REPORTS@SCM

AN ELECTRONIC JOURNAL OF THE SOCIETAT CATALANA DE MATEMÀTIQUES

Factor analysis: Existence of solution to factor models

*Adrià Prior Rovira
Universitat Autònoma de Barcelona
adriaprior2@gmail.com

∗Corresponding author

Resum *(CAT)*
Un dels resultats principals de l'anàlisi factorial afirma que si el model factorial se satisfà per a un vector aleatori X, aleshores la matriu de covariància del vector admet una descomposició en termes de les matrius que caracteritzen el model, i que el recíproc també és cert sota condicions força generals. La implicació directa es troba demostrada en moltes referències, però la prova del recíproc sembla difícil de trobar en els textos disponibles. La raó d'aquest article és compartir una demostració del recíproc concebuda per l'autor, primerament pel cas del model factorial ortogonal, àmpliament usat en anàlisi factorial exploratòria i, en segon lloc, per un model factorial que generalitza l'ortogonal i està pensat per a ser utilitzat en anàlisi factorial confirmatòria.

Abstract *(ENG)*
One of the main results in factor analysis states that if the factor model holds for a random vector X, then the covariance matrix of the vector admits a decomposition in terms of the matrices that characterize the model, and that the converse is also true under quite general conditions. The direct implication can be found proved in many references but the proof of the converse seems difficult to find in the available texts. The reason of this article is to share an original proof of the converse, first for the case of the orthogonal factor model, widely used in exploratory factor analysis, and secondly for a factor model that generalizes the orthogonal one, and which is meant to be used in confirmatory factor analysis.

Keywords: *factor analysis, exploratory factor analysis, confirmatory factor analysis, factor model, fundamental theorem of factor analysis.*

MSC (2010): *Primary 55M25, 57P10. Secondary 55P15, 57R19, 57N15.*

Received: *March 8, 2021.*

Accepted: *May 17, 2021.*

Acknowledgement

I would like to thank the anonymous referees for carefully reading the manuscript and for their useful corrections and suggestions.

Societat Catalana de Matemàtiques

Institut d'Estudis Catalans

1. Introduction

Factor analysis is a statistical theory that allows, under certain conditions, to express each variable of an observable random vector as a linear combination of a few new variables called factors, through a stochastic model. The so called factor model is adjusted to the observable vector, which we shall call initial vector, using a set of observations of it.

This technique is used in different fields such as psychology, sociology, economy or political sciences, and also in the physical sciences and biosciences. The factors may explain the initial variables in a simplified way, and they should be of theoretical interest in the specific research setting. In fact, factor analysis is usually applied to investigate concepts that can't be measured directly, like intelligence, social status or sustainable progress, by collapsing a large number of variables related to those concepts into a reduced group of latent factors.

There exists a vast literature in factor analysis that treats the applied aspects of the technique, including a variety of examples; see Mardia et al. ([5, pp. 255–280]) for a consistent introduction. One of the main results in this frame says that if the factor model holds for the initial random vector, then its covariance matrix can be decomposed in a particular way, and that the converse is also true under quite general conditions. The direct implication is proven in many references, but the proof of the converse cannot be easily found in the available texts, despite it is stated in several ones. The author proved this result after unsuccessfully seeking a complete detailed proof, following a hint given in Mardia et al. [5]. The reason of this paper is to share this proof, so that it can be accessible to anybody interested in factor analysis from the mathematical point of view.

First, we will prove the converse for the case of the orthogonal factor model, which is commonly used in exploratory factor analysis. Secondly, we will extend the result to a factorial model in which correlations between factors will be allowed, being this way adequate to use in confirmatory factor analysis, we will call this second model generalized factor model. Exploratory factor analysis is probably the most known version of the technique, and it is used to find a factor model that fits the initial vector, whereas the confirmatory variant is usually performed after an exploratory analysis, with the aim of fitting a specific factor model such that some of the parameters values are predetermined in advance by the researchers.

