PSYCHOMETRIKA—VOL. 77, NO. 4, 634–648
OCTOBER 2012
DOI: 10.1007/S11336-012-9281-5

THE INFINITESIMAL JACKKNIFE WITH EXPLORATORY FACTOR ANALYSIS

GUANGJIAN ZHANG

UNIVERSITY OF NOTRE DAME

KRISTOPHER J. PREACHER

VANDERBILT UNIVERSITY

ROBERT I. JENNRICH

UNIVERSITY OF CALIFORNIA AT LOS ANGELES

The infinitesimal jackknife, a nonparametric method for estimating standard errors, has been used to obtain standard error estimates in covariance structure analysis. In this article, we adapt it for obtaining standard errors for rotated factor loadings and factor correlations in exploratory factor analysis with sample correlation matrices. Both maximum likelihood estimation and ordinary least squares estimation are considered.

Key words: exploratory factor analysis, standard error, EFA, infinitesimal jackknife, IJK, nonparametric standard error estimates, model misspecification, nonnormal data, minimum deviance methods.

1. Introduction

Exploratory Factor analysis (EFA) is one of the most widely used statistical procedures in the social and behavioral sciences. Standard errors are essential in performing important tasks in EFA such as identifying salient factor loadings and understanding interrelationships among latent factors (Cudeck & O'Dell, 1994).

In this article, we use the infinitesimal jackknife (IJ) to produce standard error estimates for EFA parameters. IJ methods may be used to produce standard error estimates for parameters in EFA with either sample covariance or sample correlation matrices. Because of its popularity in applications, the specific IJ methods developed here will be in the context of sample correlation matrices. Arbitrary deviance functions may be used for initial factor loading extraction. Specific formulas are given for ordinary least squares (OLS) and maximum likelihood (ML) extraction.

When sampling from a normal population, analytic standard errors for EFA using ML extraction were developed by Archer and Jennrich (1973) for orthogonal rotation, Jennrich (1973) for oblique rotation, and Jennrich (1974) for the bordered information matrix method. When sampling from a normal population, analytic standard errors for EFA using OLS extraction were developed in Browne and Tateneni (2008). When sampling from a nonnormal population, analytic standard errors for EFA using ML extraction were developed by Jennrich and Clarkson (1980), Ogasawara (2007a), and Yuan, Marshall, and Bentler (2002). All these developments require the EFA model to be correctly specified or to have no more than moderate levels of lack of fit (Browne, 1984; Satorra, 1989).

Ordinary jackknife standard error estimates in EFA with nonnormal data were developed by Clarkson (1979). This method does not require a correctly specified EFA model and uses ML extraction.

Requests for reprints should be sent to Guangjian Zhang, Psychology Department, Haggar Hall, University of Notre Dame, Notre Dame, IN 46556, USA. E-mail: gzhang3@nd.edu

The bootstrap has also been used to estimate standard errors in EFA: see Lambert, Wild, and Durand (1991) and Ichikawa and Konishi (1995) for ML extraction and Zhang, Preacher, and Luo (2010) for OLS extraction. The bootstrap does not require a correctly specified EFA model.

Jennrich (2008) has shown how to use the IJ to produce standard error estimates for the values of an arbitrary function of a sample covariance matrix for a sample from an arbitrary population with finite fourth moments. He proved that the standard error estimates produced are consistent. He applied his method to covariance structure analysis and produced standard error estimates for its parameter estimates. The covariance structure did not need to be correctly specified.

Here we apply the IJ to EFA and produce consistent standard error estimates for the estimates of factor loadings and factor correlations. The population sampled need not be normal, the EFA model need not be correctly specified, and arbitrary deviance functions may be used for extraction. Details are given for EFA of a correlation matrix.

An advantage of the IJ method over the ordinary jackknife method and bootstrap methods is that it requires only one EFA, while the ordinary jackknife requires one for each jackknife sample, and the bootstrap requires one for each bootstrap sample. In the case of EFA, the ordinary jackknife and bootstrap loading and correlation matrices from each subsample must be aligned before the standard errors are computed. In the case of the IJ only one loading matrix and one factor correlation matrix are produced and no alignment is required. The first advantage of the IJ method applies to any covariance structure model, but the second advantage is specifically for EFA.

In summary, what is new in this article is the use of the IJ to produce consistent standard error estimates in EFA. These do not require normal sampling or correctly specified EFA models and can be used for EFA of sample correlation matrices.

The rest of the article is organized as follows. We first provide background information on the EFA model and its estimation with sample correlation matrices. ML estimation and OLS estimation are presented as special cases of minimum deviance methods. We then discuss how to use the IJ method to estimate standard errors in exploratory factor analysis. Implementation details on OLS estimation and ML estimation with sample correlation matrices are given. We illustrate the IJ method for both OLS estimation and ML estimation with both normal and elliptical data. IJ standard error estimates and analytic standard error estimates are further compared with an empirical example. We conclude the article with several comments.

2. Background

Let $\boldsymbol{x}_1, \ldots, \boldsymbol{x}_i, \ldots, \boldsymbol{x}_n$ be a sample of n observations. Let $\boldsymbol{S}$ be the biased form of the sample covariance matrix, where the sum of cross-products is divided by n rather than $n-1$. Standardizing $\boldsymbol{S}$

$$\boldsymbol{R}(\boldsymbol{S}) = \boldsymbol{D}_s^{-\frac{1}{2}} \boldsymbol{S} \boldsymbol{D}_s^{-\frac{1}{2}} \tag{1}$$

produces the sample correlation matrix $\boldsymbol{R}$. Here $\boldsymbol{D}_s$ is a diagonal matrix of sample variances.

The oblique factor analysis model specifies that

$$\boldsymbol{\Sigma}(\boldsymbol{\theta}) = \boldsymbol{\Lambda}\boldsymbol{\Phi}\boldsymbol{\Lambda}' + \boldsymbol{\Psi}. \tag{2}$$

Here $\boldsymbol{\Lambda}$ is a $p \times m$ factor loading matrix, $\boldsymbol{\Phi}$ is an $m \times m$ factor correlation matrix, $\boldsymbol{\Psi}$ is a $p \times p$ diagonal matrix of unique variances, and $\boldsymbol{\theta}$ is a column vector of length $q = pm + m(m-1)/2 + p$. The first pm components of $\boldsymbol{\theta}$ are components of $\boldsymbol{\Lambda}$ read column-wise, the next $m(m-1)/2$ components are the lower diagonal components of $\boldsymbol{\Phi}$ read column-wise, and the last p components are the diagonal components of $\boldsymbol{\Psi}$.

The EFA model of Equation (2) is not identified. Let $\boldsymbol{t}_1, \ldots, \boldsymbol{t}_m$ be any m independent vectors scaled so $\boldsymbol{t}_i'\boldsymbol{\Phi}\boldsymbol{t}_i = 1$, and let $\boldsymbol{T} = (\boldsymbol{t}_1, \ldots, \boldsymbol{t}_m)'$. Then $\boldsymbol{T}$ varies over an $m^2 - m$ manifold. For any $\boldsymbol{T}$ on this manifold, let $\widetilde{\boldsymbol{\Lambda}} = \boldsymbol{\Lambda}\boldsymbol{T}^{-1}$ and $\widetilde{\boldsymbol{\Phi}} = \boldsymbol{T}\boldsymbol{\Phi}\boldsymbol{T}'$. Then $\boldsymbol{\Lambda}\boldsymbol{\Phi}\boldsymbol{\Lambda}' = \widetilde{\boldsymbol{\Lambda}}\widetilde{\boldsymbol{\Phi}}\widetilde{\boldsymbol{\Lambda}}'$. Thus, there are parameter vectors $\boldsymbol{\theta}$ and $\widetilde{\boldsymbol{\theta}}$ such that $\boldsymbol{\Sigma}(\boldsymbol{\theta}) = \boldsymbol{\Sigma}(\widetilde{\boldsymbol{\theta}})$ but $\boldsymbol{\theta} \neq \widetilde{\boldsymbol{\theta}}$. This means without additional assumptions the covariance structure $\boldsymbol{\Sigma}(\boldsymbol{\theta})$ is not identified. This lack of identification in EFA is referred to as rotational indeterminacy. Factor rotation (Browne, 2001) is conducted to select one solution from these infinitely many solutions. Factor rotation often involves minimizing a scalar function of rotated factor loadings $Q(\boldsymbol{\Lambda})$, for example, CF-varimax (Crawford & Ferguson, 1970),

