PSYCHOMETRIKA—VOL. 21, NO. 3
SEPTEMBER, 1956

"BEST POSSIBLE" SYSTEMATIC ESTIMATES OF COMMUNALITIES*

LOUIS GUTTMAN

THE ISRAEL INSTITUTE OF APPLIED SOCIAL RESEARCH

At least four approaches have been used to estimate communalities that will leave an observed correlation matrix R Gramian and with minimum rank. It has long been known that the square of the observed multiple-correlation coefficient is a lower bound to any communality of a variable of R. This lower bound actually provides a "best possible" estimate in several senses. Furthermore, under certain conditions basic to the Spearman-Thurstone common-factor theory, the bound must equal the communality in the limit as the number of observed variables increases. Otherwise, this type of theory cannot hold for R.

I. *Introduction*

One of the intriguing problems of factor analysis has been to find a formula for communalities that will minimize the rank of an arbitrary correlation matrix R. More explicitly, the problem is to find a diagonal matrix U such that $R - U^2$ is Gramian and of minimum rank.

Let n denote the order of R (and of U), and m the minimum rank for Gramian $R - U^2$. At least four approaches have been used to estimate communalities that will yield m:

(a) trial-and-error exact formulas
(b) exact formulas for special cases of R
(c) successive approximations
(d) lower bounds.

The main thesis of this paper is that, in certain senses, the last-mentioned of these four approaches provides "best possible" estimates of communalities for an arbitrary R, even though biased in general by being underestimates.

Let u_j be the jth diagonal element of any U that leaves $R - U^2$ Gramian (whether or not with minimum rank), and let h_j^2 be the corresponding communality:

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

Let ρ_j denote the multiple correlation coefficient of the jth variable in R on the remaining $n - 1$ variables, and σ_j the corresponding standard error

*This research was facilitated by a grant from the Lucius N. Littauer Foundation to the American Committee for Social Research in Israel in order to promote methodological work of the Israel Institute of Applied Social Research.

of estimate (assuming all observed variables to have unit variances):

$$\rho_j^2 = 1 - \sigma_j^2 \qquad (j = 1, 2, \cdots, n). \tag{2}$$

Then it has been shown (**2**, 92*f*; **3**, 293) that always

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

No better general lower bound to h_j^2 has yet been established than ρ_j^2.

We shall prove here that there exist many nonsingular matrices R for which the equality in (3) holds for $n - m$ of the minimizing communalities—all but m of the ρ_j^2 are actual communalities (the remaining communalities equal unity). Such matrices R, however, are of a very restricted type.

A more generally useful result that we shall establish applies to the typical R postulated in the Spearman-Thurstone theory. This school of thought believes a common-factor analysis is meaningful only if m is small compared with n. We shall prove that if the ratio of m to n tends to zero as $n \to \infty$, then all except possibly a zero proportion of the ρ_j^2 must tend to the rank-minimizing h_j^2 . *If the Spearman-Thurstone hypothesis is correct for a given* R, *then the* ρ_j^2 *must almost always be very good approximations to the* h_j^2 *when* n *is large.* (Conversely, if the approximation is bad for many ρ_j , then the Spearman-Thurstone hypothesis of a limited number of common factors must be false.)

An even more general result refers to all R, regardless of the ratio of m to n. If there is to be one and only one unique-factor variable that can yield the uniqueness u_j^2 , then it must be that the limit of σ_j^2 must be u_j^2 as $n \to \infty$ (or it must be that $\rho_j^2 \to h_j^2$). Conversely, if σ_j^2 does not tend to u_j^2 as $n \to \infty$, then there is more than one "unique" variable that can provide the same loading u_j^2 (and satisfy all other algebraic requirements of common-factor theory); the larger the difference between σ_j^2 and u_j^2 (or between ρ_j^2 and h_j^2), the larger the possible difference between alternative "unique" parts for the same jth observed variable of R.

Other important properties of the lower bounds ρ_j^2 will be established. Before going on to our new results, it may be helpful to review briefly the four approaches listed above.

(*a*) *Trial and Error*

Assuming that sampling error and rounding-off errors in computations are nonexistent, trial and error is bound to yield an exact numerical answer when $m < n/2$; the diagonal elements of U^2 in such cases are *rational functions* of the non-diagonal elements of R (cf. **8**). It may turn out, of course, that U^2 is not uniquely determined; two or more different U^2 for the same R may yield m in many cases. When $m \geqq n/2$, trial and error can lead again to an expression for each communality, although in non-rational form in general. Again, multiple solutions for minimizing U^2 may occur.

