Applied Mathematics E-Notes, 23(2023), 477-483 © ISSN 1607-2510
Available free at mirror sites of http://www.math.nthu.edu.tw/∼amen/

On Nonnegative Loading Matrices: Two-Factor Case*

Peter Zizler[†], Shoba Ittyipe[‡]

Received 30 December 2022

Abstract

Factor analysis describes variability among observed variables in terms of a fewer unobserved variables, referred to as factors. In the case of two independent factors we provide a straightforward condition when the factor loading matrix has all of its entries nonnegative. Results on such loading matrices being unique are established along with algorithms as how to find them.

1 Introduction

In general, factor loadings associated with the factor analysis problem are not unique which adds to the richness of the subject. In our paper we address the case of a factor analysis problem with any number of (observed) variables having two independent (unobserved) factor variables. We provide a straightforward answer when such a factor analysis problem, under certain conditions on the covariance matrix, yields unique loading matrices with all of their entries nonnegative with at least one entry being zero. This classification condition, drawn from the singular value decomposition of the covariance matrix, is easily checked. We provide an algorithm as how to construct these (unique) loading matrices.

For reader's convenience, we introduce the main ideas behind the statistical procedure called the factor analysis, an area to which our new results apply. For more detailed information on the foundations of factor analysis we refer the reader to [1] or [2]. A study of nonnegativity of factor loadings can be found, among others, in [4]. We motivate the reader by an example of a specific size (three variables and two factors) for simplicity of notation. The general case of m variables and n factors follows readily. Let X_1, X_2 and X_3 be given random variables assumed to be of zero mean and variance one, in particular, the variables are the z scores. Suppose the variability of these three variables are to be explained by two independent normalized factors f_1 and f_2. In particular, we assume for $i, j \in \{1, 2\}$

$$\langle f_i, f_j\rangle = 0 \text{ if } i \neq j \text{ and } \langle f_i, f_i\rangle = 1.$$

We determine the loadings on the factors f_1 and f_2 from the covariances among the variables $\{X_i\}_{i=1}^3$. In particular, we write

$$X_1 = \lambda_{11} f_1 + \lambda_{12} f_2 + \epsilon_1,$$
$$X_2 = \lambda_{21} f_1 + \lambda_{22} f_2 + \epsilon_2,$$
$$X_3 = \lambda_{31} f_1 + \lambda_{32} f_2 + \epsilon_3.$$

The random variables $\{\epsilon_i\}_{i=1}^3$ are assumed to be independent, normally distributed random variables with a certain variance $\langle \epsilon_i, \epsilon_i\rangle$ for each $i \in \{1, 2, 3\}$. We define

$$\Lambda = \begin{pmatrix} \lambda_{11} & \lambda_{12} \\ \lambda_{21} & \lambda_{22} \\ \lambda_{31} & \lambda_{32} \end{pmatrix}, \quad \mathbf{X} = \begin{pmatrix} X_1 \\ X_2 \\ X_3 \end{pmatrix}, \quad \mathbf{f} = \begin{pmatrix} f_1 \\ f_2 \end{pmatrix} \quad \text{and} \quad \boldsymbol{\epsilon} = \begin{pmatrix} \epsilon_1 \\ \epsilon_2 \\ \epsilon_3 \end{pmatrix}.$$

*Mathematics Subject Classifications: 42C10, 62P10.

[†]Department of Mathematics and Computing, Mount Royal University, Calgary, Canada
[‡]Department of Mathematics and Computing, Mount Royal University, Calgary, Canada

With this notation the above reads as

$$\mathbf{X} = \Lambda\mathbf{f} + \boldsymbol{\epsilon}.$$

The matrix Λ is referred to as the loading matrix. To determine the loadings λ_{ij} we will use the knowledge of the covariances of the variables X_i and X_j, in particular, we form the matrix

$$\begin{pmatrix} \langle X_1, X_1\rangle & \langle X_1, X_2\rangle & \langle X_1, X_3\rangle \\ \langle X_2, X_1\rangle & \langle X_2, X_2\rangle & \langle X_2, X_3\rangle \\ \langle X_3, X_1\rangle & \langle X_3, X_2\rangle & \langle X_3, X_3\rangle \end{pmatrix}.$$

We assume the random variables $\{\epsilon_i\}_{i=1}^3$ are independent of the factors, in particular we assume $\langle \epsilon_i, f_j\rangle = 0$ for all $i \in \{1, 2, 3\}$ and $j \in \{1, 2\}$.

