PSYCHOMETRIKA—VOL. 52, NO. 1, 125–135
MARCH 1987

THE RANK OF REDUCED DISPERSION MATRICES

PAUL A. BEKKER

DEPARTMENT OF ECONOMETRICS, TILBURG UNIVERSITY

JAN DE LEEUW

DEPARTMENT OF DATA THEORY FSW, UNIVERSITY OF LEIDEN

Psychometricians working in factor analysis and econometricians working in regression with measurement error in all variables are both interested in the rank of dispersion matrices under variation of the diagonal elements. Psychometricians concentrate on cases in which low rank can be attained, preferably rank one, the Spearman case. Econometricians concentrate on cases in which the rank cannot be reduced below the number of variables minus one, the Frisch case. In this paper we give an extensive historial discussion of both fields, we prove the two key results in a more satisfactory and uniform way, we point out various small errors and misunderstandings, and we present a methodological comparison of factor analysis and regression on the basis of our results.

Key words: factor analysis, errors of measurement, structural regression, functional models, communalities, errors in variables.

1. Introduction

Suppose Σ is a symmetric positive definite matrix of order m. In this paper we study the function rank $(\Sigma - \Omega)$ as Ω varies over the diagonal matrices satisfying $0 \leq \Omega \leq \Sigma$. This inequality notation is convenient shorthand for the requirement that both Ω and $\Sigma - \Omega$ must be positive semidefinite. In particular we are studying:

$$mr(\Sigma) \equiv \min \{\text{rank}(\Sigma - \Omega) \mid 0 \leq \Omega \leq \Sigma; \Omega \text{ diagonal}\}. \tag{1}$$

Investigation of this matrix function is important in at least two data analytic fields. The first field, which is very familiar for most readers of the psychometric literature, is factor analysis. In this context $mr(\Sigma)$ corresponds to the number of common factors. The older factor analysis literature concentrated on studying conditions for $mr(\Sigma) = 1$, while later contributions were mainly concerned with finding bounds or estimates of $mr(\Sigma)$. In our first historical section we shall review the most important algebraic results from the factor analysis literature.

The second field, which has had far less attention in the psychometric literature, is regression with errors of measurement in the variables. This model has been mainly studied in econometrics, with the major emphasis on conditions for $mr(\Sigma) = m - 1$. We shall also review the most important contributions from econometrics. This will also give us the opportunity to contrast the factor analysis model with the regression model.

2. The Spearman Model

As we remarked in the introduction most of the early factor analysis literature concentrated on the characterization of matrices for which the Spearman model with a

Financial support by the Netherlands Organization for the Advancement of Pure Research (ZWO) is gratefully acknowledged.

Requests for reprints should be sent to Jan de Leeuw, Department of Data Theory FSW/RUL, Middelstegracht 4, 2312 TW Leiden, THE NETHERLANDS.

0033-3123/87/0300-8005$00.75/0

single common factor was appropriate. In the present context we might say that the early literature concentrated on finding conditions for $mr(\Sigma) = 1$. We briefly review this work which is riddled with errors and imprecisions.

In his famous paper on general intelligence Spearman (1904, p. 274) used the *hierarchy of correlations* as a criterion. Tests should have the property that, after rearrangement, correlations were decreasing in each row, and that rows were proportional. This turned out to be somewhat too subjective and informal.

Krueger and Spearman (1907, p. 84–85) derived a new criterion from the correction for attenuation formula. It was

$$(\rho_{ik}\rho_{i\ell}\rho_{jk}\rho_{j\ell})^{1/4} = (\rho_{ij}\rho_{k\ell})^{1/2}, \tag{2}$$

which had to hold true for all quadruples ($i \neq j, k, \ell; j \neq k, \ell; k \neq \ell$). Observe that we have formulated the criterion in terms of the correlation matrix R, with elements ρ_{ij}, which are tacitly assumed to be nonnegative. From the Krueger and Spearman formula it is easy to derive

$$\frac{\rho_{ik}}{\rho_{i\ell}} = \frac{\rho_{jk}}{\rho_{j\ell}}, \qquad (i \neq j, k, \ell; j \neq k, \ell; k \neq \ell). \tag{3}$$

This formula was published for the first time by Burt (1909, p. 159). He did not publish a proof, but he indicated that he derived it from the Krueger-Spearman formula, probably with help from Spearman. The actual (one-line) proof was not published until Spearman (1927, appendix, p. ii). Hart and Spearman (1912, p. 58, footnote) derived (3) from the partial correlation formula of Yule. Garnett (1919a, 1919b) referred to the conditions as *Burt's equations*, and he stated that there were only $\frac{1}{2}m(m-3)$ independent equations among the $m!/(m-4)!$ possible ones. This was proved in Garnett (1920, p. 245), where the name he had proposed for the conditions was formally withdrawn. Perhaps this was one of the seeds that grew into Burt's later attempts to rewrite the history of factor analysis (Hearnshaw, 1981, chap. 9). The conditions (3) were called the vanishing of the *tetrad* differences by Spearman and Holzinger (1924, 1925), who also wrote them in the more convenient form

$$\rho_{ik}\rho_{j\ell} - \rho_{i\ell}\rho_{jk} = 0, \qquad (i \neq j, k, \ell; j \neq k, \ell; k \neq \ell). \tag{4}$$

