$$= \mathrm{tr}\{(\sum_{i=1}^{s} \frac{\partial t_i}{\partial B'_i}\Lambda_i - LT')dT\},$$

where $t_i = t_i(B_i)$. From $\partial f_t / \partial T = O$ and the above equation,

$$\sum_{i=1}^{s} \frac{\partial t_i}{\partial B'_i}\Lambda_i - LT' = O$$

follows. Multiplying the above equation on the right by T, we have

$$\sum_{i=1}^{s} \frac{\partial t_i}{\partial B'_i} B_i = L.$$

From this and $L - L' = O$, (11) is obtained.

Proof 2. The partial derivative of f_t with respect to $\mathrm{vec}T$ is directly obtained by using the chain rule in the following way:

$$\frac{\partial f_t}{\partial \mathrm{vec}T} = \sum_{i=1}^{s} \frac{\partial(\mathrm{vec}B_i)'}{\partial \mathrm{vec}T} \frac{\partial t_i}{\partial \mathrm{vec}B_i} - \mathrm{vec}(TL)$$

$$= \sum_{i=1}^{s} \frac{\partial(\mathrm{vec}T)'(I_k \otimes \Lambda_i')}{\partial \mathrm{vec}T} \frac{\partial t_i}{\partial \mathrm{vec}B_i} - \mathrm{vec}(TL)$$

$$= \sum_{i=1}^{s} (I_k \otimes \Lambda'_i) \frac{\partial t_i}{\partial \mathrm{vec}B_i} - \mathrm{vec}(TL)$$

$$= \sum_{i=1}^{s} \mathrm{vec}(\Lambda_i' \frac{\partial t_i}{\partial B_i}) - \mathrm{vec}(TL)$$

$$= \mathrm{vec}(\sum_{i=1}^{s} \Lambda_i' \frac{\partial t_i}{\partial B_i} - TL),$$

where $\otimes$ denotes the right Kronecker product. Since the above equation must be zero,

$$\sum_{i=1}^{s} \Lambda_i' \frac{\partial t_i}{\partial \mathrm{B}_i} = TL.$$

Multiplying on the left by T' and using $L = L'$, (11) follows.

Appendix 2. The Restriction for Orthomax Solution with Kaiser's Normalization (Ogasawara,1996)

For simplicity we deal with the case s=1. The criterion for the orthomax rotation with Kaiser's normalization is described as

$$t(B)=\frac{1}{4}\sum_{j=1}^{k}[\sum_{i=1}^{p}\frac{\beta_{ij}^{4}}{(\sum_{m=1}^{k}\beta_{im}^{2})^{2}}-\frac{w}{p}(\sum_{i=1}^{p}\frac{\beta_{ij}^{2}}{\sum_{m=1}^{k}\beta_{im}^{2}})^{2}]$$

where w is the orthomax weight. From (11) and above equation we have

$$g_{rs}=(B'\frac{\partial t(B)}{\partial B}-\frac{\partial t(B)}{\partial B'}B)_{rs}$$

$$=\sum_{i=1}^{p}\left\{\frac{\beta_{ir}\beta_{is}(\beta_{is}^{2}-\beta_{ir}^{2})}{(\sum_{m=1}^{k}\beta_{im}^{2})^{2}}-\frac{w}{p}\frac{\beta_{is}\beta_{ir}}{\sum_{m=1}^{k}\beta_{im}^{2}}\sum_{j=1}^{p}\frac{\beta_{js}^{2}-\beta_{jr}^{2}}{\sum_{m=1}^{k}\beta_{jm}^{2}}\right\},$$

$$(r,s=1,\ldots,k;r>s).$$

Thus, the partial derivatives of g_{rs} with respect to B are obtained as

$$\frac{\partial g_{rs}}{\partial \beta_{ir}} = \frac{\beta_{is}^{3} - 3\beta_{is}\beta_{ir}^{2}}{(\sum_{m=1}^{k}\beta_{im}^{2})^{2}} - \frac{4\beta_{ir}^{2}\beta_{is}(\beta_{is}^{2} - \beta_{ir}^{2})}{(\sum_{m=1}^{k}\beta_{im}^{2})^{3}}$$

