A NOTE ON NON-NEGATIVE ESTIMATES FOR THE LOADINGS IN THE ONE-FACTOR MODEL

Wim P. Krijnen[1]

Abstract

The loadings in the one-factor model for factor analysis are the covariances between the observable variables and the latent factor. In case the observable variables correlate non-negatively, it is hypothesized that "mental factors are positive quantities" in the sense that the "loadings are non-negative" (Anderson & Rubin, 1956). Sufficient conditions for estimates of the loadings to be non-negative are given.

[1]Department of Marketing, P.O. Box 800, 9700 AV Groningen, The Netherlands. Tel. 0503636288. The author is obliged to the Netherlands Organization of Science for their post-doc grant, and to Paul Bekker and the referee for comments on an earlier version of the paper.

1 Introduction

It is well-known that the one-factor model for factor analysis assumes that m observations are generated by

$$\boldsymbol{x} = \boldsymbol{\mu} + \boldsymbol{\lambda} f + \boldsymbol{\varepsilon}, \tag{1}$$

where $\boldsymbol{x}$ is the vector with observable variables of order m, $\boldsymbol{\mu}$ its expectation ($\boldsymbol{\mu}$ =E$\boldsymbol{x}$), ($\boldsymbol{\mu}$= E$\boldsymbol{x}$), f the (unobservable) factor score, $\boldsymbol{\varepsilon}$ the error vector of order m, and $\boldsymbol{\lambda}$ the loadings vector of order m. Let $\mathrm{Var}(\boldsymbol{x}) = \boldsymbol{\Sigma}$. From the assumptions $\mathrm{E}(\boldsymbol{\varepsilon}) = \mathbf{0}$, $\mathrm{E}(\boldsymbol{\varepsilon} f) = \mathbf{0}$, and $\mathrm{E}(\boldsymbol{\varepsilon}\boldsymbol{\varepsilon}') = \boldsymbol{\Psi}$ diagonal positive definite, it follows that $\mathrm{Cov}(\boldsymbol{x}, f) = \boldsymbol{\lambda}$ and

$$\boldsymbol{\Sigma} = \boldsymbol{\lambda}\boldsymbol{\lambda}' + \boldsymbol{\Psi}. \tag{2}$$

The first motivation for this study is that in many empirical studies it has been found that the sample variance matrix $\boldsymbol{S}$ has no negative elements. Among other situations, such a finding is commonly made when the observed variables are measuring human intelligence. The second motivation is the conjecture by Anderson and Rubin (1956, p. 113) that "mental factors are positive quantities" which "implies" that the elements of $\boldsymbol{\lambda}$ are non-negative. Indeed, when $\boldsymbol{\Sigma}$ has non-negative elements, the loadings are either all negative or all positive. Hence, without loss of generality, they may be taken non-negative. The third motivation for this study is the usefulness of the one-factor model as a measurement model (Jöreskog, 1971).

The purpose of the current note is to give sufficient conditions for estimates of the loadings to be non-negative. For this purpose three methods for estimation of the parameters $\boldsymbol{\lambda}$ and $\boldsymbol{\Psi}$ will be considered: maximum likelihood (ML), unweighted least squares (ULS), and generalized least squares (GLS). From the consistency of $\boldsymbol{S}$ for $\boldsymbol{\Sigma}$, it follows that these methods yield consistent estimates for the parameters (Anderson & Rubin, 1956; Browne, 1974; Dijkstra, 1981; Shapiro, 1983). That is, when the sample size is sufficiently large and the true values of the loadings are non-negative, their estimates will be non-negative with large probability. The sufficient conditions given below, however, hold for all sample sizes.

2 The sufficient conditions

Let $|\boldsymbol{\Sigma}|$ be the determinant of $\boldsymbol{\Sigma}$, $||\boldsymbol{\Sigma}||$ the Euclidian norm of $\boldsymbol{\Sigma}$, $\boldsymbol{S}$ the usual positive definite sample variance matrix, and $\boldsymbol{S}^{\frac{1}{2}}$ the unique symmetric square root of $\boldsymbol{S}$. $\boldsymbol{S}$ will be called "non-negative" when its elements are. The estimates for the parameters $\boldsymbol{\lambda}$ and $\boldsymbol{\Psi}$ are obtained by optimizing the below defined functions ML, ULS, and GLS over these parameters. For optimal parameters this implies that $\boldsymbol{\lambda}$ is optimal given $\boldsymbol{\Psi}$. It will be assumed that the optimmal parameters are in the parameter space. Specifically, it will be assumed that $\boldsymbol{\Psi}$ is diagonal positive definite and that the diagonal of $\boldsymbol{S} - \boldsymbol{\Psi}$ contains no negative elements. The latter means that the size of the error variance is less than the total variance of the variable. We will make frequently use of a simple version of Perron's theorem (Gantmacher, 1959, pp. 64-75): If a symmetric matrix is non-negative, then its first eigen value is strictly positive and strictly larger than all others and its corresponding eigenvector has no negative elements. Now we have the following.

