

Article

Not peer-reviewed version

On Asymptotic Covariance of Unrotated Factor Solutions

[Xingwei Hu](#)*, [Caihong Hu](#), [Cheng-Kuang Wu](#)

Posted Date: 24 March 2026

doi: 10.20944/preprints202603.1901.v1

Keywords: factor analysis; asymptotic covariance; standard errors; factor loadings; uniqueness; multivariate implicit differentiation; structural VAR



Preprints.org is a free multidisciplinary platform providing preprint service that is dedicated to making early versions of research outputs permanently available and citable. Preprints posted at Preprints.org appear in Web of Science, Crossref, Google Scholar, Scilit, Europe PMC.

Copyright: This open access article is published under a [Creative Commons CC BY 4.0 license](#), which permit the free download, distribution, and reuse, provided that the author and preprint are cited in any reuse.

Disclaimer/Publisher's Note: The statements, opinions, and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions, or products referred to in the content.

Article

On Asymptotic Covariance of Unrotated Factor Solutions

Xingwei Hu ^{1,*}, Caihong Hu ² and Cheng-Kuang Wu ³

¹ International Monetary Fund, 700 19th St NW, Washington, DC 20431, USA
² Zhangjiajie College, Zhangjiajie, Hunan 427000; China
³ School of Artificial Intelligence, Zhejiang Industry and Trade Vocational College, Wenzhou, Zhejiang 325700, China
* Correspondence: xhu@imf.org; Tel.: +1-202-492-5271 (X.H.)

Abstract

This paper derives closed-form expressions for the asymptotic covariance matrices of unrotated factor loading and uniqueness estimators for several widely used non-maximum-likelihood factor extraction methods. These include least squares, principal factor, iterative principal component, alpha factor, and image factor analysis. By expressing these results explicitly in terms of the asymptotic covariance of the sample covariance or correlation matrix, the proposed formulas facilitate straightforward computation of standard errors. When combined with the delta method from rotation criteria, they further yield analytically tractable standard errors for rotated factor loadings. Monte Carlo simulations demonstrate accurate finite-sample performance, and an empirical application illustrates practical implementation of the proposed approach.

Keywords: factor analysis; asymptotic covariance; standard errors; factor loadings; uniqueness; multivariate implicit differentiation; structural VAR

MSC: 62H25; 62F12; 62H12

JEL Classification: C38; C13; C19

1. Introduction

Let $X_1, X_2, \dots, X_p$ be random variables, collectively denoted by the random vector Y . Common or exploratory factor analysis postulates the existence of k independent common factors, with $k < p$, that underlie the observed variables Y . Let these factors be jointly represented by the random vector F . The relationship between the observed variables and the common factors is specified by

$$Y = c + \Lambda F + \epsilon, \tag{1}$$

where c is a constant vector, Λ is a $p \times k$ factor-loading matrix, and ϵ is a vector of p mutually independent idiosyncratic disturbances.

Let Σ denote the variance-covariance (or correlation) matrix of Y , and assume that the common factors have unit variance. Under these assumptions, the covariance structure of Y is given by

$$\Sigma = \Lambda \Lambda^\top + \Psi, \tag{2}$$

where the superscript $\top$ denotes matrix transpose, and Ψ is a $p \times p$ diagonal matrix whose entries correspond to the variances of the idiosyncratic disturbances, often referred to as *uniquenesses*.

Eq. (2) implies that the off-diagonal covariances of Y are entirely determined by the factor loadings. In contrast, the diagonal variances reflect contributions from both the common and idiosyncratic components: the common factors capture the variation shared across variables Y , while the idiosyncratic

disturbances represent variable-specific noise. Let λ_{it} denote the element in the i -th row and t -th column of Λ . The quantity

$$h_i = \sum_{t=1}^k \lambda_{it}^2, \quad (3)$$

represents the proportion of the variance of X_i explained by the common factors and is known as the *communality* of variable X_i .

A wide range of factor-extraction methods has been proposed in the literature (e.g., Harman [1976] [7]; Lawley and Maxwell [1971] [15]; SAS [2015] [19]), including principal-factor, maximum-likelihood, and least-squares approaches. These methods aim to recover the factor model in Eq. (2) from a given sample covariance or correlation matrix $\hat{\Sigma}$. From a practical perspective, it is important not only to obtain point estimates of factor loadings and uniquenesses, but also to assess the precision of these estimates and to identify substantively meaningful—often rotated—factor structures. Such assessments require knowledge of the asymptotic covariance matrices (acov) of the estimators, as well as an understanding of how these covariances are affected by the chosen rotation criterion.

The asymptotic covariance of the estimators depends on both the factor-extraction method and the covariance of $\hat{\Sigma}$. Let n denote the sample size, and suppose that the estimator of the loading matrix can be expressed as $\Lambda = f(\Sigma)$ for a given extraction method f . Assume that f is differentiable at Σ , with differential df , which defines a linear mapping from the space of $p \times p$ symmetric matrices to the space of $p \times k$ matrices. By the delta method (see Rao [2009] [17]), the difference between $n^{1/2}d\Lambda$ and $df(n^{1/2}d\Sigma)$ converges to zero in probability as $n \rightarrow \infty$. In this setting, $d\Lambda \approx \hat{\Lambda} - \Lambda$ and $d\Sigma \approx \hat{\Sigma} - \Sigma$ denote deviations from the unobservable population parameters Σ and Λ , respectively.

Let $\Sigma = [\sigma_{ij}]$ and define $d\sigma_{ij} = \hat{\sigma}_{ij} - \sigma_{ij}$, where $\hat{\sigma}_{ij}$ denotes the sample covariance (or correlation) between X_i and X_j . The asymptotic covariance between elements of the estimated loading matrix $\hat{\Lambda}$ can then be expressed as

$$\text{acov}(\hat{\lambda}_{ir}, \hat{\lambda}_{js}) = \sum_{m,n,x,y=1}^p \frac{\partial \lambda_{ir}}{\partial \sigma_{mx}} \text{acov}(\hat{\sigma}_{mx}, \hat{\sigma}_{ny}) \frac{\partial \lambda_{js}}{\partial \sigma_{ny}}.$$

Similarly, let $\psi = (\psi_1, \dots, \psi_p)^\top$ denote the diagonal elements of the uniqueness matrix Ψ . The asymptotic covariance of the uniqueness estimators can be written as

$$\text{acov}(\hat{\psi}_i, \hat{\psi}_j) = \sum_{m,n,x,y=1}^p \frac{\partial \psi_i}{\partial \sigma_{mx}} \text{acov}(\hat{\sigma}_{mx}, \hat{\sigma}_{ny}) \frac{\partial \psi_j}{\partial \sigma_{ny}}.$$

In this paper, we derive explicit expressions for the differentials $d\Lambda$ and $d\psi$ in terms of $d\Sigma$ for several widely used factor extraction methods. The motivation for this research is twofold. First, given the covariance of $\hat{\Sigma}$, our results enable practitioners to compute standard errors for estimated factor loadings $\hat{\Lambda}$ and uniquenesses $\hat{\psi}$. When the data are drawn from a multivariate normal population, for example, the covariance of $\hat{\Sigma}$ is well-known (see Girshick [1939] [6]). Second, delta-method-based inference for rotated factor loadings (e.g., Archer and Jennrich [1973] [2]; Jennrich [1973] [10]; Yung and Hayashi [2001] [21]) requires these differentials through the chain rule of differentiation. At present, readily available standard error formulas exist primarily for maximum-likelihood factor analysis, as implemented in software such as SAS [2015] [19], reflecting a general property of maximum-likelihood estimators.

