PSYCHOMETRIKA—VOL. 19, NO. 2
JUNE, 1954

SOME NECESSARY CONDITIONS FOR COMMON-FACTOR ANALYSIS*

LOUIS GUTTMAN

THE ISRAEL INSTITUTE OF APPLIED SOCIAL RESEARCH

Let R be any correlation matrix of order n, with unity as each main diagonal element. Common-factor analysis, in the Spearman-Thurstone sense, seeks a diagonal matrix U^2 such that $G = R - U^2$ is Gramian and of minimum rank r. Let s_1 be the number of latent roots of R which are greater than or equal to unity. Then it is proved here that $r \geqq s_1$. Two further lower bounds to r are also established that are better than s_1. Simple computing procedures are shown for all three lower bounds that avoid any calculations of latent roots. It is proved further that there are many cases where the rank of all diagonal-free submatrices in R is small, but the minimum rank r for a Gramian G is nevertheless very large compared with n. Heuristic criteria are given for testing the hypothesis that a finite r exists for the infinite universe of content from which the sample of n observed variables is selected; in many cases, the Spearman-Thurstone type of multiple common-factor structure cannot hold.

1. *The Problem.* One of the fundamental problems of common-factor analysis—in the sense of Spearman, Thurstone, and others—is as follows. Given the Gramian matrix R of the intercorrelations among n observed variables, with each main diagonal element equal to unity. Let U^2 be an arbitrary diagonal matrix, with the jth main diagonal element denoted by u_j^2, subject to the restrictions that:

$$0 \leqq u_j^2 \leqq 1 \qquad (j = 1, 2, \cdots, n). \tag{1}$$

Let G be the symmetric matrix defined by

$$G = R - U^2. \tag{2}$$

Find a U^2 which will leave G Gramian but with the smallest possible rank.

An analogous algebraic problem holds for a certain type of latent structure hypothesis for dichotomies (9). The results of the present paper hold, with only a change in vocabulary, for obtaining lower bounds to the number of latent classes possible according to a given matrix of joint frequencies. For brevity, we shall refer here only to the common-factor problem.

The importance of this problem lies in the fact that the minimum rank for a Gramian G in (2) equals the minimum number of common-factors in a system that can account exactly for the non-diagonal intercorrelations in R (5).

*This research was made possible in part by an uncommitted grant-in-aid from the Behavioral Sciences Division of the Ford Foundation.

That the elements of a U^2 satisfy restrictions (1) is no guarantee that the corresponding G in (2) is Gramian. It is an essential part of the problem that G be restricted to being Gramian. Non-Gramian G's can often be found with smaller rank than any possible Gramian G for the same R. Indeed, we shall show an example below where a non-Gramian G of rank unity can be found for a certain R, while the lowest possible rank for a Gramian G for the same R is $n - 1$.

2. *The Unknown Communalities and Uniquenesses.* The jth diagonal element of a Gramian G in (2) is called a "communality" of the jth observed variable, and is denoted by h_j^2. From (2) we have—considering the respective main diagonal elements of G and R—

$$h_j^2 = 1 - u_j^2 \qquad (j = 1, 2, \cdots, n). \tag{3}$$

From (1)—which is actually a consequence of the restriction of G to being Gramian—it must be that $0 \leqq h_j^2 \leqq 1$.

The quantity u_j^2 is called a "uniqueness" of the jth observed variable, when G is Gramian.

Conventional empirical techniques for attempting to find a Gramian G of minimum rank usually proceed as follows. A trial matrix U^2 is first used to define a G as in (2), and one or more common-factors is "extracted"—usually by modifying the trial values of U^2 in the course of the computations—until a matrix is built up which differs from R in the non-diagonal elements only by "small" residuals.

3. *Which Null Hypotheses Should be Tested?* A common tendency in practice is to try to stop the factor "extractions" as quickly as possible. This implies that a series of null hypotheses is being tested. The first hypothesis is that the minimum rank for a Gramian G is unity. If this hypothesis is rejected, the next one tested is that the minimum rank is two, etc.