It is clear that in the earlier formulations the possibility of negative and zero correlations was sometimes overlooked. In fact one often has the impression that positivity of the correlations was treated as part of the definition of the Spearman hierarchy. There is a second imprecision, which is perhaps more serious. If we define

$$mr_*(\Sigma) \equiv \min \{\text{rank}(\Sigma - \Omega) \mid \Omega \text{ diagonal}\}, \tag{5}$$

then the vanishing of all tetrads (or of $\frac{1}{2}m(m-3)$ independent tetrads) is necessary and sufficient for $mr_*(\Sigma) \leq 1$. However, in general,

$$mr_*(\Sigma) \leq mr(\Sigma), \tag{6}$$

and there is no guarantee of equality. Remarks to this effect were already made by Garnett, but it was pointed out for the first time by Wilson (1928) and Camp (1932) that the conditions

$$\rho_{jk} \geq \rho_{ik}\rho_{ji}, \tag{7}$$

were necessary as well for $mr(\Sigma) = 1$.

It is remarkable that the formulation of the conditions for $mr(\Sigma) = 1$ took about thirty years. In fact, the results can be summarized in a single comprehensive theorem.

Once it is formulated, the proof is almost immediate. In order to do so we assume, without loss of generality, that Σ is irreducible, i.e. Σ cannot be brought, by permutations, into block-diagonal form. We also say that Σ is a *Spearman matrix* if $mr(\Sigma) = 1$.

Theorem 1. A positive definite, irreducible matrix Σ is a Spearman matrix if, and only if, after sign changes of rows and corresponding columns, all its elements are positive and such that

$$\sigma_{ik}\sigma_{j\ell} - \sigma_{i\ell}\sigma_{jk} = 0, \tag{8}$$

and

$$\sigma_{ik}\sigma_{ji} - \sigma_{ii}\sigma_{jk} \leq 0, \tag{9}$$

for all quadruples $(i \neq j, k, \ell; j \neq k, \ell; k \neq \ell)$.

Proof. Necessity is obvious. For sufficiency, let $q = (q_1, \ldots, q_m)'$ be an m-vector. It will be shown that q exists such that $\Sigma - qq'$ is diagonal with nonnegative elements. As we may assume that all elements of Σ are positive, it follows from (8) that $q_i \equiv (\sigma_{ik}\sigma_{ji}/\sigma_{jk})^{1/2}$ exists independently of the choice of k and j, $(i \neq j, k; j \neq k)$. Hence, $\sigma_{ij} = q_i q_j$. It follows from (9) that $\sigma_{ii} \geq q_i^2$. Q.E.D. □

There have been various attempts to generalize the approach of this theorem, and the kind of result we have obtained, to various other combinations of m and p (with $p = mr(\Sigma)$: the numbers of common factors). Kelley (1928), Wilson (1929), Wilson and Worcester (1934, 1939) study special cases such as (m, p) equal to (4, 2), (5, 2), or (6, 3). Many different special cases must be distinguished, and very little has been achieved in terms of general results.

3. The Ledermann Bound

Kelley (1928) also tries to provide much more general results, which are true for all (m, p). If we write the factor model as $\Sigma = \Lambda\Lambda' + \Omega$, then we find $\frac{1}{2}m(m+1)$ equations in $m + mp - \frac{1}{2}p(p-1)$ unknowns (taking rotational indeterminacy into account). The number of unknowns exceeds the number of equations if

$$p > p(m) \equiv \tfrac{1}{2}\{2m + 1 - (8m + 1)^{1/2}\}. \tag{10}$$

Kelley (1928), and later Thurstone (1935), therefore suggest that $mr(\Sigma) \leq p(m)$ for all Σ. This despite a warning from Wilson "There is perhaps no more tricky part of mathematics than that involved in counting equations and variables to determine whether or not the equations can in general be solved. Today this kind of mathematics is, among pure mathematicians, taboo except as a heuristic device" (Wilson, 1929, p. 156).

Ledermann (1937) has tried to put the bound $mr(\Sigma) \leq p(m)$ on a somewhat more rigorous footing. Not with much success though. "Nous offrons un pétale de rose a quiconque énoncera clairement et démontrera surement les resultats que Ledermann a voulu nous communiquer" (We offer a rose petal to anyone who clearly enunciates and positively demonstrates the results that Ledermann has wanted to communicate to us; Hakim, Lochard, Olivier, & Terouanne, 1976, p. 24). We shall make an attempt. Ledermann writes $mr_*(\Sigma) \leq p$ as a system of $\frac{1}{2}(m-p)(m+1-p)$ determinantal equations with the elements of Ω as the m unknowns. Of course each determinant can be expanded, which gives a system of $\frac{1}{2}(m-p)(m+1-p)$ polynomial equations in m unknowns. Again the number of unknowns exceeds the number of equations if $p > p(m)$. Ledermann proves, in addition, that at least for some choices of Σ these equations are independent, that is, none of them is a consequence of the others. Of course this still does not imply much about their solvability, or about the number of solutions they have.

Perhaps the most general result about the Lederman bound, which has been proved rigorously, is due recently to Shapiro (1982a). He proves that $mr_*(\Sigma) \geq p(m)$ almost surely, that is, the dispersion matrices Σ for which $mr_*(\Sigma) < p(m)$ form a set of (Lebesgue) measure zero. Although this result is theoretically of some interest, it does not give any valuable information for specific matrices Σ.