The matrix $\Lambda\Lambda^T$ has at most two nonzero (positive) eigenvalues, since the size of the matrix Λ is 3×2. This observations comes from the singular value decomposition, see [3] for example. We seek to best approximation in the least squares sense

$$\begin{pmatrix} \langle X_1, X_1\rangle & \langle X_1, X_2\rangle & \langle X_1, X_3\rangle \\ \langle X_2, X_1\rangle & \langle X_2, X_2\rangle & \langle X_2, X_3\rangle \\ \langle X_3, X_1\rangle & \langle X_3, X_2\rangle & \langle X_3, X_3\rangle \end{pmatrix} \approx \Lambda\Lambda^T + \begin{pmatrix} \langle \epsilon_1, \epsilon_1\rangle & 0 & 0 \\ 0 & \langle \epsilon_2, \epsilon_2\rangle & 0 \\ 0 & 0 & \langle \epsilon_3, \epsilon_3\rangle \end{pmatrix}$$

or equivalently

$$\Lambda\Lambda^T \approx \begin{pmatrix} \langle X_1, X_1\rangle - \langle \epsilon_1, \epsilon_1\rangle & \langle X_1, X_2\rangle & \langle X_1, X_3\rangle \\ \langle X_2, X_1\rangle & \langle X_2, X_2\rangle - \langle \epsilon_2, \epsilon_2\rangle & \langle X_2, X_3\rangle \\ \langle X_3, X_1\rangle & \langle X_3, X_2\rangle & \langle X_3, X_3\rangle - \langle \epsilon_3, \epsilon_3\rangle \end{pmatrix} = C.$$

To illustrate the above, set $i = 2$, $j = 3$, and we compute

$$\begin{aligned} \langle X_2, X_3\rangle &= \langle \lambda_{21}f_1 + \lambda_{22}f_2 + \epsilon_2, \lambda_{31}f_1 + \lambda_{32}f_2 + \epsilon_3\rangle \\ &= \lambda_{21}\lambda_{31}\langle f_1, f_1\rangle + \lambda_{22}\lambda_{32}\langle f_2, f_2\rangle + \langle \epsilon_2, \epsilon_3\rangle \\ &= \lambda_{21}\lambda_{31} + \lambda_{22}\lambda_{32}. \end{aligned}$$

Similarly, set $i = 1$, $j = 1$, and we compute

$$\begin{aligned} \langle X_1, X_1\rangle &= \langle \lambda_{11}f_1 + \lambda_{12}f_2 + \epsilon_1, \lambda_{11}f_1 + \lambda_{12}f_2 + \epsilon_1\rangle \\ &= \lambda_{11}\lambda_{11}\langle f_1, f_1\rangle + \lambda_{12}\lambda_{12}\langle f_2, f_2\rangle + \langle \epsilon_1, \epsilon_1\rangle . \end{aligned}$$

