subfaculteit der econometrie

RESEARCH MEMORANDUM

TILBURG UNIVERSITY

DEPARTMENT OF ECONOMICS

Postbus 90153 - 5000 LE Tilburg
Netherlands

IDENTIFICATION IN FACTOR ANALYSIS

Arie Kapteyn[*]
Tom Wansbeek[*]

Abstract

It is shown that a simple counting rule provides necessary and sufficient conditions for identification of the parameters in an unrestricted factor analysis model.

1. Introduction

In a classic article, Anderson and Rubin (1956) discuss identification and estimation in a factor analysis model. One of their results is a general, sufficient condition for the identification of the parameters in the unrestricted model (apart from some inevitable indeterminacies and some necessary and sufficient conditions for special cases). They note that their general condition is unnecessarily strong, in that model parameters may be identified if the condition is not met.

Although their article appeared over a quarter of a century ago, we are not aware of a later publication stating necessary and sufficient conditions for identification. The apparent non-existence of such a publication may be due to the fact that a formal proof requires results on zero-one matrices and matrix stacking operators that have only recently been developed. This note provides necessary and sufficient conditions for local identification of the parameters. The model is presented in section 2, as well as the identification result. Section 3 presents the proof of the result.

2. The Model and the Identification Result

In factor analysis a population variance covariance matrix Σ of order pxp is assumed to have the following structure

$$\Sigma = H \Phi H' + \Lambda, \qquad (2.1)$$

where H is a p × m-matrix (m ≤ p) of full column rank, Φ a m × m-matrix, and Λ a p × p-diagonal matrix of unknown positive parameters. Under the common assumption of multivariate normality of all factors, equation (2.1) completely determines the identification of the parameters. That is, if Σ is known (and it can be estimated consistently from second order sample moments), a parameter is identified if and only if it can be uniquely determined from (2.1). It is well-known that normality is the least favorable assumption with respect to identification. Under different assumptions we could supplement (2.1) with equations involving higher order moments, which would give additional information on the parameters.

Assuming normality is therefore a conservative assumption from the viewpoint of identification. If a parameter is identified under normality, it is also identified under different distributional assumptions. Moreover, the identification result derived under normality also applies to the functional factor analysis model, where the common factors are considered to be unknown constants rather than random variables [cf. Wald (1948), Anderson and Rubin (1956), Aigner et al. (1982, sec. 2)]. In this note we assume normality throughout.

Obviously, replacing H by $H^* \equiv H\Phi^{\frac{1}{2}}$ and Φ by the identity matrix has no observable consequences for Σ. So no elements of Φ are identified. Following common practice, we set Φ equal to the identity matrix. This reduces (2.1) to

$$\Sigma = HH' + \Lambda \ . \tag{2.2}$$

Furthermore, if Q is an m × m-orthogonal matrix, i.e. $QQ' = I_m$, replacing H by $\tilde{H} = HQ$ does not influence Σ either. Obviously, $QQ' = I_m$ imposes m(m+1)/2 restrictions on Q and hence Q has m(m−1)/2 free elements. This in turn implies that H can only be determined up to m(m−1)/2 indeterminacies.

In this paper we prove the following proposition:

Proposition. Apart from $m(m-1)/2$ indeterminacies in H, equation (2.2) has locally unique solutions for H and Λ if and only if

$$(p-m)^2 \geqslant p+m \; . \tag{2.3}$$

Remark 1. By contrast, the sufficient condition given by Anderson and Rubin (1956) requires:

$$p \geqslant 2m+1 \; . \tag{2.4}$$

Both conditions (2.3) and (2.4) are represented graphically in Fig. 1. The white dots indicate the (p,m)-combinations for which (2.4) holds, and the black dots indicate the (p,m)-combinations for which (2.3) holds, but not (2.4).