$$Q(\boldsymbol{\Lambda}) = \left(1 - \frac{1}{p}\right)\sum_{i=1}^{p}\sum_{j=1}^{m}\sum_{l\neq j}^{m}\lambda_{ij}^2\lambda_{il}^2 + \frac{1}{p}\sum_{j=1}^{m}\sum_{i=1}^{p}\sum_{k\neq i}^{p}\lambda_{ij}^2\lambda_{kj}^2. \tag{3}$$

When oblique rotation is conducted, factor rotation can be regarded as imposing $m^2 - m$ constraints on the rotated factor loading matrix $\boldsymbol{\Lambda}$ and the factor correlation matrix $\boldsymbol{\Phi}$ (Jennrich, 1973, Equation (28)),

$$\boldsymbol{\varphi}(\boldsymbol{\theta}) = \mathrm{vecoff}\left(\boldsymbol{\Lambda}'\frac{dQ}{d\boldsymbol{\Lambda}}\boldsymbol{\Phi}^{-1}\right) = \mathbf{0}. \tag{4}$$

Here vecoff$(\boldsymbol{A})$ is an operator extracting off-diagonal elements of a square matrix and placing them into an $m^2 - m$ component vector. Components of $\boldsymbol{\varphi}$ are indexed by double subscripts: $\varphi_{uv} = (\boldsymbol{\Lambda}'\frac{dQ}{d\boldsymbol{\Lambda}}\boldsymbol{\Phi}^{-1})_{uv}$. The $p \times m$ matrix $\frac{dQ}{d\boldsymbol{\Lambda}}$ contains partial derivatives $\frac{\partial Q}{\partial \lambda_{ij}}$,

$$\frac{dQ}{d\boldsymbol{\Lambda}} = 4\boldsymbol{\Lambda} * \boldsymbol{M}. \tag{5}$$

Here $\boldsymbol{\Lambda} * \boldsymbol{M}$ denotes the element-by-element multiplication of the two matrices. A typical element of $\boldsymbol{M}$ is

$$M_{ir} = \left(\left(1 - \frac{1}{p}\right)\sum_{r=1}^{m}\lambda_{ir}^2\right) + \left(\frac{1}{p}\sum_{i=1}^{p}\lambda_{ir}^2\right) - \lambda_{ir}^2. \tag{6}$$

Here $1 \le i \le p$ and $1 \le r \le m$. This is a special case of the result derived in Jennrich (1973, Equation (50)) for the generalized Crawford–Ferguson family.[1] The $m^2 - m$ rotation constraints in Equation (4) are sufficient to identify the EFA model.

The EFA covariance structure of Equation (2) is routinely estimated with a sample correlation matrix $\boldsymbol{R}$ using minimum deviance methods

$$F(\boldsymbol{\theta}, \boldsymbol{R}) = D\big(\boldsymbol{R}, \boldsymbol{\Sigma}(\boldsymbol{\theta})\big). \tag{7}$$

The minimum deviance function $D(\boldsymbol{R}, \boldsymbol{\Sigma}(\boldsymbol{\theta})) \ge 0$, and it is zero if $\boldsymbol{\Sigma}(\boldsymbol{\theta}) = \boldsymbol{R}$. The two most frequently used minimum deviance methods in EFA are ML,

$$F_{\mathrm{ML}}(\boldsymbol{\theta}, \boldsymbol{R}) = \log\big|\boldsymbol{\Sigma}(\boldsymbol{\theta})\big| + \mathrm{tr}\big(\boldsymbol{\Sigma}(\boldsymbol{\theta})^{-1}\boldsymbol{R}\big) - \log|\boldsymbol{R}| - p, \tag{8}$$

and OLS,

$$F_{\mathrm{OLS}}(\boldsymbol{\theta}, \boldsymbol{R}) = \mathrm{tr}\big(\boldsymbol{R} - \boldsymbol{\Sigma}(\boldsymbol{\theta})\big)^2. \tag{9}$$

Here tr$(\boldsymbol{A})$ is the trace of a square matrix $\boldsymbol{A}$.

[1]The generalized Crawford–Ferguson family requires specifying four parameters κ_1, κ_2, κ_3, and κ_4. Oblique CF-varimax rotation of Equation (3) is equivalent to the generalized Crawford–Ferguson family when $\kappa_1 = 0$, $\kappa_2 = 1 - 1/p$, $\kappa_3 = 1/p$, and $\kappa_4 = -1$. The equivalence of the oblique CF-varimax rotation to the generalized Crawford–Ferguson family was also presented in Clarkson and Jennrich (1988, Table 1).

3. The IJ Method of Estimating Standard Errors in EFA

We first present IJ standard error estimation in EFA for an arbitrary minimum deviance method in Equation (7). We then give details of IJ standard error estimation for the two most widely used minimum deviance methods: OLS and ML.

3.1. Minimum Deviance Methods of EFA with Sample Correlations

To facilitate derivation, factor extraction and factor rotation are treated in a single step. Estimates of rotated factor loadings and factor correlations $\hat{\boldsymbol{\theta}}$ are solutions to the following estimating equations:

$$g(\boldsymbol{\theta}, \boldsymbol{S}) = \begin{bmatrix} \frac{\partial}{\partial \boldsymbol{\theta}} F(\boldsymbol{\theta}, \boldsymbol{R}(\boldsymbol{S})) \\ \boldsymbol{\varphi}(\boldsymbol{\theta}) \end{bmatrix} = \boldsymbol{0}. \tag{10}$$

Here $\boldsymbol{R}(\boldsymbol{S})$ is an operator defined in Equation (1), which maps a positive definite matrix into a standardized matrix. These estimating equations are a direct adaptation of the estimating equations used in Jennrich (2008, Equation (6)) for covariance structure analysis. The adaptation takes into account two features of EFA of a sample correlation matrix. In the covariance structure analysis of Jennrich (2008) $\boldsymbol{\Sigma}(\boldsymbol{\theta})$ is assumed to be identified, while here it is not true for EFA in general. The $\boldsymbol{\varphi}(\boldsymbol{\theta})$ in Equation (10) is used to identify the EFA mode. Thus, the dimension of $g(\boldsymbol{\theta}, \boldsymbol{S})$ increases to $q^* = q + m(m-1)$. Another difference is the use of $\boldsymbol{R}$ rather than $\boldsymbol{S}$ in the definition of $F(\boldsymbol{\theta}, \boldsymbol{S})$. This makes the derivatives of $F(\boldsymbol{\theta}, \boldsymbol{S})$ with respect to the components of $\boldsymbol{S}$ somewhat more difficult. Computing IJ standard errors for EFA proceeds as follows.