$$-\frac{w}{p}\left\{\left(\frac{\beta_{is}}{\sum_{m=1}^{k}\beta_{im}^{2}} - \frac{2\beta_{ir}^{2}\beta_{is}}{(\sum_{m=1}^{k}\beta_{im}^{2})^{2}}\right)\left(\sum_{j=1}^{p}\frac{\beta_{js}^{2} - \beta_{jr}^{2}}{\sum_{m=1}^{k}\beta_{jm}^{2}}\right)\right.$$

$$\left.-\left(\sum_{i=1}^{p}\frac{\beta_{ir}\beta_{is}}{\sum_{m=1}^{k}\beta_{im}^{2}}\right)\left(\frac{2\beta_{ir}}{\sum_{m=1}^{k}\beta_{im}^{2}} + \frac{2(\beta_{is}^{2} - \beta_{ir}^{2})\beta_{ir}}{(\sum_{m=1}^{k}\beta_{im}^{2})^{2}}\right)\right\}$$

$$\frac{\partial g_{rs}}{\partial \beta_{is}} = -\frac{\beta_{ir}^{3} - 3\beta_{ir}\beta_{is}^{2}}{(\sum_{m=1}^{k}\beta_{im}^{2})^{2}} + \frac{4\beta_{is}^{2}\beta_{ir}(\beta_{ir}^{2} - \beta_{is}^{2})}{(\sum_{m=1}^{k}\beta_{im}^{2})^{3}}$$

$$+\frac{w}{p}\left\{\left(\frac{\beta_{ir}}{\sum_{m=1}^{k}\beta_{im}^{2}} - \frac{2\beta_{is}^{2}\beta_{ir}}{(\sum_{m=1}^{k}\beta_{im}^{2})^{2}}\right)\sum_{j=1}^{p}\frac{\beta_{jr}^{2} - \beta_{js}^{2}}{\sum_{m=1}^{k}\beta_{jm}^{2}}\right.$$

$$\left.-\left(\sum_{i=1}^{p}\frac{\beta_{is}\beta_{ir}}{\sum_{m=1}^{k}\beta_{im}^{2}}\right)\left(\frac{2\beta_{is}}{\sum_{m=1}^{k}\beta_{im}^{2}} + \frac{2(\beta_{ir}^{2} - \beta_{is}^{2})\beta_{is}}{(\sum_{m=1}^{k}\beta_{im}^{2})^{2}}\right)\right\}$$

$$\frac{\partial g_{rs}}{\partial \beta_{it}}_{\begin{pmatrix} t\neq r \\ t\neq s \end{pmatrix}} = -4\beta_{it}\beta_{ir}\beta_{is}\frac{\beta_{is}^{2}-\beta_{ir}^{2}}{(\sum_{m=1}^{k}\beta_{im}^{2})^{3}}$$

$$+\frac{w}{p}\left\{\frac{2\beta_{it}\beta_{ir}\beta_{is}}{(\sum_{m=1}^{k}\beta_{im}^{2})^{2}}\sum_{j=1}^{p}\frac{\beta_{js}^{2}-\beta_{jr}^{2}}{\sum_{m=1}^{k}\beta_{jm}^{2}}\right.$$

$$\left.+\left(\sum_{i=1}^{p}\frac{\beta_{ir}\beta^{\#}_{is}}{\sum_{m=1}^{k}\beta_{im}^{2}}\right)\frac{2(\beta_{is}^{2}-\beta_{ir}^{2})\beta_{it}}{(\sum_{m=1}^{k}\beta_{im}^{2})^{2}}\right\},$$

$$(r,s,t=1,\ldots,k;r>s;i=1,\ldots,p),$$

where $\beta^{\#}_{is}$ in the right-hand side of the last equation is equivalent to β_{is}, which was mistyped as β_{it} in Ogasawara (1996, p.124), while the computation for numerical examples was correct.