Remark 2. Since Σ is symmetric, the number of unconstrained elements in Σ is $\frac{1}{2}p(p+1)$. The number of indeterminacies in H is $\frac{1}{2}m(m-1)$ so we need as many extra restrictions to determine the elements of H uniquely. There are mp unknown parameters in H and p in Λ. Simply counting the number of unknown parameters, $(m+1)p$, and the number of equations provided by (2.2) plus the extra restrictions needed to remove the indeterminacies in H, $\frac{1}{2}p(p+1) + \frac{1}{2}m(m-1)$, gives as a counting rule for identification:

$$\tfrac{1}{2}p(p+1) + \tfrac{1}{2}m(m-1) \geqslant (m+1)p, \tag{2.5}$$

which is equivalent to (2.4). So, the proposition states that necessary and sufficient conditions for identification can be verified by simply counting the number of equations and the number of unrestricted parameters. Although the counting rule is encountered in the literature, e.g., in testing specification (2.2) [cf., e.g., Jöreskog (1978)], we are not aware of a proof that this rule provides a sufficient condition for identification.

1. Identification conditions
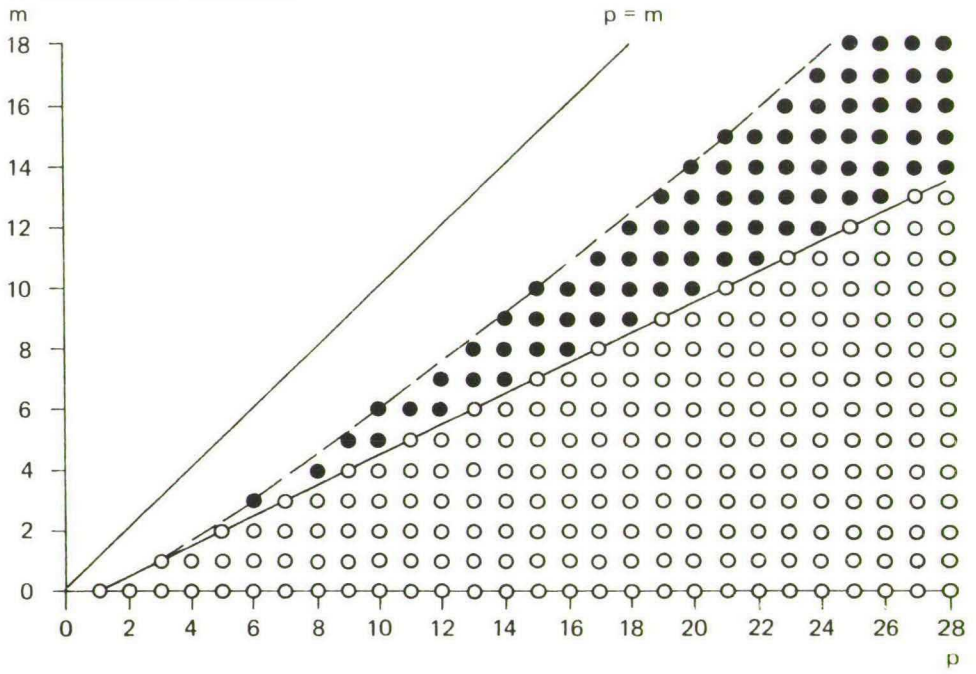
Remark 3. The proposition only provides a local identification result. Since (2.2) is a quadratic equation system no globally unique solution exists. A simple example is that the sign of a column of H can be changed without changing HH'. A more intricate example is given by Wilson and Worcester (1939).

3. Proof of the proposition

The proof rests on an analysis of the Jacobian of the equation system (2.2), taking into account the redundancy caused by the symmetry of Σ. To deal with symmetry we employ a number of useful matrix operators P, D, L, N defined as follows [cf. Balestra (1976), Magnus and Neudecker (1979,1980), Henderson and Searle (1981) for details]. Let A be an arbitrary r × s-matrix, then the permuted identity matrix $P_{r,s}$ of order rs × rs is defined by

$$P_{r,s} \; vec \; (A') \equiv vec \; A, \qquad (3.1)$$

where vec is the matrix stacking operator which "vectorizes" a matrix column by column. $P_{r,s}$ consists of s × r blocks of order r × s each. The (i,j)-th block has a unit element in position (j,i) and zeros elsewhere. The rs × rs-matrix $N_{r,s}$ is defined by

$$N_{r,s} \equiv \tfrac{1}{2}(I_{rs} + P_{r,s}). \qquad (3.2)$$

If no confusion can arise, the subscripts of P and N are omitted.