(b) *Special Exact Formulas*

Some special cases of R make possible exact and rational formulas that need no apparent resort to trial and error. The known cases are for $m < n/2$, the most celebrated being Spearman's where $m = 1$. Thurstone has summarized a number of such formulas (**8**, ch. XIII). A caution should be added to Thurstone's discussion to the effect that not all the apparent solutions may yield Gramian U^2 nor leave $R - U^2$ Gramian. Actually these formulas beg the question, for it is generally not known in advance whether or not $m < n/2$. A specialized formula in effect must be tried on the given R to see if it works. Use of specialized formulas thus seems to be but a modified type of trial and error.

(c) *Successive Approximations*

Attempts have been made to avoid a direct exact solution for U^2 by taking recourse instead to successive approximations. An approximation U_1^2 is guessed, and $R - U_1^2$ is "factored" until residuals are considered small enough, leading to a second approximation U_2^2, etc. It has been claimed that such a procedure generally converges to a satisfactory U^2 (cf. **8**, p. 295). Algebraic proof of such convergence has never been published to our knowledge. For many iterative processes, the value to which convergence takes place depends on the initial trial value. That this may be the case for the above procedure seems evident when one recalls that there are many correlation matrices which do not have a unique set of communalities. Also, unless proof is given to the contrary, there is no reason to believe that successive approximations may not converge to some U where $R - U^2$ is *not* of minimum rank, if convergence takes place at all.

The issue of successive approximations is further beclouded by sampling considerations. Lawley's maximum likelihood solution seems the most appropriate put forward to date, as Rao points out (**7**). To attain precision in the sampling theory, apparently some restrictions have been introduced as to the nature of the population R, else the possibility of equally minimizing alternative solutions would remain. Again, it is not clear when a given R obeys these restrictions or when the sampling theory is valid in practice. [After the above was written, the writer received a copy of reference (**1**) in which a numerical example is given of the failure of Lawley's iterative procedure to converge properly.]

(d) *Lower Bounds*

If we again ignore sampling and rounding-off errors, it is always possible to establish useful lower bounds to communalities without any trial and error and without any hypothesis about or restrictions on R. The best of the lower bounds thus far established are the ρ_i^2, according to inequality (3) above. It is often more convenient to discuss uniqueness rather than communalities, or to use inequality (4) rather than (3):

$$\sigma_j^2 \geqq u_j^2 \qquad (j = 1, 2, \cdots, n). \tag{4}$$

An important feature of the bounds in (3) and (4) is that they hold whether or not there is a multiple solution for U^2; they hold for all possible solutions simultaneously. Indeed, they lead to a criterion for choosing among multiple solutions, as indicated in the next section.

II. *Relationship to the Determinacy of Unique-Factor Scores*

Let r_j denote the multiple-correlation coefficient on the n observed variables of a unique-factor variable hypothesized to yield the uniqueness u_j^2. It has been shown in (6) that

$$r_j^2 = \frac{u_j^2}{\sigma_j^2} \qquad (j = 1, 2, \cdots, n). \tag{5}$$

Since (5) holds for all solutions U^2, it suggests that when a choice is necessary that which makes the inequalities (4) as small as possible is most desirable; the denominator on the right of (5) is fixed for j, so that such a choice makes the individual scores on the unique factor as determinate as possible from the observed data, or the r_j^2 as close as possible to unity. It has been shown that this also often tends to make individual scores on the *common-factor* variables as determinate as possible (6).

Should the approximations (4) for U^2 turn out *not* to be close in a given case, then the factor analysis itself may be regarded as not very useful or definitive. For it has been shown in (6) that determining factor *loadings* alone—common and unique—can be far from sufficient for pinning down *scores* on the hypothesized factors. *Alternative sets of scores for a given hypothesized factor can exist which yield identical loadings and yet correlate negligibly with each other*, according to formula (6),

$$r_j^* = 2r_j^2 - 1 \qquad (j = 1, 2, \cdots, n), \tag{6}$$

where r_j is given by (5) and r_j^* is the minimal correlation always attainable between two alternative sets of scores for the same unique factor hypothesized to underlie u_j^2. [According to (6), if $r_j^2 = .5$, then $r_j^* = 0$, or alternative score solutions for the same jth unique factor always exist that correlate *zero* with each other. Even if r_j^2 is as large as .9, this raises r_j^* only to .8. An equation parallel to (6) holds for common factors.]