Appendix 3. The Proof of Corollary 2

Let L_d be the diagonal matrix consisting of Lagrange multipliers. The matrix T optimizes the following function:

$$f_u = u(B_1,\ldots,B_s) + \frac{1}{2}\,\mathrm{tr}(\mathrm{Diag}(T'T - I_k)\,L_d)$$

$$= \sum_{i=1}^{s} u_i(B_i) + \frac{1}{2}\mathrm{tr}((T'T - I_k)\,L_d).$$

Taking the derivative of f_u with respect to T, we have

$$d f_u = \mathrm{tr}\sum_{i=1}^{s}\left(\frac{\partial u_i}{\partial B'_i} d B_i\right) + \mathrm{tr}(L_d T' dT)$$

$$= -\mathrm{tr}\sum_{i=1}^{s}\left(\frac{\partial u_i}{\partial B'_i}\Lambda_i T'^{-1}(dT')T'^{-1}\right) + \mathrm{tr}(L_d T' dT)$$

$$= \mathrm{tr}\{(-\sum_{i=1}^{s} T^{-1}\Lambda'_i \frac{\partial u_i}{\partial B_i} T^{-1} + L_d T')dT\}.$$

where $u_i = u_i(B_i)$. Since the partial derivatives of f_u with respect to T are zero,

$$\sum_{i=1}^{s} T^{-1}\Lambda'_i \frac{\partial u_i}{\partial B_i} T^{1} = L_d T'$$

follows. Multiplying the above equation on the right by $T(T'T)^{-1}$, (15) is obtained.

Appendix 4. The Proof of Theorem 3

Proof. Let L_d be the diagonal matrix consisting of Lagrange multipliers. The matrix T optimizes the following function:

$$f_v = v(\Gamma_1,\dots,\Gamma_s) + \frac{1}{2}\,\mathrm{tr}(\mathrm{Diag}(T'T - I_k)\,L_d)$$

$$= \sum_{i=1}^{s} v_i(\Gamma_i) + \frac{1}{2}\mathrm{tr}((T'T - I_k)\,L_d).$$

Taking the derivative of f_v with respect to T, we have

$$d\,f_v = \mathrm{tr}\sum_{i=1}^{s}\left(\frac{\partial v_i}{\partial \Gamma'_i}d\Gamma_i\right) - \mathrm{tr}(L_d T' dT)$$

$$= \mathrm{tr}\left\{\left(\sum_{i=1}^{s}\frac{\partial v_i}{\partial \Gamma'_i}\Lambda_i - L_d T'\right)dT\right\}.$$

where $v_i = v_i(\Gamma_i)$. Since the partial derivatives of f_v with respect to T are zero,

$$\sum_{i=1}^{s}\frac{\partial v_i}{\partial \Gamma'_i}\Lambda_i = L_d T'$$

follows. Multiplying the above equation on the right by $T(T'T)^{-1}$, (16) is obtained.

Appendix 5. The Information Matrix for the Correlation Model When B is Simplified

The possible values of subscripts are omitted, which are obvious from the context.

Model: $\Sigma = D(B\Phi B' + \mathrm{Diag}(I_p - B\Phi B'))D = DRD,$

$\mathrm{Diag}\Phi = I_k,$ and D is diagonal.

Let $r^{ij} = (R^{-1})_{ij}$ and $\sigma_i = (D)_{ii}$.

$$
\begin{aligned}
I(\beta_{ij}, \beta_{mn}) &= \frac{n}{2}\mathrm{tr}\{R^{-1}(I_{ij}\Phi B' + B\Phi I_{ji} - 2I_{ii}(B\Phi)_{ij}) \\
&\quad \times R^{-1}(I_{mn}\Phi B' + B\Phi I_{nm} - 2I_{mm}(B\Phi)_{mn})\} \\
&= n\{r^{im}(\Phi B' R^{-1} B\Phi)_{jn} + (R^{-1}B\Phi)_{mj}(R^{-1}B\Phi)_{in} \\
&\quad - 2(B\Phi)_{ij} r^{im}(R^{-1}B\Phi)_{in} - 2(B\Phi)_{mn} r^{im}(R^{-1}B\Phi)_{mj} \\
&\quad + 2(B\Phi)_{ij}(B\Phi)_{mn}(r^{im})^2\},
\end{aligned}
$$

$$
\begin{aligned}
I(\phi_{ij}, \beta_{mn}) &= \frac{n}{2}\mathrm{tr}\{R^{-1}(B(I_{ij} + I_{ji})B' - \mathrm{Diag}(B(I_{ij} + I_{ji})B')) \\
&\quad \times R^{-1}(I_{mn}\Phi B' + B\Phi I_{nm} - \mathrm{Diag}(I_{mn}\Phi B' + B\Phi I_{nm}))\} \\
&= n\mathrm{tr}\{R^{-1}(B(I_{ij} + I_{ji})B' - 2\mathrm{Diag}(BI_{ij}B')) \\
&\quad \times R^{-1}(B\Phi I_{nm} - I_{mm}(B\Phi)_{mn})\},
\end{aligned}
$$