Let e_i be the i-th unit vector and let $E_{ij} \equiv e_i e_j'$, i.e., a matrix with a single 1 in position (i,j) and zeros elsewhere, then P and N satisfy the following properties:

$$P'_{r,s} = P_{s,r} \qquad (3.3)$$

$$P_{r,s} P_{s,r} = I_{rs} \tag{3.4}$$

$$P_{r,1} = P_{1,r} = I_r \tag{3.5}$$

$$P_{r,s} = \sum_{i=1}^{s} \sum_{j=1}^{r} (E_{ij} \otimes E'_{ij}) \tag{3.6}$$

$$P_{r,k}(A \otimes B)P_{\ell,s} = B \otimes A \text{ for any } r \times s\text{-matrix } A \text{ and } k \times \ell\text{-matrix } B \tag{3.7}$$

$$N_{s,s}^2 = N_{s,s} = N'_{s,s} \text{ has rank } \tfrac{1}{2}s(s+1) . \tag{3.8}$$

Next, let S be a symmetric pxp-matrix; v(S) is obtained from vec(S) by eliminating all elements corresponding to the supradiagonal elements of S. The $\tfrac{1}{2}p(p+1) \times p^2$-matrix L, called the __elimination__ matrix, is defined by

$$L \text{ vec } S \equiv v(S). \tag{3.9}$$

The $p^2 \times \tfrac{1}{2}p(p+1)$-matrix D, called the __duplication matrix__, is defined by

$$D \equiv N_{p,p} L'(LN_{p,p}L')^{-1} . \tag{3.10}$$

From now on, E_{ij} is meant to be a pxp-matrix and e_i a p-vector.

Let $u_{ij} \equiv v(E_{ij})$, $i \geqslant j$, i.e., the $\tfrac{1}{2}p(p+1)$-vector with a unit element in position $(j-1)p+i-\tfrac{1}{2}j(j-1)$ and zeros elsewhere, then L and D satisfy:

$$L = \sum_{i \geqslant j} u_{ij}(\text{vec } E_{ij})' \tag{3.11}$$

$$LP L' = \sum_{i} u_{ii}u'_{ii} \tag{3.12}$$

$$Dv(A) = \text{vec } A \text{ for any symmetric } p \times p\text{-matrix } A \tag{3.13}$$

$$L \text{ has full row rank} \tag{3.14}$$

D has full column rank $\qquad$ (3.15)

$DLN = N$ $\qquad$ (3.16)

$PD = D = ND$ $\qquad$ (3.17)

$$L'L = \sum_{i \geqslant j} E_{jj} \otimes E_{ii} .$$ (3.18)

For a proof, see Magnus and Neudecker (1980). Next, define

$$\Psi \equiv \begin{pmatrix} E_{11} \\ \vdots \\ E_{pp} \end{pmatrix} \qquad p^2 \times p$$ (3.19)

$$K \equiv I - L\Psi (\Psi 'L'L\Psi)^{-1}\Psi 'L' \qquad \tfrac{1}{2}p(p+1) \times \tfrac{1}{2}p(p+1)$$ (3.20)

$$M \equiv I - \Psi\Psi ' \qquad p^2 \times p^2 .$$ (3.21)

