

ON THE STANDARD ERRORS OF ROTATED FACTOR LOADINGS WITH WEIGHTS FOR OBSERVED VARIABLES

Haruhiko Ogasawara*

The asymptotic standard errors of the estimates of rotated factor loadings and factor correlations are derived for the cases with weights for observed variables such as those for Kaiser's normalization. The factor analysis models employed in this paper are the exploratory ones which have orthogonal or oblique common factors and unstandardized or standardized observed variables. The asymptotic standard errors are given from an augmented information matrix. As an application, the result for the direct oblique rotation by general quartic criteria with Kaiser's normalization is derived. The results of simulation show that the theoretical standard errors are close to simulated ones.

1. Introduction

In factor analysis we have various rotation methods. Most of them are based on analytic criteria $h(B)$ for a rotated loading matrix $B(p \times q)$, where p is the number of observed variables and q the number of common factors which may or may not be orthogonal. In these cases, unrotated loadings are frequently recaled with weights for observed variables. One of the most popular weights is that of Kaiser's (1958) normalization. In the statistical packaged programs such as SAS (SAS Institute Inc., 1990), SPSS (SPSS Inc., 1990) and BMDP (Dixon, 1992), Kaiser's normalization is employed as the default value for the option of normalization for observed variables when factor rotation is performed. In the case of Kaiser's normalization, unrotated loadings are divided by the square roots of the communalities of the corresponding observed variables. The rotated loadings are denormalized, that is, multiplied by the square roots of the corresponding original communalities.

The derivation of the asymptotic standard errors for rotated loadings with Kaiser's normalization becomes more involved than that without the normalization. Ogasawara (1996, 1999) derived the asymptotic standard errors for the orthomax and direct oblimin solutions with Kaiser's normalization. There are based on the asymptotic distribution of the estimators with restrictions for rotated parameters satisfying some analytical rotation criteria under the assumption of a multivariate normal distribution for observed variables. The general forms of the restrictions for rotated parameters were obtained by Archer and Jennrich (1973) and Jennrich

Key Words and Phrases: standard errors, Kaiser's normalization, factor rotation, the augmented information matrix, quartic criteria.

*Department of Information and Management Science, Otaru University of Commerce, 3-5-21, Midori, Otaru, 047-8501 Japan. E mail: hogasa@res.otaru uc.ac.jp

(1973) for orthogonal and oblique rotations, respectively.

In behavioral sciences, exploratory factor analysis is usually applied to standardized observed variables. That is, sample correlation matrices are used as input data instead of sample variance-covariance matrices. The standard errors of rotated solutions in such cases have been derived by Jennrich (1974) and Ogasawara (1998b; see also 1998a) for the cases of orthogonal and oblique rotations, respectively, by using the restrictions given by Archer and Jennrich (1973) and Jennrich (1973). These methods have been used for deriving the standard errors of rotated parameters with Kaiser's normalization (Ogasawara, 1996, 1999). The standard errors for rotated parameters with Kaiser's normalization have also been obtained by using the method of numerical derivatives (Browne & Du Toit, 1992; Cudeck & O'Dell, 1994).

Kaiser's normalization can be treated as one of the cases with weights for observed variables in factor rotation. Among the weights other than Kaiser's one, we have Cureton and Mulaik's (1975) weights which are functions of the loadings of the first principal factor. In addition, as a simple method of practical use, the method of selecting observed variables which are included for a rotation criterion may be used. In this method, the values of the weights are unities for the variables which are included in the rotation criterion and zeros for other variables. In the case of Kaiser's normalization, the weights are the reciprocals of the square roots of communalities.

The purpose of this paper is to derive the general results for the restrictions for rotated loadings and factor correlations with weights for observed variables. As an application of the results, the exact asymptotic standard errors of the oblique varimax solution (Crawford, 1975; Browne & Du Toit, 1992) with Kaiser's normalization will be derived.

2. Restrictions of Parameters and their Derivatives in Orthogonal Rotation

With the assumption of orthogonal common factors, the covariance structure for unstandardized observed variables is described as

$$\Sigma = BB' + \Psi, \quad (1)$$

where B is a rotated loading matrix; Ψ is the diagonal matrix with the diagonal elements being the variances of unique factors; and Σ is the variance-covariance matrix for observed variables. The corresponding covariance structure for standardized observed variables is

$$\Sigma = DPD = D(\tilde{B}\tilde{B}' + \text{Diag}(I_p - \tilde{B}\tilde{B}'))D, \quad (2)$$

where $\tilde{B}(p \times q)$ is the rotated loading matrix for standardized observed variables; $D = \text{diag}(d_1, \dots, d_p)$ is the diagonal matrix with the diagonal elements d_i , ($i = 1, \dots, p$) being the standard deviations of unstandardized variables; P is the population

correlation matrix of observed variables; $\text{Diag}(\cdot)$ denotes the diagonal matrix whose diagonal elements are those of the matrix in parentheses; I_p is the $p \times p$ identity matrix. As is well known, we have a rotational indeterminacy in the exploratory factor analysis model, which is removed by imposing $(q^2 - q)/2$ restrictions for a loading matrix. The matrices B in (1) and \tilde{B} in (2) are assumed to have these restrictions. As was previously mentioned, the general form of the restrictions was derived by Archer and Jennrich (1973). However, their method is based on implicit differentiation and is not a simple one. So, here, Ogasawara's (1998b, Corollary 1) derivation using usual Lagrange multipliers is shown as the following lemma.

Lemma 1 (Archer & Jennrich, 1973; Ogasawara, 1998b). Under the model of (1) with associated assumptions, let $h_1(B)$ be a rotation criterion to be optimized by B , then the restrictions for B are

$$\left(B \frac{\partial h_1(B)}{\partial B} - \frac{\partial h_1(B)}{\partial B'} B \right)_{rs} = 0, \quad (q \geq r > s \geq 1), \quad (3)$$

where $(\cdot)_{rs}$ stands for the (r, s) th element of the parenthesized matrix.

