Proper communality estimates minimizing the sum of the smallest eigenvalues for a covariance matrix

ten Berge, Jos M.F.; Kiers, H.A.L.

Published in:
The many faces of multivariate analysis (Vol.I)

Publication date:
1988

Citation for published version (APA):
ten Berge, J. M. F., & Kiers, H. A. L. (1988). Proper communality estimates minimizing the sum of the smallest eigenvalues for a covariance matrix. In M. G. H. Jansen, & W. H. van Schuur (Eds.), The many faces of multivariate analysis (Vol.I): Proceedings of the SMABS-88 Conference held in Groningen, December 18-21 (pp. 30-37). Groningen: RION, Groningen.

Proper communality estimates minimizing the sum of the smallest eigenvalues for a covariance matrix

Jos M.F. ten Berge and Henk A.L. Kiers[1)]

Abstract

The communality problem of factor analysis has a long history. Efforts to determine the exact minimum rank of a covariance (or correlation) matrix have failed to yield results of practical value. Practitioners tend to guess the approximate minimum rank, or try out various values, and then apply some computational method to obtain loadings, with communalities as by-products. However, the latter communalities are typically inadequate because they yield an indefinite covariance matrix. The present paper offers a method of communality estimation that preserves the non-negativity of eigenvalues of the covariance matrix, and, subject to this constraint, minimizes the sum of the r smallest eigenvalues, for any fixed integer r. This method can be interpreted as 'approximate low-rank factor analysis', and is a generalization of constrained minimum-trace factor analysis. Applying approximate low-rank factor analysis consecutively for various values of r may provide a numerical solution for the classical minimum-rank problem.

The authors are obliged to Henk Camstra for providing a computer program and for other computational assistance, and to Paul Bekker for useful and stimulating comments.

[1)]Vakgroep Psychologie, R.U. Groningen, Grote Markt 31/32, 9712 HV Groningen
Tel.: 050-636349
(EARN) JOSTENB@HGRRUG5

The communality problem has plagued factor analysts for more than half a century. Let Σ_x be any given nxn covariance or correlation matrix. Then the communality problem is the problem of finding a diagonal matrix U^2 such that $(\Sigma_x - U^2)$ has the smallest possible rank, subject to the constraint

$$\Sigma_x \geq U^2 \geq 0 \tag{1}$$

which means that both $(\Sigma_x - U^2)$ and U^2 have to be positive semi-definite. Communalities are defined as the diagonal elements of $(\Sigma_x - U^2)$. They are called 'proper' only if the constraint (1) is satisfied.

The communality problem has been approached from a mathematical and from an empirical point of view. As far as the mathematical approach is concerned, there is a long history of attempts to determine lower and upper bounds for the exact minimum rank of $(\Sigma_x - U^2)$, see Bekker and De Leeuw (1987) for a review. Efforts to determine the exact minimum rank lost much of their interest when Guttman (1958) showed that, at least for a certain class of (tridiagonal) matrices, the exact minimum rank is as high as $(n-1)$, a value far too high to permit low-rank approximations of Σ_x by factor analysis. Shapiro (1982b, Theorem 3.2) and Bekker and De Leeuw (1987, Theorem 5) have further extended Guttman's result. Except for such special cases, general solutions for determining the exact minimum rank are not available.

The most popular empirical approach to the communality problem is the Minres method of factor analysis (Harman & Jones, 1966; see also Jöreskog, 1977, and Zegers & Ten Berge, 1983). Given a fixed value of $r < n$, the Minres method determines an nxr factor loading matrix A such that the function

$$f(A) = ||\Sigma_x - U^2 - AA'||^2 , \tag{2}$$

with $U^2 = \text{Diag}(\Sigma_x - AA')$, is minimized, possibly subject to the constraint $U^2 \geq 0$ in order to prevent the occurrence of so-called Heywood cases. One may obtain an empirical approximation of minimum rank by trying out various values of r and adopting the lowest value of r for which the Minres method yields a sufficiently small value of f.

Neither the mathematical nor the empirical (Minres) approach to minimum rank is satisfactory. The mathematical approach is not attractive in practical applications (see above), and the empirical (Minres) approach is mathematically messy, because it typically produces a matrix $(\Sigma_x - U^2)$ with negative eigenvalues except for the first r. That is, Minres tends to yield improper communalities, violating the constraint (1).

The present paper is aimed at finding communalities that do satisfy (1) and that, subject to this constraint, minimize the sum of the $(n-r)$ smallest eigenvalues, where r is any fixed rank parameter. In this way, a proper communality solution is obtained such that $(\Sigma_x - U^2)$ has r relatively large eigenvalues and $(n-r)$ relatively small yet non-negative eigenvalues. A solution for this problem can be obtained in the following manner.