It may be argued that the sequence of hypotheses should be reversed. In general, an arbitrary set of n intercorrelated variables may have many factors in common. A small number of common-factors may be the exception, rather than the rule. From this point of view, the following should be the sequence of null hypotheses to be tested. First, that the minimum rank is n. This can always be rejected, as is well known (and as we shall see again below); a Gramian G of rank $n - 1$ can always be found for any non-singular R, usually in many different ways. The next hypothesis then is that the minimum rank is $n - 2$, which may or may not be true; we shall show examples where this is not true, the minimum rank being $n - 1$.

The new theory of order-factors, or the *radex*, shows vividly the danger of stopping at too few common-factors. For example, if R is a *simplex* matrix, then three or four "common-factors"—as computed by conventional techniques—will often leave non-diagonal residuals that will be regarded as "small" by any existing criterion. But stopping at this small number of

common-factors will obscure the actual mathematical and psychological structure of the data (6, 7). [For example, in the case of a perfect simplex, if r_{jk} is the typical element of R, then $r_{jk} = a_j/a_k$ $(j \leqq k)$, where a_j is a certain parameter belonging to the jth observed variable. The first four principal components can be shown often to account for about 90% of the total variance of such a set of observations, and each remaining principal component accounting for but a small fraction of the remainder. The parsimony of the perfect simplex lies not in the number of common-factors implied, but in that but *one parameter* per test is required to reproduce the observed correlation coefficients.]

In the present paper, we shall present partial safeguards against stopping factor "extractions" too soon.

For finite n, we shall establish some *lower bounds* to the minimum rank possible for a Gramian G, making no assumptions whatsoever. To ascertain the minimum rank *exactly* seems to require solving the communality problem, and this has not yet been shown to have a general solution beyond trial-and-error (systematic trials will always yield a solution in any given case for finite n).

Some authors have modified the communality problem to be: find a U^2 for (2) such that G will be Gramian, and the rank of G shall be equal to the order of the largest non-singular submatrix in R that does not involve a main diagonal element. We shall prove that such a U^2 cannot exist for many cases, or the revised problem can have no general solution. [This modified problem has been suggested by Thurstone (12), and exact solutions have been proposed by Albert (1, 2) and Rosner (10). Unfortunately, these suggested solutions do not maintain the essential restriction that G be Gramian.]

For the infinite universe which the finite sample of n variables represents, we shall present heuristic criteria as to whether a finite common-factor structure exists at all. The fact that a finite number of common-factors can be found to fit a finite number of variables can be a mathematical artifact due purely to the finiteness of the sample of variables. Inference about the universe is the basic scientific problem for common-factor theory. [This point was raised before in the different context of image analysis (8), where a different type of heuristic criteria was proposed. The two types of criteria are supplementary to each other.]

In the present paper, we do not treat the problem of ordinary sampling error, that is, of sampling a population of respondents. We assume throughout that population parameters are used, and not sample statistics. The only sampling problem we shall discuss is that of selecting variables from the universe of content, which cannot usually be treated by the ordinary theory of random sampling.

4. *Definition of the Non-negative Index of a Real Symmetric Matrix;*

Sylvester's "Law of Inertia". Let S be any real symmetric matrix of order n and rank t. Then it is well known that S is congruent to many different real diagonal matrices. That is, there are many real non-singular matrices P such that, if P' is the transpose of P, then

$$PSP' = D \tag{4}$$

where D is some real diagonal matrix.

Let p and q be the number of positive and of negative diagonal elements respectively of D. Then $p + q = t$, for the rank of D must be that of S since P and P' are non-singular; and $n - t$ is the number of zero diagonal elements in D. Thus, we have a frequency distribution of the signs of the main diagonal elements of D, and:

$$p + q + (n - t) \equiv n. \tag{5}$$

Sylvester's "law of inertia" (cf. **3**), states that each of the three frequencies on the left of (5) is invariant under congruence transformations of S. That is, if P in (4) is replaced by any real non-singular P^* that results in D being replaced by any other diagonal matrix D^*, then the frequencies of positive, negative, and zero main diagonal elements of D^* will remain p, q, and $n - t$ respectively. Sylvester's theorem is basic to the theory and practice of the present paper.