Definition 1. ML estimates (Jöreskog & Lawley, 1968) minimize, over $\boldsymbol{\lambda}$ and $\boldsymbol{\Psi}$, the function

$$\ln|\boldsymbol{\lambda}\boldsymbol{\lambda}' + \boldsymbol{\Psi}| + \mathrm{tr}(\boldsymbol{\lambda}\boldsymbol{\lambda}' + \boldsymbol{\Psi})^{-1}\boldsymbol{S}. \tag{3}$$

Proposition 1. If $\boldsymbol{S}$ is non-negative, then the $\boldsymbol{\lambda}$ which minimizes (3) given $\boldsymbol{\Psi}$ is non-negative.

Proof. From the first order equations (Magnus & Neudecker, 1991, pp. 366-373), it follows that $\boldsymbol{\lambda} = d_1^{\frac{1}{2}}\boldsymbol{k}_1$, where d_1 and $\boldsymbol{k}_1$ are the first eigenvalue and eigenvector of $\boldsymbol{\Psi}^{-\frac{1}{2}}\boldsymbol{S}\boldsymbol{\Psi}^{-\frac{1}{2}}$, respectively. From $\boldsymbol{S}$ and $\boldsymbol{\Psi}$ non-negative, it follows that $\boldsymbol{\Psi}^{-\frac{1}{2}}\boldsymbol{S}\boldsymbol{\Psi}^{-\frac{1}{2}}$ is non-negative. It follows from Perron's theorem that $\boldsymbol{k}_1$ is non-negative. Hence, $\boldsymbol{\lambda}$ is non-negative. Q.E.D.

Definition 2. ULS estimates (Harman & Jones, 1966) minimize, over $\boldsymbol{\lambda}$ and $\boldsymbol{\Psi}$, the function

$$||\boldsymbol{S} - (\boldsymbol{\lambda}\boldsymbol{\lambda}' + \boldsymbol{\Psi})||^2. \tag{4}$$

Proposition 2. If $\boldsymbol{S}$ is non-negative, then the $\boldsymbol{\lambda}$ which minimizes (4) given $\boldsymbol{\Psi}$ is non-negative.

Proof. Given $\boldsymbol{\Psi}$ it follows that $\boldsymbol{\lambda}$ minimizes $\|(\boldsymbol{S}-\boldsymbol{\Psi})-\boldsymbol{\lambda\lambda}'\|^2$. It follows (cf Rao, 1973, p. 63) that $\boldsymbol{\lambda} = d_1^{\frac{1}{2}}\boldsymbol{k}_1$, where d_1 and $\boldsymbol{k}_1$ are the first eigenvalue and eigenvector of $\boldsymbol{S}-\boldsymbol{\Psi}$, respectively. From $\boldsymbol{S}-\boldsymbol{\Psi}$ non-negative and Perron's theorem, it follows that both $\boldsymbol{k}_1$ and d_1 are non-negative. Hence, $\boldsymbol{\lambda}$ is non-negative. Q.E.D.

Definition 3. GLS estimates (Jöreskog & Goldberger, 1972; Browne, 1974) minimize, over $\boldsymbol{\lambda}$ and $\boldsymbol{\Psi}$, the function

$$\mathrm{tr}[(\boldsymbol{S}-(\boldsymbol{\lambda\lambda}'+\boldsymbol{\Psi}))\boldsymbol{S}^{-1}]^2. \tag{5}$$

Proposition 3. If $\boldsymbol{S}^{\frac{1}{2}}$ is non-negative, then the $\boldsymbol{\lambda}$ which minimizes (5) given $\boldsymbol{\Psi}$ is non-negative.