Let M denote the set of diagonal $n \times n$ matrices which satisfy (1) for a given Σ_x, and let U^2 be an element of this set. It is desired to minimize

$$g(U^2) = \sum_{i=r+1}^{n} \lambda_i(\Sigma_x - U^2), \tag{3}$$

where λ_i is the i-th eigenvalue of $(\Sigma_x - U^2)$, for $U^2 \in M$. This function can equivalently be expressed as the minimum of

$$h(X,U^2) = \text{tr } X'(\Sigma_x - U^2)X \tag{4}$$

for $U^2 \in M$ and X constrained to be a column-wise orthonormal matrix of order $n \times (n-r)$. It will now be explained how (4) can be minimized by a Gauss-Seidel type of algorithm.

For any fixed $U^2 \in M$, the minimizing X for h is well-known to be the matrix containing the $(n-r)$ eigenvectors, associated with the smallest $(n-r)$ eigenvalues of $(\Sigma_x - U^2)$. Conversely, for any fixed X the $U^2 \in M$ that minimizes (4) is the element of M which maximizes

$$\text{tr } X'U^2X = \text{tr } XX'U^2 = \text{tr } W^2U^2 \tag{5}$$

where $W^2 \equiv \text{Diag }(XX')$. Finding this U^2 is a weighted constrained minimum-trace factor analysis problem (Shapiro, 1982b). To verify this, consider the problem of finding the maximum-trace diagonal matrix Σ_e of uniqueness estimates for the weighted covariance matrix

$$W\Sigma_xW = (W\Sigma_xW - \Sigma_e) + \Sigma_e, \tag{6}$$

subject to the constraint

$$W\Sigma_xW \geq \Sigma_e \geq 0 \tag{7}$$

or, equivalently, subject to the constraint

$$\Sigma_x \geq W^{-2}\Sigma_e \geq 0 \tag{8}$$

assuming non-singularity of W. The Σ_e^* that has the maximum trace subject to (7) can be obtained by the modified Bentler-Woodward procedure of Ten Berge, Snijders and Zegers (1981), applied to $W\Sigma_xW$. If we now take

$$U^2 = W^{-2}\Sigma_e^* \tag{9}$$

then U^2 satisfies (8) for $\Sigma_e = \Sigma_e^*$, hence $U^2 \in M$, and tr Σ_e^* = tr E^2U^2 has been maximized, subject to $U^2 \in M$. It follows that the maximizing U^2 for (5) can be obtained from (9).

After updating U^2 by (9), we may again update X by computing eigenvectors of $(\Sigma_x - U^2)$, and so on, until $h(X,U^2)$ converges to a stable value. In practice, one may use the orthonormality property of eigenvectors to obtain

$$W^2 = \text{Diag}(XX') = I - \text{Diag}(YY') \tag{10}$$

where Y is the $n \times r$ matrix containing the r eigenvectors associated with the first r eigenvalues of $(\Sigma_x - U^2)$.Computing W^2 by (10) instead of (5) implies that Y instead of X is to be computed. This will save computing time because Y is typically much smaller than X.

The iterative method outlined above yields, for any given Σ_x, those proper communality estimates for which the sum of the smallest $(n-r)$ eigenvalues is a minimum. This means that the associated matrix of common parts of the variables is 'as much as possible' similar to a rank r matrix. The method enables one to ignore small eigenvalues with better justification than before (for instance, using Minres), because now the sum of these 'insignificant' eigenvalues has been minimized, subject to a non-negativity constraint for each eigenvalue separately.

It should be noted that the value of r has to be known before the method can be applied. This is quite analogous to Minres and many other methods of factor analysis. In the absence of theoretical reasons for specifying r, one may try out various values of r consecutively, and adopt the lowest value for which the obtained sum of the smallest $(n-r)$ eigenvalues is considered to be sufficiently close to zero.

While the method described calls for solving (iteratively) a weighted constrained minimum-trace factor analysis problem, the method itself generalizes the unweighted constrained minimum-trace factor analysis. That is, if we set $r=0$ then the sum of all eigenvalues is minimized, hence the sum of the communalities is minimized. It is thus seen that the case $r = 0$ of our method is simply the unweighted constrained minimum-trace factor analysis. In that case the solution is well-known to be unique. On the other hand, if we set $r = (n-1)$ then there is an infinite number of solutions, with smallest eigenvalue zero.

Furthermore, it should be noted that the squared multiple covariance of any variable with the remaining $(n-1)$ variables is a lower bound to any proper communality estimate (Roff, 1936; Harris, 1978), including the estimates obtained from the method proposed above.

Harman's five socio-economic variables

In order to illustrate the method of communality estimation proposed above, it was applied to the five socio-economic variables described in Harman (1967, pp.14, 137, 162, 163, 204, 229). The 5x5 correlation matrix Σ_x for these variables is contained in Table 1, along with squared multiple correlations, Minres factor loadings for $r = 2$, and communality estimates h_m^2 obtained from these loadings.

Table 1

Correlations, squared multiple correlations, and Minres solution for Harman's five socio-economic variables.