Let

$$\boldsymbol{J}(\boldsymbol{\theta}, \boldsymbol{S}) = \frac{\partial}{\partial \boldsymbol{\theta}'} g(\boldsymbol{\theta}, \boldsymbol{S}) = \begin{bmatrix} \frac{\partial^2}{\partial \boldsymbol{\theta}' \partial \boldsymbol{\theta}} F(\boldsymbol{\theta}, \boldsymbol{R}(\boldsymbol{S})) \\ \frac{\partial}{\partial \boldsymbol{\theta}'} \boldsymbol{\varphi}(\boldsymbol{\theta}) \end{bmatrix} \tag{11}$$

be the partial Jacobian matrix of $g(\boldsymbol{\theta}, \boldsymbol{S})$ with respect to its first argument $\boldsymbol{\theta}$, and let

$$\partial_2 g_{(\boldsymbol{\theta}, \boldsymbol{S})}(d\boldsymbol{S}) \tag{12}$$

be the partial differential of $g(\boldsymbol{\theta}, \boldsymbol{S})$ at $(\boldsymbol{\theta}, \boldsymbol{S})$ and $d\boldsymbol{S}$. For each of the observations $\boldsymbol{x}_1, \ldots, \boldsymbol{x}_i, \ldots, \boldsymbol{x}_n$, let

$$\boldsymbol{y}_i = \partial_2 g_{(\hat{\boldsymbol{\theta}}, \boldsymbol{S})}\big((\boldsymbol{x}_i - \bar{\boldsymbol{x}})(\boldsymbol{x}_i - \bar{\boldsymbol{x}})'\big) \tag{13}$$

be the partial differential of the estimating equation $g(\boldsymbol{\theta}, \boldsymbol{S})$ with respect to its second argument $\boldsymbol{S}$ evaluated at $(\hat{\boldsymbol{\theta}}, \boldsymbol{S})$ and $(\boldsymbol{x}_i - \bar{\boldsymbol{x}})(\boldsymbol{x}_i - \bar{\boldsymbol{x}})'$. Because the last $m^2 - m$ rows of $g(\boldsymbol{\theta}, \boldsymbol{S})$ are zero for all applications, the last $m^2 - m$ components of $\boldsymbol{y}_i$ are always zero. The only components of $\boldsymbol{y}_i$ that change are the first q components.

Let $\boldsymbol{\theta}_i^+$ be the solution to the linear equation (Jennrich, 2008, Equation (9))

$$\boldsymbol{J}(\hat{\boldsymbol{\theta}}, \boldsymbol{S})\boldsymbol{\theta}_i^+ = -\boldsymbol{y}_i. \tag{14}$$

Let $\boldsymbol{\theta}_\cdot^+$ is the mean of all $\boldsymbol{\theta}_i^+$. From Jennrich (2008)

$$\mathrm{acov}^{\mathrm{IJ}}(\hat{\boldsymbol{\theta}}) = \frac{1}{n} \sum_{i=1}^{n} \big(\boldsymbol{\theta}_i^+ - \boldsymbol{\theta}_\cdot^+\big)\big(\boldsymbol{\theta}_i^+ - \boldsymbol{\theta}_\cdot^+\big)'$$

is the IJ estimate of the asymptotic covariance matrix for the estimator $\hat{\boldsymbol{\theta}}$, and

$$\mathrm{se}^{\mathrm{IJ}}(\hat{\theta}_r) = \big(\mathrm{acov}^{\mathrm{IJ}}(\hat{\boldsymbol{\theta}})\big)_{rr} / \sqrt{n} \tag{15}$$

is the IJ standard error estimate for $\hat{\theta}_r$.

The partial Jacobian matrix $\boldsymbol{J}(\boldsymbol{\theta}, \boldsymbol{S})$ involves partial derivatives of rotation constraints of $\boldsymbol{\varphi}(\boldsymbol{\theta})$ with respect to rotated factor loadings and factor correlations. When CF-varimax rotation is considered, the partial derivatives of rotation constraints with respect to rotated factor loadings (Jennrich, 1973, Equation (53)) are

$$\frac{\partial \varphi_{uv}}{\partial \lambda_{ir}} = \delta_{ur}\left(\frac{dQ}{d\boldsymbol{\Lambda}}\boldsymbol{\Phi}^{-1}\right)_{iv} + 4M_{ir}\lambda_{iu}\phi^{rv} + \left(8 - \frac{8}{p}\right)\lambda_{ir}\lambda_{iu}\left(\boldsymbol{\Lambda}\boldsymbol{\Phi}^{-1}\right)_{iv} + \frac{8}{p}\lambda_{ir}\left(\boldsymbol{\Lambda}'\boldsymbol{\Lambda}\right)_{ur}\phi^{rv} - 8\lambda_{ir}^2\lambda_{iu}\phi^{rv}. \tag{16}$$

Here $1 \le u \ne v \le m$, $1 \le i \le p$, $1 \le r \le m$, $\delta_{ur} = 1$ if $u = r$ and 0 otherwise, M_{ir} is given in Equation (6), and $\phi^{rv} = (\boldsymbol{\Phi}^{-1})_{rv}$. The corresponding partial derivatives of rotation constraints with respect to rotated factor correlations (Jennrich, 1973, Equation (54)) are

$$\frac{\partial \varphi_{uv}}{\partial \phi_{xy}} = -\left(\delta_{ux}\phi^{yv} + \delta_{uy}\phi^{xv}\right)\left(\boldsymbol{\Lambda}'\frac{dQ}{d\boldsymbol{\Lambda}}\boldsymbol{\Phi}^{-1}\right)_{uu}. \tag{17}$$

Here $1 \le u \ne v \le m$ and $1 \le x \ne y \le m$. The original formula presented in Jennrich (1973, Equation (54)) contains a typo (Jennrich, Personal Communication) that is corrected in Equation (17). A proof of the formula is given in Appendix A.

Analytic derivatives of rotation constraints derived in Jennrich (1973) were for the general symmetric family of quartic criteria (Clarkson & Jennrich, 1988, Table 1), which include the Crawford–Ferguson criteria as their special case. Additional results on derivatives of rotation constraints can be found in Tateneni (1998), Ogasawara (1998), Yung and Hayashi (2001), and Hayashi and Yung (1999).

We next discuss how to compute the Jacobian matrix $\boldsymbol{J}$ and the vector of differentials $\boldsymbol{y}_i$ for OLS estimation and ML estimation. The computations involve first- and second-order derivatives of the EFA covariance structure with respect to model parameters $\frac{\partial \boldsymbol{\Sigma}}{\partial \theta_a}$ and $\frac{\partial^2 \boldsymbol{\Sigma}}{\partial \theta_a \partial \theta_b}$. Here θ_a and θ_b are arbitrary model parameters, for example, rotated factor loadings, rotated factor correlations, or unique variances.

3.2. *OLS as a Minimum Deviance Method*

When the minimum deviance method is chosen to be the OLS discrepancy function of Equation (9), an element of the gradient vector in the estimating equations in Equation (10) is given by

$$\frac{\partial F_{\text{OLS}}}{\partial \theta_a} = -2\,\text{tr}\left(\frac{\partial \boldsymbol{\Sigma}}{\partial \theta_a}(\boldsymbol{R} - \boldsymbol{\Sigma})\right). \tag{18}$$

The partial Jacobian matrix $\boldsymbol{J}(\boldsymbol{\theta}, \boldsymbol{S})$ in Equation (11) includes the second-order derivatives of the OLS discrepancy function. Differentiating the gradient vector in Equation (18) provides such second-order derivatives. An example of these second-order derivatives is

$$\frac{\partial^2 F_{\text{OLS}}}{\partial \theta_a \partial \theta_b} = -2\,\text{tr}\left(\frac{\partial^2 \boldsymbol{\Sigma}}{\partial \theta_a \partial \theta_b}(\boldsymbol{R} - \boldsymbol{\Sigma}) - \frac{\partial \boldsymbol{\Sigma}}{\partial \theta_a}\frac{\partial \boldsymbol{\Sigma}}{\partial \theta_b}\right). \tag{19}$$