Proof. Given an unrotated loading matrix $A(p \times q)$, the rotated loading matrix is $B = AT$, where $T(q \times q)$ is the transformation matrix satisfying $T'T = I_q$. Let L_1 be a symmetric matrix consisting of Lagrange multipliers, then B is obtained from the optimization of the following function with respect to L_1 and T :

$$f_1 = h_1(B) - (1/2)\text{tr}\{L_1(T'T - I_q)\}. \quad (4)$$

Taking the differential of f_1 with respect to T , we have

$$\begin{aligned} df_1 &= \text{tr}\left(\frac{\partial h_1(B)}{\partial B'} dB\right) - \text{tr}(L_1 T' dT) \\ &= \text{tr}\left\{\left(\frac{\partial h_1(B)}{\partial B'} A - L_1 T'\right) dT\right\}. \end{aligned} \quad (5)$$

Since the partial derivatives of f_1 with respect to T must be zero,

$$\frac{\partial h_1(B)}{\partial B'} A - L_1 T' = O \quad (6)$$

is obtained. Post-multiplying (6) by T , we have

$$\frac{\partial h_1(B)}{\partial B'} B = L_1. \quad (7)$$

From (7) and $L_1 - L_1' = O$, (3) follows. \square

It is obvious that B can be replaced by \tilde{B} in Lemma 1. Let W be a diagonal matrix with the diagonal elements being the weights for observed variables, that is

$$W = \text{diag}(w_1, \dots, w_p), \quad (8)$$

where w_i is the weight for the i -th observed variable. For instance, in the case of Kaiser's normalization, $w_i = 1/\sqrt{c_i}$, where $c_i = \sum_{j=1}^q \beta_{ij}^2$ with $\beta_{ij} = (B)_{ij}$. For the

method of selecting observed variables for a rotation criterion, w_i is 1 or 0 depending on inclusion or exclusion of the i -th observed variable with respect to the criterion. With W of (8) the unrotated loading matrix becomes WA , which is rotated to

$$\Gamma = WA T \quad (9)$$

by the transformation matrix T . The rotated loading matrix is followed by the denormalization which gives $B = AT$. Note that the optimization is not in terms of B , but of Γ , though Γ is a function of B .

The restrictions for the rotated loadings with the weights W for observed variables are given by the following theorem.

Theorem 1. In the case of orthogonal rotation, let $h_1^*(\Gamma)$ be a function to be optimized by Γ with associated assumptions in (9), then the restrictions for the rotated loadings are

$$g_{1rs} = \left(\Gamma^r \frac{\partial h_1^*(\Gamma)}{\partial \Gamma^r} - \frac{\partial h_1^*(\Gamma)}{\partial \Gamma^s} \Gamma \right)_{rs} = 0, \quad (q \geq r > s \geq 1), \quad (10)$$

Proof. Replacing B and $h_1(B)$ by Γ and $h_1^*(\Gamma)$, respectively, in Lemma 1, we have (10). \square

Under the assumption of multivariate normality, the asymptotic covariance matrix for the maximum likelihood estimators with restrictions can be obtained in the following way. Let $\boldsymbol{\theta}(t \times 1)$ be a vector of t parameters with u restrictions represented by $\mathbf{g}(\boldsymbol{\theta}) = \mathbf{0}(u \times 1)$. Then, the maximum likelihood estimator $\hat{\boldsymbol{\theta}}$ is given by maximizing the following log Wishart likelihood

$$\ln L = -(n/2)\{\ln \Sigma + \text{tr}(\Sigma^{-1}S)\} + \text{const.} \quad (11)$$

with the restrictions $\mathbf{g}(\hat{\boldsymbol{\theta}}) = \mathbf{0}$, where $n+1 = N$ is the number of observations and S is an unbiased sample covariance matrix. The asymptotic covariance matrix for $\hat{\boldsymbol{\theta}}$ is obtained as the submatrix I^* of the inverse of the augmented information matrix I_A (see e.g., Silvey, 1975):

$$I_A^{-1} = \begin{bmatrix} I(\boldsymbol{\theta}) & \partial \mathbf{g}' / \partial \boldsymbol{\theta} \\ \partial \mathbf{g}' / \partial \boldsymbol{\theta}' & O \end{bmatrix}^{-1} = \begin{bmatrix} I^* & \# \\ \# & \# \end{bmatrix}, \quad (12)$$

where $I(\boldsymbol{\theta})$ is the information matrix with respect to $\boldsymbol{\theta}$ whose (i, j) th element is $E(-\partial^2 \ln L / \partial \theta_i \partial \theta_j)$ with $\theta_i = (\boldsymbol{\theta})_{i1}$, $(i = 1, \dots, t)$; and #'s denote matrices which will not be used. In the case of the model represented by (1), $\boldsymbol{\theta}$ consists of the non-fixed parameters in B and Ψ , while for the model of (2) the elements of $\boldsymbol{\theta}$ are the non-fixed parameters in \tilde{B} and D . The vector $\mathbf{g}(\boldsymbol{\theta})$ is given by (10) which is expressed by the parameters and arranged in a vector with $u = (q^2 - q)/2$ elements.