When the assumption of normality is violated, the results developed here can be combined with existing asymptotic results for sample covariances or correlations (e.g., Hsu [1949] [9]; de Leeuw [1983] [4]) to obtain more general asymptotic covariance formulas for unrotated factor loadings. While previous studies have examined the asymptotic properties of principal component analysis (e.g., Girshick [1939] [6]; Anderson [1963] [1]) and derived standard error formulas for maximum-likelihood factor analysis (Lawley [1967] [14], Jennrich and Thayer [1973] [11]), our focus is on least-squares, iter-

ative principal component, principal, alpha, and image factor solutions, as implemented in SAS [2015] [19].

Despite the availability of maximum-likelihood (ML) factor analysis, non-ML extraction methods remain widely used due to their computational simplicity, robustness, distribution-free restrictions, and broad implementation in standard statistical software (e.g., Lin et al. [2023] [16]; Raykov et al. [2026] [18]). In many empirical applications, factor analysis is applied to binary, categorical, or qualitative data that may be high-frequency, high-dimensional, or heavy-tailed. Although recent research has emphasized robustness, high-dimensional asymptotics, and software-based inference for rotated solutions, explicit closed-form asymptotic covariance formulas for non-ML factor extraction methods remain largely unavailable. This gap substantially limits formal statistical inference and model comparison. The present paper addresses this limitation by deriving analytically tractable asymptotic covariance expressions, thereby enabling delta-method inference under both normal and non-normal settings.

Our approach is based on multivariate implicit differentiation of the estimating equations that define the factor loadings $\hat{\Lambda}$ and uniquenesses $\hat{\Sigma}$ under each extraction method, following techniques widely used in factor-analysis literature (e.g., Archer and Jennrich [1973] [2]; Girshick [1939] [6]; Jennrich [1973] [10]; Yung and Hayashi [2001] [21]). We assess the accuracy of the proposed standard errors for the uniqueness estimator $\hat{\psi}$ through a simulation study and present an empirical application illustrating their use for the least-squares factor solution.

Accordingly, the main contributions of this paper are as follows:

- (i) Derivation of explicit asymptotic covariance formulas for several non-maximum-likelihood factor solutions;
- (ii) A unified implicit-differentiation framework applicable across multiple factor extraction methods;
- (iii) Closed-form standard error formulas for uniqueness estimators;
- (iv) A practical extension to rotated factor loadings via the delta method; and
- (v) Simulation evidence demonstrating finite-sample accuracy.

2. Iterative Principal Component and Principal Factor Analysis

We begin by introducing notation. For a $p \times k$ matrix Λ , let $\vec{\Lambda}$ denote its vectorization obtained by stacking the columns of Λ , that is,

$$\vec{\Lambda} = (\lambda_{11}, \dots, \lambda_{p1}, \lambda_{12}, \dots, \lambda_{p2}, \dots, \lambda_{1k}, \dots, \lambda_{pk})^\top.$$

When no confusion arises, we continue to use double subscripts to refer to individual elements of $\vec{\Lambda}$. In particular, the (i, r) -th element of Λ , denoted by λ_{ir} , corresponds to the $((r-1)p + i)$ -th element of $\vec{\Lambda}$. For a square matrix M , let $\text{diag}(M)$ denote the vector formed by the diagonal elements of M , and let $\text{Diag}(M)$ denote the diagonal matrix whose diagonal is given by $\text{diag}(M)$. We also use δ_{ij} to denote the Kronecker delta, which equals 1 if $i = j$ and 0 otherwise.

Both iterative principal component factor analysis and principal factor analysis construct the loading matrix Λ from the principal eigenvectors of $\Sigma - \Psi$:

$$\Sigma - \Psi = \Lambda \Lambda^\top = \sum_{r=1}^k \lambda_r \lambda_r^\top, \quad (4)$$

where λ_r denotes the r -th column of $\Lambda = [\lambda_1, \lambda_2, \dots, \lambda_k]$. These columns are ordered according to the eigenvalues of $\Sigma - \Psi$, with λ_r corresponding to the r -th largest eigenvalue. Without loss of generality, we assume that these eigenvalues are distinct, which is typically the case for a sample

variance-covariance matrix. Consequently, the inner product between any two distinct columns of Λ is zero. In particular, Eq. (4) implies

$$\sigma_{ii} - \psi_i = \sum_{t=1}^k \lambda_{it}^2, \quad \forall i = 1, \dots, p. \quad (5)$$

For the sample variance-covariance matrix $\hat{\Sigma}$, Eq. (4) may, in general, include additional items of the form $\hat{\lambda}_r \hat{\lambda}_r^\top$ for $r > k$. These items are omitted in accordance with the model specification in Eq. (2).

Using Eq. (4), we obtain

$$(\Sigma - \Psi)\lambda_r = \Lambda \Lambda^\top \lambda_r = \sum_{s=1}^k \lambda_s \lambda_s^\top \lambda_r = \sum_{s=1}^k \lambda_s (\lambda_s^\top \lambda_r) = (\lambda_r^\top \lambda_r) \lambda_r. \quad (6)$$

This result holds for any $1 \leq r \leq k$. Now fix indices $1 \leq i \leq p$ and $1 \leq r \leq k$. Applying Eq. (5) to the i -th components of both sides of Eq. (6) yields

$$\sum_{j \neq i} \sigma_{ij} \lambda_{jr} + \lambda_{ir} \sum_{t=1}^k \lambda_{it}^2 = \lambda_{ir} \sum_{j=1}^p \lambda_{jr}^2.$$