4. Beyond The Ledermann Bound

We first mention, as an important step ahead, the work of Albert (1944a, 1944b). He defined $\mu(\Sigma)$, the *ideal rank* of Σ, to be the rank of the largest nonsingular square off-diagonal submatrix. Obviously $mr_*(\Sigma) \geq \mu(\Sigma)$. Albert gives necessary and sufficient conditions for equality in his first paper, and he gives a sufficient condition for equality in the second paper. Tumura and Fukutomi (1968) give another sufficient condition.

Guttman (1954, p. 160) observed that "merely studying the minors outside the main diagonal was not sufficient" for determining $mr(\Sigma)$, and he showed that $mr(\Sigma) = m - 1$ for correlation matrices with two different latent roots, the largest of which with a multiplicity of $m - 1$.

Guttman (1956, Theorem 1) argued that $\underline{mr}(\Sigma) \leq m - k$ if Σ^{-1} has a $k \times k$ diagonal principal submatrix. Guttman (p. 283) also gave a lower bound, which was a correction of an earlier given bound (Guttman, 1954, p. 153). Let D^{-1} be the diagonal matrix formed by the diagonal elements of Σ^{-1} and let $\pi(\cdot)$ denote the number of positive roots. Then a lower bound is given by $mr(\Sigma) \geq \pi(\Sigma - D)$. In the same paper he presented an interesting inequality which may be reformulated as follows:

Theorem 2. If Σ is irreducible, then

$$mr(\Sigma) + mr(\Sigma^{-1}) \geq m. \tag{11}$$

Proof. Guttman (1956, p. 283) showed that $mr(R) + mr(SR^{-1}S) \geq m$, where R is an $(m \times m)$-correlation matrix and S in a diagonal matrix such that the diagonal elements in $SR^{-1}S$ equal one. Obviously, if R is a correlation matrix corresponding to Σ, then $mr(\Sigma) = mr(R)$, and $mr(\Sigma^{-1}) = mr(SR^{-1}S)$. Q.E.D. □

In order to prove that his inequality was the "best possible" one, Guttman provided examples for which the inequality becomes an equality. Curiously, he did not use Spearman matrices for this purpose. Obviously, as $\underline{mr}(\Sigma) \leq m - 1$ for any Σ (Guttman, 1954, p. 159), it must hold true that (11) becomes an equality if $mr(\Sigma) = 1$, or $mr(\Sigma^{-1}) = 1$. As an interesting consequence of (11) we thus have that $mr(\Sigma) = m - 1$ if Σ^{-1} is an irreducible Spearman matrix.

The fallacy behind interpreting the Ledermann bound as providing an upper bound to the number of common factors was further discussed in Guttman (1958). It was remarked that a symmetric tridiagonal Σ, with all subdiagonal elements nonzero, had $mr_*(\Sigma) = m - 1$. Guttman was mistaken in his assertions made in the same paper that for the "perfect simplex" $mr_*(\Sigma) = m - 2$, and for the "quasi-simplex" $mr_*(\Sigma) = m - 3$. He tried to prove that if Σ^{-1} is tridiagonal, as is true for the perfect simplex, then $mr_*(\Sigma) = m - 2$. This is very strange, since application of the aforementioned Theorem 1 in Guttman (1956) shows that if Σ^{-1} is tridiagonal, and $m \geq 5$, then even $mr(\Sigma)$ is smaller than $m - 2$.

Tumura and Fukutomi (1968) proved that the tridiagonality of Σ, with nonzero subdiagonal elements, is not only sufficient for $mr_*(\Sigma) = m - 1$, it is, after permutation, also necessary. Their proof is difficult to understand. Hakim et al. (1976, p. 26) give another proof, which is more solid but a bit complicated. Fairly simple proofs are available in Fiedler (1969) and Rheinboldt and Sheperd (1974), who discovered the theorem in an entirely different context.

A slightly more interesting theorem, which was proved by Shapiro (1982b), implies the existence of a set of nonzero (Lebesgue) measure satisfying $mr(\Sigma) = m - 1$. Let

$$\mathrm{mr}_{\S}(\Sigma) \equiv \min\{\mathrm{rank}(\Sigma - \Omega) \mid \Omega \leq \Sigma;\ \Omega \text{ diagonal}\}, \tag{16}$$

so that

$$\underline{\mu(\Sigma)} \leq mr_*(\Sigma) \leq mr_{\S}(\Sigma) \leq mr(\Sigma). \tag{17}$$

Shapiro proved that a necessary and sufficient condition for $mr_{\S}(\Sigma) = m - 1$ is that Σ is irreducible and all its off-diagonal elements can be made nonpositive by sign changes of rows and corresponding columns.

So far we have found conditions which can be considered as sufficient conditions for $mr(\Sigma) = m - 1$. Necessary and sufficient conditions are given in the following theorem, which is similar to a result that has been proved by Hakim et al. (1976, p. 14, Corollaire 2.4).

Theorem 3. $mr(\Sigma) = m - 1$, if and only if, for each vector $\gamma' = (\gamma_1, \ldots, \gamma_m) \neq 0$ such that $(\Sigma - \Omega)\gamma = 0$, where Ω is diagonal and $0 \leq \Omega \leq \Sigma$, $\gamma_i \neq 0$ for all $i = 1, \ldots, m$.