Let $\partial_2 g^{(a)}_{(\boldsymbol{\theta},\boldsymbol{S})}(d\boldsymbol{S})$ denotes the ath component of the differentials $\partial_2 g_{(\boldsymbol{\theta},\boldsymbol{S})}(d\boldsymbol{S})$ in Equation (13) for OLS estimation. For $a = 1, \ldots, q$,

$$\partial_2 g^{(a)}_{(\boldsymbol{\theta},\boldsymbol{S})}(d\boldsymbol{S}) = -2\,\text{tr}\left(\frac{\partial \boldsymbol{\Sigma}}{\partial \theta_a} d\boldsymbol{R}_{\boldsymbol{S}}(d\boldsymbol{S})\right) \tag{20}$$

where $d\boldsymbol{R}_{\boldsymbol{S}}(d\boldsymbol{S})$ is the differential of the operator $\boldsymbol{R}(\boldsymbol{S})$ at $\boldsymbol{S}$ and $d\boldsymbol{S}$. For $a = q + 1, \ldots, q^*$,

$$\partial_2 g^{(a)}_{(\boldsymbol{\theta},\boldsymbol{S})}(d\boldsymbol{S}) = 0.$$

A formula for $d\boldsymbol{R}_{\boldsymbol{S}}(d\boldsymbol{S})$ is derived in Appendix B.

3.3. *ML as a Minimum Deviance Method*

When the minimum deviance method is chosen to be the ML discrepancy function of Equation (8), an element of the gradient vector in the estimating equations is given by

$$\frac{\partial F_{\mathrm{ML}}}{\partial \theta_a} = -\operatorname{tr}\left(\boldsymbol{\Sigma}^{-1}(\boldsymbol{R} - \boldsymbol{\Sigma})\boldsymbol{\Sigma}^{-1}\frac{\partial \boldsymbol{\Sigma}}{\partial \theta_a}\right). \tag{21}$$

The partial Jacobian matrix $\boldsymbol{J}(\boldsymbol{\theta}, \boldsymbol{S})$ in Equation (11) includes the second-order derivatives of the ML discrepancy function. Differentiating the gradient vector in Equation (21) provides such second-order derivatives. These second-order derivatives are

$$\begin{aligned}\frac{\partial^2 F_{\mathrm{ML}}}{\partial \theta_a \partial \theta_b} &= 2\operatorname{tr}\left(\boldsymbol{\Sigma}^{-1}\frac{\partial \boldsymbol{\Sigma}}{\partial \theta_b}\boldsymbol{\Sigma}^{-1}(\boldsymbol{R} - \boldsymbol{\Sigma})\boldsymbol{\Sigma}^{-1}\frac{\partial \boldsymbol{\Sigma}}{\partial \theta_a}\right)\\ &\quad + \operatorname{tr}\left(\boldsymbol{\Sigma}^{-1}\frac{\partial \boldsymbol{\Sigma}}{\partial \theta_b}\boldsymbol{\Sigma}^{-1}\frac{\partial \boldsymbol{\Sigma}}{\partial \theta_a}\right)\\ &\quad - \operatorname{tr}\left(\boldsymbol{\Sigma}^{-1}(\boldsymbol{R} - \boldsymbol{\Sigma})\boldsymbol{\Sigma}^{-1}\frac{\partial^2 \boldsymbol{\Sigma}}{\partial \theta_a \partial \theta_b}\right).\end{aligned} \tag{22}$$

Let $\partial_2 g^{(a)}_{(\boldsymbol{\theta},\boldsymbol{S})}(d\boldsymbol{S})$ denotes the ath component of the differentials $\partial_2 g_{(\boldsymbol{\theta},\boldsymbol{S})}(d\boldsymbol{S})$ in Equation (13) for ML estimation. For $a = 1, \ldots, q$,

$$\partial_2 g^{(a)}_{(\boldsymbol{\theta},\boldsymbol{S})}(d\boldsymbol{S}) = -\operatorname{tr}\left(\boldsymbol{\Sigma}^{-1}\frac{\partial \boldsymbol{\Sigma}}{\partial \theta_a}\boldsymbol{\Sigma}^{-1} d\boldsymbol{R}_{\boldsymbol{S}}(d\boldsymbol{S})\right), \tag{23}$$

where $d\boldsymbol{R}_{\boldsymbol{S}}(d\boldsymbol{S})$ is the differential of the operator $\boldsymbol{R}(\boldsymbol{S})$ at $\boldsymbol{S}$ and $d\boldsymbol{S}$. For $a = q + 1, \ldots, q^*$,

$$\partial_2 g^{(a)}_{(\boldsymbol{\theta},\boldsymbol{S})}(d\boldsymbol{S}) = 0.$$

Derivations of the differential for ML estimation are provided in Appendix B.

4. Illustrative Examples

The illustrative examples are divided into two parts. The first part compares IJ standard error estimates using a nearly correctly specified EFA model and sampling from a normal and from an elliptical distribution. The second part compares IJ and analytic standard error estimates when the EFA model is misspecified.

4.1. *IJ Standard Error Estimates of ML and OLS Estimates with Normal and Nonnormal Data*

Simulation samples were used in the illustrations. The population covariance matrix was chosen to be the sample correlation matrix of Holzinger's unpublished data (Harman, 1960, p. 227). A three-factor EFA model is satisfactory for this 9×9 correlation matrix. A 90 % confidence interval for the root mean square error of approximation (the RMSEA, Browne & Cudeck, 1993; Steiger & Lind, 1980, June) produced by OLS estimation is (0.000, 0.036); a 90 % confidence interval for the RMSEA produced by ML estimation is (0.000, 0.037). The analyses were carried out using EFA software (Browne, Cudeck, Tateneni, & Mels, 2008). Table 1 presents OLS rotated factor loadings and factor correlations. The rotation criterion is oblique CF-varimax. ML rotated factor loadings and factor correlations are similar to those reported in Table 1. The original sample size was 696.

Two simulation samples were generated from the correlation matrix of Holzinger's unpublished data. Let $\boldsymbol{R}_H$ be this correlation matrix. In the first sample, 696 nine-variate vectors were

TABLE 1.
OLS rotated factor loadings and factor correlations from Holzinger's unpublished study.

	Factor loadings		
	Verbal	Arithmetic	Spatial
Word meaning	0.90	0.00	0.02
Sentence completion	0.73	0.18	0.02
Odd words	0.79	0.05	0.13
Mixed arithmetic	0.01	0.95	0.00
Remainders	0.08	0.77	0.11
Missing numbers	0.18	0.72	0.11
Gloves	−0.05	0.17	0.54
Boots	0.06	0.04	0.72
Hatchets	0.02	−0.03	0.89

	Factor correlations		
	Verbal	Arithmetic	Spatial
Verbal	1.00		
Arithmetic	0.48	1.00	
Spatial	0.34	0.37	1.00

simulated from a multivariate normal distribution $N(\mathbf{0}, \boldsymbol{R}_H)$; in the second sample, 696 nine-variate vectors were simulated from an elliptical distribution, which is a mixture of two multivariate normal distributions (Ichikawa & Konishi, 1995),

$$0.7N(\mathbf{0}, \boldsymbol{R}_H) + 0.3N(\mathbf{0}, 3\boldsymbol{R}_H).$$

The kurtosis of the elliptical distribution is 1.328 times larger than that of the multivariate normal distribution.

In both the normal sample and the elliptical sample, the three-factor EFA model was estimated with a correlation matrix using both OLS and ML. Model estimation, factor rotation, and analytic standard error estimation for OLS and ML estimates were carried out using the computer program CEFA (Browne et al., 2008). IJ standard error estimation was implemented using the software package R (R Development Core Team, 2007).[2]

Analytic standard error estimates for ML rotated factor loadings and factor correlations implemented in CEFA are an extension of the bordered information matrix method (Jennrich, 1974). Analytic standard error estimates for OLS rotated factor loadings and factor correlations implemented in CEFA are an extension of a sandwich standard error estimator (Browne, 1984, Proposition 4).