Taking differentials on both sides, we obtain

$$\sum_{j \neq i} (\sigma_{ij} d\lambda_{jr} + \lambda_{jr} d\sigma_{ij}) + \left(\sum_{t=1}^k \lambda_{it}^2 \right) d\lambda_{ir} + 2\lambda_{ir} \sum_{t=1}^k \lambda_{it} d\lambda_{it} = \left(\sum_{j=1}^p \lambda_{jr}^2 \right) d\lambda_{ir} + 2\lambda_{ir} \sum_{j=1}^p \lambda_{jr} d\lambda_{jr}.$$

After algebraic simplification, this expression reduces to

$$\sum_{j \neq i} (\sigma_{ij} - 2\lambda_{ir} \lambda_{jr}) d\lambda_{jr} + \left(\sum_{t=1}^k \lambda_{it}^2 - \sum_{j=1}^p \lambda_{jr}^2 \right) d\lambda_{ir} + 2\lambda_{ir} \sum_{t \neq r} \lambda_{it} d\lambda_{it} = - \sum_{j \neq i} \lambda_{jr} d\sigma_{ij}.$$

Collecting all linear equations derived above, the system can be written compactly in matrix form as

$$Ad\vec{\Lambda} = Bd\vec{\Sigma},$$

where the coefficient of $d\lambda_{jt}$ in the (i, r) -th equation is given by

$$A_{ir,jt} = \begin{cases} \sum_{s=1}^k \lambda_{is}^2 - \sum_{s=1}^p \lambda_{sr}^2, & \text{if } j = i \text{ and } t = r, \\ 2\lambda_{ir} \lambda_{it}, & \text{if } j = i \text{ but } t \neq r, \\ \sigma_{ij} - 2\lambda_{ir} \lambda_{jr}, & \text{if } j \neq i \text{ but } t = r, \\ 0, & \text{otherwise,} \end{cases}$$

for all $1 \leq i, j \leq p$ and $1 \leq r, t \leq k$. Similarly, the coefficient of $d\sigma_{xy}$ in the same equation is

$$B_{ir,xy} = -\delta_{ix}(1 - \delta_{xy})\lambda_{yr}, \quad \forall 1 \leq i, x, y \leq p, 1 \leq r \leq k.$$

Finally, we obtain

$$d\vec{\Lambda} = A^{-1}Bd\vec{\Sigma}$$

and, by the delta method, the asymptotic covariance matrix of $\vec{\Lambda}$ is given by

$$\text{acov}(\vec{\Lambda}) = (A^{-1}B)\text{acov}(\vec{\Sigma})(A^{-1}B)^\top.$$

3. Least-Squares Factor Analysis

In least-squares factor analysis (see Harman [1976] [7]), the objective is to minimize the sum of squared off-diagonal elements of the residual covariance matrix $\hat{\Sigma} - \hat{\Lambda}\hat{\Lambda}^\top$. Define the objective function

$$g(\Lambda) = \sum_{1 \leq y < x \leq p} \left(\sigma_{xy} - \sum_{t=1}^k \lambda_{xt} \lambda_{yt} \right)^2,$$

that is, the sum of squares of the strictly lower-triangular elements of $\Sigma - \Lambda\Lambda^\top$.

The first-order optimality conditions yield $p \times k$ constraints on the loading matrix Λ . Specifically, we require

$$\frac{\partial g(\Lambda)}{\partial \lambda_{ir}} = 0$$

for all $1 \leq i \leq p$ and $1 \leq r \leq k$. Straightforward differentiation gives

$$\begin{aligned} -\frac{1}{2} \frac{\partial g}{\partial \lambda_{ir}}(\Lambda) &= \sum_{1 \leq y < x \leq p} \left(\sigma_{xy} - \sum_{t=1}^k \lambda_{xt} \lambda_{yt} \right) \sum_{t=1}^k (\delta_{ix} \delta_{rt} \lambda_{yt} + \delta_{iy} \delta_{rt} \lambda_{xt}) \\ &= \sum_{1 \leq y < x \leq p} (\delta_{ix} \lambda_{yr} + \delta_{iy} \lambda_{xr}) \left(\sigma_{xy} - \sum_{t=1}^k \lambda_{xt} \lambda_{yt} \right) \\ &= \sum_{1 \leq y < i} \lambda_{yr} \left(\sigma_{iy} - \sum_{t=1}^k \lambda_{it} \lambda_{yt} \right) + \sum_{i < x \leq p} \lambda_{xr} \left(\sigma_{xi} - \sum_{t=1}^k \lambda_{xt} \lambda_{it} \right) \\ &= \sum_{j \neq i} \lambda_{jr} \left(\sigma_{ij} - \sum_{t=1}^k \lambda_{it} \lambda_{jt} \right). \end{aligned}$$

For any fixed $1 \leq i \leq p$ and $1 \leq r \leq k$, the first-order condition reduces to

$$\sum_{j \neq i} \lambda_{jr} \sigma_{ij} = \sum_{j \neq i} \lambda_{jr} \sum_{t=1}^k \lambda_{it} \lambda_{jt}. \quad (7)$$

The differential form of Eq. (7) is given by

$$\sum_{j \neq i} (\lambda_{jr} d\sigma_{ij} + \sigma_{ij} d\lambda_{jr}) = \sum_{j \neq i} \left(\sum_{t=1}^k \lambda_{it} \lambda_{jt} \right) d\lambda_{jr} + \sum_{j \neq i} \lambda_{jr} \sum_{t=1}^k (\lambda_{jt} d\lambda_{it} + \lambda_{it} d\lambda_{jt}).$$

Grouping terms and simplifying yield

$$\sum_{j \neq i} \lambda_{jr} d\sigma_{ij} = \sum_{j \neq i} \left(\sum_{t=1}^k \lambda_{it} \lambda_{jt} - \sigma_{ij} + \lambda_{jr} \lambda_{ir} \right) d\lambda_{jr} + \sum_{t=1}^k \left(\sum_{j \neq i} \lambda_{jr} \lambda_{jt} \right) d\lambda_{it} + \sum_{j \neq i} \sum_{t \neq r} \lambda_{jr} \lambda_{it} d\lambda_{jt},$$

which completes the differential characterization of the first-order condition.