$$I(\phi_{ij},\phi_{mn})=\frac{n}{2}\mathrm{tr}\{R^{-1}(B(I_{ij}+I_{ji})B'-\mathrm{Diag}(B(I_{ij}+I_{ji})B'))$$
$$\times R^{-1}(B(I_{mn}+I_{nm})B'-\mathrm{Diag}(B(I_{mn}+I_{nm})B'))\}$$
$$=n\,\mathrm{tr}\{R^{-1}(B(I_{ij}+I_{ji})B'-2\mathrm{Diag}(BI_{ij}B'))$$
$$\times R^{-1}(BI_{mn}B'-\mathrm{Diag}(BI_{mn}B'))\},$$

$$I(\sigma_i,\beta_{mn})=\frac{n}{2}\mathrm{tr}\{R^{-1}D^{-1}(I_{ii}RD+DRI_{ii})D^{-1}R^{-1}$$
$$\times(I_{mn}\Phi B'+B\Phi I_{nm}-2I_{mm}(B\Phi)_{mn})\}$$
$$=\frac{n}{\sigma_i}\{r^{im}(B\Phi)_{in}+\delta_{im}(R^{-1}B\Phi)_{mn}-2\delta_{im}(B\Phi)_{mn}r^{mm}\},$$

$$I(\sigma_i,\phi_{mn})=\frac{n}{2}\mathrm{tr}\{R^{-1}D^{-1}(I_{ii}RD+DRI_{ii})D^{-1}R^{-1})$$
$$\times(B(I_{mn}+I_{nm})B'-\mathrm{Diag}(B(I_{mn}+I_{nm})B'))\}$$
$$=\frac{n}{\sigma_i}\{(R^{-1}B)_{im}\beta_{in}+(R^{-1}B)_{in}\beta_{im}-2r^{ii}\beta_{im}\beta_{in}\},$$

$$I(\sigma_i,\sigma_j)=\frac{n}{\sigma_i\sigma_j}(r^{ij}r_{ij}+\delta_{ij}),$$

where I_{ij} is the matrix of an appropriate size in which the $(i, j,)$th element is one otherwise zero; and δ_{im} is the Kronecker delta ($\delta_{im}=1,\ i=m$; $\delta_{im}=0,\ i\neq m$).

Appendix 6. The Information Matrices When Γ is Simplified

The possible values of subscripts are omitted, which are obvious from the context.

1. The Covariance Model

Model: $\Sigma = \Gamma\Phi^{-1}\Gamma' + \Psi$, $\mathrm{Diag}\Phi = I_k$.

Let $\psi_i = (\Psi)_{ii}$.

$$I(\gamma_{ij}, \gamma_{mn}) = \frac{n}{2}\mathrm{tr}\{\Sigma^{-1}(I_{ij}\Phi^{-1}\Gamma' + \Gamma\Phi^{-1}I_{ji})\Sigma^{-1}(I_{mn}\Phi^{-1}\Gamma' + \Gamma\Phi^{-1}I_{nm})\}$$

$$= n\{(\Sigma^{-1}\Gamma\Phi^{-1})_{in}(\Sigma^{-1}\Gamma\Phi^{-1})_{mj} + \sigma^{mi}(\Phi^{-1}\Gamma'\Sigma^{-1}\Gamma\Phi^{-1})_{nj}\},$$

$$I(\phi_{ij}, \gamma_{mn}) = \frac{n}{2}\mathrm{tr}\{\Sigma^{-1}\Gamma(-\Phi^{-1}(I_{ij} + I_{ji})\Phi^{-1})\Gamma'$$

$$\times\Sigma^{-1}(I_{mn}\Phi^{-1}\Gamma' + \Gamma\Phi^{-1}I_{nm})\}$$

$$= -n\{(\Sigma^{-1}\Gamma\Phi^{-1})_{mi}(\Phi^{-1}\Gamma'\Sigma^{-1}\Gamma\Phi^{-1})_{nj}$$

$$+ (\Sigma^{-1}\Gamma\Phi^{-1})_{mj}(\Phi^{-1}\Gamma'\Sigma^{-1}\Gamma\Phi^{-1})_{ni}\},$$