We will name *Theorem of existence of solution to the orthogonal factor model* and *Theorem of existence of solution to the generalized factor model* the two results, which are precisely stated and proved in Sections 2 and 3, respectively. Both theorems are important in practice because they ensure that the factor model is valid when the sample covariance matrix fits well to a given pattern, which describes the relations between the initial variables and the latent unobserved factors, as well as the relations between the factors.

2. Existence of solution to the orthogonal factor model

We begin with the orthogonal case, by defining what is a solution to the orthogonal factor model for a random vector X.

Definition 2.1. Let $X^t = (X_1, \dots, X_p)$ be a $p \times 1$ random vector with $\mathrm{E}[X] = 0_{p\times 1}$ and $\mathrm{E}[X^2] < \infty$[1]. We say that the orthogonal factor model holds for X if there exist two random vectors $f^t = (f_1, \dots, f_m)$ with $m < p$ and $u^t = (u_1, \dots, u_p)$ and a matrix $Q = (q_{ij})_{ij} \in M_{p\times m}(\mathbb{R})$, such that

$$\begin{cases} X_1 = q_{11} f_1 + q_{12} f_2 + \cdots + q_{1m} f_m + u_1 \\ X_2 = q_{21} f_1 + q_{22} f_2 + \cdots + q_{2m} f_m + u_2 \\ \qquad\qquad \vdots \\ X_p = q_{p1} f_1 + q_{p2} f_2 + \cdots + q_{pm} f_m + u_p \end{cases} \tag{1}$$

and satisfying the following conditions:

(i) $\mathrm{E}[f] = 0_{m\times 1}$, $\mathrm{Cov}(f) = I_m$, with I_m the identity matrix on $\mathbb{R}^m$.

(ii) $\mathrm{E}[u] = 0_{p\times 1}$, $\mathrm{Cov}(u) = \Psi$, with Ψ a diagonal matrix in $M_p(\mathbb{R})$.

(iii) $\mathrm{Cov}(f, u) = 0_{m\times p}$, where $\mathrm{Cov}(f, u)$ denotes the cross-covariance matrix between f and u.

In this case we say that the triplet (Q, f, u) is a solution to the orthogonal factor model for X, $(f_1, \dots, f_m)$ are called the common factors of the model, $(u_1, \dots, u_p)$ are called the specific factors, the matrix Q is called the loadings matrix and the elements on the diagonal of Ψ are named specific variances.

The model equations system (1) can be written as $X = Qf + u$. In practice, X is the observed random vector for which we want to fit the model, the assumption $\mathrm{E}[X] = 0_{p\times 1}$ is not restrictive since data can be centered to get the model and translated to the original center at the end, if necessary. We demand $m < p$ because one of the objectives of factor analysis is explaining the initial variables in a simplified way with a few common factors. The specific factors can be understood as the stochastic error terms in regression. Condition (i) asks the common factors to be uncorrelated and have unit variance. Is in this sense that we call the model orthogonal, considering the covariance as a scalar product. Conditions (ii) and (iii) ask the specific factors to be uncorrelated one to each other and uncorrelated to the common factors. These last two assumptions seem natural, in the sense that the common factors capture and explain a part of the variability of each initial variable, letting the remaining amount to the specific factors.

To clarify notation, we will use Σ_X as well as $\mathrm{Cov}(X)$ to denote the covariance matrix of a random vector X, depending on the situation, that is, $\Sigma_X = \mathrm{Cov}(X)$.

The next proposition is sometimes called *The fundamental theorem of factor analysis* and it shows us a necessary condition for the model to have solution in the sense of Definition 2.1. Similar proofs as the given below can be found in the literature, for example in Härdle and Simar ([2, p. 310]).