Let m, n be natural numbers with $m > n \geq 2$. We will assume C is a $m \times m$ symmetric matrix whose n largest eigenvalues in absolute value are positive and distinct, listed in descending order $\left\{\sigma_i^2\right\}_{i=1}^n$. Let $\{\mathbf{u}_i\}_{i=1}^n$ be any choice of orthonormal eigenvectors corresponding to these eigenvalues, in particular $C\mathbf{u}_i = \sigma_i^2\mathbf{u}_i$ for $i \in \{1, 2, \ldots, n\}$. Form the matrix

$$U_n = [\sigma_1\mathbf{u}_1, \sigma_2\mathbf{u}_2, \ldots, \sigma_n\mathbf{u}_n]$$

the columns of U_n are the vectors $\{\mathbf{u}_i\}_{i=1}^n$ scaled by σ_i. The eigenvectors are unique up to a sign switch due to the assumption of distinct eigenvalues. It turns out the loading matrix can be attained as

$$\Lambda = U_n V^T = \sum_{i=1}^{n} \sigma_i \mathbf{u}_i \mathbf{v}_i^T$$

where V is an arbitrary $n \times n$ orthogonal matrix with $\{\mathbf{v}_i\}_{i=1}^n$, an orthonormal set of vectors in $\mathbf{R}^n$, as its columns. In many applications it is imperative we find factor loadings that are all nonnegative for the factor analysis to be useful. We emphasize that when such factor loadings are found, if they exists, they need not be unique. In the next section we address the question of existence and uniqueness of loading matrices with all entries nonnegative pertaining to factor analysis problems with two independent factors and any number of (observed) variables. The order of the factors $\{f_1, f_2\}$ can be permuted which results in the column permutation in the matrix U_2. When we say the matrix U_2 is associated with the factor analysis problem we understand the uniqueness of this matrix up to a column sign switch or column permutation. We will assume the variances $\langle \epsilon_i, \epsilon_i \rangle$ are known for all $i \in \{1, 2, \ldots, m\}$, possibly estimated using available statistical techniques, see [1] or [2].

2 Main Results

Consider a $m \times n$ real matrix $\Lambda = \Lambda_{m \times n}$. We say that Λ is a $m \times n$ optimal loading matrix for C if the following is minimized

$$\min_\Lambda ||C - \Lambda\Lambda^T||_F.$$

Here the Frobenius norm of the matrix $C = [c_{ij}]_{i,j=1,2,\ldots,m}$ is given by $||C||_F^2 = \sum_{i,j=1,2,\ldots,m} |c_{ij}|^2$. We observe that optimal loading matrices are not unique. The following is a well known result, we include its proof for reader's convenience. We refer the reader to [3] or [5] for a very nice exposition on the topic of a singular value decomposition of a matrix.

Let C be a $m \times m$ symmetric matrix whose n largest eigenvalues in absolute value are positive and distinct, listed in descending order $\left\{\sigma_i^2\right\}_{i=1}^n$. Any optimal $m \times n$ loading matrix for C is given by

$$\Lambda = \sum_i^n \sigma_i \mathbf{u}_i \mathbf{v}_i^T$$

with $\{\mathbf{u}_i\}_{i=1}^n$ being the corresponding normalized eigenvectors, unique up to a sign. The vectors $\{\mathbf{v}_i\}_{i=1}^n$ form an arbitrary orthonormal basis for $\mathbf{R}^n$.

To see this consider the singular value decomposition of the $m \times n$ matrix Λ written as $\Lambda = UDV^T$ where U is a $m \times m$ orthogonal matrix, V is a $n \times n$ orthogonal matrix and D is a $m \times n$ diagonal matrix. Observe that

$$\Lambda\Lambda^T = \left(UDV^T\right)\left(UDV^T\right)^T = UDV^TVD^TU^T = UD_1U^T$$

with D_1 being a $m \times m$ diagonal matrix with the only possible nonzero entries being the first n diagonal entries. To obtain a loading matrix that we seek, it follows from the properties of the singular value decomposition that the optimal choice for the diagonal entries in D_1 are the values $\left\{\sigma_i^2\right\}_{i=1}^n$. The vectors $\{\mathbf{u}_i\}_{i=1}^n$ are the corresponding normalized eigenvectors. These normalized eigenvectors are unique up to a sign since the matrix C is assumed to have the eigenvalues in question distinct.