The matrix $I(\boldsymbol{\theta})$ is easily obtained (see e.g., Jennrich, 1974, p. 125, p. 129), while $\partial \mathbf{g}(\boldsymbol{\theta}) / \partial \boldsymbol{\theta}'$ depends on the form of $\mathbf{g}(\boldsymbol{\theta})$ or $h_1^*(\Gamma)$. Ogasawara (1996, p. 124) derived $\partial \mathbf{g}(\boldsymbol{\theta}) / \partial \boldsymbol{\theta}'$ for the orthomax solution with Kaiser's normalization. Here, we derive

$\partial \mathbf{g}(\boldsymbol{\theta})/\partial \boldsymbol{\theta}'$ for the general case of rotation criteria with Kaiser's normalization. Let $\boldsymbol{\gamma}_{(i) \cdot}$ be the i -th row of Γ , then $\Gamma' \partial h_1^*(\Gamma)/\partial \Gamma = \sum_{i=1}^p \boldsymbol{\gamma}_{(i) \cdot} \partial h_1^*(\Gamma)/\partial \boldsymbol{\gamma}_{(i) \cdot}$. Noting that $\boldsymbol{\gamma}_{ij}(\boldsymbol{\Gamma})_{ij} = \beta_{ij}/\sqrt{c_i} = \beta_{ij}/\sqrt{\sum_{k=1}^q \beta_{ik}^2}$, we have

$$\frac{\partial \mathbf{g}_{rs}}{\partial \beta_{ij}} = (A_{ij} - A_{ij}')_{rs}, \quad (i=1, \dots, p; \quad r, s, j=1, \dots, q; \quad r > s), \quad (13)$$

where

$$\begin{aligned} A_{ij} &= \partial \left(\Gamma' \frac{\partial h_1^*(\Gamma)}{\partial \Gamma} \right) / \partial \beta_{ij} \\ &= \frac{\partial \boldsymbol{\gamma}_{(i) \cdot}}{\partial \beta_{ij}} \frac{\partial h_1^*(\Gamma)}{\partial \boldsymbol{\gamma}_{(i) \cdot}} + \Gamma' \frac{\partial^2 h_1^*(\Gamma)}{\partial \Gamma \partial \beta_{ij}} \\ &= \frac{1}{\sqrt{c_i}} (\mathbf{e}_j - \boldsymbol{\gamma}_{(i) \cdot} \boldsymbol{\gamma}_{ij}) \frac{\partial h_1^*(\Gamma)}{\partial \boldsymbol{\gamma}_{(i) \cdot}} + \Gamma' \frac{\partial^2 h_1^*(\Gamma)}{\partial \Gamma \partial \beta_{ij}} \end{aligned} \quad (14)$$

with \mathbf{e}_j being the q -dimensional vector whose j -th element is one and others zero. The second term in the right-hand side of (14) can also be described as

$$\begin{aligned} \Gamma' \frac{\partial^2 h_1^*(\Gamma)}{\partial \Gamma \partial \beta_{ij}} &= \sum_{k=1}^q \Gamma' \frac{\partial^2 h_1^*(\Gamma)}{\partial \Gamma \partial \gamma_{ik}} \frac{\partial \gamma_{ik}}{\partial \beta_{ij}} \\ &= \sum_{k=1}^q \Gamma' \frac{\partial^2 h_1^*(\Gamma)}{\partial \Gamma \partial \gamma_{ik}} \frac{(\delta_{kj} - \gamma_{ik} \gamma_{ij})}{\sqrt{c_i}} \\ &= \Gamma' \frac{\partial^2 h_1^*(\Gamma)}{\partial \Gamma \partial \gamma_{ij}} \frac{1}{\sqrt{c_i}} - \sum_{k=1}^q \Gamma' \frac{\partial^2 h_1^*(\Gamma)}{\partial \Gamma \partial \gamma_{ik}} \frac{\gamma_{ik} \gamma_{ij}}{\sqrt{c_i}}, \end{aligned} \quad (15)$$

where δ_{kj} is the Kronecker delta. The expression of (15) may be useful in some cases since it is described only by Γ and c_i .

For the orthomax rotation with Kaiser's normalization,

$$h_1^*(\Gamma) = \frac{1}{4} \sum_{j=1}^q \left\{ \sum_{i=1}^p \gamma_{ij}^4 - \frac{w}{p} \left(\sum_{i=1}^p \gamma_{ij}^2 \right)^2 \right\}, \quad (16)$$

where w is the weight for the orthomax rotation. In this case, we directly use (14) instead of (15): the partial derivatives in (14) are

$$\frac{\partial h_1^*(\Gamma)}{\partial \gamma_{uv}} = \gamma_{uv}^3 - \frac{w}{p} \left(\sum_{a=1}^p \gamma_{av}^2 \right) \gamma_{uv} \quad (17)$$

and from (17)

$$\begin{aligned} \left(\frac{\partial^2 h_1^*(\Gamma)}{\partial \Gamma \partial \beta_{ij}} \right)_{uv} &= \frac{\partial^2 h_1^*(\Gamma)}{\partial \gamma_{uv} \partial \beta_{ij}} \\ &= \left\{ \delta_{ui} \left(3\gamma_{uv}^2 - \frac{w}{p} \sum_{a=1}^p \gamma_{av}^2 \right) - \frac{2w}{p} \gamma_{iv} \gamma_{uv} \right\} \frac{\partial \gamma_{iv}}{\partial \beta_{ij}}, \end{aligned} \quad (18)$$

($i, u=1, \dots, p; \quad j, v=1, \dots, q$).

From (17) and (18), $(A_{ij})_{rs}$ in (14) is

$$\begin{aligned}
(A_{ij})_{rs} = & \left\{ \gamma_{is}^3 - \frac{w}{p} \left(\sum_{a=1}^p \gamma_{as}^2 \right) \gamma_{is} \right\} \frac{\delta_{rj} - \gamma_{ir} \gamma_{ij}}{\sqrt{c_i}} \\
& + \left\{ \gamma_{ir} \left(3\gamma_{is}^3 - \frac{w}{p} \sum_{a=1}^p \gamma_{as}^2 \right) - \frac{2w}{p} \gamma_{is} \sum_{a=1}^p \gamma_{ar} \gamma_{as} \right\} \frac{\delta_{sj} - \gamma_{is} \gamma_{ij}}{\sqrt{c_i}}, \\
& (q \geq r > s \geq 1).
\end{aligned} \tag{19}$$

From (19) and $(A'_{ij})_{rs} = (A_{ij})_{sr}$, $\partial g_{1rs} / \partial \beta_{ij}$ in (13) is obtained. Though this result gives the same value as that of Ogasawara (1996), (19) is described only by Γ and c_i , and is simpler than that of Ogasawara (1996) where B is used.