Proof: let Σ and Ω be partitioned as

$$\Sigma = \begin{bmatrix} \sigma_{11} & \sigma_{12} \\ \sigma_{21} & \Sigma_{22} \end{bmatrix}, \qquad \Omega = \begin{bmatrix} \omega_{11} & 0 \\ 0 & \Omega_{22} \end{bmatrix}.$$

(i) Sufficiency. If rank $(\Sigma - \Omega) \leq m - 2$, then rank $(\sigma_{21}, \Sigma_{22} - \Omega_{22}) \leq m - 2$, and thus there exists an $(m - 1)$-vector $\alpha \neq 0$, such that $(\sigma_{21}, \Sigma_{22} - \Omega_{22})'\alpha = 0$. Let $\gamma' = (0, \alpha')$.

(ii) Necessity. Without loss of generality we assume that $\gamma_1 = 0$. Let β and δ be $(m - 1)$-vectors such that $\sigma_{21} = (\Sigma_{22} - \Omega_{22})\beta + \delta$, where $(\Sigma_{22} - \Omega_{22})\delta = 0$. Then $(0, \delta')(\Sigma - \Omega)(0, \delta')' = 0$, and as $\Sigma - \Omega \geq 0$, $(\Sigma - \Omega)(0, \delta')' = 0$. Hence $\sigma_{12}\delta = \delta'\delta = 0$, so that $\delta = 0$. Consequently, $\Sigma - \Omega = (\beta, I)'(\Sigma_{22} - \Omega_{22})(\beta, I) + (1, -\beta')(\Sigma - \Omega)(1, -\beta')'e_1e_1'$, where $e_1' = (1, 0, \ldots, 0)$. Let $\tilde{\Omega} = \Omega + (1, -\beta')(\Sigma - \Omega)(1, -\beta')' e_1e_1'$, so that $\tilde{\Omega}$ is diagonal and $0 \leq \tilde{\Omega} \leq \Sigma$. As $(\Sigma - \tilde{\Omega})\gamma = 0$ and $(\Sigma - \tilde{\Omega})(1, -\beta')' = 0$, the independence of γ and $(1, -\beta')'$ implies that $mr(\Sigma) \leq m - 2$. This contradicts the assumption $mr(\Sigma) = m - 1$. □

Of course, the conditions in this theorem are not very satisfactory. However, it is an important step towards the formulation of simple necessary and sufficient conditions for $mr(\Sigma) = m - 1$. These conditions, and also the conditions in Theorem 3 (Reiersøl, 1941, Theorem 12), have been derived in an entirely different tradition, which we will now review.

5. Structural Regression

Consider the following "errors-in-variables model"

$$\xi_i'\gamma = 0, \qquad i = 1, \ldots, n, \tag{18}$$

$$x_i' = \xi_i + \varepsilon_i, \tag{19}$$

where the stochastic vectors $(x_i', \xi_i', \varepsilon_i')$, $i = 1, \ldots, n$, are i.i.d. with zero expectation. The m variables in ξ_i are not observed, instead the m variables in x_i are observed. It is assumed that the disturbances, or measurement errors, in ε_i are mutually independent and also independent of the systematic parts in ξ_i. The (fixed) m-vector γ is called the *structural vector*.

Obviously, the model can be considered as a regression model where all variables are subject to measurement error. Indeed, if only one of the errors in ε has a nonzero variance, so that $m - 1$ errors equal zero almost surely, then the model represents an elementary regression, where one of the variables is regressed on the other variables. As we can do this for each variable, we can also find m different *elementary regression vectors* γ.

Let Σ and Ω denote the covariance matrices of x_i and ε_i, respectively, where Σ is assumed to be nonsingular. Then the covariance matrix of ξ_i is given by $\Sigma - \Omega$, where Ω is diagonal and

$$0 \leq \Omega \leq \Sigma. \tag{20}$$

Furthermore, γ satisfies the moment equations

$$(\Sigma - \Omega)\gamma = 0. \tag{21}$$

If Ω is known up to a proportionality factor, that is, $\Omega = \mu\Omega^*$, where Ω^* is some known fixed positive semidefinite matrix, then the equations in (20) and (21) can be used for estimation purposes. That is to say, Σ should be replaced by its sample estimate $\hat{\Sigma}$, μ should be set equal to the smallest root $\hat{\mu}$ of the determinantal equation $|\hat{\Sigma} - \hat{\mu}\Omega^*| = 0$, and γ may be estimated by $(\hat{\Sigma} - \hat{\mu}\Omega^*)\hat{\gamma} = 0$. This gives consistent estimates. As an example we may consider "orthogonal regression," where $\Omega^* = I$, which was introduced by Pearson (1901), or we may consider the m different elementary regressions, where each time Ω has only one nonzero diagonal element. It follows immediately from (21) that the i-th elementary regression vector γ must be proportional to the i-th column of Σ^{-1}; the estimate of the i-th elementary regression vector is thus proportional to the i-th column of $\hat{\Sigma}^{-1}$.

However, in general, the dispersion matrix of the measurement errors is not known up to a proportionality factor, and many diagonal matrices Ω and vectors γ satisfy both (20) and (21). In other words, if all variables are normally distributed, so that the moment equations (21) contain all information with respect to γ, then the model is not identified; which corresponds to the underidentification of a factor model with $m - 1$ factors. Other distributional assumptions may result in identification. For example Bekker (1986) gives a condition that is sufficient for the present model to be identified. However, in this paper it is assumed that the equations in (21) contain all information.