III. *The "Best Possible" Estimates*

Can inequality (4) be improved on without recourse to some form of trial and error or use of specialized hypotheses? This does not seem possible. According to (5) this would imply some advance information on the r_j^2; there is no apparent way of getting such information on the r_j^2 in a universally systematic manner. The situation seems to be the reverse: r_j^2 is determined by u_j^2 rather than conversely.

The rest of this paper will be devoted largely to showing that (4) is actually a "best possible" inequality in the sense that the phrase "best possible" is usually used mathematically for inequalities. The essential characteristics are that (a) many correlation matrices R exist for which the *equality* in (4) is actually attained at the same time that minimum rank m is attained, and (b) the inequalities in (4) must tend to equalities as n increases, under certain general conditions important to the theory of common-factor analysis. The bounds improve systematically in general as n increases, or as there is more information available from more observed variables. Furthermore, inequality (4) leads to inequalities for m that are also "best possible," and is closely related to the problem of estimating individual scores on the unique factors without any rank assumptions, via image analysis.

In virtually all attempts to solve the communality problem—whether exactly or by successive approximations—the problem is stated as for a fixed and finite n, or where R is from a finite number of n observed variables. It seems appropriate to ask also what happens to communalities as n increases or decreases.

While this issue is not discussed very explicitly by most writers, it usually seems implied that if the additional variables retain the same general kind of content as the initial ones, communalities of the initial ones should remain constant for all n sufficiently large. This would imply that for n small enough we should generally have $m > n/2$, or easy exact computations for U^2 (even ignoring sampling error) should be the exception rather than the rule. Having $m > n/2$ for relatively small n does not preclude m from remaining constant—and hence becoming relatively small—as n increases. It does imply that multiple solutions should be quite prevalent for finite n in practice. Furthermore, it cautions that an apparently exact solution for finite n may be but an artifact due to the finiteness of the number of variables observed.

It would be desirable, in view of all the preceding considerations, to have a systematic way of getting information about communalities with no assumptions whatsoever about R, yet without resorting to trial and error. Furthermore, this information should remain valid as n increases.

One of the virtues of the bounds (3) and (4) is that they possess these qualities in a simple and direct manner. This seems to be another type of "best possible" property from that usually considered, and one which appears peculiarly relevant to the problem of factor analysis.

IV. *Attaining Equality When n Is Finite*

If $\sigma_j^2 = 0$ for some j (so that $\rho_j^2 = 1$), then it must be that $u_j^2 = 0$ from (4) and the fact that a uniqueness cannot be negative. Here is one kind of special circumstance wherein our bound becomes an exact estimate even when n is finite. In practice, this is not to be expected, since having one observed variable perfectly predictable from all the rest makes R singular.

Many cases of nonsingular R also exist for which the equality in (4) holds and n is finite. We shall exhibit some now. To this end, let us first recall that the σ_j^2 are the reciprocals of the corresponding main diagonals of R^{-1}. The following notation will be useful here and also later. Let S^{-2} (the inverse of S^2) be the diagonal matrix with the same main diagonal elements as R^{-1}. Then the jth main diagonal element of S^2 itself is simply σ_j^2 $(j = 1, 2, \cdots, n)$:

$$S^2 = [\sigma_1^2, \sigma_2^2, \cdots, \sigma_n^2]. \tag{7}$$

If R is nonsingular, there exists a nonsingular matrix F such that

$$R = FF'. \tag{8}$$

F can be chosen in infinitely many ways for (8) when $n \geqq 2$, but always we can rearrange variables to find an F of the form

$$F = \begin{Vmatrix} A & 0 \\ B & C \end{Vmatrix}, \tag{9}$$

where A is a nonsingular square submatrix of order m, B is of order $(n - m) \times m$, and C is nonsingular and of order $n - m$. From (8) and (9),

$$R = \begin{Vmatrix} AA' & AB' \\ BA' & BB' + CC' \end{Vmatrix}. \tag{10}$$

It is easily verified that the inverse of F is given by

$$F^{-1} = \begin{Vmatrix} A^{-1} & 0 \\ -C^{-1}BA^{-1} & C^{-1} \end{Vmatrix}. \tag{11}$$

