COWLES COMMISSION DISCUSSION PAPER: STATISTICS: No. 334

On the identification problem in factor analysis.

July 25, 1949

by Olav Reiersøl

This paper is part of the paper "The Identification of Structural Characteristics" by Koopmans and Reiersøl. The notation has been chosen to correspond in some way to the notation in the econometric example in the same paper. We shall here give a table showing the correspondence between these notations and those used by Thurstone.

	Notations in the present paper		Notations used by Thurstone
	y_i		s_j
	z_k		x_m
	λ_{ik}		a_{jm}
Variable	v_i	Standard deviation	u_j
	G		m
	ρ		r
	M_{yy}		R_1
	M_{zz}		R_{pq}
	Δ		$R_1 - R$
	$(\Delta)_{ii}$		u_j^2

4. An example from factor analysis.[1]

Factor analysis has been presented in different forms by different authors. We shall here consider the multiple factor analysis of Thurstone only. (Thurstone 1935 and Thurstone 1947)

The factor analysis methods were developed primarily for the purpose of analyzing intelligence tests, but they have also been used for other psychological problems and in other sciences.

Suppose that a person is given a battery of G tests. Let his score in test i be y_i. The fundamental assumption in factor analysis is that these scores can be explained in terms of a relatively small number of hypothetical primary factors. Let $z_1, z_2, \ldots, z_\rho$ denote the hypothetical scores of the person in the common factors, i.e., those primary factors which are common to at least two tests in the battery. We assume that y_i is a homogeneous linear function of the scores z_k plus a unique part v_i, which may be thought of as consisting of an error term plus the contribution of a specific factor. The coefficients π_{ik} in the linear function just mentioned are called factor loadings. The factor loading π_{ik} expresses the relative importance of the common factor k in the answering of test i.

We shall introduce the row vectors $y = [y_i]$, $z = [z_i]$, $v = [v_i]$ and the matrix $\Pi = [\pi_{ik}]$. The covariance matrices of the sets of variables y, z, and v will be denoted by M_{yy}, M_{zz}, and Δ, respectively.

In contrast with the preceding example, the variables y are the only observed variables. The variables v and z are latent variables.

Our model will be given by the following specifications:

(4.1) $y' = \Pi z' + v'$

(4.2) $E(z) = 0$ and $E(v) = 0$

(4.3) The set of variables z is stochastically independent of the set of variables v.

---

1. Proofs of the statements in this section will be found in a separate paper by one of the authors (Reiersøl 1949).

(4.4) Δ is diagonal and different from 0.

(4.5) The elements of z and v are jointly normally distributed.

(4.6) Each y_i is correlated with at least one of the other y's.

(4.7) The rank of Π equals its number of columns.

(4.8) M_{zz} is nonsingular.

(4.9) ρ is the smallest number of variables z which is compatible with joint probability distribution of the observed variables and specifications (4.1)-(4.8).

(4.10) Each column of Π contains at least ρ zeros.

(4.11) A normalization rule fixing the units of the variables z and a rule fixing the order of the columns of Π.

Denote by Π_k the matrix consisting of all the rows of Π which have a zero in the k'th column. Let the number of rows in the matrix Π_k be p_k. Let Π_{ki} denote the submatrix of Π_k which we get when deleting the i'th row of Π_k. Using these notations we shall formulate the final specification of our model.

(4.12) The rank of each of the matrices Π_{ki} ($k=1, 2, \ldots, \rho$; $i=1, 2, \ldots p_k$) is $\rho-1$.

Specification (4.10) means that the experimenter thinks he can construct tests where at least one of the common primary factors is absent.

We shall first consider a model $\mathcal{S}$ containing Specifications (4.1)-(4.9) only. From (4.9) follows that ρ is uniformly identifiable.

Let $\rho_G = \frac{1}{2}(2G + 1 - \sqrt{8G + 1})$. If $\rho > \rho_G$, the matrix Δ is generally not identifiable. If $\rho < \rho_G$, Δ generally is identifiable. When $\rho = \rho_G$, the number of values of Δ, which correspond to a given covariance matrix M_{yy}, is usually finite, and may be equal to one or greater than one. The matrices Π and M_{zz} are never identifiable in this model. The set of all structures $\{\Pi^*, M^*_{zz}, \Delta\}$ equivalent to the structure $\{\Pi, M_{zz}, \Delta\}$ is given by the set of all matrices

(4.13) $\quad \Pi^* = \Pi \Psi$

and

(4.14) $\quad M^*_{zz} = \Psi^{-1} M_{zz} (\Psi')^{-1}$

where Ψ is any square, ρ-rowed and nonsingular matrix.

In the following we shall confine our discussion to the case $\rho < \rho_G$, and to structures in which the matrix M_{yy} is such that Δ is identifiable in G.

We shall now consider the model G' defined by Specifications (4.1)-(4.11). In this model a necessary and sufficient condition for the identifiability of Π is that any square ρ-rowed minor of Π which is of rank ρ-1, is contained in one of the matrices Π_k. This result excludes the possibility that all elements belonging to the intersection of ρ-1 rows and two columns of Π are all equal to zero. In order to be able to use this result, the experimenter would have to be able to construct tests where one, but not more than one, common factor would be absent. Therefore the result is not particularly useful. In order not to exclude the case where two common factors occur in more than ρ-2 tests, we have introduced Specification (4.12).

We shall finally consider the model G'' defined by Specifications (4.1)-(4.12). Assuming M_{yy} known, we can determine some value Π^* of Π which satisfies Specifications (4.1)-(4.9). Since, by assumption, Δ is identifiable in G, Π^* must be of the form $\Pi_0 \Psi$, where Π_0 is the true factor loadings matrix and Ψ is non-singular. Let Π^*_k be a submatrix of Π^* containing all the columns of Π^* and satisfying the following conditions

(4.15) The rank of Π_k^* is $\rho-1$

(4.16) The addition to Π_k^* of a row contained in Π^* increases the rank to ρ.

(4.17) Each submatrix of Π_k^* obtained by deleting one row of Π_k^* has rank $\rho-1$.

A necessary and sufficient condition for the identifiability of Π in the complete model $\mathfrak{S}''$ is that there exist exactly ρ submatrices Π_k^* of Π^* which satisfy Conditions (4.15)-(4.17), and that the ρ vectors q_k, satisfying the equations $\Pi_k^* q_k = 0$ when $k = 1, 2, \ldots, \rho$, are linearly independent.

It should be noted that Specifications (4.10) and (4.12) are observationally restricting, i.e., they are in principle subject to statistical test.