If the matrix Λ, a $m \times n$ optimal loading matrix for C, has all of its entries nonnegative then we say we have a nonnegative $m \times n$ optimal loading matrix for C. Given two vectors $\mathbf{u}$ and $\mathbf{v}$ in $\mathbf{R}^p$ the inner product between these two vectors is denoted by $\langle \mathbf{u}, \mathbf{v} \rangle$.

Lemma 1 *Let $m > 2$ be an arbitrary natural number and $n = 2$. Consider vectors $\{\mathbf{u}_i\}_{i=1}^m$ in $\mathbf{R}^2$. The property $\langle \mathbf{u}_i, \mathbf{u}_j \rangle \geq 0$ for all $i, j \in \{1, 2, \ldots, m\}$ holds if and only if there exists an orthonormal basis $\{\mathbf{v}_1, \mathbf{v}_2\}$ for $\mathbf{R}^2$ with the property that $\langle \mathbf{u}_i, \mathbf{v}_j \rangle \geq 0$ for all $i \in \{1, 2, \ldots, m\}$ and $j \in \{1, 2\}$. Furthermore, suppose $\langle \mathbf{u}_i, \mathbf{u}_j \rangle \geq 0$ for all $i, j \in \{1, 2, \ldots, m\}$. Then the above orthonormal basis is unique if and only if there exists $\mathbf{u}_i$ and $\mathbf{u}_j$ such that $\langle \mathbf{u}_i, \mathbf{u}_j \rangle = 0$. In such a case we have $\{\mathbf{v}_1, \mathbf{v}_2\} = \{\mathbf{u}_i, \mathbf{u}_j\}$.*

Proof. Without loss of generality we can assume $\{\mathbf{u}_i\}_{i=1}^m$ are normalized. Suppose the property $\langle \mathbf{u}_i, \mathbf{u}_j \rangle \geq 0$ for all $i, j \in \{1, 2, \ldots, m\}$ holds. Choose any $\{\mathbf{u}_{i_0}, \mathbf{u}_{j_0}\}$ for which

$$\langle \mathbf{u}_{i_0}, \mathbf{u}_{j_0} \rangle = \min_{i,j \in \{1,2,\ldots,m\}} \left\{\langle \mathbf{u}_i, \mathbf{u}_j \rangle\right\}.$$

Let $\mathbf{v}_1 = \mathbf{u}_{i_0}$. The vector $\mathbf{v}_1$ is normalized and it will be the first vector in the orthonormal set we construct. Note $\langle \mathbf{u}_i, \mathbf{v}_1 \rangle \geq 0$ for all i. Define

$$V_1 = \mathbf{v}_1^\perp = \left\{\mathbf{u} \in \mathbf{R}^2 \mid \langle \mathbf{u}, \mathbf{v}_1 \rangle = 0\right\} = \text{span}\left\{\mathbf{w}\right\}$$

for some unit vector $\mathbf{w} \in \mathbf{R}^2$. Now observe that for all $i \in \{1, 2, \ldots, m\}$ either

$$\langle \mathbf{u}_i, \mathbf{w} \rangle \geq 0 \text{ or } \langle \mathbf{u}_i, -\mathbf{w} \rangle \geq 0.$$

Choose $\mathbf{v}_2 = \mathbf{w}$ or $\mathbf{v}_2 = -\mathbf{w}$ as the second vector in the orthonormal set we construct, select the choice that gives all inner products nonnegative. On the other hand, suppose there exists and orthonormal basis $\{\mathbf{v}_1, \mathbf{v}_2\}$ for $\mathbf{R}^2$ with the property that $\langle \mathbf{u}_i, \mathbf{v}_j \rangle \geq 0$ for all $i \in \{1, 2, \ldots, m\}$ and $j \in \{1, 2\}$. Then we can choose a rotation transformation to rotate the basis $\{\mathbf{v}_1, \mathbf{v}_2\}$ to coincide with the standard basis $(1, 0)^T$ and $(0, 1)^T$. The inner products $\langle \mathbf{u}_i, \mathbf{u}_j \rangle$ for all $i, j \in \{1, 2, \ldots, m\}$ are preserved upon rotation. The new rotated vectors have all their entries nonnegative and hence the inner products among them are nonnegative. The uniqueness result readily follows. ■