From (8), $R^{-1} = (F^{-1})'F^{-1}$, or using (11)

$$R^{-1} = \begin{Vmatrix} G & H' \\ H & (CC')^{-1} \end{Vmatrix}, \tag{12}$$

where

$$G = (AA')^{-1} + (A^{-1})'B'(CC')^{-1}BA^{-1} \tag{13}$$

and

$$H = -(CC')^{-1}BA^{-1}. \tag{14}$$

Now consider the special case where CC' is a diagonal matrix. Then $(CC')^{-1}$ is also diagonal. According to (12) and (7), $(CC')^{-1}$ constitutes the lower right-hand submatrix of S^{-2}, or CC' constitutes the corresponding submatrix of S^2 and defines the σ_j^2 for $j = m + 1, m + 2, \cdots, n$. If we subtract this submatrix CC' from the lower right-hand corner of R in (10), we are clearly left with a reduced R that is Gramian and of rank m, it being the product of $[A\ B]'$ and its transpose. Thus we have:

Theorem 1. *If R can be factored into an F of the form* (9) *where CC' is diagonal* (*and A and C are nonsingular*), *then the main diagonal elements of CC' are the respective σ_j^2 for $j = m + 1, m + 2, \cdots, n$. If these $n - m$ σ_j^2 in CC' are subtracted from the corresponding main diagonal elements of R, the resulting matrix will be Gramian and of rank m.*

According to Theorem 1, when m is the actual minimal rank possible for Gramian $R - U^2$, then the first m diagonal elements of U^2 can be set equal to zero, and the last $n - m$ diagonal elements equal to the corresponding σ_j^2 as given by CC'. Thus, the last $n - m$ of the σ_j^2 serve exactly as rank-minimizing uniquenesses, or the equality in (4) holds for $j = m + 1, m + 2, \cdots, n$.

Notice that the *first* m uniquenesses implied by Theorem 1 are zero and not equal to the σ_j^2. If the first m σ_j^2 were also subtracted out from the main diagonal of R, then the resulting $R - S^2$ would in general not be Gramian, nor of minimum rank (cf. **4**).

Theorem 1 holds even when the m in it is not minimal. It is always possible to use the Theorem for the case where C is of order 1, and hence CC' is *necessarily* a diagonal matrix. This provides:

Corollary. For any nonsingular R, if any one σ_j^2 is subtracted from the corresponding main diagonal element of R, then the resulting matrix is of rank $n - 1$.

This result was partly indicated by Thurstone in his discussion of the "diagonal" method of matrix factoring (better known to mathematicians as the Schmidt or Gram-Schmidt process of orthogonalization), but without noticing apparently that his implied uniqueness was exactly σ_n^2 (**8**, p. 308).

We have thus completed showing that there are many matrices for which many of the σ_j^2 can serve as rank-minimizing uniquenesses. Also, we have the curious result that any one of the σ_j^2 alone will reduce nonsingular R to a Gramian matrix of rank $n - 1$.

V. *Equality in the Limit as $n \to \infty$*

We have already seen in Part III how, if a "unique"-factor variable is really to be uniquely determined for a given u_j^2, then we must have $u_j^2/\sigma_j^2 \to 1$ as $n \to \infty$, according to (5) and (6). This conclusion does not depend on the size of m, nor in particular on whether m remains finite or becomes infinite as $n \to \infty$. It thus applies to ordered factor theories—such as the radex, with its simplexes and circumplexes (**5**)—as well as to limited common-factor theories like those of Spearman and Thurstone, whenever the δ-law of deviation (**5**, p. 308) holds for the unique-factor variables.

Thus, a general sufficient condition for σ_j^2 to tend to u_j^2 when $u_j^2 > 0$ is that $r_j^2 \to 1$ or $r^* \to 1$ as $n \to \infty$. This holds for each j separately.

A less general sufficient condition, and one that does not necessarily hold for any one j but only for "almost all" j, is given in

Theorem 2. *If R is nonsingular for all n, and if* $\lim_{n\to\infty} m/n = 0$, *then*

$$\lim_{n\to\infty} \frac{1}{n} \sum_{j=1}^{n} \frac{u_j^2}{\sigma_j^2} = 1.$$

For all except possibly a zero proportion of the j it must be that $\lim_{n\to\infty} u_j^2/\sigma_j^2 = 1$.