The frequency p is usually called the *index* of S. We shall find it more convenient to call p the *positive index* of S, to distinguish it from another index we wish to use as well. Let s be the number of *non-negative* diagonal elements of a D in (4), or:

$$s = p + (n - t) = n - q. \tag{6}$$

Then we shall call s the *non-negative* index of S; it is clearly also invariant under congruence transformations. The practical use of our results below requires the computing of s for certain real symmetric matrices derived from R.

An important special case of a congruence transformation of S is where P is an orthogonal matrix in (4). Then the main diagonal element of D are the *latent roots* of S. This suggests a practical way of computing s. Compute the *characteristic polynomial* of S; this immediately determines t. Then p and q can be determined merely by inspection of the coefficients of the polynomial, using Descartes' rule of signs.

Usually the easiest way of computing s may be simply to reduce S to diagonal form by successive elementary transformations, and then just counting the non-negative diagonal elements in the resulting diagonal matrix. For example, let P_1 be the elementary transformation that subtracts a constant times the first row of S from the second to yield zero as the first element of the second row, and compute P_1SP_1'. Then let P_2 be another

elementary transformation to induce another zero element below the main diagonal of P_1SP_1', etc. Then if we let $P = \cdots P_2P_1$, we have PSP' a diagonal matrix from which s can be ascertained by inspection.

The practical results of the present paper are presented in the next two sections. The proof underlying these results is given in §8 below, and further points of interest are then established in the final two sections.

5. *The Three Lower Bounds.* Let r be the (unknown) minimum rank possible for a Gramian G in (2) for the given R. Let s_1 be the number of latent roots of R which are greater than or equal to unity. Then we shall show that s_1 is necessarily a lower bound to r:

$$r \geqq s_1 . \tag{6}$$

Another way of defining s_1 is as follows. Let S_1 be the symmetric matrix defined by

$$S_1 = R - I, \tag{7}$$

where I is the unit matrix. That is, S_1 is obtained from R by replacing each main diagonal element of R by zero. Let s_1 be the non-negative index of S_1. Then s_1 is also the number of non-negative latent roots of S_1, or is the same as the number of roots of R which are not less than unity, namely the s_1 defined in the preceding paragraph. Hence (6) holds also for this definition of s_1.

An advantage of determining s_1 from S_1 rather than from R is to avoid computing latent roots. Any elementary transformations of S_1 to diagonal form, maintaining congruence, will suffice to determine s_1, as noted in the preceding section.

The second lower bound to r, s_2, is computed as follows. Let r_j be the multiple correlation coefficient of the jth observed variable on the remaining $n - 1$ observed variables. Let D_2 be the diagonal matrix whose jth diagonal element is $1 - r_j^2$ $(j = 1, 2, \cdots, n)$, and let S_2 be the following difference matrix:

$$S_2 = R - D_2 . \tag{8}$$

Let s_2 be the non-negative index of S_2. Then we shall show that s_2 is necessarily a lower bound to r, or:

$$r \geqq s_2 . \tag{9}$$

The diagonal matrix D_2 in (8) is easily computed from the inverse of R when R is non-singular—which is the general case in practice. The jth diagonal element of D_2 is simply the reciprocal of the corresponding element of R^{-1}. The main diagonal elements of S_2 itself are, of course, the r_j^2, in place of the unity elements of R.

The calculation of D_2 may be prohibitive for large n. A weaker lower

bound than s_2—but better than s_1—is obtainable without such labor. Let $\bar{r}_j$ be the largest zero order correlation (in absolute value) that the jth observed variable has with any of the $n - 1$ remaining observed variables; that is, $\bar{r}_j$ is the largest non-diagonal element in the jth row (column) of R. Let D_3 be the diagonal matrix whose jth main diagonal element is $1 - \bar{r}_j^2$ $(j = 1, 2, \cdots, n)$ and let S_3 be the symmetric matrix defined by:

$$S_3 = R - D_3 . \tag{10}$$

Let s_3 be the non-negative index of S_3. Then it is necessarily true that

$$r \geqq s_3 . \tag{11}$$

It will be further shown that, of the three lower bounds, s_2 is the best and s_1 the weakest, or:

$$r \geqq s_2 \geqq s_3 \geqq s_1 . \tag{12}$$

In practice, s_3 may be the most convenient to use. It is better than s_1, and is not essentially more difficult to calculate. S_3 differs from R only by replacing the latter's diagonal elements of unity by the $\bar{r}_j^2$. The $\bar{r}_j$ are ascertainable by mere inspection of R. It should be remarked that it is the $\bar{r}_j^2$ and not the $\bar{r}_j$ that are in the main diagonal of S_3. In Thurstone's centroid method (**12**), he uses the $\bar{r}_j$ as *estimates* of the communalities; they may be too high or too low, and in a manner which is not ascertainable in advance. Here, we use the $\bar{r}_j^2$ because they are necessarily lower bounds to, or always *under*estimates of, the communalities.