Theorem 1 *Let C be a $m \times m$ symmetric matrix whose two largest eigenvalues in absolute value are positive and distinct, denoted by $\left\{\sigma_1^2, \sigma_2^2\right\}$ in decreasing order. Let $\{\mathbf{u}_1, \mathbf{u}_2\}$ be any corresponding normalized eigenvectors in $\mathbf{R}^m$. Define a $m \times 2$ matrix*

$$U_2 = \begin{bmatrix} \sigma_1 \mathbf{u}_1, \sigma_2 \mathbf{u}_2 & \\ & \end{bmatrix}.$$

Then there exists a $m \times 2$ optimal loading matrix for C with all of its entries nonnegative if and only if the matrix $U_2 U_2^T$ has all of its entries nonnegative. In this case the nonnegative $m \times 2$ optimal loading matrix for C is unique (up to permutation of columns) if and only if the matrix $U_2 U_2^T$ has a zero entry.

Proof. Since we assume the matrix C has the eigenvalues in question distinct, the corresponding normalized eigenvectors $\{\mathbf{u}_1, \mathbf{u}_2\}$ are unique up to a sign. As a result the matrix $U_2 U_2^T$ is the same matrix regardless of the (sign) choice of $\{\mathbf{u}_1, \mathbf{u}_2\}$. We note the $m \times 2$ optimal loading matrix for C is given by $\Lambda = U_2 V^T$ where V^T is an arbitrary 2×2 orthogonal matrix. Invoking the Lemma above we observe the matrix $U_2 U_2^T$ has all of its entries nonnegative if and only if there exists an orthogonal 2×2 matrix V^T such that the matrix $U_2 V^T$ has all of its entries nonnegative. The uniqueness result follows readily. ■

Corollary 1 *Consider the $m \times 2$ factor analysis problem*

$$\mathbf{X} = \Lambda \mathbf{f} + \boldsymbol{\epsilon}$$

as above. Let

$$C = \begin{pmatrix} \langle X_1, X_1 \rangle - \langle \epsilon_1, \epsilon_1 \rangle & \langle X_1, X_2 \rangle & \langle X_1, X_3 \rangle \\ \langle X_2, X_1 \rangle & \langle X_2, X_2 \rangle - \langle \epsilon_2, \epsilon_2 \rangle & \langle X_2, X_3 \rangle \\ \langle X_3, X_1 \rangle & \langle X_3, X_2 \rangle & \langle X_3, X_3 \rangle - \langle \epsilon_3, \epsilon_3 \rangle \end{pmatrix}.$$

Assume C is a $m \times m$ symmetric matrix whose two largest eigenvalues in absolute value are positive and distinct listed as $\left\{\sigma_1^2, \sigma_2^2\right\}$ in decreasing order. Let $\{\mathbf{u}_1, \mathbf{u}_2\}$ be any corresponding normalized eigenvectors in $\mathbf{R}^m$. Then there exists a loading matrix Λ with all of its entries nonnegative if and only if the matrix $U_2 U_2^T$ has all of its entries nonnegative where

$$U_2 = \begin{bmatrix} \sigma_1 \mathbf{u}_1, \sigma_2 \mathbf{u}_2 & \\ & \end{bmatrix}.$$

Suppose the matrix $U_2 U_2^T$ has all of its entries nonnegative. Then

- *We have a unique loading matrix Λ with all of its entries nonnegative (up to a factor permutation) if and only if the matrix $U_2 U_2^T$ has a zero entry.*
- *If the matrix $U_2 U_2^T$ has all entries positive then there are just two possible loading matrices (up to a factor permutation) with all of its entries nonnegative along with the further property of having at least one zero entry in the loading matrix.*