Another problem is that there may exist diagonal matrices Ω satisfying (20) such that rank $(\Sigma - \Omega) < m - 1$. In that case one may not exclude the possibility that there exist two, or even more, linear relations between the systematic parts in ξ_i. Consequently, as has been noted by Frisch (1934, p. 191), it would be sheer nonsense, in such cases, to look for significant elementary regression coefficients.

Frisch (1934) was the first to study these problems in some depth in his "confluence analysis." In particular he proved that for two observed variables, as in simple regression, the structural regression vector must be proportional to a convex linear combination of the two elementary regression vectors. As a result the correct regression line is located between the two elementary regression lines.

Although Frisch conjectured that similar conditions held in the general m-variables case, Koopmans (1937, p. 98–101) was the first to present an m-variable analog of Frisch's result. It says that the structural regression vector is proportional to a convex linear combination of the elementary regression vectors, subject to the condition that all elements of Σ^{-1} are strictly positive. It is clear that, as the columns of Σ^{-1} are proportional to the elementary regression vectors, the condition in the theorem can be satisfied, after sign changes, if all elementary regression vectors are located in the same orthant.

Koopmans' proof is complicated. Reiersøl (1941, p. 8) noted the applicability of a theorem by Frobenius, and all later proofs given by Reiersøl (1945), Dhondt (1960), Patefield (1981), Kalman (1982a) and Klepper and Leamer (1984) make use of the Perron-Frobenius theorem, just as Shapiro (1982b) did when proving his result on $mr_{\S}(\Sigma) = m - 1$.

The second part of Koopmans' theorem says that if all elements of Σ^{-1} are strictly positive, then each vector in the cone (by which we mean the union of all scalar multiples of the convex hull) of the elementary regression vectors is a structural vector for some diagonal error dispersion matrix Ω satisfying both (20) and (21). Koopmans (1937, p. 103) also claimed to have proved this second part of the theorem. However, as has been pointed out by Kalman (1982a, p. 152), Koopmans' proof was wrong. Later proofs were given by Reiersøl (1945), Kalman, and Klepper and Leamer (1984); again all authors use the Perron-Frobenius theorem.

Here a formulation of the theorem will be presented which is slightly more general than previous formulations. The theorem will be proved without using the Perron-Frobenius theorem. Furthermore, the two parts of the theorem will be proved almost simultaneously, thereby emphasizing the if and only if argument in the theorem.

It will be convenient to use the following lemma. Let A be a symmetric matrix with strictly positive off-diagonal elements, $A_{ij} > 0$ if $i \neq j$. Let Λ be a diagonal matrix, $\Lambda = \text{diag}(\lambda)$, and let u be a vector of ones, $u' = (1, \dots, 1)$, q is an arbitrary vector.

Lemma 1. $\text{diag}(\Lambda A \Lambda u) \geq \Lambda A \Lambda$, if and only if, $\lambda_i \lambda_j \geq 0$ for all i, j.

Proof.

$$q'\{\text{diag}(\Lambda A \Lambda u) - \Lambda A \Lambda\} q = \sum_i \sum_j A_{ij} \lambda_i \lambda_j (q_i^2 - q_i q_j)$$
$$= \sum_{i} \sum_{< j} A_{ij} \lambda_i \lambda_j (q_i - q_j)^2.$$

(i) If $\lambda_i \lambda_j \geq 0$, for all i, j, then $A_{ij} \lambda_i \lambda_j (q_i - q_j)^2 \geq 0$, for all i, j.
(ii) If for some i, j $\lambda_i \lambda_j < 0$, then choose $q_i = \text{sign}(\lambda_i)$, so that

$$\sum_{i} \sum_{< j} A_{ij} \lambda_i \lambda_j (q_i - q_j)^2 < 0. \text{ Q.E.D.}$$ □

We will also use the result that $0 \leq \Omega \leq \Sigma$ is equivalent to $\Omega \geq \Omega\Sigma^{-1}\Omega$ if Σ is positive definite. The proof is simple. If $0 \leq \Omega \leq \Sigma$, then $\Omega - \Omega\Sigma^{-1}\Omega = (I - \Omega\Sigma^{-1})\Omega(I - \Sigma^{-1}\Omega) + \Omega\Sigma^{-1}(\Sigma - \Omega)\Sigma^{-1}\Omega \geq 0$. If $\Omega \geq \Omega\Sigma^{-1}\Omega \geq 0$, then $\Sigma - \Omega = \Omega - \Omega\Sigma^{-1}\Omega + (I - \Omega\Sigma^{-1})\Sigma(I - \Sigma^{-1}\Omega) \geq 0$. A more general result, where also Σ is allowed to be singular, has been given by Bekker, Kapteyn and Wansbeek (1984, p. 88).

Contrary to most other proofs we do not assume that the error dispersion matrix Ω is nonsingular. Without loss of generality we assume

$$\Omega \equiv \begin{bmatrix} \Omega_1 & 0 \\ 0 & 0 \end{bmatrix}, \tag{22}$$

where Ω_1 is a $k \times k$ diagonal matrix, $k \leq m$. Σ and Σ^{-1} are partitioned analogously, that is,

$$\Sigma \equiv \begin{bmatrix} \Sigma_{11} & \Sigma_{12} \\ \Sigma_{21} & \Sigma_{22} \end{bmatrix}; \qquad \Sigma^{-1} \equiv \begin{bmatrix} (\Sigma^{-1})_{11} & (\Sigma^{-1})_{12} \\ (\Sigma^{-1})_{21} & (\Sigma^{-1})_{22} \end{bmatrix} \equiv (\Sigma_1^{-1}, \Sigma_2^{-1}), \tag{23}$$

and also the vector γ has a corresponding partitioning $\gamma' = (\gamma_1', \gamma_2')$.