6. *The Frequency Distribution of the Latent Roots of R.* Even if r were known exactly for a given sample of n variables, this by itself would give little information as to what r might be for a larger n from the same universe of content, nor what the limit of r might be as $n \to \infty$.

We shall propose here some heuristic considerations concerning the limit of r as $n \to \infty$. These are based on the latent roots of R. Alternatively they could be based on the latent roots of S_2 or of S_3 (the roots of S_1 are exactly related to those of R, so S_1 needs no extra treatment). For brevity, we shall state the case only for R, and the modifications for S_2 and S_3 can be seen to follow immediately.

As is well known, the sum of the latent roots of any matrix is always equal to the trace, or the sum of the elements in the main diagonal. Since each main diagonal element of R is unity, the sum of its latent roots is always n. Therefore, the arithmetic mean of the latent roots of R is always unity, for there are always n latent roots. Hence, except for the trivial case where all roots of R are equal—or $R = I$ and all variables are uncorrelated—at least one root of R is greater than unity and at least one is less than unity. Furthermore, since R is Gramian, all its roots are non-negative, so they must all be in the interval from zero to n. Hence, one or more roots of R is between

zero and unity, and all the rest are between unity and n, the arithmetic mean of all n roots being unity.

In general, then, the n latent roots of R have an asymmetric frequency distribution about the mean of unity. By recalling the definition above of our first lower bound to r, we see that s_1 is the number of roots of R that are greater than or equal to the mean. Conversely $n - s_1$ is the number of roots less than the mean. Our first lower bound is thus one feature of the frequency distribution of the latent roots of R.

What will happen to the frequency distribution of latent roots as n increases? This is a crucial question if we are concerned about the entire universe of content. The arithmetic mean, of course, will remain invariant at unity. But will s_1 change? Evidently, no general answer can be given; it depends on the nature of the empirical data. However, the following heuristic considerations may be proposed. If, for a given n, s_1 is relatively small while the distribution of latent roots is clearly *bimodal*—with one cluster well above unity, another cluster clearly below unity, and no roots in the neighborhood of unity—then it may be expected that this bimodality will be maintained as n increases, and in particular that s_1 will not increase much, if at all, as n increases. Indeed, if n is not insubstantial but s_1 is very small, it might be hypothesized that this s_1 was essentially the limiting value of r as $n \to \infty$.

This heuristic rule, of course, cannot be stated easily in more precise terms as to what a "substantial" value for n is, nor what a "relatively small" s_1 is, nor how large the empty interval around unity should be for the roots. Regardless, it may be of some value in practice.

7. *Inference About an Infinite Universe of Variables.* One assumption underlying the above heuristic rule is that there may be in practice a certain inertia for bimodality of the latent roots as n increases, if such bimodality exists at all. A more rigorous consideration is the fact that, if r—the minimum rank of a Gramian G for R—remains constant as n increases—then the largest latent root of R must in general become infinite as $n \to \infty$. More precisely, the following theorem is true.

Theorem 1. *Let the minimum rank r of Gramian G for R be constant for all n sufficiently large*. Let $s_1(n)$ be the number of latent roots of R not less than unity for the indicated value of n. Suppose there exists a fixed non-negative number δ such that*

$$\delta < 1, \tag{13}$$

and that $n - s_1(n)$ of the latent roots of R are not greater than δ for all n. Then the largest latent root of R must increase without limit as $n \to \infty$.

*One can be more precise by specifying a constant r for all $n > 2r$, for then the communalities that are implied to exist here must be unique, as shown for example in (2), or the smallest common-factor space is the same for all $n > 2r$.

