The factor analysis model is rewritten as a system of linear structural relations with errors in variables. The method of instrumental variables is applied to this revised form of the model to obtain estimates of the factor loading matrix. The relation between this method and interbattery analysis, proportional profile analysis, and canonical factor analysis is pointed out. In addition, an estimation procedure based on replicated sampling different from proportional profile analysis is given.

1. Introduction

The factor analysis model [2] is given by

\[ Z = FX + U, \]

where

- \( Z \) is a vector of \( n \) components (\( n \) test scores, say),
- \( X \) is a vector of \( r \) (\( \leq n \)) components (the common factor scores),
- \( U \) is an \( n \)-vector (the unique part of the test scores),
- \( F \) is an \( n \times r \) matrix (of factor loadings).

It is assumed that \( U \) and \( X \) are independent random vectors, with \( \varepsilon U = \varepsilon X = 0, \varepsilon UU' = D, \) a diagonal matrix, and \( \varepsilon XX' = M. \) From these assumptions, one deduces that the covariance matrix \( \Sigma \) of \( Z \) is given by

\[ \Sigma = \varepsilon ZZ' = FMF' + D. \]

After assuming that \( M = I, \) the identity matrix, the usual attack in determining \( F \) (or estimating \( F \) if \( \Sigma \) is unknown and only estimated) is to proceed from this equation, or equivalently from the equation corresponding to (2) based on the correlation matrix of \( Z. \) Certainly if \( D \) were known, the problem of finding \( F \) would be trivial, as it is given, up to rotation, by “factoring” the matrix \( \Sigma - D. \) Thus, the usual factor analytic techniques can be characterized as attempts either at successively better approximating \( D \) (or equivalently the diagonal terms of \( FF' \), i.e., the communalities) or at assuming away the problem of estimating \( D \) by imposing additional restrictions on the problem which, in addition, eliminate the indeterminacy of \( F \) due to rotation (for example, the restriction that \( F'D^{-1}F \) be diagonal, as in Rao’s work [16]).
In this paper we shall instead proceed directly from (1) to the problem of estimating $F$. We shall see that the problem is equivalent to that of fitting a linear relation when both the independent and dependent variables are subject to error, as has been well known to some factor analysts, notably Burt [3] (see also [6]), and thus that the difficulty in estimating $F$ without some additional restrictions on the problem (such as Rao's alluded to above) is just a reflection of the difficulties in the equivalent problem (cf. [14] for a catalogue of these). We shall in particular look into one of the approaches taken by econometricians when confronted with the "errors-in-variables" problem, namely, the use of instrumental variables, and see that in essence this approach is that which is being taken by Tucker and Gibson [9, 10, 11, 21] in the interbattery method, but that they are working with (2), in the tradition of factor analysts, rather than with (1). We shall then point out connections between this approach and that of proportional profile analysis, analysis of variance in factor analysis, and canonical factor analysis.

2. The "Errors-in-Variables" Model

As the rank of $F$ is $r$, there is some $r \times r$ submatrix of $F$ which is nonsingular. Suppose for convenience that it is the submatrix consisting of the first $r$ rows. Write

$$Z = \begin{bmatrix} Z_1 \\ Z_2 \end{bmatrix} = \begin{bmatrix} F_1 \\ F_2 \end{bmatrix}X + \begin{bmatrix} U_1 \\ U_2 \end{bmatrix},$$

where $Z_1$, $F_1$, and $U_1$ have $r$ rows. Then

$$Z_1 = F_1X + U_1,$$

or

$$X = F_1^{-1}(Z_1 - U_1).$$

Thus

$$Z_2 = F_2X + U_2 = F_2F_1^{-1}(Z_1 - U_1) + U_2.$$

Let $B = F_2F_1^{-1}$. Then the above equation is a linear structural relation $Z_2^* = BZ_1^*$ between the "true" variables $Z_1^* = Z_1 - U_1$ and $Z_2^* = Z_2 - U_2$, where the respective true variables are observed with "error" $U_1$ and $U_2$. As the matrix $F$ is unique up to a right-multiplication by a nonsingular matrix, we might as well adopt as the canonical form for the matrix $F$ the form

$$\tilde{F} = \begin{bmatrix} I \\ B \end{bmatrix},$$

for right-multiplication by the nonsingular matrix $F_1$ yields the original matrix of interest, $F$. Our problem is then to estimate $B$. 