Theorem 4. Let $(\Sigma^{-1})_{11}$ have strictly positive elements.

(i) If Ω is of the form (22) such that $0 \leq \Omega \leq \Sigma$ and $(\Sigma - \Omega)\gamma = 0$, then γ lies in the cone of Σ_1^{-1}.

(ii) For each $\gamma \neq 0$ in the cone of Σ_1^{-1} there exists one and only one Ω of the form (22) such that $0 \leq \Omega \leq \Sigma$, and that $(\Sigma - \Omega)\gamma = 0$.

Proof. Define $\lambda \equiv \Omega_1\gamma_1$, $\Lambda \equiv \text{diag}(\lambda)$, $C \equiv \text{diag}(\gamma_1)$, so that $\gamma = \Sigma_1^{-1}\lambda$, $\Lambda = \Omega_1 C$, and $C = \text{diag}((\Sigma^{-1})_{11}\Lambda u)$.

(i) Let,

$$0 \leq \Omega \leq \Sigma. \tag{24}$$

As has been proved above, this is equivalent to

$$\Omega_1 \geq \Omega_1(\Sigma^{-1})_{11}\Omega_1. \tag{25}$$

Pre- and post-multiplication by C implies

$$C\Omega_1 C \geq C\Omega_1(\Sigma^{-1})_{11}\Omega_1 C. \tag{26}$$

Setting $\Lambda = \Omega_1 C$ and $C = \text{diag}\,((\Sigma^{-1})_{11}\Lambda u)$ gives an equivalent inequality

$$\text{diag}(\Lambda(\Sigma^{-1})_{11}\Lambda u) \geq \Lambda(\Sigma^{-1})_{11}\Lambda, \tag{27}$$

which is, according to Lemma 1, equivalent to

$$\lambda_i\lambda_j \geq 0, \qquad i, j = 1, \ldots, k. \tag{28}$$

As a result, all elements of λ have the same sign, so $\gamma = \Sigma_1^{-1}\lambda$ lies in the cone of Σ_1^{-1}.

(ii) If $\gamma \neq 0$ lies in the cone of Σ_1^{-1}, then γ_1 must have strictly positive or negative elements, so C is nonsingular. Hence $\Omega_1 = \Lambda C^{-1}$ is unique and (25) and (26) are equivalent, so that (28) implies (24). □

Thus we have proved that if $(\Sigma^{-1})_{11}$ has strictly positive elements and only the first k variables are subject to measurement error, then the set of all structural vectors *is* the cone of the first k elementary regression vectors.

This important theorem can be used to derive necessary and sufficient conditions for $mr(\Sigma) = m - 1$.

Theorem 5. $mr(\Sigma) = m - 1$, if and only if, Σ^{-1} has strictly positive elements, possibly after sign changes of rows and corresponding columns.

Proof. (i) If Σ^{-1} has strictly positive elements, then, by theorem 4(i), the null-space of $\Sigma - \Omega$ is contained within the cone of Σ^{-1}. Hence all elements of this null-space lie in the strictly positive or in the strictly negative orthant. Consequently, this null-space can be at most one-dimensional.

(ii) If not all elements of Σ^{-1} have compatible signs, then there are two columns of Σ^{-1}, the i-th and j-th say, such that the (ij)-th element of Σ^{-1} is positive, possibly after sign changes of rows and columns, while the i-th and j-th columns do not lie in the same orthant (or $\Sigma_{ij}^{-1} = 0$). That is to say that in the cone of these columns (or in the cone of the i-th column) there is a vector γ with a zero element. As the 2×2 submatrix of Σ^{-1}, formed by its i-th and j-th rows and columns, has strictly positive elements (or the 1×1 submatrix of Σ^{-1} formed by its ii-th element is positive), we may apply theorem 4(ii).

Hence the vector γ is a vector in the null-space of $\Sigma - \Omega$ for some diagonal Ω satisfying (20). Then, according to theorem 3, $mr(\Sigma) < m - 1$. Q.E.D. □

The result in this theorem has been proved before by Reiersøl (1941, Theorem 14) and by Kalman (1982a). However, Reierøl's proof is not very transparant. Kalman does not use the result in Theorem 3, instead he reduces the μ-variable problem to a three variable problem in order to prove the second part of the theorem. Klepper and Leamer (1984) presented Theorem 5 in a disguised form. They stated that the coefficients in the normalized structural vector are bounded if and only if all elementary regression vectors lie in the same orthant. Their proof uses also a reduction to a 3-variable problem.

Using Theorem 4(i) it is easy to derive two different generalizations of Theorem 5(i). Again, let Σ_{11} be a $k \times k$ principal submatrix of Σ, and let $(\Sigma^{-1})_{11}$ be the corresponding $k \times k$ submatrix of Σ^{-1}, $k \leq m$.

Theorem 6. (i) If $(\Sigma^{-1})_{11}$ has strictly positive elements then $mr(\Sigma) \geq k - 1$.
(ii) If $(\Sigma_{11})^{-1}$ has strictly positive elements then $mr(\Sigma) \geq k - 1$.