3. Restrictions of Parameters and their Derivatives in Oblique Rotation

The covariance structures for unstandardized and standardized observed variables with obliquely rotated common-factors are

$$\Sigma = B\Phi B' + \Psi, \text{Diag } \Phi = I_q, \tag{20}$$

and

$$\Sigma = D(\tilde{B}\tilde{\Phi}\tilde{B}' + \text{Diag}(I_p - \tilde{B}\tilde{\Phi}\tilde{B}'))D, \text{Diag } \tilde{\Phi} = I_q, \tag{21}$$

respectively. The equations (20) and (21) correspond to (1) and (2) with orthogonal rotation, respectively. The matrix D is equivalent to that for orthogonal rotation. The $p \times q$ loading matrices B and \tilde{B} and obliquely rotated ones. The $q \times q$ matrices Φ and $\tilde{\Phi}$ are factor correlation matrices.

For oblique rotations, $q^2 - q$ restrictions are required for the exploratory factor analysis models (20) and (21) to be identified. We have two kinds of optimization criteria for oblique rotation, i.e., those for $B(\tilde{B})$ (direct method) and for $B\Phi(\tilde{B}\tilde{\Phi})$ (indirect method). Since presently the direct method is usually used, only the direct one will be dealt with. (For the standard errors of obliquely rotated parameters given by the indirect method, see Ogasawara, 1998b.) The general form of the restrictions for obliquely rotated parameters by the direct method was derived by Jennrich (1973). Because Jennrich's derivation involves implicit differentiation as was the case for Archer and Jennrich (1973) and is rather complicated, we show the result using Lagrange multipliers derived by Ogasawara (1998b, Corollary 2) as the following lemma:

Lemma 2 (Jennrich, 1973; Ogasawara, 1998b). Let $h_2(B)$ be an optimization criterion for B in (20), then the restrictions for B and Φ are

$$\left(B' \frac{\partial h_2(B)}{\partial B} \Phi^{-1} \right)_{rs} = 0, \quad (r, s = 1, \dots, q; r \neq s). \tag{22}$$

Proof. Let A be an unrotated loading matrix for orthogonal common factors and T be a transformation matrix for oblique rotation. Then, the obliquely rotated loading matrix $B = AT^{-1}$ is obtained by optimizing $h_2(B)$ with respect to B subject to $\text{Diag}(T'T) = I_q$. Let $L_2(q \times q)$ be a diagonal matrix with the diagonal elements being Lagrange multipliers, then B is given by optimizing the following f_2

with respect to T and L_2 :

$$f_2 = h_2(B) + (1/2)\text{tr}\{(T'T - I_q)L_2\}. \quad (23)$$

Taking the differential of f_2 with respect to T ,

$$\begin{aligned} df_2 &= \text{tr}\left\{\frac{\partial h_2(B)}{\partial B'} dB\right\} + \text{tr}(L_2 T' dT) \\ &= -\text{tr}\left\{\frac{\partial h_2(B)}{\partial B'} \Lambda T'^{-1} (dT') T'^{-1}\right\} + \text{tr}(L_2 T' dT) \\ &= -\text{tr}\left\{\left(T^{-1} \Lambda' \frac{\partial h_2(B)}{\partial B} T^{-1} - L_2 T'\right) dT\right\} \end{aligned} \quad (24)$$

follows. Since the partial derivatives of f_2 with respect to T are zero, we have

$$B' \frac{\partial h_2(B)}{\partial B} T^{-1} = L_2 T'. \quad (25)$$

Post-multiplying (25) by $T(T'T)^{-1} = T\Phi^{-1}$,

$$B' \frac{\partial h_2(B)}{\partial B} \Phi^{-1} = L_2 \quad (26)$$

follows. Because L_2 is a diagonal matrix, (22) is obtained. \square

Obviously, B and Φ in (22) can be replaced by \tilde{B} and $\tilde{\Phi}$ in Lemma 2. Let W be a diagonal matrix with the diagonal elements w_i , ($i=1, \dots, p$) being weights for observed variables as was the case for orthogonal rotation. For Kaiser's normalization, $w_i = 1/\sqrt{c_i}$ where

$$c_i = (B\Phi B')_{ii} = \sum_{j=1}^q \sum_{k=1}^q \beta_{ij} \phi_{jk} \beta_{ik}, \quad (i=1, \dots, p), \quad (27)$$

with $\phi_{jk} = (\Phi)_{jk}$, ($j, k=1, \dots, q$). In a manner similar to the case of orthogonal rotation,

$$\Gamma = W \Lambda T'^{-1} = W B \quad (28)$$

is defined.

The restrictions for obliquely rotated parameters with the weights W for observed variables are obtained as the following theorem.

Theorem 2. Let $h_2^*(\Gamma)$ be an optimization criterion of Γ for oblique rotation with the weights W in (28), then the restrictions for an obliquely rotated solution are

$$g_{2rs} = \left(\Gamma' \frac{\partial h_2^*(\Gamma)}{\partial \Gamma} \Phi^{-1}\right)_{rs} = 0, \quad (r, s=1, \dots, q; r \neq s). \quad (29)$$

Proof. Replacing B and $h_2(B)$ by Γ and $h_2^*(\Gamma)$, respectively in Lemma 2, (29) follows. \square

Let θ be the vector of the non-fixed non-duplicated parameters in B (or \tilde{B}), Φ (or $\tilde{\Phi}$) and Ψ (or D), then the asymptotic covariance matrix of the estimates of the parameters can be obtained as I^* in (12) with the restrictions $\mathbf{g}(\theta) = \mathbf{0}$, where the $(q^2 - q) \times 1$ vector $\mathbf{g}(\theta)$ is composed of the elements of (29). The information matrix