The condition that $m/n \to 0$ holds in particular for the Spearman-Thurstone approach to factor analysis, which postulates that the number of common factors should be small compared to the number of observed variables.

Since $u_j^2/\sigma_j^2 \leqq 1$ for all j, according to (4), we must have the mean ratio also bounded above by unity:

$$\frac{1}{n} \sum_{j=1}^{n} \frac{u_j^2}{\sigma_j^2} \leqq 1. \tag{15}$$

The hypothesis that R is nonsingular for all n ensures that no observed variable is perfectly predictable from the rest, or that $\sigma_j^2 > 0$ for all j and n, so that division by σ_j^2 in (15) is always justified. The first conclusion of Theorem 2 is that the limit of the left member of (15) as $n \to \infty$ is actually the right member. But clearly, the mean value of a sequence cannot tend to an upper bound to each member of the sequence unless almost all members of the sequence also tend to this upper bound. Hence the second conclusion of Theorem 2 follows from the first. We need only to establish the first part of the theorem now.

As is well known, if $R - U^2$ is Gramian and of rank m, we can write

$$R = AA' + U^2, \tag{16}$$

where A is some matrix of order $n \times m$ and of rank m. Let Q be defined as the symmetric matrix of order m:

$$Q = I_m + A'U^{-2}A, \tag{17}$$

where I_m is the unit matrix of order m. It has been shown in (**2**, 92) that Q is Gramian and nonsingular, and furthermore

$$R^{-1} = U^{-2} - U^{-2}AQ^{-1}A'U^{-2}. \tag{18}$$

It is easily verified further, from (18) and (17), that

$$A'R^{-1}A = I_m - Q^{-1}. \tag{19}$$

Since the left member of (19) is clearly Gramian, so must the right member be. Indeed, it is known that $I_m - Q^{-1}$ is the covariance matrix of the predicted values (from the observed n variables) of any m orthogonal common-factor scores underlying loading matrix A (6). Let q^{kk} denote the kth main diagonal element of Q^{-1}, or the variance of estimate of the kth common factor, and let p_k^2 be defined as

$$p_k^2 = 1 - q^{kk} \qquad (k = 1, 2, \cdots, m). \tag{20}$$

Then p_k^2 is the square of the multiple-correlation coefficient of the kth common factor from the n observed variables, and

$$0 \leqq p_k^2 \leqq 1 \qquad (k = 1, 2, \cdots, m). \tag{21}$$

Therefore, the *trace*—or sum of the main diagonal elements—of $I_m - Q^{-1}$ satisfies

$$\operatorname{tr}(I_m - Q^{-1}) = \sum_{k=1}^{m} p_k^2 \leqq m. \tag{22}$$

We are particularly interested in the trace of U^2R^{-1}, for clearly—remembering (7)—

$$\operatorname{tr}(U^2R^{-1}) = \operatorname{tr}(U^2S^{-2}) = \sum_{j=1}^{n} \frac{u_j^2}{\sigma_j^2}. \tag{23}$$

Since the trace of a product is unchanged if order of multiplication is reversed,

$$\operatorname{tr}(A'R^{-1}A) = \operatorname{tr}(AA'R^{-1}) = \operatorname{tr}(I_n - U^2R^{-1}), \tag{24}$$

the last member following from the middle member by recalling (16). Therefore, taking traces of both members of (19) and using (23), (24), and (22),

$$\sum_{j=1}^{n} \frac{u_j^2}{\sigma_j^2} = n - \sum_{k=1}^{m} p_k^2 \geqq n - m. \tag{25}$$

Dividing (25) through by n and prefixing inequality (15),

$$1 \geqq \frac{1}{n} \sum_{j=1}^{n} \frac{u_j^2}{\sigma_j^2} = 1 - \frac{1}{n} \sum_{k=1}^{m} p_k^2 \geqq 1 - \frac{m}{n}. \tag{26}$$

Clearly, if $m/n \to 0$ in the last member of (26), the middle members must tend to unity, or Theorem 2 is established.