	Correlations					SMC	Minres		h_m^2
1	1.000					.969	.625	.766	.978
2	.010	1.000				.822	.714	-.555	.818
3	.972	.154	1.000			.969	.714	.679	.972
4	.439	.691	.515	1.000		.786	.880	-.159	.799
5	.022	.863	.122	.778	1.000	.847	.741	-.578	.882
								Sum	4.449

The results of Table 1 serve as a basis for comparison with results obtained from our method, which are reported in Table 2. In Table 2 h_o^2 refers to the $r = 0$ communalities (the constrained minimum-trace solution), and h_1^2, h_2^2 and h_3^2 similarly refer to the communalities obtained for $r = 1$, 2 and 3, respectively. The loadings reported have been obtained from a principal factor solution based on the $r = 2$ case. In addition, the sums of the communalities and the values of $g(U^2)$ are reported for $r = 0,1,2,3$.

Table 2

New communalities estimates for r=0,1,2 and 3, and loadings for r=2.

	h_o^2	h_1^2	h_2^2	h_3^2	Loadings	
1	1.000	1.000	1.000	.977	.618	.782
2	.831	.831	.831	.996	.717	-.536
3	1.000	1.000	1.000	.992	.708	.697
4	.795	.795	.795	.820	.877	-.140
5	.982	.983	.982	.947	.781	-.600
sum	4.609	4.609	4.609	4.733		
$g(U^2)$	4.609	1.832	.068	.000		

Various inferences can be made from Tables 1 and 2. Firstly, it can be seen that the Minres communality .818 for Variable 2 is below its lower bound .822, thus reflecting an improper solution. No such violations of lower bounds occurs in Table 2.

Secondly, the average (improper) Minres communality for r=2 is lower than its (proper) counterpart in Table 2, yet the sizes of the associated loadings are highly similar. This implies that, for all practical purposes, Minres yields the same factors as our (proper) alternative method, for the data set at hand. The generality of this result is yet to be investigated.

Thirdly, it is clear that the communalities for $r = 0$, 1 and 2 are essentially the same, but differ from those obtained for r=3. However, inspection of the eigenvalues obtained for r=0, 1 and 2 reveils two zero eigenvalues. It follows that this solution is also optimal for r=3. The fact that different values were obtained for r=3 merely reflects non-uniqueness of the solution for r=3. It seems, therefore, that we have obtained a set of communalities (for r=0,1,2) that is optimal for every value of r, $r = 0,1,2,3,4$, and that has thus acquired 'uniqueness' in a new sense. This phenomenon, however, must be interpreted as a peculiarity of the data set at hand. It has been known since Ledermann (1939) that minimum rank and minimum trace solutions do not generally coincide.

Discussion

The newly proposed method of communality estimation seems to be a major improvement over existing methods, because it yields proper communalities that minimize the sum of those eigenvalues that will be ignored.

In addition, however, the method may be used to determine the exact minimum rank of a covariance matrix, see Bekker and De Leeuw (1987). That is, the smallest value of r for which $g(U^2)$ has a global minimum of zero is precisely the (exact) minimum rank. If our method is applied for various values of r, and if the smallest positive value of $g(U^2)$ is not a local minimum, then one can simply determine the minimum rank upon inspection of the values of $g(U^2)$. For instance, in the example of the socio-economic variables reported above, we have $g(U^2) = 0$ for $r \geq 3$, and $g(U^2) > 0$ for $r < 3$, which *guarantees* that the minimum rank is at most 3, and *indicates* that the minimum rank is no less than 3. In fact, the minimum rank for this data is 3, because the necessary condition for minimum rank 4 (Bekker & De Leeuw, 1987) is not satisfied, and, on the other hand, the minimum rank is almost surely at or above the Ledermann bound (Shapiro, 1982a, Theorem 2.2). It follows that, at least in this example, our method of communality estimation yields the exact minimum rank as a by-product. The generality of this, again, will be a matter of continued research.

References

Bekker, P.A. & De Leeuw, J. (1987). The rank of reduced dispersion matrices. Psychometrika, 52, 125–135.

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

Harman, H.H. & Jones, W.H. (1966). Factor analysis by minimizing residuals (Minres). Psychometrika, 31, 351–369.

Harman, H.H. (1967). Modern factor analysis. Chicago: The University of Chicago Press.

Harris, C.W. (1978). Note on the squared multiple correlation as a lower bound to communality. Psychometrika, 43, 283–284.

Jöreskog, K.G. (1977). Methods. In: K. Enslein, A. Ralston & H.S. Wilf (Eds.), Mathematical methods for digital computers (Vol. 3). New York:

Wiley.

Ledermann, W. (1939). On a problem concerning matrices with variable diagonal elements. Proceedings of the Royal Society of Edinburgh, 60, 1–17.

Roff, M. (1936). Some properties of the communality in multiple factor theory. Psychometrika, 1, 1–6.

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.

Ten Berge, J.M.F., Snijders, T.A.B. & Zegers, F.E. (1981). Computational aspects of the Greatest Lower Bound to the Reliability and Constrained Minimum Trace factor analysis. Psychometrika, 46, 201–213.

Zegers, F.E. & Ten Berge, J.M.F. (1983). A fast and simple method of Minimum Residual Factor Analysis. Multivariate Behavioral Research, 18, 331–340.