Ψ is a matrix with ones in positions $((i-1)p+i,i)$, $i = 1,\dots,p$, and zeros elsewhere. K projects perpendicular to the space spanned by the columns of $L\Psi$, M to that of Ψ (as $\Psi '\Psi = I$). The following properties of Ψ, K and M can be verified straightforwardly:

$$L\Psi = \sum_j u_{jj} e'_j$$ (3.22)

$$\Psi 'L'L\Psi = I_p$$ (3.23)

$$L\Psi\Psi 'L' = LP\ L'$$ (3.24)

$$L'L\Psi = \Psi$$ (3.25)

$$K = I - L\Psi\Psi'L' \tag{3.26}$$

$$KL = LM \tag{3.27}$$

$$MN = NM \quad . \tag{3.28}$$

Now, write (2.2) in vector form:

$$v(\Sigma) = v(HH') + v(\Lambda) = L \text{ vec } HH' + L \text{ vec } \Lambda \; . \tag{3.29}$$

Let λ be the p-vector of diagonal elements of Λ and define

$$\varphi \equiv \binom{\text{vec } H}{\lambda}. \tag{3.30}$$

Lemma 1. The Jacobian $J \equiv \partial v(\Sigma)/\partial\varphi'$ of the equation system is:

$$J \equiv L[2N(H\otimes I_p), \Psi] \quad . \tag{3.31}$$

Proof: Let h_i be the i-th column of H. Then

$$\frac{\partial v(HH')}{\partial(\text{vec } H)'} = L\frac{\partial \text{ vec } HH'}{\partial(\text{vec } H)'} = L\sum_{i=1}^{m}\frac{\partial(h_i\otimes h_i)}{\partial(\text{vec } H)'} =$$

$$= L\,(h_1\otimes I_p + I_p\otimes h_1,\ldots,h_m\otimes I_p + I_p\otimes h_m) = L\,(I+P_{p,p})(H\otimes I_p) = 2LN_{p,p}(H\otimes I_p).$$

$$\tag{3.32}$$

Furthermore it can be checked directly that

$$\frac{\partial v(\Lambda)}{\partial\lambda'} = L\frac{\partial \text{ vec } \Lambda}{\partial\lambda'} = L\Psi \quad . \tag{3.33}$$

Lemma 2. The rank of $LN(H\otimes I_p)$ is $mp-\frac{1}{2}m(m-1)$.

Proof: We first consider the rank of $N(H \otimes I_p)$. Let the $p \times (p-m)$-matrix G satisfy $G'G = I_{p-m}$ and $G'H = 0$. The space of dimension pm has a basis (R_1, R_2, R_3), where

$$R_1 \equiv (I_m \otimes H)(I_{m^2} - P_{m,m}) \tag{3.34}$$

$$R_2 \equiv (I_m \otimes H)(I_{m^2} + P_{m,m}) \tag{3.35}$$

$$R_3 = I_m \otimes G \quad . \tag{3.36}$$

Let x and y be arbitrary m^2-vectors. Then, using (3.7), we see that

$$N(H \otimes I_p)(I_m \otimes H)(I_{m^2} - P_{m,m})x = 0 \tag{3.37}$$

Thus, post-multiplying $N(H \otimes I_p)$ by any vector in R_1 yields zero. Since $\frac{1}{2}(I_{m^2} - P_{m,m})$ is idempotent its rank equals its trace, which is equal to $\frac{1}{2}m(m-1)$. Hence R_1 spans a $\frac{1}{2}m(m-1)$-dimensional space, so that according to (3.37) the columns of $N(H \otimes I_p)$ satisfy $\frac{1}{2}m(m-1)$ independent restrictions. There are no more than $\frac{1}{2}m(m-1)$ independent restrictions because postmultiplication of $N(H \otimes I_p)$ by a vector contained in R_2 or R_3 only yields zero if that vector is the zero vector. This can be verified as follows: Consider the equation

$$N(H \otimes I_p)R_2 y = \frac{1}{2}(H \otimes H)(I_{m^2} + P_{m,m})y = 0 \quad , \tag{3.38}$$

which, in view of the full column rank of H, implies

$$(I_{m^2} + P_{m,m})y = 0 \quad . \tag{3.39}$$

This only holds if y is zero or if y is of the form $(I_{m^2} - P_{m,m})x$ in which case $R_2 y = 0$. In both cases we would be postmultiplying $N(H \otimes I_p)$ by the zero vector.