Proof. Using the symmetry of the matrices involved, it follows that, given $\boldsymbol{\Psi}$, $\boldsymbol{\lambda}$ minimizes

$$\|\boldsymbol{S}^{-\frac{1}{2}}(\boldsymbol{S}-\boldsymbol{\Psi})\boldsymbol{S}^{-\frac{1}{2}}-\boldsymbol{S}^{-\frac{1}{2}}\boldsymbol{\lambda\lambda}'\boldsymbol{S}^{-\frac{1}{2}}\|^2. \tag{6}$$

It follows (cf Rao, 1973, p. 63) that $\boldsymbol{S}^{-\frac{1}{2}}\boldsymbol{\lambda} = d_1^{\frac{1}{2}}\boldsymbol{k}_1$, where d_1 and $\boldsymbol{k}_1$ are the first eigenvalue and eigenvector of $\boldsymbol{S}^{-\frac{1}{2}}(\boldsymbol{S}-\boldsymbol{\Psi})\boldsymbol{S}^{-\frac{1}{2}} = \boldsymbol{I}-\boldsymbol{S}^{-\frac{1}{2}}\boldsymbol{\Psi}\boldsymbol{S}^{-\frac{1}{2}}$, respectively. The first eigenvalue d_1 corresponds to the smallest eigenvalue of $\boldsymbol{S}^{-\frac{1}{2}}\boldsymbol{\Psi}\boldsymbol{S}^{-\frac{1}{2}}$. The latter eigenvalue corresponds to the largest eigenvalue of $\boldsymbol{S}^{\frac{1}{2}}\boldsymbol{\Psi}^{-1}\boldsymbol{S}^{\frac{1}{2}}$. From $\boldsymbol{S}^{\frac{1}{2}}$ and $\boldsymbol{\Psi}$ non-negative, it follows that $\boldsymbol{S}^{\frac{1}{2}}\boldsymbol{\Psi}^{-1}\boldsymbol{S}^{\frac{1}{2}}$ is non-negative. Hence, by Perron's theorem, $\boldsymbol{k}_1$ is non-negative. From this, $\boldsymbol{S}^{\frac{1}{2}}$ non-negative, and $\boldsymbol{\lambda} = d_1^{\frac{1}{2}}\boldsymbol{S}^{\frac{1}{2}}\boldsymbol{k}_1$, it follows that $\boldsymbol{\lambda}$ is non-negative. Q.E.D.

It may be noted that $\boldsymbol{S}^{\frac{1}{2}}$ non-negative implies that $\boldsymbol{S}^{\frac{1}{2}}\boldsymbol{S}^{\frac{1}{2}} = \boldsymbol{S}$ is non-negative. Thus the condition $\boldsymbol{S}^{\frac{1}{2}}$ non-negative is stronger than the condition S non-negative.

3 Conclusions

It can be concluded that, under weak assumptions, the conjecture that the loadings are non-negative in the one-factor model when the the sample variance matrix is non-negative is sustained. The non-negativity of the loadings makes interpretation simpler than when the loadings would constrast in sign. It is remarkable that the above results hold irrespective the sample size.

4 References

Anderson, T.W., & Rubin, H. (1956). Statistical inference in factor analysis. *Proceedings of the Third Berkeley Symposium*, 5, 111-150.

Browne, M.W. (1974). Generalized least squares estimates in the analysis of covariance structures. *South African Statistical Journal*, 8, 1-24.

Dijkstra, T.K. (1981). *Latent Variables in Linear Stochastic Models.* Amsterdam: Sociometric Research Foundation.

Gantmacher, F.R. (1959). *Applications of the theory of matrices.* New York: Interscience Publishers.

Harman, H.H., & Jones, W.H. (1966). Factor analysis by minimizing residuals (MinRes). *Psychometrika*, 31, 351-368.

Jöreskog, K.G. & Lawley, D.N. (1968). New methods in maximum likelihood factor analysis. *British Journal of Mathematical and Statistical Psychology*, 21, 85-96.

Jöreskog, K.G. (1971). Statistical analysis of sets of congeneric tests. *Psychometrika*, 36, 109-133.

Jöreskog, K.G. & Goldberger, A.S. (1972). Factor analysis by generalized least squares. *Psychometrika*, 37, 243-260.

Magnus, J.R. and Neudecker, H. (1991). *Matrix differential calculus with applications in statistics and economics.* Chichester: John Wiley and Sons.

Shapiro, A. (1983). Asymptotic distribution theory in the analysis of covariance structures (A unified approach) *South African Statistical Journal*, 17, 33-81.

Ontvangen: 12-10-1995
Geaccepteerd: 3-6-1996