$I(\boldsymbol{\theta})$ for unstandardized observed variables is given by various literatures (e.g., Jöreskog, 1969, p. 189), while $I(\boldsymbol{\theta})$ for standardized observed variables is described in Ogasawara (1998b, Appendix 5). The partial derivatives of (29) with respect to β_{ij} and ϕ_{hl} with Kaiser's normalization are derived as follows:

$$\begin{aligned} \frac{\partial g_{2rs}}{\partial \beta_{ij}} &= \sum_{k=1}^q \frac{\partial g_{2rs}}{\partial \gamma_{ik}} \frac{\partial \gamma_{ik}}{\partial \beta_{ij}} \\ &= \left(\frac{\partial h_2^*(\Gamma)}{\partial \Gamma} \boldsymbol{\Phi}^{-1} \right)_{is} \frac{\partial \gamma_{ir}}{\partial \beta_{ij}} + \sum_{k=1}^q \left(\Gamma' \frac{\partial^2 h_2^*(\Gamma)}{\partial \Gamma \partial \gamma_{ik}} \boldsymbol{\Phi}^{-1} \right)_{rs} \frac{\partial \gamma_{ik}}{\partial \beta_{ij}}, \end{aligned} \quad (30)$$

$(i=1, \dots, p; r, s, j=1, \dots, q; r \neq s),$

where

$$\frac{\partial \gamma_{iu}}{\partial \beta_{ij}} = \frac{1}{\sqrt{c_i}} \{ \delta_{ju} - \gamma_{iu}(\Gamma \boldsymbol{\Phi})_{ij} \} \quad (31)$$

and

$$\begin{aligned} \frac{\partial g_{2rs}}{\partial \phi_{ef}} &= \sum_{i=1}^p \sum_{k=1}^q \frac{\partial g_{2rs}}{\partial \gamma_{ik}} \frac{\partial \gamma_{ik}}{\partial \phi_{ef}} - \left(\Gamma' \frac{\partial h_2^*(\Gamma)}{\partial \Gamma} \boldsymbol{\Phi}^{-1} (I_{ef} + I_{fe}) \boldsymbol{\Phi}^{-1} \right)_{rs} \\ &= \sum_{i=1}^p \left\{ \left(\frac{\partial h_2^*(\Gamma)}{\partial \Gamma} \boldsymbol{\Phi}^{-1} \right)_{is} \frac{\partial \gamma_{ir}}{\partial \phi_{ef}} + \sum_{k=1}^q \left(\Gamma' \frac{\partial^2 h_2^*(\Gamma)}{\partial \Gamma \partial \gamma_{ik}} \boldsymbol{\Phi}^{-1} \right)_{rs} \frac{\partial \gamma_{ik}}{\partial \phi_{ef}} \right\} \\ &\quad - \left(\Gamma' \frac{\partial h_2^*(\Gamma)}{\partial \Gamma} \boldsymbol{\Phi}^{-1} \right)_{rr} (\delta_{re} \phi^{fs} + \delta_{rf} \phi^{es}), \end{aligned} \quad (32)$$

$(r, s, e, f=1, \dots, q; r \neq s; e > f),$

where

$$\frac{\partial \gamma_{iu}}{\partial \phi_{ef}} = -\frac{\beta_{iu}}{2c_i^{3/2}} \frac{\partial c_i}{\partial \phi_{ef}} = -\gamma_{iu} \gamma_{ie} \gamma_{ij}; \quad (33)$$

I_{ef} is the matrix whose (e, f) th element is one and others zero; and $\phi^{fs} = (\boldsymbol{\Phi}^{-1})_{rs}$.

4. Standard Errors for the Rotated Solution with the General Symmetric Family of Quartic Criterion with Kaiser's Normalization

Based on Crawford and Ferguson's (1970) rotation criterion:

$$K_1 \sum_{e \neq f} \sum_{i=1}^p \beta_{ie}^2 \beta_{if}^2 + K_2 \sum_{i+j=k=1}^q \beta_{ik}^2 \beta_{jk}^2, \quad (34)$$

Jennrich (1973) proposed a generalized rotation criterion:

$$\begin{aligned} 4h_2(B) &= k_1 \left(\sum_{i=1}^p \sum_{j=1}^q \beta_{ij}^2 \right)^2 + k_2 \sum_{i=1}^p \left(\sum_{j=1}^q \beta_{ij}^2 \right)^2 + k_3 \sum_{j=1}^q \left(\sum_{i=1}^p \beta_{ij}^2 \right)^2 \\ &\quad + k_4 \sum_{i=1}^p \sum_{j=1}^q \beta_{ij}^4. \end{aligned} \quad (35)$$

The criterion (34) is equivalent to that of the orthomax rotation when the weight in (16) is

$$w = pK_2 / (K_1 + K_2) \quad (36)$$

(Crawford & Ferguson, 1970). The criterion (35) is equivalent to (34) when

$$k_1=0, k_2=K_1, k_3=K_2, k_4=-(K_1+K_2), \quad (37)$$

and is equivalent to the oblimin criterion when

$$k_1=-\gamma/p, k_2=1, k_3=\gamma/p, k_4=-1, \quad (38)$$

where γ is the weight for the oblimin rotation. The equation (37) was first incorrectly introduced by Jennrich (1973, p. 599). The correct equation was given by Clarkson and Jennrich (1988, Table 1), although the equation for the covarimin method has still typographical errors: $k_3=-1/p$ and $k_4=1$ in their Table 1 should be $k_3=1/p$ and $k_4=-1$. The criterion (35) was later called the general symmetric family of quartic criterion (Clarkson & Jennrich, 1988). In orthogonal rotation the added usefulness of (35) is unknown (note that (34) is equivalent to that of the orthomax rotation). On the other hand, in the case of oblique rotation, by employing (34) as a special case of (35) we can use criteria other than the oblimin ones. One of the examples of practical use is Crawford's (1975) criterion for his primary parsimony rotation method whose special case is called the oblique varimax method (Browne & Du Toit, 1992). The primary parsimony method minimizes (34) with K_1 and K_2 which correspond to w in the orthomax criterion (see (36)). For instance, $w=1$ for the varimax method corresponds to $K_1=p-1$ and $K_2=1$ (i.e., $k_1=0, k_2=p-1, k_3=1, k_4=-p$). For the parsimax rotation in the orthomax family, $w=p(q-1)/(p+q-2)$ corresponds to $K_1=p-1$ and $K_2=q-1$. Consequently, when $q=2$, the oblique varimax criterion becomes equivalent to that of the oblique parsimax method, while in orthogonal rotation the varimax method is equivalent to the equamax method ($w=q/2$) when $q=2$. Note that the oblique varimax method does not maximize (16), but that it minimizes (34) with $K_1=p-1$ and $K_2=1$. In case of oblique rotation, these two optimization criteria are different, while they are equivalent in orthogonal rotation.