Proposition 2.2. *Let X be a random vector with $\mathrm{E}[X] = 0_{p\times 1}$ and $\mathrm{E}[X^2] < \infty$. If (Q, f, u) is a solution to the orthogonal factor model for X, with $\mathrm{Cov}(u) = \Psi$, then*

$$\Sigma_X = QQ^t + \Psi. \tag{2}$$

Proof. Using basic properties of the covariance matrix and Definition 2.1, it is clear that

$$\begin{aligned} \Sigma_X = \mathrm{Cov}(X) = \mathrm{Cov}(Qf + u) &= \mathrm{Cov}(Qf) + \mathrm{Cov}(Qf, u) + \mathrm{Cov}(u, Qf) + \mathrm{Cov}(u) \\ &= QQ^t + \Psi. \end{aligned}$$

□

[1] Let $X^t = (X_1, \dots, X_p)$ be a $p \times 1$ random vector. We say that $\mathrm{E}[X^2] < \infty$ if $\mathrm{E}[X_j^2] < \infty$ for all $j \in \{1, \dots, p\}$. We demand $\mathrm{E}[X^2] < \infty$ to ensure existence of the covariances between the variables in X.

The theorem of existence of solution we are interested in is the converse of the above proposition, and it states that condition (2) is sufficient provided that Ψ is positive definite. The proof of this result seems difficult to find in the existing literature. An original proof is given below, which uses some results on block matrices and basic knowledge on linear algebra and probability.

Theorem 2.3 (Existence of solution to the orthogonal factor model)**.** *Let X be a $p \times 1$ random vector with $\mathrm{E}[X] = 0_{p\times 1}$ and $\mathrm{E}[X^2] < \infty$. Assume that there exist two matrices $Q \in M_{p\times m}(\mathbb{R})$, with $m < p$, and $\Psi \in M_p(\mathbb{R})$ diagonal and positive definite, such that $\Sigma_X = QQ^t + \Psi$. Then, there exist two random vectors $f^t = (f_1, \dots, f_m)$ and $u^t = (u_1, \dots, u_p)$ satisfying the orthogonal factor model for X, with loadings matrix Q and $\mathrm{Cov}(u) = \Psi$.*

Proof. Following the hint given in Mardia et al. ([5, p. 276]), we will show first that there exists a multivariate normal random vector $Y^t = (Y_1, \dots, Y_m)$ with $Y \sim N_m(0_{m\times 1}, I_m + Q^t\Psi^{-1}Q)$, and then we will show that the pair of random vectors defined by

$$\begin{pmatrix} u \\ f \end{pmatrix} := \underbrace{\begin{pmatrix} I_p & Q \\ -Q^t\Psi^{-1} & I_m \end{pmatrix}^{-1}}_{A^{-1}} \begin{pmatrix} X \\ Y \end{pmatrix} \tag{3}$$

are a solution to the orthogonal factor model.

Take $W := I_m + Q^t\Psi^{-1}Q$, which is well defined since $\Psi = \mathrm{diag}(\psi_1, \dots, \psi_p)$ with $\psi_i > 0, \forall\, i \in \{1, \dots, p\}$. First of all, let's see that W is symmetric and positive definite, and therefore we can consider a multivariate normal vector Y with covariance matrix W.

Indeed, $Q^t\Psi^{-1}Q$ and I_m are symmetric and hence W is. Let $v \in M_{m\times 1}(\mathbb{R})$ be any vector and take $y := Qv$. We have $v^t\, Q^t\Psi^{-1}Qv = y^t\Psi^{-1}y \geq 0$. Therefore $v^tWv = v^tv + v^t\, Q^t\Psi^{-1}Qv > 0$ for any non null $v \in M_{m\times 1}(\mathbb{R})$, and W is positive definite.

Now, let's see that the matrix A in (3) is invertible. A is a square matrix and I_m is invertible, hence *Schur's determinant formula* (Schur, [6]) applies to obtain

$$\det(A) = \det\begin{pmatrix} I_p & Q \\ -Q^t\Psi^{-1} & I_m \end{pmatrix} = \det(I_m)\det(A/I_m) = \det(I_p + QQ^t\Psi^{-1}) = \det(\Psi + QQ^t)\det(\Psi^{-1}),$$

where $A/I_m = I_p + QQ^t\Psi^{-1}$ is the Schur complement of I_m in A. Then, A is invertible if and only if $\Psi + QQ^t = \Sigma_X$ is. As QQ^t is positive semidefinite and Ψ is positive definite, $\Psi + QQ^t$ is positive definite, so A is invertible.