Figure 1 displays comparisons between IJ standard error estimates and analytic standard error estimates for rotated factor loadings and factor correlations. When manifest variables are normally distributed, IJ standard error estimates and analytic standard error estimates are expected to be close. Thus the scatter plot displaying these two kinds of standard error estimates should be close to the line $y = x$. The upper left plot of Figure 1 shows that the IJ standard error estimates and analytic standard error estimates are close for ML rotated factor loadings and factor correlations. The lower left plot of Figure 1 shows the same pattern for OLS parameter estimates.

When manifest variables are elliptically distributed, normal theory standard error estimates will be inappropriate. This is shown by the right two plots of Figure 1. All the points are above

[2]R code implementing the IJ standard error estimation with EFA is available upon request.

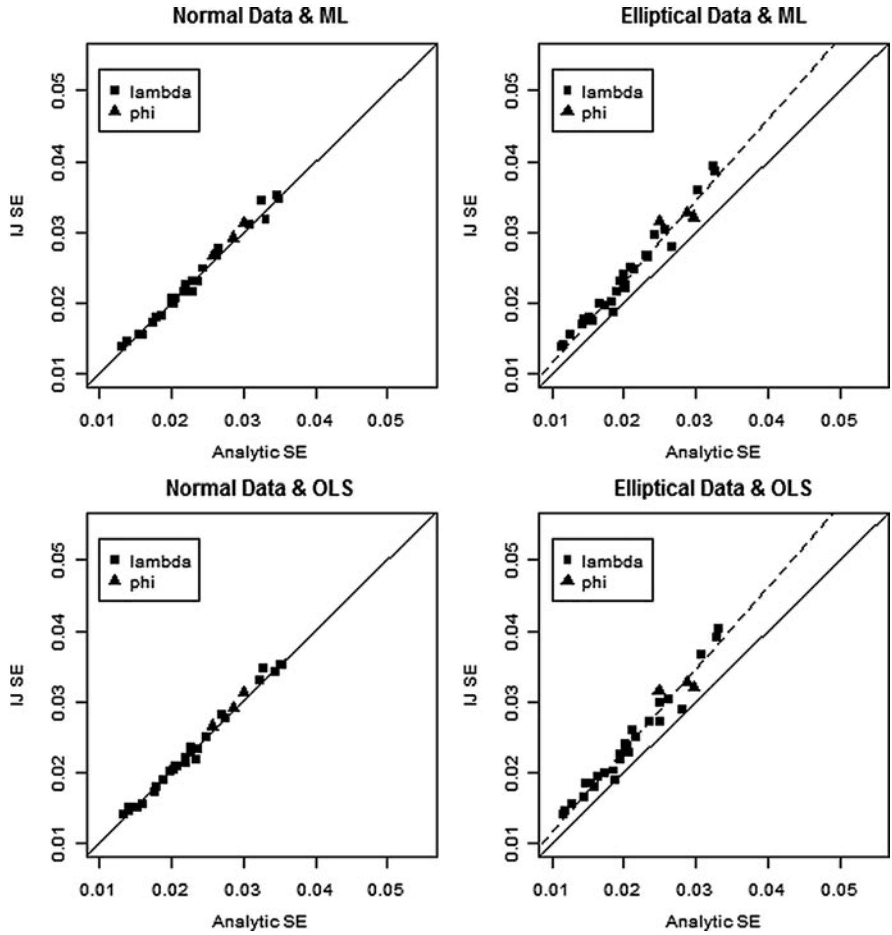

FIGURE 1.
Comparisons of IJ SE estimates and analytic SE estimates. "IJ SE", "Analytic SE", "OLS", and "ML" represent infinitesimal jackknife standard error estimates, analytic standard error estimates, ordinary least squares, and maximum likelihood, respectively.

the line $y = x$: IJ standard error estimates are larger than the corresponding analytic standard error estimates. Browne (1984, Proposition 5) proposed that appropriate standard errors for ML estimates of SEM models with elliptical data could be obtained from scaling standard errors for ML estimates of SEM models with normal data. This scaling factor is the ratio between the kurtosis of the elliptical distribution and that of normal distributions. The scaling factor is 1.328 for the elliptical distribution in the current illustration. Thus analytic standard error estimates for ML estimation with the elliptical distribution need to be multiplied by $\sqrt{1.328} = 1.152$. The upper right plot of Figure 1 shows that the IJ standard error estimates and scaled analytic standard error estimates are close for ML rotated factor loadings and factor correlations: all the points are close to the dash line $y = 1.152x$. We applied the same scaling method to analytic standard error estimates for OLS estimates with the elliptical distribution. As shown at the lower right plot of Figure 1, IJ standard error estimates and scaled analytic standard error estimates are close for OLS rotated factor loadings and factor correlations.

4.2. *IJ Standard Error Estimates for Misspecified EFA Models*

Luo, Chen, Yue, Zhang, Zhaoyang, and Xu (2008) reported a study on marital satisfaction. The participants of the study were 537 newlywed urban Chinese couples. We compare IJ SE estimates and normal theory ML standard error estimates using husbands' ratings on 28 personality facet scores. Table 2 presents ML rotated factor loadings and factor correlations of a four-factor EFA model. The factor rotation procedure is oblique CF-varimax rotation. The RMSEA for the four-factor model is 0.049 with a 90 % CI of (0.044, 0.054). The test of perfect fit is rejected at the 0.05 level; the test of close fit is retained at the 0.05 level (Browne & Cudeck, 1993). Although the amount of model misspecification is nonignorable, the four-factor EFA model still provides a useful approximation to the manifest variable correlation matrix. The sample multivariate kurtosis is 865.37; the corresponding multivariate kurtosis for the normal distribution is 840 (Mardia, 1970). Thus the sample multivariate kurtosis is close to the corresponding normal one.

Also presented in Table 2 are IJ standard error estimates and normal theory ML standard error estimates. Most IJ standard error estimates are larger than the corresponding normal theory ML standard error estimates. No clear patterns on the relative magnitude of these two types of standard error estimates are detected, however. For example, rotated factor loadings of the manifest variable "'veraciousness versus slickness" on "social potency", "dependability", "accommodation", and "relatedness" are −0.06, 0.18, 0.54, and 0.27, respectively. IJ standard error estimates and normal theory ML standard error estimates are very similar for the factor loadings on "social potency" (0.05 versus 0.04) and "relatedness" (0.06 versus 0.06). IJ standard error estimates are almost twice as large as normal theory ML standard error estimates for the factor loadings on "dependability" (0.09 versus 0.05) and "accommodation" (0.10 versus 0.05).

Normal theory ML standard error estimates are appropriate for normal data and correctly specified models. These two assumptions are often violated in practice. On the other hand, IJ standard error estimates are consistent regardless of data distributions and model misspecification. As indicated by the confidence interval for the RMSEA, a nonignorable amount of model error is present in the empirical illustration. Although IJ standard error estimates and normal theory ML standard error estimates are close in many places, these two types of standard error estimates can substantially differ from each other. IJ standard error estimates can be twice as large as their normal theory ML counterparts.

5. Concluding Comments

We adapted the IJ method in covariance structure analysis to estimate standard errors in EFA with minimum deviance methods and sample correlation matrices. Factor rotation was treated as the imposition of constraints on rotated factor loadings and factor correlations. Two frequently used estimation methods, ML and OLS, were presented as examples of minimum deviance methods. The empirical illustrations showed that IJ standard error estimates were close to analytic standard error estimates for both ML and OLS rotated factor loadings and factor correlations for normal data and IJ standard error estimates were close to scaled analytic standard error estimates for both ML and OLS rotated factor loadings and factor correlations for elliptical data. Additionally, IJ standard error estimates and analytic standard error estimates for some factor loadings were different when the EFA model was misspecified.