Proof. Let, analogous to the partitioning of Σ, $\Omega = \Omega^{(1)} + \Omega^{(2)}$, such that

$$\Omega^{(1)} = \begin{bmatrix} \Omega_1 & 0 \\ 0 & 0 \end{bmatrix}, \qquad \Omega^{(2)} = \begin{bmatrix} 0 & 0 \\ 0 & \Omega_2 \end{bmatrix},$$

where Ω_1 and Ω_2 are diagonal.

(i) If $0 \leq \Omega^{(1)} \leq \Sigma$, then it follows from Theorem 4(i) and a reasoning analogous to the one used in the proof of Theorem 5(i), that rank $(\Sigma - \Omega^{(1)}) \geq m - 1$. Therefore, an elementary inequality (Marsaglia & Styan, 1974), rank $(\Sigma - \Omega) \geq$ rank $(\Sigma - \Omega^{(1)})$ − rank $(\Omega^{(2)})$, implies that rank $(\Sigma - \Omega) \geq k - 1$.

(ii) If $0 \leq \Omega^{(1)} \leq \Sigma$, then $0 \leq \Omega_1 \leq \Sigma_{11}$. Consequently, according to Theorem 5(i), rank $(\Sigma_{11} - \Omega_1) \geq k - 1$. As $\Sigma_{11} - \Omega_1$ is a submatrix of $\Sigma - \Omega$, the result follows. Q.E.D. □

6. Discussion

Both factor analysis and structural regression analysis are extremes of a common model which simply says that, apart from unique components or error components, the variables in the analysis are linearly related. In other words, the model says that the covariance matrix $\Sigma - \Omega$ has a deficient rank. In factor analysis attention centres around the low-dimensional range-space of $\Sigma - \Omega = \Lambda\Lambda'$. In structural regression the model is formulated in terms of the one-dimensional null-space: $(\Sigma - \Omega)\gamma = 0$.

Evaluation of $mr(\Sigma)$, or $m - mr(\Sigma)$, is important in both fields. In fact, the number $mr(\Sigma)$ tells us whether the common model should be considered as a factor analysis model or as a structural regression model, or even, as some model in between. Consequently, if a structural regression model is justified, that is, if the necessary and sufficient conditions for $mr(\Sigma) = m - 1$ are satisfied, then applying a factor model to the data would be nonsense. Just as Frisch thought it nonsense to look for a single linear relation in case there are two or more linear relations between the variables.

These latter models, where there exist a number, albeit a small number, of linear relations between the variables, have had relatively little attention in the literature. It is only recently that Kalman (1982a, 1982b, 1983, 1984) discussed these models in some detail. In his 1983 paper he claims, "It is impossible to avoid the conclusion that the lack of progress on and the present confusion surrounding Frisch's ideas are due to mathematical rather than conceptual difficulties" (p. 119). Indeed, as we have seen, there are necessary and sufficient conditions for $mr(\Sigma) = 1$ and $mr(\Sigma) = m - 1$, however, no such conditions are available for intermediate values of $mr(\Sigma)$.

Although no complete solutions are known for intermediate values of $mr(\Sigma)$, there are sufficient conditions. For example, using theorems 5 and 6, it is easy to derive sufficient conditions for $mr(\Sigma) = m - 2$. On the other hand, there does exist a necessary and sufficient condition for an intermediate value of $mr(\Sigma)$ in case m is small. Clearly if $m = 3$, then necessary and sufficient conditions are available for all values of $mr(\Sigma)$. If $m = 4$, then $mr(\Sigma) = 1$ and $mr(\Sigma) = 3$ are completely characterized by application of theorems 1 and 5 respectively. Consequently, if $m = 4$, also $mr(\Sigma) = 2$ is completely characterized, since all (irreducible) matrices Σ that do not satisfy the necessary and sufficient conditions for $mr(\Sigma) = 1$ and $mr(\Sigma) = 3$, and only those matrices Σ, must have $mr(\Sigma) = 2$.

Kalman (1984, p. 118) also claims that "it is possible to given a (rigorous) closed-form solution" for the case $m = 5$ and $mr(\Sigma) = 3$. As we have difficulty in arriving at this closed-form solution, we would like to offer a *pétale de rose* to anyone who is able to produce that solution.

References

Albert, A. A. (1944a). The matrices of factor analysis. *Proceedings of the National Academy of Sciences, 30*, 90–95.

Albert, A. A. (1944b). The minimum rank of a correlation matrix. *Proceedings of the National Academy of Sciences, 30*, 144–148.

Bekker, P. A. (1986). Comment on identification in the linear errors in variables model. *Econometrica, 54*, 215–217.

Bekker, P. A., Kapteyn, A., & Wansbeek, T. J. (1984). Measurement error and endogeneity in regression: Bounds for ML and VI estimates. In T. K. Dijkstra (ed.), *Misspecification analysis*. Berlin: Springer Verlag.

Burt, C. (1909). Experimental tests of general intelligence. *British Journal of Psychology, 3*, 94–177.

Camp, B. H. (1932). The converse of Spearman's two-factor theorem. *Biometrika, 24*, 418–428.

Dhondt, A. (1960). Sur une généralisation d'un theorème de R. Frisch en analyse de confluence [On a generalization of a theorem by R. Frisch about confluence analysis]. *Cahiers du Centre d'Etude de Recherche Operationelle, 2*, 37–46.

Fiedler, M. (1969). A characterization of tridiagonal matrices. *Linear Algebra and its Applications, 2*, 191–197.

Frisch, R. (1934). *Statistical confluence analysis by means of complete regression systems* (Publication No. 5). Oslo: University of Oslo, Economic Institute.