After this technical details, we are ready to prove that the factors in (3) give a solution to the model. Since A is invertible, we have:

$$\begin{pmatrix} u \\ f \end{pmatrix} = \begin{pmatrix} I_p & Q \\ -Q^t\Psi^{-1} & I_m \end{pmatrix}^{-1} \begin{pmatrix} X \\ Y \end{pmatrix} \iff \begin{pmatrix} X \\ Y \end{pmatrix} = \underbrace{\begin{pmatrix} I_p & Q \\ -Q^t\Psi^{-1} & I_m \end{pmatrix}}_{A} \begin{pmatrix} u \\ f \end{pmatrix}.$$

Clearly, $X = Qf + u$, we must show that f and u satisfy conditions (i), (ii) and (iii) of Definition 2.1. To see that (ii) holds observe that

$$\mathrm{E}\begin{pmatrix} u \\ f \end{pmatrix} = \mathrm{E}A^{-1}\begin{pmatrix} X \\ Y \end{pmatrix} = A^{-1}\mathrm{E}\begin{pmatrix} X \\ Y \end{pmatrix} = 0_{(p+m)\times 1}$$

thus $\mathrm{E}[u] = 0_{p\times 1}$ and $\mathrm{E}[f] = 0_{m\times 1}$. Now, take $M := I_p + QQ^t\Psi^{-1} = A/I_m$. Since A and I_m are invertible, M is too by the *Schur's determinant formula* and we can apply a *Banachiewicz inversion formula* (Banachiewicz, [1]) to obtain the inverse

$$A^{-1} = \begin{pmatrix} I_p & Q \\ -Q^t\Psi^{-1} & I_m \end{pmatrix}^{-1} = \begin{pmatrix} M^{-1} & -M^{-1}Q \\ Q^t\Psi^{-1}M^{-1} & I_m - Q^t\Psi^{-1}M^{-1}Q \end{pmatrix}.$$

Thus,

$$\begin{pmatrix} u \\ f \end{pmatrix} = A^{-1}\begin{pmatrix} X \\ Y \end{pmatrix} = \begin{pmatrix} M^{-1} & -M^{-1}Q \\ Q^t\Psi^{-1}M^{-1} & I_m - Q^t\Psi^{-1}M^{-1}Q \end{pmatrix}\begin{pmatrix} X \\ Y \end{pmatrix}$$

and therefore,

$$u = M^{-1}(X - QY). \tag{4}$$

Now, $\mathrm{Cov}(u) = \mathrm{Cov}(M^{-1}(X - QY)) = M^{-1}\,\mathrm{Cov}(X - QY)(M^{-1})^t$, and developing the covariance:

$$\begin{aligned} \mathrm{Cov}(X - QY) &= \mathrm{Cov}(X) + \mathrm{Cov}(X, -QY) + \mathrm{Cov}(-QY, X) + \mathrm{Cov}(-QY) \\ &= \Sigma_X + Q\Sigma_Y Q^t \\ &= (\Psi + QQ^t) + Q(I_m + Q^t\Psi^{-1}Q)Q^t \\ &= (I_p + QQ^t\Psi^{-1})(\Psi + QQ^t) \\ &= (I_p + QQ^t\Psi^{-1})\Psi(I_p + \Psi^{-1}QQ^t) = M\Psi M^t. \end{aligned}$$

Where we have used $\Sigma_X = QQ^t + \Psi$ by hypothesis, and $\mathrm{Cov}(X, Y) = 0_{p\times m}$ since Y is taken independently of X, hence $\mathrm{Cov}(u) = M^{-1}M\Psi M^t(M^t)^{-1} = \Psi$ and (ii) is proven.