The consistency of IJ standard error estimates in covariance structure analysis was proved in Jennrich (2008, Theorem 1). We have used this to obtain consistent standard error estimates for an EFA of a correlation matrix. We have not assumed that the EFA model was correctly specified, so our results apply to both correctly and incorrectly specified models. Covariance structure analysis methods described in Shapiro (1983) and Ogasawara (2007b) are also appropriate for

TABLE 2.
Rotated factor loadings and factor correlations of Luo et al.'s (2008) survey data.

	Factor loadings			
	Social potency	Dependability	Accommodation	Relatedness
Novelty	0.65(0.05, 0.04)	0.05(0.07, 0.05)	0.08(0.05, 0.04)	0.12(0.12, 0.08)
Diversity	0.49(0.12, 0.08)	−0.15(0.06, 0.04)	0.23(0.06, 0.05)	0.40(0.15, 0.10)
Diverse thinking	0.44(0.08, 0.06)	−0.03(0.06, 0.05)	−0.04(0.06, 0.05)	0.27(0.09, 0.07)
Leadership	0.65(0.04, 0.04)	0.03(0.07, 0.05)	−0.35(0.05, 0.04)	0.04(0.08, 0.07)
Logical versus affective	0.33(0.09, 0.06)	0.12(0.06, 0.05)	0.06(0.07, 0.05)	0.39(0.08, 0.06)
Aesthetics	0.35(0.10, 0.07)	−0.25(0.08, 0.06)	−0.11(0.08, 0.05)	0.23(0.09, 0.08)
Extroversion–introversion	0.58(0.05, 0.04)	0.03(0.11, 0.06)	0.00(0.07, 0.05)	−0.09(0.11, 0.07)
Enterprise	0.53(0.09, 0.07)	0.45(0.09, 0.08)	0.11(0.05, 0.05)	−0.22(0.05, 0.05)
Responsibility	0.06(0.05, 0.04)	0.57(0.08, 0.05)	−0.09(0.07, 0.05)	0.23(0.09, 0.06)
Emotionality	0.01(0.08, 0.06)	−0.75(0.05, 0.04)	0.00(0.05, 0.04)	0.02(0.07, 0.05)
Inferiority versus self-acceptance	−0.21(0.04, 0.04)	−0.51(0.05, 0.04)	−0.37(0.04, 0.04)	−0.11(0.07, 0.04)
Practical mindedness	−0.06(0.04, 0.04)	0.56(0.08, 0.05)	0.06(0.09, 0.05)	0.29(0.09, 0.06)
Optimism versus pessimism	0.21(0.07, 0.06)	0.59(0.07, 0.05)	0.14(0.05, 0.04)	−0.01(0.05, 0.04)
Meticulousness	−0.11(0.06, 0.05)	0.40(0.10, 0.07)	−0.11(0.09, 0.06)	0.30(0.10, 0.07)
Face	0.06(0.07, 0.06)	−0.26(0.08, 0.06)	−0.29(0.06, 0.05)	0.14(0.07, 0.06)
Internal versus external control	0.11(0.06, 0.05)	0.19(0.07, 0.06)	0.42(0.05, 0.05)	−0.13(0.07, 0.06)
Family orientation	0.07(0.04, 0.04)	0.39(0.07, 0.05)	0.25(0.07, 0.05)	0.28(0.06, 0.05)
Defensiveness	0.07(0.03, 0.03)	−0.19(0.05, 0.04)	−0.73(0.04, 0.03)	−0.07(0.08, 0.06)
Graciousness versus meanness	0.00(0.05, 0.04)	0.39(0.06, 0.04)	0.54(0.06, 0.04)	0.05(0.06, 0.05)
Interpersonal tolerance	0.22(0.05, 0.04)	0.16(0.05, 0.05)	0.53(0.04, 0.04)	0.10(0.10, 0.06)
Self versus social orientation	0.08(0.05, 0.04)	−0.06(0.08, 0.05)	−0.63(0.06, 0.04)	0.10(0.10, 0.06)
Veraciousness versus slickness	−0.06(0.05, 0.04)	0.18(0.09, 0.05)	0.54(0.10, 0.05)	0.27(0.06, 0.06)
Traditionalism versus modernity	−0.06(0.06, 0.04)	0.09(0.07, 0.06)	−0.52(0.06, 0.05)	0.10(0.12, 0.07)

TABLE 2.
(*Continued*)

	Factor loadings			
	Social potency	Dependability	Accommodation	Relatedness
Relationship orientation	0.13(0.10, 0.08)	0.05(0.06, 0.05)	−0.02(0.09, 0.06)	0.64(0.06, 0.04)
Social sensitivity	0.36(0.11, 0.09)	−0.04(0.05, 0.04)	−0.15(0.09, 0.06)	0.56(0.08, 0.07)
Discipline	0.00(0.06, 0.05)	0.18(0.08, 0.06)	−0.58(0.11, 0.07)	0.47(0.12, 0.08)
Harmony	−0.02(0.08, 0.06)	0.24(0.06, 0.06)	0.31(0.09, 0.07)	0.57(0.07, 0.05)
Thrift versus extravagance	−0.22(0.08, 0.07)	0.00(0.11, 0.07)	−0.15(0.11, 0.07)	0.45(0.09, 0.06)

	Factor correlations			
	Social potency	Dependability	Accommodation	Relatedness
Dependability	0.22(0.05, 0.04)			
Accommodation	0.09(0.06, 0.05)	0.45(0.06, 0.04)		
Relatedness	0.29(0.04, 0.03)	0.31(0.08, 0.05)	−0.01(0.04, 0.04)	

Note. Point estimates are followed, in parentheses, by IJ standard error estimates and ML standard error estimates, respectively.

misspecified models, but their foci are not on standard errors for EFA parameters. Consistent analytic standard error estimates for an arbitrary minimum deviance method in EFA with nonnormal data and model misspecification are currently unavailable. Bootstrap standard error estimation does not require a correctly specified model either. A systematic investigation of IJ standard error estimates, bootstrap standard error estimates, and analytic standard error estimates under different levels of model misspecification will be highly informative but beyond the scope of the present article.

Although analytic derivatives are used in the present article, implementation of the IJ procedure can be substantially simplified if derivatives are approximated numerically. Complex computer code evaluating analytic derivatives is then replaced by simple generic code computing numerical derivatives. Efforts on deriving and programming analytic derivatives can be avoided. Numerical derivatives involve heavier computation costs, but they have been found to be feasible and accurate in the context of factor analysis. Using a confirmatory factor analysis model involving nine manifest variables and three factors, Jennrich (2008) demonstrated that IJ standard error estimates with analytic derivatives and IJ standard error estimates with numerical derivatives agreed to at least two decimal places. Cudeck and O'Dell (1994) employed numerical derivatives of rotated factor loadings and factor correlations with regard to unrotated factor loadings when they estimated standard errors in EFA using the delta method. Tateneni (1998) compared different numerical derivatives in EFA. The computer program CEFA (Browne et al., 2008) allows users to specify whether analytic derivatives and numerical derivatives are used to compute standard errors for rotated factor loadings and factor correlations.