Secondly, let Z be an arbitrary $(p-m) \times m$-matrix and $z \equiv$ vec Z. Then consider the equation

$$N(H \otimes I_p)R_3 z = N(H \otimes G)\text{vec } Z = N \text{ vec } GZH' = \tfrac{1}{2} \text{ vec } (GZH' + HZ'G') = 0 \ .$$

(3.40)

This implies

$$GZH' + HZ'G' = 0 \quad .$$

(3.41)

Premultiplication by G' gives:

$$ZH' + 0 = 0$$

(3.42)

Since H has full column rank this implies that $Z = 0$ is the only solution of (3.40).

As a result we have that the columns of the $p^2 \times mp$-matrix $N(\otimes I_p)$ satisfy exactly $\tfrac{1}{2}m(m-1)$ independent restrictions, so its rank is $mp - \tfrac{1}{2}m(m-1)$.

Turning to an analysis of the rank of $LN(H \otimes I_p)$ we have:

$$\text{Rank}[LN(H \otimes I_p)] \leqslant \text{Min}\{ \text{Rank}(L), \ \text{Rank}[N(H \otimes I_p)]\} =$$

$$= \text{Min}\{ \tfrac{1}{2}p(p+1), \ mp - \tfrac{1}{2}m(m-1)\} = mp - \tfrac{1}{2}m(m-1),$$

(3.43)

since $m \leqslant p$. In addition, using (3.16), we observe that

$$mp - \tfrac{1}{2}m(m-1) = \text{Rank}[N(H \otimes I_p)] = \text{Rank}[DLN(H \otimes I_p)]$$
$$\leqslant \text{Min}\{ \text{Rank}(D), \ \text{Rank } LN(H \otimes I_p)\} = \text{Rank}[LN(H \otimes I_p)] \ .$$

(3.44)

Inequalities (3.43) and (3.44) establish the Lemma.

Remark 4. $2LN(H \otimes I_p)$ is the Jacobian of the equation system:

$$c = v(HH'), \tag{3.45}$$

with c a known vector. As discussed in the previous section, there are $\frac{1}{2}m(m-1)$ indeterminacies in H. This is brought out by the rank of the Jacobian which is $\frac{1}{2}m(m-1)$ less than the number of parameters in H.

Remark 5. From the proof of Lemma 2 it is clear that $N(H \otimes I_p)$ and $LN(H \otimes I_p)$ have the same rank. This is according to expectation, as the elimination matrix L simply eliminates superfluous equations from the system.

Remark 6. In view of (3.45) it is clear that if Λ is identified, H will be identified up to $\frac{1}{2}m(m-1)$ indeterminacies, since in that case c can be taken equal to $v(\Sigma) - v(\Lambda)$ which are then both observable. So the elements of H are determined up to $\frac{1}{2}m(m-1)$ indeterminacies if and only if Λ is identified. Thus we can concentrate on necessary and sufficient conditions for the identification of Λ. The elements of Λ are identified if and only if there are no linear dependencies between $LN(H \otimes I_p)$ and $L\Psi$. This requires the rank of $KLN(H \otimes I_p)$ to be equal to the rank of $LN(H \otimes I_p)$.

Lemma 3. The rank of $KLN(H \otimes I_p)$ equals $\text{Min} \left\{ \frac{1}{2}p(p-1), pm - \frac{1}{2}m(m-1) \right\}$.

Proof. From (3.27) it follows that

$$KLN(H \otimes I_p) = LMN(H \otimes I_p). \tag{3.46}$$

We first study the rank of $MN(H \otimes I_p)$. From (3.28), MN = NM. Obviously MN is idempotent. Its rank is equal to its trace, which can easily be seen

to be $\frac{1}{2}p(p-1)$. So the pm column vectors of $MN(H\otimes I_p)$ lie in the $\frac{1}{2}p(p-1)$-dimensional space spanned by the columns of MN. From (3.37) it is once again clear that they satisfy $\frac{1}{2}m(m-1)$ restrictions. This establishes the rank of $MN(H\otimes I_p)$ as $\min\{\frac{1}{2}p(p-1),\ pm-\frac{1}{2}m(m-1)\}$. Finally, using an argument similar to (3.43) and (3.44), $LNM(H\otimes I_p)$ has the same rank as $NM(H\otimes I_p)$, because $DLN = N$.