We rewrite the preceding system of equations in matrix form as

$$-Bd\vec{\Sigma} = Dd\vec{\Lambda},$$

where the coefficient of $d\lambda_{jt}$ in the (i, r) -th equation is defined as

$$D_{ir,jt} = \begin{cases} \sum_{z \neq i} \lambda_{zr} \lambda_{zt}, & \text{if } j = i, \\ \sum_{z=1}^k \lambda_{iz} \lambda_{jz} - \sigma_{ij} + \lambda_{ir} \lambda_{jr}, & \text{if } j \neq i \text{ and } t = r, \\ \lambda_{jr} \lambda_{it}, & \text{if } j \neq i \text{ and } t \neq r. \end{cases}$$

These definitions apply for all $1 \leq i, j \leq p$ and $1 \leq r, t \leq k$. Finally, solving for $d\vec{\Lambda}$ yields

$$d\vec{\Lambda} = -D^{-1}Bd\vec{\Sigma}.$$

4. Alpha Factor Analysis

Let H be the diagonal matrix whose i -th diagonal element is $\sqrt{h_i}$. Clearly, $h_i > 0$; otherwise, Eq. (3) would imply that the i -th row of the loading matrix Λ is a zero vector, rendering the corresponding factor redundant. Moreover, $H^{-1}(\Sigma - \Psi)H^{-1}$ is the correlation matrix for ΛF and Eq (3) implies that

$$dh_i = 2 \sum_{t=1}^k \lambda_{it} d\lambda_{it}. \quad (8)$$

To obtain the spectral decomposition of this correlation matrix, let θ_r denote its r -th largest eigenvalue, and let β_r be the corresponding eigenvector with length $\theta_r^{1/2}$. Thus,

$$\theta_r = \beta_r^\top \beta_r,$$

and

$$H^{-1}(\Sigma - \Psi)H^{-1}\beta_r = \theta_r\beta_r, \quad \forall 1 \leq r \leq k.$$

In alpha factor analysis (Kaiser and Caffrey [1965] [13]), the loading matrix $\Lambda = [\lambda_1, \lambda_2, \dots, \lambda_k]$ is defined as

$$\Lambda = H[\beta_1, \beta_2, \dots, \beta_k].$$

Additionally, let $\Gamma = [\gamma_1, \gamma_2, \dots, \gamma_k]$, where $\gamma_r = H^{-1}\beta_r$. Then, for each $r = 1, \dots, k$, the following relationships—corresponding to Eqs. (9)-(11)—hold:

$$\lambda_r = H\beta_r = H^2\gamma_r, \quad (9)$$

$$\theta_r = \beta_r^\top \beta_r = \gamma_r^\top H^2\gamma_r = \gamma_r^\top \lambda_r, \quad (10)$$

$$(\Sigma - \Psi)\gamma_r = (\Sigma - \Psi)H^{-1}\beta_r = HH^{-1}(\Sigma - \Psi)H^{-1}\beta_r = \theta_r H\beta_r = \gamma_r^\top \lambda_r \lambda_r. \quad (11)$$

In the remainder of this section, we first express $d\vec{\Gamma}$ in terms of $d\vec{\Lambda}$ using Eq. (9). We then derive a matrix equation relating $d\vec{\Sigma}$, $d\vec{\Gamma}$, and $d\vec{\Lambda}$ from Eq. (11).

From Eq. (9), for any $1 \leq i \leq p$ and $1 \leq r \leq k$, we have

$$\lambda_{ir} = h_i \gamma_{ir},$$

whose differential is given by

$$d\lambda_{ir} = h_i d\gamma_{ir} + 2\gamma_{ir} \sum_{t=1}^k \lambda_{it} d\lambda_{it}.$$

Collecting these expressions yields the following matrix representation:

$$d\vec{\Gamma} = E d\vec{\Lambda}, \quad (12)$$

where the elements of the matrix E are defined by

$$E_{ir,jt} = \frac{\partial \gamma_{ir}}{\partial \lambda_{jt}} = \delta_{ij}(\delta_{rt} - 2\gamma_{ir}\lambda_{it})/h_i,$$

for all $1 \leq i, j \leq p$ and $1 \leq r, t \leq k$.

For any fixed $1 \leq i \leq p$ and $1 \leq r \leq k$, the i -th component of Eq. (11) can be written as

$$\sum_{j \neq i} \sigma_{ij} \gamma_{jr} + h_i \gamma_{ir} = \lambda_{ir} \sum_{j=1}^p \lambda_{jr} \gamma_{jr}. \quad (13)$$

Taking differentials of Eq. (13) yields

$$\sum_{j \neq i} (\sigma_{ij} d\gamma_{jr} + \gamma_{jr} d\sigma_{ij}) + h_i d\gamma_{ir} + \gamma_{ir} dh_i = \lambda_{ir} \sum_{j=1}^p (\lambda_{jr} d\gamma_{jr} + \gamma_{jr} d\lambda_{jr}) + \left(\sum_{j=1}^p \lambda_{jr} \gamma_{jr} \right) d\lambda_{ir}.$$

Substituting Eq. (8) into the above expression and rearranging terms, we obtain

$$\sum_{j \neq i} \gamma_{jr} d\sigma_{ij} = \sum_{j \neq i} (\lambda_{ir} \lambda_{jr} - \sigma_{ij}) d\gamma_{jr} + (\lambda_{ir}^2 - h_i) d\gamma_{ir} + \left(\sum_{j \neq i} \lambda_{jr} \gamma_{jr} \right) d\lambda_{ir} - 2\gamma_{ir} \sum_{t \neq r} \lambda_{it} d\lambda_{it} + \lambda_{ir} \sum_{j \neq i} \gamma_{jr} d\lambda_{jr}.$$

We express all the differential forms above in matrix notation as

$$Fd\vec{\Sigma} = Gd\vec{\Gamma} + Jd\vec{\Lambda}. \quad (14)$$

The coefficient matrices F , G , and J are defined as follows.

- Matrix F : The coefficient of $d\sigma_{xy}$ in the (i, r) -th equation of Eq. (14) is

$$F_{ir,xy} = \delta_{ix}(1 - \delta_{iy})\gamma_{yr}.$$

- Matrix G : The coefficient of $d\gamma_{jt}$ is given by

$$G_{ir,jt} = \begin{cases} \lambda_{ir}^2 - h_i, & \text{if } j = i \text{ and } t = r, \\ \lambda_{ir} \lambda_{jr} - \sigma_{ij}, & \text{if } t = r \text{ but } j \neq i, \\ 0, & \text{otherwise.} \end{cases}$$

- Matrix J : The coefficient of $d\lambda_{jt}$ is

$$J_{ir,jt} = \begin{cases} \sum_{s \neq i} \lambda_{sr} \gamma_{sr}, & \text{if } j = i \text{ and } t = r, \\ -2\gamma_{ir} \lambda_{it}, & \text{if } j = i \text{ but } t \neq r, \\ \lambda_{ir} \gamma_{jr}, & \text{if } t = r \text{ but } j \neq i, \\ 0, & \text{otherwise.} \end{cases}$$

Combining Eqs. (12) and (14), we obtain

$$Fd\vec{\Sigma} = Gd\vec{\Gamma} + Jd\vec{\Lambda} = G(Ed\vec{\Lambda}) + Jd\vec{\Lambda} = (GE + J)d\vec{\Lambda}.$$

Consequently,

$$d\vec{\Lambda} = (GE + J)^{-1}Fd\vec{\Sigma}.$$

5. Image Factor Analysis

Let $\Delta = (\text{Diag}(\Sigma^{-1}))^{-1}$. In image factor analysis (see Jöreskog [1969] [12]), the covariance matrix Σ is modeled as

$$\Sigma = \Lambda\Lambda^\top + \tau\Delta, \quad (15)$$

where $\tau > 0$ is an unknown scalar parameter. Within this framework, one may write $\Psi = \tau\Delta$. Jöreskog [1969] [12] proposed a maximum-likelihood procedure for estimating this model, under

which the covariance matrix of the estimators follows directly from standard maximum-likelihood theory.