The primary parsimony rotation method usually gives less oblique rotated factors than the oblimin method with $\gamma=0$ (the quartimin method) (see Crawford, 1975, Tables 4 and 5) and seems to be relatively free from the non-convergence problem which frequently occurs for the oblimin method with positive γ (Jennrich, 1979). The primary parsimony method can be used easily by replacing the coefficients of the cubic equation to be solved in the direct oblimin method with other appropriate ones (Clarkson & Jennrich, 1988).

The restrictions and their derivatives with respect to obliquely rotated parameters when the general quartic criterion (35) is used were given by Jennrich (1973) for the case without Kaiser's normalization. In this section, the derivatives will be given for the case with Kaiser's normalization. Replacing B and $h_2(\cdot)$ in (35) by Γ and $h_2^*(\cdot)$, the derivatives of $h_2^*(\Gamma)$ in (30) and (32) with respect to Γ are

$$\frac{\partial h_2^*(\Gamma)}{\partial \gamma_{uv}} = \gamma_{uv} M_{uv}, \quad (u=1, \dots, p; v=1, \dots, q), \quad (39)$$

where

$$M_{uv} = k_1 \sum_{a=1}^p \sum_{b=1}^q \gamma_{ab}^2 + k_2 \sum_{b=1}^q \gamma_{ub}^2 + k_3 \sum_{a=1}^p \gamma_{av}^2 + k_4 \gamma_{uv}^2, \quad (40)$$

and

$$\begin{aligned} \frac{\partial^2 h_2^*(\Gamma)}{\partial \gamma_{uv} \partial \gamma_{ik}} &= \delta_{ui} \delta_{vk} M_{ik} + \gamma_{ik} \frac{\partial M_{ik}}{\partial \gamma_{uv}} \\ &= \delta_{ui} \delta_{vk} M_{ik} + 2\gamma_{ik} \gamma_{uv} (k_1 + k_2 \delta_{iu} + k_3 \delta_{kv} + k_4 \delta_{iu} \delta_{kv}), \end{aligned} \quad (41)$$

$(u, i = 1, \dots, p; v, k = 1, \dots, q).$

From (39), (40) and (41), the derivatives (30) and (32) become

$$\begin{aligned} \frac{\partial \underline{g}_{2rs}}{\partial \beta_{ij}} &= \sum_{k=1}^q \left\{ \gamma_{ik} M_{ik} \phi^{ks} \frac{\partial \gamma_{ir}}{\partial \beta_{ij}} + (\gamma_{ir} M_{ik} \phi^{ks} + 2k_1 \gamma_{ik} (\Gamma' \Gamma \Phi^{-1})_{rs} \right. \\ &\quad \left. + 2k_2 \gamma_{ik} \gamma_{ir} (\Gamma \Phi^{-1})_{is} + 2k_3 \gamma_{ik} (\Gamma' \Gamma)_{rk} \phi^{ks} + 2k_4 \gamma_{ik}^2 \gamma_{ir} \phi^{ks} \right) \frac{\partial \gamma_{ik}}{\partial \beta_{ij}} \Big\}, \end{aligned} \quad (42)$$

$(i = 1, \dots, p; r, s, j = 1, \dots, q; r \neq s),$

and

$$\begin{aligned} \frac{\partial \underline{g}_{2rs}}{\partial \phi_{ef}} &= \sum_{i=1}^p \sum_{k=1}^q \left\{ \gamma_{ik} M_{ik} \phi^{ks} \frac{\partial \gamma_{ir}}{\partial \phi_{ef}} + (\gamma_{ir} M_{ik} \phi^{ks} + 2k_1 \gamma_{ik} (\Gamma' \Gamma \Phi^{-1})_{rs} \right. \\ &\quad \left. + 2k_2 \gamma_{ik} \gamma_{ir} (\Gamma \Phi^{-1})_{is} + 2k_3 \gamma_{ik} (\Gamma' \Gamma)_{rk} \phi^{ks} + 2k_4 \gamma_{ik}^2 \gamma_{ir} \phi^{ks} \right) \frac{\partial \gamma_{ik}}{\partial \phi_{ef}} \\ &\quad - \gamma_{ir} \gamma_{ik} M_{ik} \phi^{kr} (\delta_{re} \phi^{fs} + \delta_{rf} \phi^{es}), \end{aligned} \quad (43)$$

$(i = 1, \dots, p; r, s, e, f = 1, \dots, q; r \neq s; e > f),$

respectively (see also Jennrich, 1973). Replacing $\partial \underline{g}(\boldsymbol{\theta}) / \partial \boldsymbol{\theta}'$ in (12) by (42) and (43) and taking the inverse of the augmented information matrix, we have the asymptotic covariance matrix of the estimators of the obliquely rotated parameters with Kaiser's normalization.