Proof. Suppose $U_2U_2^T$ has all of its entries nonnegative. The matrix $U_2U_2^T$ has a zero entry if and only if there exist two vectors in $\mathbf{R}^2$ extracted from the rows in U_2 which when normalized are the unique orthonormal columns in the matrix V^T, unique up to a permutation of columns. Multiple zero entries in $U_2U_2^T$ yield the same matrix V^T up to a column permutation. Now suppose the matrix $U_2U_2^T$ has all of its entries positive. Extract $\mathbf{u}_{i_0}$ and $\mathbf{u}_{j_0}$, two rows in U_2, such that

$$\left\langle \mathbf{u}_{i_0}^T, \mathbf{u}_{j_0}^T \right\rangle = \min_{i,j\in\{1,2,\ldots,m\}} \left\{ \left\langle \mathbf{u}_i^T, \mathbf{u}_j^T \right\rangle \right\}$$

where $\mathbf{u}_i, \mathbf{u}_j$ are any rows in U_2. Set $\mathbf{v}_1 = \mathbf{u}_{i_0}^T$. Let $\mathbf{v}_2 = (\mathbf{v}_1(2), -\mathbf{v}_1(1))^T$. If $\langle \mathbf{v}_2, \mathbf{u}_i \rangle < 0$ for some i set $\mathbf{v}_2 = -\mathbf{v}_2$. Let V^T be the matrix whose columns are $\mathbf{v}_1$ and $\mathbf{v}_2$. The loading matrix is obtained as $\Lambda_1 = U_2V^T$. Similarly, set $\mathbf{v}_1 = \mathbf{u}_{j_0}$ and generate $\Lambda_2 = U_2V^T$. These two loading matrices Λ_1 and Λ_2 are unique matrices (up to column permutation) with all entries nonnegative with at least one zero entry. ■

3 Algorithm

The following is an algorithm to compute the loading matrices with all of their entries nonnegative with at least one zero entry. We assume we have generated the matrix U_2 from the singular value decomposition of the covariance matrix under the assumptions in the Theorem.

- If the matrix $U_2U_2^T$ has a negative entry, then no loading matrix with all entries nonnegative exists.

- Suppose the matrix $U_2U_2^T$ has all of its entries nonnegative and some entries are zero. Pick any (i, j) location for the zero entry. Select $\mathbf{u}_i$ and $\mathbf{u}_j$, the corresponding rows in U_2, normalize these vectors, keeping the same notation for these now normalized vectors. Denote $\mathbf{v}_1 = \mathbf{u}_i$ and $\mathbf{v}_2 = \mathbf{u}_j$. These are (up to a column permutation) the columns in the desired (unique) matrix V^T. The desired loading matrix is obtained as $\Lambda = U_2V^T$.

- Suppose the matrix $U_2U_2^T$ has all of its entries positive. Extract the vectors $\{\mathbf{u}_i\}_{i=1}^m$ which are the rows in U_2 and normalized them. Observe $\langle \mathbf{u}_i, \mathbf{u}_j \rangle > 0$ for all $i, j \in \{1, 2, \cdots, m\}$. We find the unique vectors $\mathbf{u}_{i_0}$ and $\mathbf{u}_{j_0}$ (up to a permutation) such that

  $$\langle \mathbf{u}_{i_0}, \mathbf{u}_{j_0} \rangle = \min_{i,j\in\{1,2,\ldots,m\}} \{\langle \mathbf{u}_i, \mathbf{u}_j \rangle\}$$

  using the following algorithm

  – Set $i_0 = 1$ and $j_0 = 2$. Select $j = 3$ and find $k, l \in \{i_0, j_0, j\}$ so that

  $$\langle \mathbf{u}_k, \mathbf{u}_l \rangle = \min \{\langle \mathbf{u}_{i_0}, \mathbf{u}_{j_0} \rangle, \langle \mathbf{u}_{i_0}, \mathbf{u}_j \rangle, \langle \mathbf{u}_{j_0}, \mathbf{u}_j \rangle\}.$$

  – Update $i_0 = k$ and $j_0 = l$. Set $j = 4$ and continue inductively until $j = m$.