For the principal-factor procedure (see SAS [2015] [19]), the loading matrix Λ is obtained from the k largest eigenvalues and their corresponding eigenvectors of $\Sigma - \tau\Delta$. The parameter τ is simultaneously estimated as the coefficient in the regression of $\text{diag}(\Sigma - \Lambda\Lambda^\top)$ on $\text{diag}(\Delta)$, namely,

$$\tau = \frac{\langle \text{diag}(\Delta), \text{diag}(\Sigma - \Lambda\Lambda^\top) \rangle}{\langle \text{diag}(\Delta), \text{diag}(\Delta) \rangle} \quad (16)$$

where $\langle \cdot, \cdot \rangle$ denotes the standard inner product.

Let σ^{xy} denote the element at the x -th row and y -th column of Σ^{-1} . Since $\Sigma\Sigma^{-1}$ is the $p \times p$ identity matrix, differentiating this identity using the product rule yields

$$d\Sigma^{-1} = -\Sigma^{-1}(d\Sigma)\Sigma^{-1}.$$

In particular, for each $i = 1, \dots, p$,

$$d\sigma^{ii} = - \sum_{x,y=1}^p \sigma^{ix}(d\sigma_{xy})\sigma^{yi} = - \sum_{x,y=1}^p \sigma^{ix}\sigma^{iy}d\sigma_{xy}.$$

Consequently,

$$d\left(\frac{1}{\sigma^{ii}}\right) = -\frac{1}{(\sigma^{ii})^2}d\sigma^{ii} = \sum_{x,y=1}^p \frac{\sigma^{ix}\sigma^{iy}}{(\sigma^{ii})^2}d\sigma_{xy}. \quad (17)$$

Eq. (16) implies that

$$\tau \langle \text{diag}(\Delta), \text{diag}(\Delta) \rangle = \langle \text{diag}(\Delta), \text{diag}(\Sigma - \Lambda\Lambda^\top) \rangle,$$

and therefore

$$\tau \sum_{i=1}^p \left(\frac{1}{\sigma^{ii}}\right)^2 = \sum_{i=1}^p \frac{1}{\sigma^{ii}}(\sigma_{ii} - h_i).$$

Differentiating both sides yields

$$\left[\sum_{i=1}^p \left(\frac{1}{\sigma^{ii}}\right)^2 \right] d\tau + \sum_{i=1}^p \frac{2\tau}{\sigma^{ii}} d\left(\frac{1}{\sigma^{ii}}\right) = \sum_{i=1}^p \left[\frac{1}{\sigma^{ii}}(d\sigma_{ii} - dh_i) + (\sigma_{ii} - h_i)d\left(\frac{1}{\sigma^{ii}}\right) \right].$$

Substituting Eq. (17) into the above expression, we obtain

$$\left[\sum_{i=1}^p \left(\frac{1}{\sigma^{ii}}\right)^2 \right] d\tau = \sum_{x,y=1}^p \left[\sum_{i=1}^p \left(\sigma_{ii} - h_i - \frac{2\tau}{\sigma^{ii}}\right) \frac{\sigma^{ix}\sigma^{iy}}{(\sigma^{ii})^2} \right] d\sigma_{xy} + \sum_{i=1}^p \frac{1}{\sigma^{ii}} \left(d\sigma_{ii} - 2 \sum_{t=1}^k \lambda_{it} d\lambda_{it} \right).$$

These expressions can be written compactly as

$$d\tau = \mu^\top d\vec{\Sigma} + \eta^\top d\vec{\Lambda}, \quad (18)$$

where μ and η are vectors of dimensions p^2 and pk , respectively. The coefficient of $d\sigma_{xy}$ is given by

$$\mu_{xy} = \frac{\frac{\delta_{xy}}{\sigma^{xx}} + \sum_{i=1}^p \left(\sigma_{ii} - h_i - \frac{2\tau}{\sigma^{ii}}\right) \frac{\sigma^{ix}\sigma^{iy}}{(\sigma^{ii})^2}}{\sum_{i=1}^p \left(\frac{1}{\sigma^{ii}}\right)^2}$$

and the coefficient of $d\lambda_{ir}$ is

$$\eta_{ir} = -\frac{2\lambda_{ir}}{\sigma^{ii} \sum_{z=1}^p \left(\frac{1}{\sigma^{zz}}\right)^2}.$$

Next, let θ_r denote the r -th largest eigenvalue of $\Sigma - \tau\Delta$, and let λ_r be the associated eigenvector with length $\theta_r^{1/2}$. Then,

$$\lambda_r^\top \lambda_r = \theta_r,$$

and

$$(\Sigma - \tau\Delta)\lambda_r = \theta_r \lambda_r = \lambda_r^\top \lambda_r \lambda_r.$$

For any fixed $1 \leq i \leq p$, the i -th component of this equation is given by

$$\sum_{j=1}^p \sigma_{ij} \lambda_{jr} - \frac{\tau \lambda_{ir}}{\sigma^{ii}} = \lambda_{ir} \sum_{j=1}^p \lambda_{jr}^2.$$

Differentiating both sides with respect to the underlying parameters yields

$$\sum_{j=1}^p (\sigma_{ij} d\lambda_{jr} + \lambda_{jr} d\sigma_{ij}) - \frac{\tau}{\sigma^{ii}} d\lambda_{ir} - \frac{\lambda_{ir}}{\sigma^{ii}} d\tau - \tau \lambda_{ir} d\left(\frac{1}{\sigma^{ii}}\right) = 2\lambda_{ir} \sum_{j=1}^p \lambda_{jr} d\lambda_{jr} + \left(\sum_{j=1}^p \lambda_{jr}^2\right) d\lambda_{ir}.$$