Garnett, J. C. M. (1919a). On certain independent factors in mental measurement. *Proceedings of the Royal Society of London, 96*, 91–111.

Garnett, J. C. M. (1919b). General ability, cleverness, and purpose. *British Journal of Psychology, 9*, 345–366.

Garnett, J. C. M. (1920). The single general factor in dissimilar mental measurements. *British Journal of Psychology, 10*, 242–258.

Guttman, L. (1954). Some necessary conditions for common factor analysis. *Psychometrika, 19*, 149–161.

Guttman, L. (1956). "Best possible" systematic estimates of communalities. *Psychometrika, 21*, 273–285.

Guttman, L. (1958). To what extent can communalities reduce rank? *Psychometrika, 23*, 297–308.

Hakim, M., Lochard, E. O., Olivier, J. P., & Terouanne, E. (1976). *Sur les traces de Spearman* [On the traces of Spearman] (Série Recherche N 25). Paris: Cahiers de Bureau Universitaire de Recherche Operationelle, Université Pierre et Marie Curie.

Hart, B., & Spearman, C. E. (1912). General ability, its existence and nature. *British Journal of Psychology, 5*, 51–84.

Hearnshaw, L. S. (1981). *Cyril Burt, psychologist*. New York: Random House.

Kalman, R. E. (1982a). System identification from noisy data. In A. R. Bednarek & L. Cesari (Eds.), *Dynamical Systems II*. New York: Academic Press.

Kalman, R. E. (1982b). Identification from real data. In M. Hazewinkel & A. H. G. Rinnooy Kan (Eds.), *Current developments in the interface economics, econometrics, mathematics*. Dordrecht: D. Reidel.

Kalman, R. E. (1983). Identifiability and modelling in econometrics. In P. R. Krishnaiah (Ed.), *Developments in Statistics 4*. New York: Academic Press.

Kalman, R. E. (1984). We can do something about multicollinearity! *Communications in Statistics—Theory and Methods, 13*, 115–125.

Kelley, T. L. (1928). *Crossroads in the mind of man: A study of differentiable mental abilities*. Stanford University Press.

Klepper, S. & Leamer, E. E. (1984). Consistent sets of estimates for regressions with errors in all variables. *Econometrica, 52*, 162–183.

Koopmans, T. C. (1937). *Linear regression analysis of economic time series.* Haarlem: De Erven F. Bohn NV.

Krueger, F. & Spearman, C. E. (1907). Die Korrelation zwischen verscheidenen geistigen Leistungsfähigkeiten [The correlation between several mental abilities]. *Zeitschrift für Psychologie, 54*, 50–114.

Ledermann, W. (1937). On the rank of the reduced correlational matrix in multiple factor analysis. *Psychometrika, 2*, 85–93.

Marsaglia, G., & Styan, G. P. H. (1974). Equalities and inequalities for ranks of matrices. *Linear and Multilinear Algebra, 2*, 269–292.

Patefield, W. M. (1981). Multivariate linear relationships: Maximum likelihood estimation and regression bounds. *Journal of the Royal Statistical Society (Series B), 43*, 342–352.

Pearson, K. (1901). On lines and planes of closest fit to points in space. *Philosophical Magazine 2*, 559–572.

Reiersøl, O. (1941). Confluence analysis by means of lag moments and other methods of confluence analysis. *Econometrica, 9*, 1–24.

Reiersøl, O. (1945). Confluence analysis by means of instrumental sets of variables. *Arkiv för Mathematik, Astronomi och Fysik, 32*(A), 1–119.

Rheinboldt, W. C., & Sheperd, R. A. (1974). On a characterization of tridiagonal matrices by M. Fiedler. *Linear Algebra and its Applications, 8*, 87–90.

Shapiro, A. (1982a). Rank reducibility of a symmetric matrix and sampling theory of minimum trace factor analysis. *Psychometrika, 47*, 187–199.

Shapiro, A. (1982b). Weighted minimum trace factor analysis. *Psychometrika, 47*, 243–264.

Spearman, C. E. (1904). General intelligence objectively measured and defined. *American Journal of Psychology, 15*, 201–299.

Spearman, C. E. (1927). *The abilities of man.* London: McMillan.

Spearman, C. E., & Holzinger, K. J. (1924). The sampling error in the theory of two factors. *British Journal of Psychology, 15*, 17–19.

Spearman, C. E., & Holzinger, K. J. (1925). Note on the sampling error of tetrad differences. *British Journal of Psychology, 15*, 86–89.

Thurstone, L. L. (1935). *The vectors of mind.* University of Chicago Press.

Tumura, Y., & Fukutomi, K. (1968). On the identification in factor analysis. *Report of Statistical Application Research, Union of Japanese Scientists and Engineers, 15*, 6–11.

Wilson, E. B. (1928). On hierarchical correlation systems. *Proceedings of the National Academy of Sciences, 14*, 283–291.

Wilson, E. B. (1929). Review of Kelley, Crossroads in the mind of man. *Journal of General Psychology, 2*, 153–169.

Wilson, E. B., & Worcester, J. (1934). The resolution of four tests. *Proceedings of the National Academy of Sciences, 20*, 189–192.

Wilson, E. B., & Worcester, J. (1939). The resolution of six tests into three general factors. *Proceedings of the National Academy of Sciences, 25*, 73–79.

Manuscript received 7/11/85
Final version received 5/22/86