- Let i_0, j_0 be the values generated above. Set $\mathbf{v}_1 = \mathbf{u}_{i_0}$. Let $\mathbf{v}_2 = (\mathbf{v}_1(2), -\mathbf{v}_1(1))^T$. If $\langle \mathbf{v}_2, \mathbf{u}_i \rangle < 0$ for some i then set $\mathbf{v}_2 = -\mathbf{v}_2$. Set V^T to be the matrix whose columns are $\mathbf{v}_1$ and $\mathbf{v}_2$. The loading matrix is obtained as $\Lambda_1 = U_2V^T$.

- Set $\mathbf{v}_1 = \mathbf{u}_{j_0}$, repeat as above, and generate $\Lambda_2 = U_2V^T$.

The loading matrices Λ_1 and Λ_2, generated above, are the two unique (up to a column permutation) loading matrices with all of their entries nonnegative and with at least one entry being zero.

4 Examples

Imagine we are given test scores, in the form of z scores, from a large group of students, their test scores in chemistry, the vector X_1, their test scores in biology, the vector X_2, and their test scores in social science, the vector X_3. Consider the following hypothetical covariance matrix for the variables $\{X_i\}_{i=1}^3$

$$\begin{pmatrix} 1.00 & 0.70 & 0.20 \\ 0.70 & 1.00 & 0.40 \\ 0.20 & 0.40 & 1.00 \end{pmatrix}.$$

For example, the entry value 0.7 in the $(2,1)$ entry in the covariance matrix measures the covariance between the biology and chemistry test scores. Similarly, the entry value 0.4 in the $(3,2)$ entry in the covariance matrix measures the covariance between the social science and biology test scores. Though hypothetical, these covariances are not surprising, due to the nature of the subjects. Assume that 80 % of the test score performance in the respective disciplines is determined by two unobserved variables (factors). These two (unobserved) variables can be thought of as the quantitative ability factor and the qualitative ability factor. The remaining 20 % of the test score performance in the respective disciplines is determined by random outcomes. In particular, we assume $\langle \epsilon_i, \epsilon_i \rangle = 0.2$ for $i \in \{1, 2, 3\}$. We have

$$C = \begin{pmatrix} 1.00 & 0.70 & 0.20 \\ 0.70 & 1.00 & 0.40 \\ 0.20 & 0.40 & 1.00 \end{pmatrix} - \begin{pmatrix} 0.20 & 0.00 & 0.00 \\ 0.00 & 0.20 & 0.00 \\ 0.00 & 0.00 & 0.20 \end{pmatrix} = \begin{pmatrix} 0.80 & 0.70 & 0.20 \\ 0.70 & 0.80 & 0.40 \\ 0.20 & 0.40 & 0.80 \end{pmatrix}.$$

The eigenvalues of C are $\{1.7023, 0.6306, 0.0671\}$. The largest two eigenvalues of C with a choice of corresponding eigenvectors are

$$\sigma_1^2 = 1.7023 \ ; \ \sigma_2^2 = 0.6306$$

$$\mathbf{u}_1 = (0.6110, 0.6646, 0.4301)^T \ ; \ \mathbf{u}_2 = (0.4663, 0.1369, -0.8740)^T.$$

We set

$$U_2 = [\sigma_1 \mathbf{u}_1, \sigma_2 \mathbf{u}_2] = \begin{pmatrix} 0.7971 & 0.3703 \\ 0.8672 & 0.1087 \\ 0.5611 & -0.6940 \end{pmatrix}$$

and obtain

$$U_2 U_2^T = \begin{pmatrix} 0.7725 & 0.7315 & 0.1903 \\ 0.7315 & 0.7638 & 0.4111 \\ 0.1903 & 0.4111 & 0.7966 \end{pmatrix}$$

which implies the existence of a loading matrix with all of its entries nonnegative. We obtain the following two loading matrices with a zero entry (unique up to factor permutation)

$$\Lambda_1 = \begin{pmatrix} 0.8789 & 0.0000 \\ 0.8327 & 0.2668 \\ 0.2165 & 0.8658 \end{pmatrix} \text{ or } \Lambda_2 = \begin{pmatrix} 0.2132 & 0.8526 \\ 0.4607 & 0.7427 \\ 0.8924 & 0.0000 \end{pmatrix}.$$