Applying Eq. (17) and collecting terms involving $d\Sigma$, $d\Gamma$, and $d\tau$, we obtain

$$\sum_{j=1}^p \lambda_{jr} d\sigma_{ij} - \tau \lambda_{ir} \sum_{x,y=1}^p \frac{\sigma^{ix} \sigma^{iy}}{(\sigma^{ii})^2} d\sigma_{xy} = \sum_{j \neq i} (2\lambda_{ir} \lambda_{jr} - \sigma_{ij}) d\lambda_{jr} + \left(2\lambda_{ir}^2 - \sigma_{ii} + \frac{\tau}{\sigma^{ii}} + \sum_{j=1}^p \lambda_{jr}^2\right) d\lambda_{ir} + \frac{\lambda_{ir}}{\sigma^{ii}} d\tau.$$

The matrix representation of the above system of equations can be written as

$$L d\vec{\Sigma} = P d\vec{\Lambda} + \pi d\tau \quad (19)$$

where L , P , and π are appropriately defined coefficient matrices. In particular, the coefficient of $d\sigma_{xy}$ in the (i, r) -th equation is given by

$$L_{ir,xy} = \delta_{ix} \lambda_{yr} - \frac{\tau \lambda_{ir} \sigma^{ix} \sigma^{iy}}{(\sigma^{ii})^2}.$$

The coefficient of $d\lambda_{jt}$ is

$$P_{ir,jt} = \begin{cases} 2\lambda_{ir}^2 - \sigma_{ii} + \frac{\tau}{\sigma^{ii}} + \sum_{z=1}^p \lambda_{zr}^2, & \text{if } j = i \text{ and } t = r, \\ 2\lambda_{ir} \lambda_{jr} - \sigma_{ij}, & \text{if } t = r \text{ but } j \neq i, \\ 0, & \text{otherwise,} \end{cases}$$

and the coefficient of $d\tau$ is

$$\pi_{ir} = \frac{\lambda_{ir}}{\sigma^{ii}},$$

for any $1 \leq i, j, x, y \leq p$ and $1 \leq r, t \leq k$. Note that π is a vector of dimension pk .

Combining Eqs. (18) and (19) yields

$$d\vec{\Lambda} = (P + \pi \eta^\top)^{-1} (L - \pi \mu^\top) d\vec{\Sigma}.$$

Substituting this expression back into Eq. (18) gives the differential of τ in terms of $d\vec{\Sigma}$:

$$d\tau = \left[\mu^\top + \eta^\top (P + \pi \eta^\top)^{-1} (L - \pi \mu^\top) \right] d\vec{\Sigma}. \quad (20)$$

6. Standard Error of Uniqueness Estimator $\hat{\psi}$

Assume that $d\vec{\Lambda} = Md\vec{\Sigma}$ for some $pk \times p^2$ matrix M . To derive the standard errors of the uniqueness estimator $\hat{\psi}$, we begin with the generic relationship from Eq. (5): $\psi_i = \sigma_{ii} - \sum_{t=1}^k \lambda_{it}^2$. Differentiating both sides yields

$$d\psi_i = d\sigma_{ii} - 2 \sum_{t=1}^k \lambda_{it} d\lambda_{it}, \quad i = 1, 2, \dots, p. \quad (21)$$

Eq. (21) can be written more compactly in matrix form as

$$d\psi = Td\vec{\Sigma} + Zd\vec{\Lambda}, \quad (22)$$

where the coefficient of $d\sigma_{xy}$ in the i -th equation of Eq. (22) is given by

$$T_{i,xy} = \delta_{ix}\delta_{iy}$$

and the coefficient of $d\lambda_{jt}$ in the same equation is

$$Z_{i,jt} = -2\delta_{ij}\lambda_{jt},$$

for all $1 \leq i, j, x, y \leq p$ and $1 \leq t \leq k$. Consequently, we obtain

$$d\psi = (T + ZM)d\vec{\Sigma}.$$

When Σ is a correlation matrix, $d\sigma_{ii} = 0$, and the expression simplifies further to $d\psi = ZMd\vec{\Sigma}$.

In image factor analysis, the relationship $\hat{\psi}_i = \hat{\sigma}_{ii} - \sum_{t=1}^k \hat{\lambda}_{it}^2$ may no longer hold. Instead, we rewrite Eq. (20) as $d\tau = \phi^\top d\vec{\Sigma}$ for some vector ϕ . Using the identity $\psi_i = \frac{\tau}{\sigma^{ii}}$ and applying Eq. (17), we obtain

$$d\psi_i = \frac{d\tau}{\sigma^{ii}} + \tau d\left(\frac{1}{\sigma^{ii}}\right) = \frac{1}{\sigma^{ii}} \sum_{x,y=1}^p \phi_{xy} d\sigma_{xy} + \tau \sum_{x,y=1}^p \frac{\sigma^{ix}\sigma^{iy}}{(\sigma^{ii})^2} d\sigma_{xy}, \quad (23)$$

for all $1 \leq i \leq p$. In matrix form, this expression can be written compactly as

$$d\psi = Qd\vec{\Sigma},$$

where the coefficient of $d\sigma_{xy}$ in the i -th equation is given by

$$Q_{i,xy} = \frac{\partial \psi_i}{\partial \sigma_{xy}} = \frac{\sigma^{ii}\phi_{xy} + \tau\sigma^{ix}\sigma^{iy}}{(\sigma^{ii})^2}, \quad \forall 1 \leq i, x, y \leq p.$$

7. An Empirical and Simulation Study

In this section, we apply the least-squares factor analysis method to a sample correlation matrix consisting of nine variables measured on 211 subjects. The data were originally reported by Lawley and Maxwell [1971, p.43] [15], and the sample is assumed to be drawn from an approximately multivariate normal population. For simplicity, we set the number of common factors to $k = 2$. This choice is supported by both Horn [1965] [8] 's parallel analysis and Velicer [1976] [20] 's minimum average partial (MAP) test. Estimates of the factor loading matrix $\hat{\Lambda}$ and the uniqueness vector $\hat{\psi}$ are obtained using the iterative procedure described by Harman [1976] [7].

To compute the asymptotic covariances of these estimators, we first apply Girshick [1939] [6]'s formula (3.23) to derive the covariances of the sample correlation coefficients. We then use the results developed in Sections 3 and 6 to compute the asymptotic covariance matrix of $\hat{\Lambda}$, as well as the

standard errors of $\hat{\psi}$. To illustrate the effect of factor rotation, we orthogonally rotate $\hat{\Lambda}$ using the varimax criterion and apply the method of Archer and Jennrich [1973] [11] to obtain the standard errors of the rotated factor loadings. Table 1 reports the parameter estimates for the two-factor least-squares model; standard errors are reported in parentheses.