For the proof, we note from (26) and (28) that

$$G = S + (D - E). \tag{30}$$

Now, from (27), $D - E$ is a diagonal matrix with all main diagonal elements positive. Therefore, $D - E$ is a non-singular Gramian matrix. Hence, we can identify (30) with (19), by letting $G = H$, and $D - E = G_0$. Then (29) is a special case of (20), since the number of positive roots of G equals its rank, G being Gramian; and Theorem 2 is established.

Our three lower bounds are each special cases of Theorem 2. In each case, $E = U^2$. For s_1 , we set $D = I$, so (27) certainly holds. Consequently (29) holds for $s = s_1$.

For s_2 , we set $D = D_2$ in (28). That (27) then holds follows from the theorem proved elsewhere (4) that $r_j^2 < h_j^2$ $(j = 1, 2, \cdots, n)$ for non-singular R. Hence (29) holds for $s = s_2$.

Since a multiple correlation coefficient is never less in absolute value than any lower order correlation, it must be that $\bar{r}_j^2 \leqq r_j^2$, so that also $\bar{r}_j^2 < h_j^2$ $(j = 1, 2, \cdots, n)$ by the same theorem referred to in the previous paragraph. Hence, setting $D = D_3$ in (28) necessarily implies that (29) holds for $s = s_3$.

9. *Proof of the Order Among the Lower Bounds.* That s_2 is a better bound than s_1 also follows from Lemma 1. We note that, from (7) and (8),

$$S_2 = S_1 + (I - D_2). \tag{31}$$

Now, $1 - r_j^2$ is always positive for all j when R is non-singular. Hence $I - D_2$ is a non-singular Gramian matrix, and can serve as G_0 in (19). Setting $S = S_1$ and $H = S_2$, and remarking that consequently

$$s_2 \geqq h, \tag{32}$$

since s_2 is the number of non-negative roots of S_2 while h is only the number of positive roots of S_2 , we have from (20) and (32) that

$$s_2 \geqq s_1 . \tag{33}$$

Similarly, it can be proved that $s_3 \geqq s_1$ and that $s_2 \geqq s_3$, by noting that $I - D_3$ and $D_3 - D_2$ are each Gramian matrices, rewriting (31) appropriately for each case, and using Lemma 1. Thus we have established the continued inequality (12) above.

10. *The Special Case of Equal Uniquenesses.* In general, only a *lower bound* can be set to the minimum value for r for a Gramian G in (2). A special case where the minimum rank can be determined exactly is that of equal uniquenesses. Assume that all n uniquenesses have a common value u^2:

$$u_j^2 = u^2 \qquad (j = 1, 2, \cdots, n). \tag{34}$$

Then U^2 is a scalar matrix, and we can write

$$U^2 = u^2 I. \tag{35}$$

Then (2) can be rewritten as

$$G = R - u^2 I. \tag{36}$$

From (36), if λ_j is the jth latent root of R, then $\lambda_j - u^2$ is a latent root of G ($j = 1, 2, \cdots, n$). But a necessary and sufficient condition for G to be Gramian is that all its latent roots be non-negative. Hence, for u^2 to leave G Gramian it must be that

$$\lambda_j \geqq u^2 \qquad (j = 1, 2, \cdots, n). \tag{37}$$

The number of zero roots of G will be the number of times the equality holds in (37). Hence, the following theorem:

Theorem 3. *For the case where all the uniquenesses are equal, the minimum rank for a Gramian G in* (2) *is obtained when the smallest latent root of G is used for each uniqueness. If r is the minimum rank, then $n - r$ is the multiplicity of this smallest root.*

We immediately have the following corollary:

Corollary. The minimum rank for a Gramian G for any R, as in (2), *is not greater than $n - 1$.*

This follows from Theorem 3, since we can always consider a set of uniquenesses all equal to the smallest latent root of R, and its multiplicity is at least unity. So $n - r$ in the theorem is not less than unity, or $r \leqq n - 1$.

11. *Cases Where r Cannot be Small Compared with n.* We have just seen how U^2 can always be determined to have $r \leqq n - 1$. But common-factor theory in the Spearman-Thurstone sense requires far more than this. It requires r to be very small compared with n, especially as n increases.