5. Numerical Examples

Numerical examples are based on two real correlation matrices and an artificial data set. In each example, an exploratory factor analysis model for standardized observed variables is assumed (see (2) and (21)). One of the real correlation matrices is that of eight physical variables (Harman, 1976, p. 22, $N = 305$), where a two-factor model is assumed. The other is the correlation matrix of twelve psychological tests (Harman, 1976, p. 401, $N = 355$), where a three-factor model is assumed. The artificial data (Ogasawara's 1999 Data B) are based on the following assumed estimates of loadings of orthogonal factors with the assumption of $N = 300$,

$$\tilde{\Lambda}' = \begin{bmatrix} .8 & .8 & .1 & .4 & .3 & .01 \\ .4 & .3 & .01 & .8 & .8 & .1 \end{bmatrix}.$$

The oblique varimax method was employed for each example. Tables 1-3 show the estimates of rotated loadings and factor correlations with their estimated asymptotic standard errors. Tables 1 and 2 contain simulated results, which have been obtained as follows. First, the correlation matrices which had been constructed by estimated parameters were regarded as population covariance matrices. Under the assumption of multivariate normality, random observations of the sample size equal to the real ones were generated. Based on these generated data, the parameters in the models were estimated by the maximum likelihood method followed by the oblique varimax rotation. There remained indeterminacies of the signs and permutations in the columns of rotated loading matrices, which were removed by choosing the loading matrices closest to the corresponding population ones. This procedure was repeated 1,000 times and we had 1,000 estimates for each parameter. The S's in Tables 1 and 2 indicate the standard deviations of these estimates.

The simulated values are close to the theoretical ones, which shows the accuracy of the theoretical values. The standard errors of the parameters with Kaiser's normalization are not so different from those without Kaiser's normalization, which corresponds to the fact that the estimates of the parameters in the two models are similar. However, we see that some of them are different both in theoretical and simulated values (e.g., Table 1, Variable 5, Factor I; Table 2, Variables 11 and 12, Factor I). Table 3 shows the results for the artificial data, where the results by simulation are not included since the artificial data, in which some of the communalities are very small, gave frequently non-convergent cases in simulation. From Table 3 we see that by Kaiser's normalization the standard errors of the estimates of the loadings for the variables with small communalities become small while the standard error of the factor correlation becomes extremely large. This

Table 1
Oblique varimax solution for eight physical variables ($N=305$)

Variable No.	Raw oblique-varimax				Normal oblique-varimax				Communalities
	I		II		I		II		
	Esti.	(N S)	Esti.	(N S)	Esti.	(N S)	Esti.	(N S)	
1	.84	(.018 .019)	.16	(.025 .026)	.85	(.018 .018)	.14	(.025 .026)	.83
2	.93	(.012 .013)	.04	(.021 .022)	.94	(.012 .013)	.02	(.021 .022)	.89
3	.90	(.015 .016)	.04	(.024 .025)	.90	(.015 .015)	.02	(.024 .025)	.83
4	.84	(.019 .019)	.12	(.027 .027)	.85	(.018 .019)	.11	(.026 .026)	.80
5	.04	(.023 .024)	.94	(.020 .021)	.06	(.027 .028)	.93	(.022 .023)	.91
6	.03	(.033 .034)	.78	(.028 .029)	.05	(.034 .034)	.78	(.028 .029)	.64
7	-.02	(.035 .035)	.77	(.030 .030)	-.01	(.034 .034)	.77	(.030 .030)	.58
8	.15	(.044 .043)	.61	(.039 .039)	.17	(.042 .042)	.60	(.039 .039)	.46
ϕ_{21}	.35 (.039 .039)				.36 (.034 .039)				

Note. Esti. = Estimates of parameters; N = Normal theory estimates of standard errors; S = Simulated standard errors.

Table 2
Oblique varimax solution for twelve psychological tests ($N = 355$)

Variable No.	Raw oblique-varimax			Normal oblique-varimax			Communalities
	I Esti. (N S)	II Esti. (N S)	III Esti. (N S)	I Esti. (N S)	II Esti. (N S)	III Esti. (N S)	
1	.14 (.047 .048)	.71 (.036 .036)	.03 (.049 .050)	.15 (.047 .049)	.71 (.036 .036)	.01 (.049 .051)	.62
2	.05 (.038 .039)	.78 (.032 .035)	.10 (.043 .043)	.06 (.042 .043)	.79 (.032 .035)	.08 (.046 .046)	.72
3	.05 (.043 .044)	.76 (.033 .033)	.03 (.047 .049)	.06 (.044 .045)	.76 (.034 .034)	.01 (.048 .050)	.62
4	-.06 (.052 .051)	.67 (.040 .041)	.11 (.060 .060)	-.05 (.046 .046)	.67 (.040 .041)	.10 (.053 .054)	.48
5	.77 (.041 .043)	.12 (.035 .036)	-.03 (.039 .042)	.79 (.042 .044)	.12 (.036 .038)	-.06 (.043 .046)	.66
6	.58 (.054 .055)	-.01 (.044 .044)	.21 (.064 .063)	.60 (.051 .051)	-.01 (.041 .042)	.18 (.059 .058)	.50
7	.73 (.043 .045)	-.01 (.032 .031)	.18 (.051 .052)	.75 (.041 .042)	-.01 (.034 .034)	.15 (.048 .049)	.71
8	.72 (.043 .042)	.08 (.034 .033)	.14 (.049 .048)	.74 (.041 .041)	.09 (.035 .035)	.11 (.047 .047)	.71
9	.21 (.067 .069)	.10 (.050 .049)	.45 (.074 .079)	.23 (.065 .066)	.11 (.050 .050)	.43 (.072 .076)	.41
10	.12 (.067 .067)	.17 (.053 .050)	.43 (.076 .076)	.15 (.065 .063)	.18 (.053 .050)	.40 (.074 .073)	.35
11	-.01 (.050 .054)	.00 (.038 .040)	.70 (.074 .087)	.03 (.064 .070)	.02 (.047 .048)	.68 (.082 .096)	.49
12	.04 (.065 .068)	-.04 (.050 .053)	.56 (.077 .084)	-.02 (.048 .053)	-.02 (.038 .041)	.54 (.074 .081)	.28
ϕ_{21}	.39 (.040 .041)			.40 (.041 .041)			
ϕ_{31}	.55 (.047 .046)			.53 (.049 .045)			
ϕ_{32}	.38 (.048 .048)			.36 (.050 .048)			

Note. Esti. = Estimates of parameters; N = Normal theory standard errors; S = Simulated standard errors.