An advantage of the IJ method is its generality. EFA consists of two major steps: factor extraction and factor rotation. We provide details of estimating standard errors for CF-varimax obliquely rotated factor loadings and factor correlations when factors are extracted from the sample correlation matrix using OLS and ML. Adaptations of the IJ method to other factor extraction and factor rotation conditions are straightforward. When factors are extracted from the sample covariance matrix, the only required changes are to replace $d\boldsymbol{R}_S(d\boldsymbol{S})$ in Equations (20) and (23) with $d\boldsymbol{S}$. When other factor extraction methods are considered, the only required changes are the first-order derivatives $\frac{\partial}{\partial\boldsymbol{\theta}}F(\boldsymbol{\theta},\boldsymbol{R})$ in Equation (10) and the second-order derivatives $\frac{\partial^2}{\partial\boldsymbol{\theta}'\partial\boldsymbol{\theta}}F(\boldsymbol{\theta},\boldsymbol{R})$ in Equation (11). When other rotation criteria or orthogonal rotation are considered, the only required changes are the rotation constraints $\boldsymbol{\varphi}(\boldsymbol{\theta})$ in Equation (10) and their derivatives $\frac{\partial}{\partial\boldsymbol{\theta}'}\boldsymbol{\varphi}(\boldsymbol{\theta})$ in Equation (11). Whenever analytic derivatives are unavailable, numerical derivatives can be used to simplify the implementation of the IJ method.

Acknowledgements

We are grateful to Michael W. Browne for numerous helpful conversations.

Appendix A. Partial Derivatives of Rotation Constraints with Respect to Factor Correlations

Let $\boldsymbol{J}_{xy}$ be an $m \times m$ matrix with a one in row x and column y and zeros elsewhere. Then

$$\begin{aligned}\frac{\partial\varphi_{uv}}{\partial\phi_{xy}} &= \frac{\partial}{\partial\phi_{xy}}\left(\boldsymbol{\Lambda}\frac{dQ}{d\boldsymbol{\Lambda}}\boldsymbol{\Phi}^{-1}\right)_{uv}\\ &= -\left(\boldsymbol{\Lambda}\frac{dQ}{d\boldsymbol{\Lambda}}\boldsymbol{\Phi}^{-1}(\boldsymbol{J}_{xy}+\boldsymbol{J}_{yx})\boldsymbol{\Phi}^{-1}\right)_{uv}\\ &= -\left(\boldsymbol{\Lambda}\frac{dQ}{d\boldsymbol{\Lambda}}\boldsymbol{\Phi}^{-1}\boldsymbol{J}_{xy}\boldsymbol{\Phi}^{-1}\right)_{uv} - \left(\boldsymbol{\Lambda}\frac{dQ}{d\boldsymbol{\Lambda}}\boldsymbol{\Phi}^{-1}\boldsymbol{J}_{yx}\boldsymbol{\Phi}^{-1}\right)_{uv}\end{aligned}$$

$$= -\left(\boldsymbol{\Lambda}\frac{dQ}{d\boldsymbol{\Lambda}}\boldsymbol{\Phi}^{-1}\right)_{ux}\phi^{yv} - \left(\boldsymbol{\Lambda}\frac{dQ}{d\boldsymbol{\Lambda}}\boldsymbol{\Phi}^{-1}\right)_{uy}\phi^{xv}$$

$$= -\delta_{ux}\left(\boldsymbol{\Lambda}\frac{dQ}{d\boldsymbol{\Lambda}}\boldsymbol{\Phi}^{-1}\right)_{uu}\phi^{yv} - \delta_{uy}\left(\boldsymbol{\Lambda}\frac{dQ}{d\boldsymbol{\Lambda}}\boldsymbol{\Phi}^{-1}\right)_{uu}\phi^{xv}$$

$$= -\left(\delta_{ux}\phi^{yv} + \delta_{uy}\phi^{xv}\right)\left(\boldsymbol{\Lambda}\frac{dQ}{d\boldsymbol{\Lambda}}\boldsymbol{\Phi}^{-1}\right)_{uu}.$$

The proof repeatedly uses the property that $(\boldsymbol{\Lambda}\frac{dQ}{d\boldsymbol{\Lambda}}\boldsymbol{\Phi}^{-1})$ is a diagonal matrix.

Appendix B. Computing the y_i Using Differentials

This appendix consists of four lemmas (B1, B2, B3, and B4). B1 expresses the differential of the sample correlation matrix as a function of the differential of the sample covariance matrix; B2 derives the partial differential of the OLS estimating equation with respect to the differential of the sample correlation matrix; B3 derives the partial differential of the ML estimating equation with respect to the differential of the sample correlation matrix; B4 computes the partial differentials of estimating equations for the ith observation.

The sample correlation matrix $\boldsymbol{R}$ can be obtained from the sample covariance matrix $\boldsymbol{S}$ using

$$\boldsymbol{R}(\boldsymbol{S}) = \boldsymbol{D}_s^{-\frac{1}{2}}\boldsymbol{S}\boldsymbol{D}_s^{-\frac{1}{2}}.$$

Here $\boldsymbol{D}_s$ is a diagonal matrix formed by extracting diagonal elements of $\boldsymbol{S}$.

Lemma B1. *The differential of the correlation matrix is*

$$d\boldsymbol{R}_S(d\boldsymbol{S}) = -\frac{1}{2}\operatorname{diag}(d\boldsymbol{S})\boldsymbol{D}_s^{-1}\boldsymbol{R} + \boldsymbol{D}_s^{-\frac{1}{2}}(d\boldsymbol{S})\boldsymbol{D}_s^{-\frac{1}{2}} - \frac{1}{2}\boldsymbol{R}\boldsymbol{D}_s^{-1}\operatorname{diag}(d\boldsymbol{S}).$$

Proof:

$$\begin{aligned}
d\boldsymbol{R}_S(d\boldsymbol{S}) &= -\frac{1}{2}\operatorname{diag}(\boldsymbol{S})^{-\frac{3}{2}}\operatorname{diag}(d\boldsymbol{S})\boldsymbol{S}\operatorname{diag}(\boldsymbol{S})^{-\frac{1}{2}} \\
&\quad + \operatorname{diag}(\boldsymbol{S})^{-\frac{1}{2}}(d\boldsymbol{S})\operatorname{diag}(\boldsymbol{S})^{-\frac{1}{2}} \\
&\quad - \frac{1}{2}\operatorname{diag}(\boldsymbol{S})^{-\frac{1}{2}}\boldsymbol{S}\operatorname{diag}(\boldsymbol{S})^{-\frac{3}{2}}\operatorname{diag}(d\boldsymbol{S}) \\
&= -\frac{1}{2}\operatorname{diag}(d\boldsymbol{S})\operatorname{diag}(\boldsymbol{S})^{-1}\boldsymbol{R} \\
&\quad + \operatorname{diag}(\boldsymbol{S})^{-\frac{1}{2}}(d\boldsymbol{S})\operatorname{diag}(\boldsymbol{S})^{-\frac{1}{2}} \\
&\quad - \frac{1}{2}\boldsymbol{R}\operatorname{diag}(\boldsymbol{S})^{-1}\operatorname{diag}(d\boldsymbol{S}).
\end{aligned}$$

Using the definition of $\boldsymbol{D}_s$ completes the proof. □

Lemma B2. *Let g^a be the ath component of g. Then for OLS estimation and $a = 1, \dots, q$,*

$$\partial_2 g_{\boldsymbol{\theta},\boldsymbol{S}}^{(a)}(d\boldsymbol{S}) = -2\operatorname{tr}\left(\frac{\partial\boldsymbol{\Sigma}}{\partial\theta_a}d\boldsymbol{R}_{\boldsymbol{S}}(d\boldsymbol{S})\right).$$

Proof: For OLS

$$g^{(a)}_{\boldsymbol{\theta},\boldsymbol{S}}(d\boldsymbol{S}) = -2\,\mathrm{tr}\left(\frac{\partial\boldsymbol{\Sigma}}{\partial\theta_a}\big(\boldsymbol{R}(S)-\boldsymbol{\Sigma}\big)\right).$$