Using similar arguments we will prove that $\mathrm{Cov}(f) = I_m$. Provided that A and I_p are invertible we now use another *Banachiewicz inversion formula* to get the following expression:

$$A^{-1} = \begin{pmatrix} I_p & Q \\ -Q^t\Psi^{-1} & I_m \end{pmatrix}^{-1} = \begin{pmatrix} I_p - QW^{-1}Q^t\Psi^{-1} & -QW^{-1} \\ W^{-1}Q^t\Psi^{-1} & W^{-1} \end{pmatrix},$$

where $W = I_m + Q^t\Psi^{-1}Q = A/I_p$ is the Schur complement of I_p in A. This way,

$$\begin{pmatrix} u \\ f \end{pmatrix} = A^{-1}\begin{pmatrix} X \\ Y \end{pmatrix} = \begin{pmatrix} I_p - QW^{-1}Q^t\Psi^{-1} & -QW^{-1} \\ W^{-1}Q^t\Psi^{-1} & W^{-1} \end{pmatrix}\begin{pmatrix} X \\ Y \end{pmatrix}$$

and we obtain

$$f = W^{-1}(Q^t\Psi^{-1}X + Y). \tag{5}$$

Then, $\mathrm{Cov}(f) = W^{-1}\,\mathrm{Cov}(Q^t\Psi^{-1}X + Y)(W^{-1})^t$ and using again that $\mathrm{Cov}(X, Y) = 0_{p\times m}$, it follows that

$$\begin{aligned} \mathrm{Cov}(Q^t\Psi^{-1}X + Y) &= Q^t\Psi^{-1}\Sigma_X\Psi^{-1}Q + \Sigma_Y \\ &= Q^t\Psi^{-1}(QQ^t + \Psi)\Psi^{-1}Q + W \\ &= (Q^t\Psi^{-1}Q + I_m)(Q^t\Psi^{-1}Q) + W \\ &= W(Q^t\Psi^{-1}Q) + W \\ &= W(Q^t\Psi^{-1}Q + I_m) = WW. \end{aligned}$$

Therefore, using that W is symmetric, we have $\mathrm{Cov}(f) = W^{-1}WW(W^{-1})^t = W^{-1}WWW^{-1} = I_m$ and (i) holds.

Finally, let us see that (iii) holds, that is, $\mathrm{Cov}(u, f) = 0_{p\times m}$. Using the expressions (4) and (5) we have:

$$\begin{aligned}
\mathrm{Cov}(u, f) &= \mathrm{Cov}(M^{-1}(X - QY), W^{-1}(Q^t\Psi^{-1}X + Y)) \\
&= M^{-1}\,\mathrm{Cov}(X - QY, Q^t\Psi^{-1}X + Y)(W^{-1})^t \\
&= M^{-1}[\mathrm{Cov}(X, X)\Psi^{-1}Q - Q\,\mathrm{Cov}(Y, Y)]W^{-1} \\
&= M^{-1}[\Sigma_X\Psi^{-1}Q - Q\Sigma_Y]W^{-1},
\end{aligned}$$

but the term in square brackets is null, that is,

$$\Sigma_X\Psi^{-1}Q - Q\Sigma_Y = (QQ^t + \Psi)\Psi^{-1}Q - QW = Q(Q^t\Psi^{-1}Q + I_m) - QW = QW - QW = 0_{p\times m}.$$

So $\mathrm{Cov}(u, f) = 0_{p\times m}$, and the proof is complete. □

Under the hypotheses of Theorem 2.3 the existence of a solution holds, but the solution is not unique, in fact, every orthogonal matrix provides another solution. Next proposition states this result (see Mardia et al. for a proof, [5, pp. 257–258]).

Proposition 2.4. *Let X be a random vector with $E[X] = 0_{p\times 1}$ and $E[X^2] < \infty$, let $m < p$ and let $G \in M_m(\mathbb{R})$ be an orthogonal matrix, that is, $G^tG = GG^t = I_m$. If (Q, f, u) is a solution to the orthogonal m-factor model for X, then (QG, G^tf, u) is a solution too.*

In particular, the result holds when $G \in M_m(\mathbb{R})$ is orthogonal and $\det(G) = 1$, that is, when G is a rotation matrix in $\mathbb{R}^m$.