Table 3
Oblique varimax solution for artificial data

Variable No.	Raw oblique-varimax		Normal oblique-varimax		Communalities
	I	II	I	II	
	Esti. (N)	Esti. (N)	Esti. (N)	Esti. (N)	
1	.80 (.24)	.14 (.19)	.76 (.22)	.25 (.26)	.80
2	.84 (.19)	.02 (.11)	.79 (.21)	.15 (.26)	.73
3	.12 (.09)	-.03 (.09)	.10 (.07)	-.01 (.04)	.01
4	.14 (.19)	.80 (.24)	.25 (.26)	.76 (.22)	.80
5	.02 (.11)	.84 (.19)	.15 (.26)	.79 (.21)	.73
6	-.03 (.09)	.12 (.09)	-.01 (.04)	.10 (.07)	.01
ϕ_{21}	.63 (.05)		.39 (.32)		

Note. Esti.=Estimates of parameters; N=Normal theory estimates of standard errors.

suggests that when we have large differences in communalities, not only the estimates of factor loadings, but also their standard errors may substantially change by employing Kaiser's normalization.

6. Discussion

In the first section, the weights $w_i=1, 0$ were mentioned. For this case the function to be optimized by rotated solutions are constructed from the loadings for the variables with $w_i=1$. The derivatives of the restrictions with respect to the loadings for the variables with $w_i=0$ are set equal to zero. (Note that the information matrix should be provided for all parameters.) For the variables with $w_i=1$, Kaiser's normalization may be imposed if necessary.

In the first section, Cureton and Mulaik's (1979, p.187) weight was also mentioned. Their weight w_i for the i -th variable is a function of the loading of the first principal factor written as

$$w_i = f(\lambda_{i1}), \quad (i=1, \dots, p), \quad (44)$$

where $\lambda_{i1}=(A)_{i1}$. These weights have a practical advantage in that they give a simple pattern to the data such as Thurstone's box problem to which the varimax method does not. However, presently this method is not used so often and the usefulness of the standard errors of the rotated results with the weights is unknown. Nevertheless, from a theoretical point of view, the weights are of interest. In Kaiser's normalization Γ was described as a function of B or that of B and Φ . With the weights w_i in (44), Γ is a function of B and A , or a function of B , $\Phi(=T'T)$ and A . Therefore, the set of parameters for which the information matrix is to be constructed should contain both unrotated and rotated ones. Ogasawara (1998a) derived the information matrix by using the reparameterization of covariance structures, which is described in the following way for standardized observed variables;

$$\Sigma = D(ATB' + \text{Diag}(I_p - ATB'))D. \quad (45)$$

By using the above method, the asymptotic standard errors of the rotated results with Cureton and Mulaik's (1975) weights may be derived.

REFERENCES

- Archer, C.O. & Jennrich, R.I. (1973). Standard errors for rotated factor loadings. *Psychometrika*, **38**, 581-592.
- Browne, M.W. & Du Toit, S.H.C. (1992). Automated fitting of nonstandard models. *Multivariate Behavioral Research*, **27**, 269-300.
- Clarkson, D.B. & Jennrich, R.I. (1988). Quartic rotation criteria and algorithms. *Psychometrika*, **53**, 251-259.
- Crawford, C. (1975). A comparison of the direct oblimin and primary parsimony methods of oblique rotation. *British Journal of Mathematical and Statistical Psychology*, **28**, 201-213.
- Crawford, C.G. & Ferguson, G.A. (1970). A general rotation criterion and its use in orthogonal rotation. *Psychometrika*, **35**, 321-332.
- Cudeck, R. & O'Dell, L.L. (1994). Applications of standard error estimates in unrestricted factor analysis: Significance tests for loadings and correlations. *Psychological Bulletin*, **115**, 475-487.
- Cureton, E.E. & Mulaik, S.A. (1975). The weighted varimax rotation and the promax rotation. *Psychometrika*, **40**, 183-195.
- Dixon, W.J. (Ed.) (1992). *BMDP statistical software manual* Vol. 1. Los Angeles: University of California Press.
- Harman, H.H. (1976). *Modern factor analysis* (3rd ed.). Chicago: University of Chicago Press.
- 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 Journal of Mathematical and Statistical Psychology*, **27**, 122-131.
- Jennrich, R.I. (1979). Admissible values of γ in direct oblimin rotation. *Psychometrika*, **44**, 173-177.
- Jöreskog, K.G. (1969). A general approach to confirmatory maximum likelihood factor analysis. *Psychometrika*, **34**, 183-201.
- Kaiser, H.F. (1958). The varimax criterion for analytic rotation in factor analysis. *Psychometrika*, **23**, 187-200.
- Ogasawara, H. (1996). Standard errors for rotated factor loadings by normalized orthomax method. *Japanese Journal of Behaviormetrics*, **23** (2), 122-129. (in Japanese with English abstract)
- Ogasawara, H. (1998a). Standard errors for rotation matrices with an application to the promax solution. *British Journal of Mathematical and Statistical Psychology*, **51**, 163-178.
- Ogasawara, H. (1998b). Standard errors of several indices for unrotated and rotated factors. *Economic Review*, **49** (1), Otaru University of Commerce, 21-69.
- Ogasawara, H. (1999). Standard errors for the direct oblimin solution with Kaiser's normalization. *Japanese Journal of Psychology*, **70**, 333-338. (in Japanese with English abstract)
- SAS Institute Inc. (1990). *SAS/STAT user's guide* Vol. 1 (version 6, 4th ed.). Cary, NC: Author.
- Silvey, S.D. (1975). *Statistical inference*. New York: Chapman & Hall.
- SPSS Inc. (1990). *SPSS reference guide*. Chicago: Author.

(Received May 1999, Revised January 2000)