$$I(\phi_{ij}, \phi_{mn}) = \frac{n}{2}\mathrm{tr}\{\Sigma^{-1}\Gamma(-\Phi^{-1}(I_{ij} + I_{ji})\Phi^{-1})\Gamma'$$

$$\times\Sigma^{-1}\Gamma(-\Phi^{-1}(I_{mn} + I_{nm})\Phi^{-1})\Gamma'\}$$

$$= n\{(\Phi^{-1}\Gamma'\Sigma^{-1}\Gamma\Phi^{-1})_{im}(\Phi^{-1}\Gamma'\Sigma^{-1}\Gamma\Phi^{-1})_{jn}$$

$$+ (\Phi^{-1}\Gamma'\Sigma^{-1}\Gamma\Phi^{-1})_{in}(\Sigma^{-1}\Gamma'\Phi^{-1}\Gamma\Phi^{-1})_{jm}\},$$

$$I(\psi_i, \gamma_{mn}) = \frac{n}{2}\mathrm{tr}\{\Sigma^{-1} I_{ii} \Sigma^{-1}(I_{mn}\Phi^{-1}\Gamma' + \Gamma\Phi^{-1} I_{nm})\}$$
$$= n(\Sigma^{-1}\Gamma\Phi^{-1})_{in}\sigma^{im},$$

$$I(\psi_i, \phi_{mn}) = \frac{n}{2}\mathrm{tr}\{\Sigma^{-1} I_{ii} \Sigma^{-1}\Gamma(-\Phi^{-1}(I_{mn} + I_{nm})\Phi^{-1})\Gamma'\}$$
$$= -n(\Sigma^{-1}\Gamma\Phi^{-1})_{im}(\Sigma^{-1}\Gamma\Phi^{-1})_{in},$$

$$I(\psi_i, \psi_j) = \frac{n}{2}(\sigma^{ij})^2.$$

2. The Correlation Model

Model:
$$\Sigma = D(\Gamma\Phi^{-1}\Gamma' + \mathrm{Diag}(I_p - \Gamma\Phi^{-1}\Gamma'))D$$
$$= DRD,$$

$\mathrm{Diag}\Phi = I_k$, and D is diagonal.

$$I(\gamma_{ij}, \gamma_{mn}) = \frac{n}{2}\mathrm{tr}\{R^{-1}(I_{ij}\Phi^{-1}\Gamma' + \Gamma\Phi^{-1} I_{ji} - 2 I_{ii}(\Gamma\Phi^{-1})_{ij})$$
$$\times R^{-1}(I_{mn}\Phi^{-1}\Gamma' + \Gamma\Phi^{-1} I_{nm} - 2 I_{mm}(\Gamma\Phi^{-1})_{mn})\}$$
$$= n\{r^{im}(\Phi^{-1}\Gamma' R^{-1}\Gamma\Phi^{-1})_{jn} + (R^{-1}\Gamma\Phi^{-1})_{mj}(R^{-1}\Gamma\Phi^{-1})_{in}$$
$$-2(\Gamma\Phi^{-1})_{ij} r^{im}(R^{-1}\Gamma\Phi^{-1})_{in} - 2(\Gamma\Phi^{-1})_{mn} r^{im}(R^{-1}\Gamma\Phi^{-1})_{mj}$$
$$+2(\Gamma\Phi^{-1})_{ij}(\Gamma\Phi^{-1})_{mn}(r^{im})^2\},$$

$$I(\phi_{ij},\gamma_{mn})=\frac{n}{2}\mathrm{tr}\{R^{-1}(-\Gamma\Phi^{-1}(I_{ij}+I_{ji})\Phi^{-1}\Gamma'$$

$$+\mathrm{Diag}(\Gamma\Phi^{-1}(I_{ij}+I_{ji})\Phi^{-1}\Gamma'))$$

$$\times R^{-1}(I_{mn}\Phi^{-1}\Gamma'+\Gamma\Phi^{-1}I_{nm}$$

$$-\mathrm{Diag}(I_{mn}\Phi^{-1}\Gamma'+\Gamma\Phi^{-1}I_{nm}))\}$$

$$=n\,\mathrm{tr}\{R^{-1}(-\Gamma\Phi^{-1}I_{ij}\Phi^{-1}\Gamma'+\mathrm{Diag}(\Gamma\Phi^{-1}I_{ij}\Phi^{-1}\Gamma'))$$