Theorem 2.3 indicates how to proceed to adjust the model to a given observed vector $X^t = (X_1, \dots, X_p)$. In practice we have a data matrix $\widetilde{X} \in M_{n\times p}(\mathbb{R})$, where each row of $\widetilde{X}$ is an observation of X, and we estimate Σ_X by the sample covariance matrix S, then, our objective is to find matrices $\widehat{Q} \in M_{p\times m}(\mathbb{R})$ and $\widehat{\Psi} \in M_p(\mathbb{R})$, with $\widehat{\Psi}$ diagonal and positive definite, such that the equality

$$S = \widehat{Q}\widehat{Q}^t + \widehat{\Psi}$$

holds, at least approximately. If we find such matrices $\widehat{Q}$ and $\widehat{\Psi}$, they can be taken as loadings and specific variances estimates and so, they give rise to an estimated solution to the orthogonal factor model.

Estimates $\widehat{Q}$ and $\widehat{\Psi}$ are found using numerical methods currently implemented in statistical software environments like **R**, being the Maximum Likelihood Estimation (MLE) and the least squares methods two popular examples (Jöreskog, [3] and [4]). In exploratory factor analysis, the solution estimated by these methods may not be useful enough for the researchers, in the sense that the factors may load on too many variables and it could be difficult to interpret them. In this case, there exist methods (varimax, orthomax and others) that aim to provide a rotation matrix G such that the rotated factors, given by Proposition 2.4, may be more relevant for the ongoing investigation. From this point of view, the non uniqueness of solution is not a drawback.

3. The generalized factor model: existence of solution

The generalized factor model will allow correlations between the common factors, which is a less restrictive and so more realistic assumption in many settings. We define:

Definition 3.1. Let $X^t = (X_1, \ldots, X_p)$ be a $p \times 1$ random vector with $\mathrm{E}[X] = 0_{p\times 1}$ and $\mathrm{E}[X^2] < \infty$. We say that the generalized factor model holds for X if there exist two random vectors $f^t = (f_1, \ldots, f_m)$ with $m < p$ and $u^t = (u_1, \ldots, u_p)$ and a matrix $Q = (q_{ij})_{ij} \in M_{p\times m}(\mathbb{R})$, such that

$$\begin{cases} X_1 = q_{11}f_1 + q_{12}f_2 + \cdots + q_{1m}f_m + u_1 \\ X_2 = q_{21}f_1 + q_{22}f_2 + \cdots + q_{2m}f_m + u_2 \\ \qquad\qquad\qquad \vdots \\ X_p = q_{p1}f_1 + q_{p2}f_2 + \cdots + q_{pm}f_m + u_p \end{cases}$$

and satisfying the following conditions:

(i) $\mathrm{E}[f] = 0_{m\times 1}$, $\mathrm{Cov}(f) = \Theta$, with $\Theta \in M_m(\mathbb{R})$ symmetric and positive semidefinite.

(ii) $\mathrm{E}[u] = 0_{p\times 1}$, $\mathrm{Cov}(u) = \Psi$, with Ψ a diagonal matrix in $M_p(\mathbb{R})$.

(iii) $\mathrm{Cov}(f, u) = 0_{m\times p}$.

In this case we say that the triplet (Q, f, u) is a solution to the generalized factor model for X. The common factors in f are also called "latent" or "hidden" factors for X.