These loading matrices can be understood as follows. If we had to assign a full quantitative loading to one discipline, it would be chemistry with the loading value of 0.8789. This yields the matrix Λ_1. As a result we have the unique consequences for the performance in biology with the quantitative loading of 0.8327 and the qualitative loading of 0.2668. The social science performance has the quantitative loading of 0.2165 and the qualitative loading of 0.8658. The smallest eigenvalue of C is significantly smaller than the largest two, the use of two factors would be appropriate here.

If, on the other hand, we had to assign a full qualitative loading to one discipline it would be social science with the loading value of 0.8924, this yields the matrix Λ_2. The factors are permuted in the loading matrix Λ_2 with the first column representing the qualitative factor and the second column the quantitative factor. In this instant we have the unique consequences for the performance in biology with the quantitative loading of 0.7427 and the qualitative loading of 0.4607. The chemistry performance has the quantitative loading of 0.8526 and the qualitative loading of 0.2132.

Now consider the same example as above except with a different covariance matrix for the variables $\{X_i\}_{i=1}^3$

$$\begin{pmatrix} 1 & 0.01 & 0.02 \\ 0.01 & 1 & 0.05 \\ 0.02 & 0.05 & 1 \end{pmatrix}.$$

The variances for the random variables are $\langle \epsilon_i, \epsilon_i \rangle = 0.02$ for all $i \in \{1, 2, 3\}$. We have

$$C = \begin{pmatrix} 0.98 & 0.01 & 0.02 \\ 0.01 & 0.98 & 0.05 \\ 0.02 & 0.05 & 0.98 \end{pmatrix}.$$

The eigenvalues are C are $\{1.0378, 0.9732, 0.9289\}$. The largest two eigenvalues of C with a choice of corresponding eigenvectors are

$$\sigma_1^2 = 1.0378 \; ; \; \sigma_2^2 = 0.9732$$

$$\mathbf{u}_1 = (0.3467, 0.6470, 0.6791)^T \; ; \; \mathbf{u}_2 = (0.9259, -0.3517, -0.1376)^T.$$

We set

$$U_2 = [\sigma_1 \mathbf{u}_1, \sigma_2 \mathbf{u}_2] = \begin{pmatrix} 0.3532 & 0.9135 \\ 0.6591 & -0.3470 \\ 0.6919 & -0.1357 \end{pmatrix}$$

and obtain

$$U_2 U_2^T = \begin{pmatrix} 0.9591 & -0.0842 & 0.1204 \\ -0.0842 & 0.5548 & 0.5031 \\ 0.1204 & 0.5031 & 0.4971 \end{pmatrix}$$

indicating the non-existence of a loading matrix with all of its entries nonnegative. The three eigenvalues of C are close to each other, the use of two factors would not be appropriate here. Still it is of importance to note, should two factors be used, the existence of a loading matrix with all its entries nonnegative is impossible here.

References

[1] A. L. Comrey and H. B. Lee, A First Course in Factor Analysis, Psychology Press, New York, 2013.

[2] J.-O. Kim and C. W. Mueller, Factor Analysis: Statistical Methods and Practical Issues, SAGE Publications, 1978.

[3] P. Lancaster and M. Tismenetsky, The Theory of Matrices, Academic Press, 1985.

[4] S. Pudney, Factor Rotation with Non-negativity Constraints, The Institute for Fiscal Studies Department of Economics, UCL, cemmap working paper CWP, 2011.

[5] G. Strang, Linear Algebra and Its Applications, Cengage, previously Brooks/Cole, 2006.