$$\times R^{-1}(I_{mn}\Phi^{-1}\Gamma'+\Gamma\Phi^{-1}I_{nm}-2I_{mm}(\Gamma\Phi^{-1})_{mn})\},$$

$$I(\phi_{ij},\phi_{mn})=\frac{n}{2}\mathrm{tr}\{R^{-1}(-\Gamma\Phi^{-1}(I_{ij}+I_{ji})\Phi^{-1}\Gamma'$$

$$+\mathrm{Diag}(\Gamma\Phi^{-1}(I_{ij}+I_{ji})\Phi^{-1}\Gamma'))$$

$$\times R^{-1}(-\Gamma\Phi^{-1}(I_{mn}+I_{nm})\Phi^{-1}\Gamma'$$

$$+\mathrm{Diag}(\Gamma\Phi^{-1}(I_{mn}+I_{nm})\Phi^{-1}\Gamma'))\}$$

$$=n\,\mathrm{tr}\{R^{-1}(-\Gamma\Phi^{-1}(I_{ij}+I_{ji})\Phi^{-1}\Gamma'+2\mathrm{Diag}(\Gamma\Phi^{-1}I_{ij}\Phi^{-1}\Gamma'))$$

$$\times R^{-1}(-\Gamma\Phi^{-1}I_{mn}\Phi^{-1}\Gamma'+\mathrm{Diag}(\Gamma\Phi^{-1}I_{mn}\Phi^{-1}\Gamma'))\},$$

$$I(\sigma_i,\gamma_{mn})=\frac{n}{2}\mathrm{tr}\{R^{-1}D^{-1}(I_{ii}RD+DRI_{ii})D^{-1}$$

$$\times R^{-1}(I_{mn}\Phi^{-1}\Gamma'+\Gamma\Phi^{-1}I_{nm}-2I_{mm}(\Gamma\Phi^{-1})_{mn})\}$$

$$=\frac{n}{\sigma_i}\{r^{im}(\Gamma\Phi^{-1})_{in}+\delta_{im}(R^{-1}\Gamma\Phi^{-1})_{mn}-2\delta_{im}(\Gamma\Phi^{-1})_{mn}r^{mm}\},$$

$$I(\sigma_i,\phi_{mn})=\frac{n}{2}\mathrm{tr}\{R^{-1}D^{-1}(I_{ii}RD+DRI_{ii})D^{-1}$$

$$\times R^{-1}(\Gamma(-\Phi^{-1}(I_{mn}+I_{nm})\Phi^{-1}\Gamma'$$

$$-\mathrm{Diag}(\Gamma(-\Phi^{-1}(I_{mn}+I_{nm})\Phi^{-1}\Gamma'))\}$$

$$=\frac{n}{\sigma_i}\{-(R^{-1}\Gamma\Phi^{-1})_{im}(\Gamma\Phi^{-1})_{in}-(R^{-1}\Gamma\Phi^{-1})_{in}(\Gamma\Phi^{-1})_{im}$$

$$+2r^{ii}(\Gamma\Phi^{-1})_{im}(\Gamma\Phi^{-1})_{in}\},$$

$$I(\sigma_i,\sigma_j)=\frac{n}{\sigma_i\sigma_i}(r^{ij}r_{ij}+\delta_{ij}).$$

Appendix 7. The Partial Derivatives for Loadings, Structures and Correlations for Oblique Factors

1. $\Gamma(=B\Phi)$ When B is Simplified

$$\frac{\partial\gamma_{ij}}{\partial\beta_{mn}}=\delta_{im}\phi_{nj},\qquad (i,m=1,\ldots,p;j,n=1,\ldots,k),$$

$$\frac{\partial\gamma_{ij}}{\partial\phi_{mn}}=(B(I_{mn}+I_{nm}))_{ij}=\delta_{jn}\beta_{im}+\delta_{jm}\beta_{in},$$

$$(i=1,\ldots,p;\ j,m,n=1,\ldots,k;\ m>n),$$

2. $B(=\Gamma\Phi^{-1})$ When Γ is Simplified

$$\frac{\partial\beta_{ij}}{\partial\gamma_{mn}}=\delta_{im}\phi^{nj},\quad (i,m=1,\ldots,p;j,n=1,\ldots,k),$$

$$\frac{\partial\beta_{ij}}{\partial\phi_{mn}}=-(\Gamma\Phi^{-1}(I_{mn}+I_{nm})\Phi^{-1})_{ij}$$

$$=-\beta_{im}\phi^{nj}-\beta_{in}\phi^{mj},$$

$$(i=1,\ldots,p;\ j,m,n=1,\ldots,k;\ m>n),$$

where $\phi^{nj}=(\Phi^{-1})_{nj}$.