Thus

$$\partial_2 g^{(a)}_{\boldsymbol{\theta},\boldsymbol{S}}(d\boldsymbol{S}) = -2\,\mathrm{tr}\left(\frac{\partial\boldsymbol{\Sigma}}{\partial\theta_a}d\boldsymbol{R}_{\boldsymbol{S}}(d\boldsymbol{S})\right).$$ □

Lemma B3. *Let g^a be the ath component of g. Then for ML estimation and $a = 1, \ldots, q$*

$$\partial_2 g^{(a)}_{\boldsymbol{\theta},\boldsymbol{S}}(d\boldsymbol{S}) = -\,\mathrm{tr}\left(\boldsymbol{\Sigma}^{-1}\frac{\partial\boldsymbol{\Sigma}}{\partial\theta_a}\boldsymbol{\Sigma}^{-1}d\boldsymbol{R}_{\boldsymbol{S}}(d\boldsymbol{S})\right).$$

Proof: For ML,

$$\begin{aligned} g^{(a)}_{\boldsymbol{\theta},\boldsymbol{S}}(d\boldsymbol{S}) &= -\,\mathrm{tr}\left(\boldsymbol{\Sigma}^{-1}\big(\boldsymbol{R}(S)-\boldsymbol{\Sigma}\big)\boldsymbol{\Sigma}^{-1}\frac{\partial\boldsymbol{\Sigma}}{\partial\theta_a}\right) \\ &= -\,\mathrm{tr}\left(\boldsymbol{\Sigma}^{-1}\frac{\partial\boldsymbol{\Sigma}}{\partial\theta_a}\boldsymbol{\Sigma}^{-1}\big(\boldsymbol{R}(S)-\boldsymbol{\Sigma}\big)\right). \end{aligned}$$

Thus,

$$\partial_2 g^{(a)}_{\boldsymbol{\theta},\boldsymbol{S}}(d\boldsymbol{S}) = -\,\mathrm{tr}\left(\boldsymbol{\Sigma}^{-1}\frac{\partial\boldsymbol{\Sigma}}{\partial\theta_a}\boldsymbol{\Sigma}^{-1}d\boldsymbol{R}_{\boldsymbol{S}}(d\boldsymbol{S})\right).$$ □

Lemma B4. *Let y_{ia} be the ath component of $\boldsymbol{y}_i$. Then*

$$y_{ia} = \begin{cases} \partial_2 g^{(a)}_{\hat{\boldsymbol{\theta}},\boldsymbol{S}}((\boldsymbol{x}_i-\bar{\boldsymbol{x}})(\boldsymbol{x}_i-\bar{\boldsymbol{x}})') & \text{if } a = 1, \ldots, q, \\ 0 & \text{if } a = q+1, \ldots, q^*. \end{cases}$$

Proof: From the definition of $g(\boldsymbol{\theta}, \boldsymbol{S})$, $\partial_2 g^{(a)}_{\boldsymbol{\theta},\boldsymbol{S}}(d\boldsymbol{S}) = 0$ for all $a > q$. Thus $y_{ia} = 0$ for all $a > q$. □

References

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

Browne, M.W. (1984). Asymptotically distribution-free methods for the analysis of covariance structures. *British Journal of Mathematical & Statistical Psychology*, *37*, 62–83.

Browne, M.W. (2001). An overview of analytic rotation in exploratory factor analysis. *Multivariate Behavioral Research*, *36*, 111–150.

Browne, M.W., & Cudeck, R. (1993). Alternative ways of assessing model fit. In K.A. Bollen & J.S. Long (Eds.), *Testing structural equation models* (pp. 136–162). Newbury Park: Sage.

Browne, M.W., Cudeck, R., Tateneni, K., & Mels, G. (2008). CEFA: comprehensive exploratory factor analysis. Retrieved from http://faculty.psy.ohio-state.edu/browne/.

Browne, M.W., & Tateneni, K. (2008). Standard errors for OLS estimates in exploratory factor analysis. Annual meeting of the society of multivariate experimental psychology, Montreal, Canada.

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

Clarkson, D.B., & Jennrich, R.I. (1988). Quartic rotation criteria and algorithm. *Psychometrika*, *53*, 251–259.

Crawford, C.B., & 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 factor loadings and correlations. *Psychological Bulletin*, *115*, 475–487.

Harman, H.H. (1960). *Modern factor analysis*. Chicago: University of Chicago Press.

Hayashi, K., & Yung, Y.F. (1999). Standard errors for the class of orthomax-rotated factor loadings: some matrix results. *Psychometrika*, *64*, 451–460.

Ichikawa, M., & Konishi, S. (1995). Application of the bootstrap methods in factor analysis. *Psychometrika*, *60*, 77–93.

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 & Statistical Psychology*, *27*, 122–131.
Jennrich, R.I. (2008). Nonparametric estimation of standard errors in covariance analysis using the infinitesimal jackknife. *Psychometrika*, *73*, 579–594.
Jennrich, R.I., & Clarkson, D.B. (1980). A feasible method for standard errors of estimate in maximum likelihood factor analysis. *Psychometrika*, *45*, 237–247.
Lambert, Z.V., Wildt, A.R., & Durand, R.M. (1991). Approximating confidence intervals for factor loadings. *Multivariate Behavioral Research*, *26*, 421–434.
Luo, S., Chen, H., Yue, G., Zhang, G., Zhaoyang, R., & Xu, D. (2008). Predicting marital satisfaction from self, partner, and couple characteristics: is it me, you, or us? *Journal of Personality*, *76*, 1231–1266.
Mardia, K.V. (1970). Measures of multivariate skewness and kurtosis with applications. *Biometrika*, *57*, 519–530.
Ogasawara, H. (1998). Standard errors of several indices for unrotated and rotated factors. *Econ. Rev. Otaru Univ. Commer.*, *49*(1), 21–69.
Ogasawara, H. (2007a). Asymptotic expansion of the distributions of the estimators in factor analysis under nonnormality. *British Journal of Mathematical & Statistical Psychology*, *60*, 395–420.
Ogasawara, H. (2007b). Higher-order approximations to the distributions of fit indexes under fixed alternatives in structural equation models. *Psychometrika*, *72*, 227–243.
R Development Core Team (2007). *R: a language and environment for statistical computing [computer software manual]*. (ISBN 3-900051-07-0). Vienna, Austria. Available form http://www.R-project.org.
Satorra, A. (1989). Alternative test criteria in covariance structure analysis: a unified approach. *Psychometrika*, *54*, 131–151.
Shapiro, A. (1983). Asymptotic distribution theory in the analysis of covariance structures. *South African Statistical Journal*, *17*, 33–81.
Steiger, J.H., & Lind, J.C. (1980, June). Statistically based tests for the number of common factors. Paper presented at the annual meeting of the psychometric society, Iowa City, IA.
Tateneni, K. (1998). *Use of automatic and numerical differentiation in the estimation of asymptotic standard errors in exploratory factor analysis*. Doctoral dissertation, Ohio State University, Columbus, OH.
Yuan, K., Marshall, L.L., & Bentler, P.M. (2002). A unified approach to exploratory factor analysis with missing data, nonnormal data, and in the presence of outliers. *Psychometrika*, *67*, 95–122.
Yung, Y.F., & Hayashi, K. (2001). A computationally efficient method for obtaining standard error estimates for the promax and related solutions. *British Journal of Mathematical & Statistical Psychology*, *54*, 125–138.
Zhang, G., Preacher, K.J., & Luo, S. (2010). Bootstrap confidence intervals for ordinary least squares factor loadings and correlations in exploratory factor analysis. *Multivariate Behavioral Research*, *45*, 104–134.

Manuscript Received: 26 JUN 2011
Final Version Received: 14 JAN 2012
Published Online Date: 31 AUG 2012