We shall now show many examples wherein a parsimonious common-factor structure in the Spearman-Thurstone sense is impossible.

Consider the case where R has only two distinct latent roots. Say that the two distinct values are λ_1 and λ_2 respectively, where $\lambda_1 > \lambda_2$. Then, since the (weighted) mean of these two values must be unity, we have

$$\lambda_1 > 1, \qquad \lambda_2 < 1. \tag{38}$$

Let p be the multiplicity of λ_1 so that $n - p$ is the multiplicity of λ_2. From the lower bound s_1, we have immediately that for any Gramian G derivable from R,

$$r \geqq p. \tag{39}$$

In particular, if $p = n - 1$, then from (39) and the Corollary to Theorem 3 above, we have

$$r = n - 1 \qquad (p = n - 1), \tag{40}$$

or r cannot at all be parsimonious compared with n.

Indeed, this is an example of a "Heywood" case (11). For when

$p = n - 1$, then all tetrads, or second-order minors, in R vanish that do not involve the main diagonal. To see this, consider the matrix $R - \lambda_1 I$. This has p zero roots and $n - p$ negative ones $(= \lambda_2 - \lambda_1)$, and so is a symmetric—but non-Gramian—matrix of rank $n - p$. Hence, all minors vanish that are of order larger than $n - p$; or in R, all minors of order $n - p + 1$ or larger, and that do not involve the main diagonal, must vanish. When $p = n - 1$, we have $n - p = 1$, or the Heywood case. All tetrads in R outside the main diagonal vanish, yet the minimum rank for a Gramian G is $n - 1$. (For a non-Gramian G, in this case, the minimal rank is unity.)

We have in effect generalized the Heywood case to $p \leqq n - 1$. We have shown that many possible correlation matrices are such that, although all diagonal-free minors of a given small order ($n - p$ small) vanish, yet no communalities exist that will yield a Gramian G of the same or comparable small rank ($r \geqq p$, p large). Therefore, merely studying the minors outside the main diagonal, as suggested by Thurstone (**12**), is not sufficient for determining communalities or the minimum possible rank for a Gramian G.

Many further examples can be exhibited, with more than two distinct latent roots for R, leading to the same conclusion.

The question as to whether a parsimonious common-factor system exists at all for a given set of data remains a fundamental one to be reexplored in each empirical case. Current computing procedures which aim at stopping at some relatively small number of common-factors may prejudge the issue. In many cases, an order-factor system, such as the radex, may be more appropriate than a finite common-factor system (**6, 7, 8**).

REFERENCES

1. Albert, A. A. The matrices of factor analysis. *Proc. Nat. Acad. Sci.*, 1944, **30**, 90–95.
2. Albert, A. A. The minimum rank of a correlation matrix. *Proc. Nat. Acad. Sci.*, **30**, 144–148.
3. Bôcher, Maxime. Introduction to Higher Algebra. New York: Macmillan, 1907. (Sixteenth printing, 1952).
4. Guttman, Louis. Multiple rectilinear prediction and the resolution into components. *Psychometrika*, 1940, **5**, 75–99.
5. Guttman, Louis, General theory and methods for matric factoring. *Psychometrika*, 1944, **9**, 1–16.
6. Guttman, Louis. A new approach to factor analysis: the radex. In Paul F. Lazarsfeld (ed.), Mathematical thinking in the social sciences. The Free Press, 1954.
7. Guttman, Louis. Two new approaches to factor analysis. Office of Naval Research, Annual technical report on project Nonr-731(00), 1953.
8. Guttman, Louis. Image theory for the structure of quantitative variates. *Psychometrika*, 1953, **19**, 277–296.
9. Lazarsfeld, Paul F. The logical and mathematical foundation of latent structure analysis. In S. A. Stouffer, et al., Measurement and prediction, Princeton Univ. Press, 1950.

10. Rosner, Burt. An algebraic solution for the communalities. *Psychometrika*, 1948, 13, 181–184.
11. Thomson, Godfrey. The factorial analysis of human ability. Univ. London Press, 1950.
12. Thurstone, L. L. Multiple-factor analysis. Univ. Chicago Press, 1947.

Manuscript received 6/11/53

Revised manuscript received 10/19/53