3. Loadings of Primary Factors When Γ is Simplified

$$\frac{\partial e_{ij}}{\partial \gamma_{mn}} = \delta_{im}\delta_{jn}(\phi^{jj})^{1/2},$$

$$(i,m=1,\ldots,p; j,n=1,\ldots,k),$$

$$\frac{\partial e_{ij}}{\partial \phi_{mn}} = \{\Gamma\frac{1}{2}(\mathrm{Diag}\Phi^{-1})^{-1/2}\mathrm{Diag}(-\Phi^{-1}(I_{mn}+I_{nm})\Phi^{-1})\}_{ij}$$

$$= -\gamma_{ij}(\phi^{jj})^{-1/2}\ \phi^{jm}\phi^{jn},$$

$$(i=1,\ldots,p;\ j,m,n=1,\ldots,k;\ m>n).$$

4. Correlations among Primary Factors When Γ is Simplified

$$\frac{\partial \rho_{ij}}{\partial \phi_{mn}} = -\frac{1}{2}\{(\mathrm{Diag}\Phi^{-1})^{-3/2}\mathrm{Diag}(-\Phi^{-1}(I_{mn}+I_{nm})\Phi^{-1}) \times \Phi^{-1}(\mathrm{Diag}\Phi^{-1})^{-1/2}\}_{ij}$$

$$+\{(\mathrm{Diag}\,\Phi^{-1})^{-1/2}(-\Phi^{-1}(I_{mn}+I_{nm})\Phi^{-1})(\mathrm{Diag}\,\Phi^{-1})^{-1/2}\}_{ij}$$

$$-\frac{1}{2}\{(\mathrm{Diag}\Phi^{-1})^{-1/2}\Phi^{-1}\mathrm{Diag}(-\Phi^{-1}(I_{mn}+I_{nm})\Phi^{-1}) \times(\mathrm{Diag}\Phi^{-1})^{-3/2}\}_{ij}$$

$$=(\phi^{ii})^{-1/2}(\frac{\phi^{im}\phi^{in}\phi^{ij}}{\phi^{ii}}-\phi^{im}\phi^{jn}-\phi^{in}\phi^{jm}+\frac{\phi^{jm}\phi^{jn}\phi^{ij}}{\phi^{jj}})(\phi^{jj})^{-1/2},$$

$$(i,j,m,n=1,\ldots,k;\ i>j;\ m>n).$$

Appendix 8. The Partial Derivatives for the Contributions of Oblique Factors

$$\frac{\partial c_{di}}{\partial \beta_{mn}} = 2\delta_{in}\beta_{mn}, (i,n=1,\ldots,k; m=1,\ldots,p),$$

$$\frac{\partial c_{ni}}{\partial \beta_{mq}} = 2\left(\sum_{l=1}^{k}\beta_{ml}\phi_{li}\right)\phi_{qi},$$

$$(i,q=1,\ldots,k; m=1,\ldots,p),$$

$$\frac{\partial c_{ni}}{\partial \phi_{mq}} = 2\sum_{j=1}^{p}\sum_{l=1}^{k}\beta_{jl}\phi_{li}(\delta_{im}\beta_{jq}+\delta_{iq}\beta_{jm}),$$

$$(i,m,q=1,\ldots,k; m>q),$$

$$\frac{\partial c_{ei}}{\partial \beta_{mn}} = 2\delta_{in}\beta_{mn}/\phi^{nn}, \quad (i,n=1,\ldots,k; m=1,\ldots,p),$$

$$\frac{\partial c_{ei}}{\partial \phi_{mn}} = \sum_{j=1}^{p} 2\beta_{ji}^{2}(\phi^{ii})^{-2}\phi^{im}\phi^{in}, \quad (i,m,n=1,\ldots,k; m>n),$$

$$\frac{\partial c_{ij}}{\partial \beta_{mn}} = 2\delta_{in}\beta_{mj}\phi_{ij} + 2\delta_{jn}\beta_{mi}\phi_{ij},$$

$$(i,j,n=1,\ldots,k;\ i>j;\ m=1,\ldots,p),$$

$$\frac{\partial c_{ij}}{\partial \phi_{mn}} = 2\delta_{im}\delta_{jn}\sum_{l=1}^{p}\beta_{li}\beta_{lj},$$

$$(i,j,m,n=1,\ldots,k;\ i>j;\ m>n).$$

References

Archer, C. O., & Jennrich, R. I. (1973). Standard errors for rotated factor loadings. *Psychometrika, 38*, 581-592.