Proposition 3.2. *Let X be a random vector with $E[X] = 0_{p\times 1}$ and $\mathrm{E}[X^2] < \infty$. If (Q, f, u) is a solution to the generalized factor model for X, with $\mathrm{Cov}(f) = \Theta$ and $\mathrm{Cov}(u) = \Psi$, then*

$$\Sigma_X = Q\Theta Q^t + \Psi. \tag{6}$$

Proof.

$$\begin{aligned} \Sigma_X = \mathrm{Cov}(X) = \mathrm{Cov}(Qf + u) &= Q\,\mathrm{Cov}(f)Q^t + Q\,\mathrm{Cov}(f, u) + \mathrm{Cov}(u, f)Q^t + \mathrm{Cov}(u) \\ &= Q\Theta Q^t + \Psi. \end{aligned}$$

□

Therefore, the necessary condition for the model to have solution is now $\Sigma_X = Q\Theta Q^t + \Psi$, with $Q \in M_{p\times m}(\mathbb{R})$, $m < p$, and Θ and Ψ covariance matrices, with the second being diagonal. This condition is also sufficient as in the orthogonal case, if we ask Ψ to be positive definite. The result is given by the next theorem, which is a corollary of Theorem 2.3.

Theorem 3.3 (Existence of solution to the generalized factor model)**.** *Let X be a $p \times 1$ random vector with $E[X] = 0_{p\times 1}$ and $\mathrm{E}[X^2] < \infty$. Assume that there exist three matrices $Q \in M_{p\times m}(\mathbb{R})$, with $m < p$, $\Psi \in M_p(\mathbb{R})$, with Ψ diagonal and positive definite, and $\Theta \in M_m(\mathbb{R})$, with Θ symmetric and positive semidefinite, such that $\Sigma_X = Q\Theta Q^t + \Psi$. Then, there exist two random vectors $f^t = (f_1, \ldots, f_m)$ and $u^t = (u_1, \ldots, u_p)$ satisfying the generalized factor model for X, with loadings matrix Q, with $\mathrm{Cov}(f) = \Theta$ and $\mathrm{Cov}(u) = \Psi$.*

Proof. Assume $\Sigma_X = Q\Theta Q^t + \Psi$, with $Q \in M_{p\times m}(\mathbb{R})$, $\Theta \in M_m(\mathbb{R})$ and $\Psi \in M_p(\mathbb{R})$, with the stated conditions. Since Θ is symmetric, we can consider its spectral decomposition $\Theta = V\Lambda V^t$. Then, $\Lambda = \operatorname{diag}(\lambda_1, \dots, \lambda_m)$ and since Θ is positive semidefinite $\lambda_1 \geq \cdots \geq \lambda_m \geq 0$. Therefore, we can take $\Lambda^{1/2} = \operatorname{diag}(\sqrt{\lambda_1}, \dots, \sqrt{\lambda_m})$ and write $\Theta = V\Lambda^{1/2}\Lambda^{1/2}V^t$. Now denote $Q_0 = QV\Lambda^{1/2}$, then:

$$\Sigma_X = Q\Theta Q^t + \Psi = QV\Lambda^{1/2}\Lambda^{1/2}V^tQ^t + \Psi = Q_0Q_0^t + \Psi.$$

Thus $\Sigma_X = Q_0Q_0^t + \Psi$, with $Q_0 \in M_{p\times m}(\mathbb{R})$ and $\Psi \in M_p(\mathbb{R})$, with Ψ diagonal and positive definite, so we are on the hypotheses of Theorem 2.3. Hence, there exists two random vectors $f_0 = (f_{01}, \dots, f_{0m})^t$ and $u_0 = (u_{01}, \dots, u_{0p})^t$ that satisfy the orthogonal factor model for X with loadings matrix Q_0 and $\operatorname{Cov}(u_0) = \Psi$, that is, satisfying $X = Q_0f_0 + u_0$, $\operatorname{Cov}(u_0) = \Psi$, $\mathrm{E}[u_0] = 0_{p\times 1}$, $\operatorname{Cov}(f_0) = I_m$, $\mathrm{E}[f_0] = 0_{m\times 1}$ and $\operatorname{Cov}(f_0, u_0) = 0_{m\times p}$. Now, define the random vectors $f := V\Lambda^{1/2}f_0$ and $u := u_0$, and let's see that these vectors give a solution to the generalized factor model for X with loadings matrix Q, $\operatorname{Cov}(f) = \Theta$ and $\operatorname{Cov}(u) = \Psi$. It holds:

$$X = Q_0f_0 + u_0 = QV\Lambda^{1/2}f_0 + u_0 = Qf + u$$

and it also holds:

$$\begin{gathered}\operatorname{Cov}(u) = \operatorname{Cov}(u_0) = \Psi,\ \mathrm{E}[u] = \mathrm{E}[u_0] = 0_{p\times 1},\\ \operatorname{Cov}(f) = \operatorname{Cov}(V\Lambda^{1/2}f_0) = V\Lambda^{1/2}\operatorname{Cov}(f_0)(V\Lambda^{1/2})^t = V\Lambda^{1/2}I_m\Lambda^{1/2}V^t = \Theta,\\ \mathrm{E}[f] = \mathrm{E}[V\Lambda^{1/2}f_0] = 0_{m\times 1},\ \operatorname{Cov}(f,u) = \operatorname{Cov}(V\Lambda^{1/2}f_0, u_0) = 0_{m\times p}.\end{gathered}$$

Thus, $X = Qf + u$ with f and u satisfying the conditions in Definition 3.1 and the proof is complete. ☐

In view of Theorem 3.3 and similarly to the orthogonal case, to fit the generalized factor model to a data matrix $\widetilde{X}$, Σ_X is replaced by the sample covariance matrix S and one tries to obtain parameters $\widehat{Q} \in M_{p\times m}(\mathbb{R})$, with $m < p$, $\widehat{\Theta} \in M_m(\mathbb{R})$, $\widehat{\Theta}$ symmetric and positive semidefinite, and $\widehat{\Psi} \in M_p(\mathbb{R})$, $\widehat{\Psi}$ diagonal and positive definite such that the equality

$$S = \widehat{Q}\widehat{\Theta}\widehat{Q}^t + \widehat{\Psi}$$

holds, at least approximately.

Factor models such as the one discussed in this section are used in confirmatory factor analysis. In this technique the value of some parameters of the model is fixed in advance, and only the free parameters are estimated. For example, it is common to fix some loadings to be zero. Values are fixed to obtain a solution with a meaningful structure for the researcher. For this reason, if the prefixed model fits the observed data, rotations may not be necessary, in contrast to the exploratory case. It is usual to perform an exploratory analysis before the confirmatory one, in order to choose a model that is not in contradiction with the data, but the exploratory and confirmatory procedures should be checked on different subsamples to honestly confirm the model.

Expression of gratitude

I would like to thank Xavier Bardina, editor of Reports@SCM, for seeing an interest in sharing the proofs given in this article, when I first exposed them in my final degree project. I would also like to thank Mercè Farré, my project advisor in the Math Department of Universitat Autònoma de Barcelona, for her precious comments about this document.

References

[1] T. Banachiewicz, "Zur Berechnung der Determinanten, wie auch der Inversen und zur darauf basierten Auflosung der Systeme linearer Gleichungen", *Acta Astronom. Ser. C* **3** (1937), 41–67.

[2] W.K. Härdle, L. Simar, "Applied Multivariate Statistical Analysis", Springer, Berlin Heidelberg, 2012.

[3] K.G. Jöreskog, "Some contributions to maximum likelihood factor analysis", *Psychometrika* **32**(4) (1967), 443–482.

[4] K.G. Jöreskog, A.S. Goldberger, "Factor analysis by generalized least squares", *Psychometrika* **37**(3) (1972), 243–260.

[5] K.V. Mardia, J.T. Kent, J.M. Bibby, "Multivariate Analysis", Probability and Mathematical Statistics: A Series of Monographs and Textbooks, Academic Press [Harcourt Brace Jovanovich, Publishers], London-New York-Toronto, Ont., 1979.

[6] J. Schur, "Über Potenzreihen, die im Innern des Einheitskreises beschränkt sind", *J. Reine Angew. Math.* **1917**(147) (1917), 205–232.