It is now obvious that $LN(H\otimes I_p)$ and $KLN(H\otimes I_p)$ have identical ranks if $pm-\frac{1}{2}m(m-1) \leq \frac{1}{2}p(p-1)$ which proves the proposition.

Remark 7. $KLN(H\otimes I_p)$ differs from $LN(H\otimes I_p)$ in that the rows corresponding to the diagonal elements of Σ are set equal to zero. So $KLN(H\otimes I_p)$ is the Jacobian corresponding to the system which equates the off-diagonal elements of Σ to the off-diagonal-elements of HH'. The diagonal elements of Σ are "reserved" to identify Λ.

REFERENCES

AIGNER, D.J., C. HSIAO, A KAPTEYN and T.J. WANSBEEK (1982). Latent Variables in Econometrics and Elsewhere. In: Intriligator, M.D. and Z. Griliches (Eds.), Handbook of Econometrics. North-Holland, Amsterdam (Forthcoming).

ANDERSON, T.W. and H. RUBIN (1956). Statistical Inference in Factor Analysis. In: Neyman, J. (Ed.), Proceedings of the Third Berkeley Symposium on Mathematical Statistics and Probability, Volume V. University of California Press, Berkeley.

BALESTRA, P. (1976). La Dériviation Matricielle. Collection de l'IME, No. 12. Sirey, Paris.

HENDERSON, H.V. and S.R. SEARLE (1981). The Vec-Permutation Matrix, the Vec-Operator and Kronecker Products: A Review. Linear and Multilinear

Algebra 9, 271-288.

JÖRESKOG, K.G. (1978). Stuctural Analysis of Covariance and Correlations Matrices. Psychometrika 43, 443-477.

MAGNUS, J.R. and H. NEUDECKER (1979). The Commutation Matrix: Some Properties and Applications. The Annals of Statistics 7, 381-394.

MAGNUS, J.R. and H. NEUDECKER (1980). The Elimination Matrix: Some Lemmas and Applications. SIAM Journal on Algebraic and Discrete Methods 1, 422-449.

WALD, A. (1948). Estimation of a Parameter When the Number of Unknown Parameters Increases Indefinitely with the Number of Observations. Annals of Mathematical Statistics 19, 220-227.

WILSON, E.B. and J. WORCESTER (1939). The resolution of six tests into three general factors. Proceedings of the National Academy of Science 25, 73-77.

IN 1981 REEDS VERSCHENEN:

96 B.B. Van der Genugten
Central limit theorems for least squares estimators in linear
regression models including lagged variables.

97 A.J.M. Swanenberg
Rationing and price dynamics in a simple market-game.

98 A.J. Hendriks
Theoretische en practische problemen van het distributieplanologisch
onderzoek.
Referaten themadag RSA op 14 mei 1981.

99 A.J.J. Talman en L. van der Heyden
Algorithms for the linear complementarity problem which allow an
arbitrary starting point.

100 J.P.C. Kleijnen
Cross-validation using the t statistic.

101 J.P.C. Kleijnen
Statistical aspects of simulation: an updated survey. Version 1.

102 Dr. Wim G.H. van Hulst
On the concept of divergence in the theory of industrial organization.

103 H. Gremmen, T. van Bergen, J. Hotterbeekx
Waar liggen de Nederlandse comparatieve voordelen? Voorlopige versie.

104 A.L. Hempenius en P.G.H. Mulder
Multiple Failure Rates and Observations of Time Dependent Covariables.
(Part 1: Theory.)

105 M.H.C. Paardekooper
A Newton-like method for error analysis. Applied to linear continuous
systems and eigenproblems.

106 G. van der Laan en A.J.J. Talman
Note of the Path following approach of equilibrium programming.