Archer, C. O., & Jennrich, R. I. (1976). A look, by simulation, at the validity of some asymptotic distribution results for rotated loadings. *Psychometrika, 41*, 537-541.

Bentler, P. M. (1989). *EQS Structural equations program manual.* Los Angels: BMDP Statistical Software Inc..

Clarkson, D. B. (1979). Estimating the standard errors of rotated factor loadings by jackknifing. *Psychometrika, 44*, 297-314.

Dixon, W. J. (Ed.) (1992). *BMDP statistical software manual.* Los Angels: University of California Press.

Efron, B., & Tibshirani, R. J. (1993). *An introduction to the bootstrap.* New York: Wiley.

Hakstian, A. R. (1976). Two-matrix orthogonal rotation procedures. *Psychometrika, 41*, 267-272.

Harman, H. H. (1976). *Modern factor analysis* (3rd ed.). Chicago: University of Chicago Press.

Horst, P. (1965). *Factor analysis of data matrices*. New York: Holt, Rinehart & Winston.

Ichikawa, M. (1992). Asymptotic distributions of the estimators of communalities in factor analysis. *Psychometrika, 57*, 399-404.

Jennrich, R. I. (1973a). Standard errors for obliquely rotated factor loadings. *Psychometrika. 38*, 593-604.

Jennrich, R. I. (1973b). On the stability of rotated factor loadings: The Wexler phenomenon. *British Journal of Mathematical and Statistical Psychology, 26*, 167-176.

Jennrich, R. I. (1974). Simplified formulae for standard errors in maximum likelihood factor analysis. *British Journal of Mathematical and Statistical Psychology, 27*, 122-131.

Jennrich, R. I., & Thayer, D. (1973). A note on Lawley's formulas for standard errors in maximum likelihood factor analysis. *Psychometrika, 38*, 571-580.

Jöreskog, K. G. (1969). A general approach to confirmatory maximum likelihood factor analysis. *Psychometrika, 34*, 183-202.

Jöreskog, K. G., & Sörbom, D. (1984). *LISREL VI User's Guide.* Mooresville, IN: Scientific Software.

Lawley, D. N. (1967). Some new results in maximum likelihood factor analysis. *Proceedings of the Royal Society of Edinburgh, A, 67*, 256-264.

Lawley, D. N., & Maxwell, A. E. (1971). *Factor analysis as a statistical method* (2nd ed.). London: Butterworths.

Magnus, J. R., & Neudecker, H. (1988). *Matrix differential calculus with applications in statistics and econometrics*. New York: Wiley.

Ogasawara, H. (1996). Standard errors for rotated factor loadings by normalized orthomax method. *Japanese Journal of Behaviormetrics, 23*(2), 122-129. (in Japanese)

Ogasawara, H. (1997). Factor rotation in bilinear models. *Economic Review, Otaru University of Commerce, 47*(2 and 3), 181-206. (in Japanese)

SAS Institute Inc. (1990). *SAS/STAT user's guide Vol.1* (version 6, 4th ed.). Cary, NC: Author.

Sherin, R. J. (1966). A matrix formulation of Kaiser's varimax criterion. *Psychometrika, 31*, 535-538.

Silvey, S. D. (1975). *Statistical inference*. New York: Chapman and Hall.

SPSS Inc. (1988). *SPSS-X user's guide* (3rd ed.). Chicago: Author.

Takeuchi, K., & Yanai. H. (1972). Foundation of multivariate analysis. Tokyo: Touyoukeizai. (in Japanese)

Yanai, H., Shigemasu, K., Mayekawa, S., & Ichikawa, M. (1990). Factor analysis: Theories and methods. Tokyo: Asakura-shoten. (in Japanese)