Table 1. Estimates and Standard Errors for Two-Factor Least-Squares Factor Analysis.

Variable	Unrotated Factors		Rotated Factors		Uniqueness
	I	II	I	II	
X_1	.6639(.0397)	.3285(.0502)	.6745(.0453)	.3063(.0538)	.4512(.0557)
X_2	.6879(.0381)	.2388(.0545)	.6202(.0487)	.3815(.0555)	.4698(.0539)
X_3	.4956(.0536)	.2831(.0657)	.5328(.0583)	.2047(.0658)	.6743(.0599)
X_4	.8470(.0249)	-.3037(.0372)	.3007(.0381)	.8481(.0295)	.1904(.0412)
X_5	.7035(.0392)	-.3179(.0637)	.1990(.0493)	.7459(.0399)	.4040(.0551)
X_6	.8037(.0297)	-.3581(.0659)	.2312(.0404)	.8490(.0353)	.2258(.0526)
X_7	.6686(.0440)	.3889(.0654)	.7242(.0412)	.2717(.0513)	.4018(.0546)
X_8	.4236(.0609)	.2552(.0813)	.4656(.0639)	.1666(.0695)	.7555(.0578)
X_9	.7718(.0347)	.4398(.0598)	.8289(.0328)	.3194(.0444)	.2109(.0443)

To verify the theoretical derivation of the asymptotic covariance, $\text{acov}(\hat{\psi})$, presented in Section 6, we conduct a simulation study. The sample correlation matrix reported above is treated as the population correlation matrix Σ . Using this matrix, we generate a large number m of random correlation matrices from a Wishart distribution with parameter $(\Sigma, 211)$. For each simulated correlation matrix, the uniqueness parameters are estimated using the least-squares method with $k = 2$.

Based on the resulting collection of simulated estimates of ψ , we compute empirical standard errors for the uniqueness estimates. For $m = 50, 100, 500, 1000, 2000$, the corresponding empirical standard errors are reported in Table 2. The final column of the table presents the theoretical asymptotic standard errors computed using the formulas developed in Section 6. The results demonstrate clear convergence of the empirical standard errors toward their theoretical counterparts, although the convergence occurs at a relatively slow rate.

Table 2. Simulation of Standard Errors for Uniqueness Estimates.

Uniqueness	Empirical S.E. of Uniqueness for Simulated Correlations					Theoretical S.E.
	$m = 50$	100	500	1000	2000	
ψ_1	.0499849	.0511051	.0549498	.0560314	.0556181	.0556623
ψ_2	.0513352	.0520701	.0545346	.0555878	.0549598	.0538818
ψ_3	.0666649	.0660852	.0599577	.0603062	.0598762	.0598473
ψ_4	.0405467	.0433619	.0432832	.0428048	.0415737	.0411420
ψ_5	.0481342	.0538724	.0524052	.0530417	.0544235	.0550669
ψ_6	.0501591	.0543790	.0541692	.0550827	.0546638	.0526150
ψ_7	.0524787	.0565299	.0547573	.0553479	.0546232	.0546297
ψ_8	.0507511	.0562004	.0576808	.0582864	.0580424	.0578393
ψ_9	.0394781	.0409714	.0430909	.0448280	.0438135	.0443326

8. Discussions

Recent years have seen renewed interest in formal statistical inference for factor-analytic models. This resurgence is driven by a range of practical considerations, including applications involving non-normal data, concerns about robustness, and the growing complexity of covariance structures encountered in empirical work. Broadly speaking, the contemporary literature can be grouped into three main strands. The first strand emphasizes likelihood-based approaches to factor analysis and covariance structure modeling, building on classical frameworks and their subsequent extensions (e.g., Lawley [1967] [14]; Jennrich and Thayer [1973] [11]; Raykov et al. [2026] [18]). A second line of research focuses on high-dimensional and dynamic factor models, in which both the cross-sectional and temporal dimensions grow, posing new challenges for estimation and inference (e.g., Bai [2003]

[3]; Fan et al. [2021] [5]). The third strand concentrates on software-oriented implementations of standard error estimation for rotated factor solutions, typically relying on numerical differentiation or sandwich-type approximations (e.g., Archer and Jennrich [1973] [2]; Yung and Hayashi [2001] [21]).

In contrast, explicit asymptotic covariance formulas for widely used non-maximum-likelihood extraction method—such as least-squares, principal, alpha, and image factor analysis—remain relatively scarce, despite their continued prevalence in empirical research and standard statistical software (e.g., Harman [1976] [7]; SAS [2015] [19]). The results developed in this paper help bridge this gap by providing closed-form, analytically tractable asymptotic covariance expressions for several classical no-ML factor solutions. By formulating these results within a unified implicit-differentiation framework, the proposed approach accommodates non-normal data and can be readily combined with existing delta-method results for factor rotation (e.g., Archer and Jennrich [1973] [2]; Jennrich [1973] [10]; Yung and Hayashi [2001] [21]). From a practical perspective, this framework enables transparent computation of standard errors without reliance on numerical differentiation or resampling. From a methodological standpoint, it complements recent advances in robust and high-dimensional factor modeling by providing a rigorous inferential foundation for classical extraction methods that remain central to applied factor analysis.

In the preceding sections, the derived asymptotic covariances of $\hat{\Lambda}$ and $\hat{\psi}$ are conditional on the assumed number of factors. Consequently, abnormal standard errors may reflect model misspecification or systematic errors—such as an incorrect choice of factor dimensionality—rather than genuine sampling variability. In particular, such distortions may arise from either over-factoring (when k is chosen too large) or under-factoring (when k is chosen too small).

For expositional simplicity, we have not explicitly imposed the symmetry condition $\sigma_{ij} = \sigma_{ji}$. In practical implementations, however, exploiting this property is important not only for reducing memory requirements but also for improving computational efficiency, especially when working with large-scale datasets.

The matrix Σ need not represent the variance-covariance matrix of a single group of observed data series; it may also correspond to the variance-covariance matrix of residuals in a system of equations. As an illustration, consider a structural vector autoregressive (SVAR) model with p time series:

$$Y_t = c + \sum_{s=1}^{\zeta} A_s Y_{t-s} + \Lambda F_t + \epsilon_t,$$

where Y_t is a p -dimensional vector of observed time series at time t , ζ denotes the lag length, F_t represents k common structural shocks, and ϵ_t captures p idiosyncratic shocks. The composite disturbance $\Lambda F_t + \epsilon_t$ therefore exhibits a two-component error structure, while the common factors F in Eq. (1) are latent and may possess a nonzero mean.

In much of the SVAR literature, idiosyncratic shocks are frequently ignored, effectively imposing the restriction $k = p$. Allowing for idiosyncratic shocks, however, substantially reduces both the number of structural shocks and the number of identification conditions required. From a computational perspective, incorporating an error-component structure—such as a common-factor model—into SVAR systems is straightforward. The formulas developed in this paper offer practical guidance for implementing such extensions and for assessing the precision of estimated coefficients associated with structural shocks.

This form of dimension reduction is particularly valuable in environments characterized by highly interconnected data. In macroeconomic modeling, structural shocks are typically assumed to be independent. Although an economy may generate a large number of interrelated observable time series, its dynamics are generally driven by only a small number of underlying shocks, such as demand, supply, and technological disturbances.

Finally, an important advantage of the derived formulas is their applicability under non-normality without requiring modification. For example, when the data are multivariate binary, the asymptotic covariance $\text{acov}(\hat{\Lambda})$ remains well defined. More generally, for categorical data—which are common

in practical factor-analysis applications—the computation of $\text{acov}(\hat{\Lambda})$ is unlikely to pose significant difficulties.

Author Contributions: All authors have equally contributed to conceptualization, methodology, and validation. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Conflicts of Interest: The authors declare no conflicts of interest.

Abbreviations

The following abbreviations are used in this manuscript:

Σ	Unobservable population variance-covariance matrix of the random vector Y
$\hat{\Sigma}$	Sample variance-covariance matrix of the random vector Y
Λ	Loading matrix in the variance-covariance structure model $\Sigma = \Lambda\Lambda^T + \Psi$
Ψ	Uniqueness, a diagonal matrix
$\hat{\Lambda}, \hat{\Psi}$	Estimates of the loading matrix and uniqueness, respectively
p and k	Dimensions of Λ ; p is also the dimension for Y and, consequently, for Σ and Ψ
δ_{ij}	The Kronecker delta, that equals 1 when $i = j$ and 0 otherwise
$\vec{\Lambda}$	The vectorization of the matrix Λ , ordered by columns. Similar for $\vec{\Sigma}$
ψ	The diagonal vector of Ψ
θ_r	The r -th largest eigenvalue of a symmetric matrix
σ_{xy}	The element at the x -th row and y -th column of Σ
σ^{xy}	The element at the x -th row and y -th column of Σ^{-1} , the inverse of Σ
$\text{diag}(M)$	The vector of diagonal elements of the square matrix M
$\text{Diag}(M)$	The diagonal matrix with diagonal elements in M
ACOV	Asymptotic covariance
SVAR	Structural Vector AutoRegressive model

References

1. Anderson, T.W. (1963). Asymptotic theory for principal components analysis. *Annals of Mathematical Statistics*, 34, 122-148. <https://doi.org/10.1214/AOMS%2F1177704248>
2. Archer, C.O., Jennrich, R.I. (1973). Standard errors for orthogonally rotated factor loadings. *Psychometrika*, 38, 581-592. <https://doi.org/10.1007/BF02291496>
3. Bai, J. (2003). Inferential theory for factor models of large dimensions. *Econometrica* 71(1), 135–171. <https://doi.org/10.1111/1468-0262.00392>
4. de Leeuw, J. (1983). "Models and methods for the analysis of correlation coefficients," *Journal of Econometrics*, 22, 113-137. [https://doi.org/10.1016/0304-4076\(83\)90096-9](https://doi.org/10.1016/0304-4076(83)90096-9)
5. Fan, J., Wang, K., Zhong, Y., Zhu, Z. (2021). Robust high-dimensional factor models with applications to statistical machine learning. *Statistical Science* 36(2), 303–327. <https://doi.org/10.1214/20-STS785>
6. Girshick, M.A. (1939). On the sampling theory of roots of determinantal equations. *Annals of Mathematical Statistics*, 10, 203-224. <https://doi.org/10.1214/aoms/1177732180>
7. Harman, H.H. (1976). *Modern Factor Analysis* (Third Edition). Chicago: University of Chicago Press.
8. Horn, J. (1965). A rationale and test for the number of factors in factor analysis. *Psychometrika*, 30, 179-185. <https://doi.org/10.1007/BF02289447>
9. Hsu, P.L. (1949). "The limiting distribution of functions of sample means and application to testing hypotheses," in *Proceedings of the First Berkeley Symposium in Mathematical Statistics and Probability*, pp.359-402.
10. Jennrich, R.I. (1973). Standard errors for obliquely rotated factor loadings. *Psychometrika*, 38, 593-604. <https://doi.org/10.1007/BF02291497>
11. Jennrich, R.I., Thayer, D.T. (1973). A note on Lawley’s formulas for standard errors in maximum likelihood factor analysis. *Psychometrika*, 38, 571-580. <https://doi.org/10.1007/BF02291495>
12. Jöreskog, K.G. (1969). Efficient estimation in image factor analysis. *Psychometrika*, 34, 51-75. <https://doi.org/10.1007/BF02290173>

13. Kaiser, H.F., Caffrey, J. (1965). Alpha factor analysis. *Psychometrika*, 30, 1-14. <https://doi.org/10.1007/BF02289743>
14. Lawley, D.N. (1967). Some new results in maximum likelihood factor analysis. *Proceedings of the Royal Society of Edinburgh*, 67A, 256-264. <https://doi.org/10.1017/S0080454100008128>
15. Lawley, D.N., Maxwell, A.E. (1971). *Factor Analysis as a Statistical Method*. New York: American Elsevier.
16. Lin, T.I., Chen, I.A., Wang, W.L. (2023). A robust factor analysis model based on the canonical fundamental skew-t distribution. *Stat Papers* 64, 367-393 (2023). <https://doi.org/10.1007/s00362-022-01318-8>
17. Rao, C.R.(2009). *Linear Statistical Inference And Its Applications* (Second Edition). New York: Wiley.
18. Raykov, T., Penev, S., Zhang, B. (2026). Fitting covariance structure models with arbitrary observed distributions: strong convergence of the asymptotically distribution-free estimator. *Structural Equation Modeling: A Multidisciplinary Journal*, 33:1, 32-38. <https://doi.org/10.1080/10705511.2025.2556028>
19. SAS (2015). *SAS/STAT® 14.1 User's Guide The FACTOR Procedure*. Cary, NC: SAS Institute Inc.
20. Velicer, W.F. (1976). Determining the number of components from the matrix of partial correlations. *Psychometrika*, 41, 321-327. <https://doi.org/10.1007/BF02293557>
21. Yung, Y.F., Hayashi, K. (2001). A computational efficient method for obtaining standard error estimates for the promax and related solutions. *British Journal of Mathematical and Statistical Psychology*, 54, 125 - 138. <https://doi.org/10.1348/000711001159465>

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.