Notice that Theorem 2 could be rephrased to say that almost all $r_j^2 \to 1$, or almost all unique-factors must be determinate in the limit. It is interesting to see this in a slightly different way. From (5) and the middle members of (26),

$$\frac{1}{n} \sum_{j=1}^{n} r_j^2 + \frac{1}{n} \sum_{k=1}^{m} p_k^2 = 1, \tag{27}$$

or

$$\overline{r^2} + \frac{m}{n}\overline{p^2} = 1, \tag{28}$$

where $\overline{r^2}$ is the mean predictability of the n unique-factors, while $\overline{p^2}$ is the mean predictability of the m common factors. When m/n is small, the average predictability of the common factors cannot influence greatly the average predictability of the unique factors: $\overline{r^2}$ must be close to unity. A further consequence is that, if both $\overline{r^2}$ and $\overline{p^2}$ tend to unity as $n \to \infty$, it must be that $m/n \to 0$. This does not require m to remain finite, of course, but only to increase at a less rapid pace than does n.

VI. *Increase in Information with n*

A desirable property of estimates of communalities is that they should improve in general as n increases. Any n variables studied empirically by a factor analysis are usually regarded as but a sample of a far larger universe of variables. The communalities sought are those of the universe.

Of the four approaches to estimates outlined in Part I above, the only one which has its estimates vary explicitly with n is that of lower bounds. In this sense, it is the only one not tied to algebraic artifacts that may arise in data due to the finiteness of n of the observed sample of variables (cf. **3** and **4**).

For fixed j, ρ_j^2 must increase with n—or at worst remain constant—for a multiple-correlation cannot become worse as the number of predictors increases. If h_j^2 is defined as for the universe of variables ($n = \infty$), then ρ_j^2 must improve in general as an estimate of h_j^2 as n increases, considering (3). The lower bounds improve as estimates as n increases, taking advantage of the increased information.

Similarly, if the jth unique-factor scores are defined uniquely as for the universe of observed variables, r_j^2 must in general increase with n. From (5), this again makes σ_j^2 an increasingly better estimate of the fixed u_j^2 as n increases.

Thus, the lower bounds automatically take advantage of whatever new information is brought in with increased n, without making any assumptions at all. In broad classes of cases, as we have seen, this new information can make $\rho_j^2 \to h_j^2$ for all or almost all j.

VII. *Further "Best Possible" Inequalities*

We have concentrated until now on the approximation of the σ_j^2 to the u_j^2. Related to this is another problem: the estimation of minimum rank m for Gramian $R - U^2$. We shall show that using the diagonal matrix S^2 of (7) as an estimate of U^2 for finite n leads also to a "best possible" inequality for m, as well as to other important inequalities.

With any nonsingular correlation matrix R is associated another nonsingular correlation matrix R^* defined by

$$R^* = SR^{-1}S. \tag{29}$$

R^* is clearly Gramian, for R^{-1} is Gramian and S is a diagonal matrix. The main diagonal elements of R^* are all unity from the definition of S and the fact that $1/\sigma_j^2$ is the jth diagonal element of R^{-1}. Indeed, R^* is the correlation matrix of the *anti-images* of the n variables of R (cf. **3**, p. 294f). Regardless of the statistical meaning of R^*, it is a perfectly good correlation matrix when n is finite, and we can seek a diagonal Gramian matrix U^* that will leave $R^* - U^{*2}$ Gramian and with minimum rank m^*. This will lead to the interesting and important inequality for the case where no σ_j^2 is a uniqueness nor equals unity:

$$m + m^* \geqq n \qquad (u_j^2 < \sigma_j^2 < 1; j = 1, 2, \cdots ; n). \tag{30}$$

The restrictions that $S^2 - U^2$ and $I - S^2$ be nonsingular are essential here (consider the counter-example where $S = R = R^* = I$). That $\sigma_j^2 \neq 1$ ($I - S^2$ be nonsingular) implies that each variable in R has at least one nonzero correlation with some other variable.

According to (30), if m/n is small, then m^*/n must be large. Conversely, if m^*/n is small, m/n must be large. This is rather paradoxical in view of the fact that R^* can *always* be reduced to rank m by subtracting out the diagonal matrix $SU^{-2}S$ ($= S^2U^{-2}$). This follows by pre- and post-multiplying (18) through by S, remembering (29), and noting that the second term on the right is of rank m. Conversely, R can always be reduced to rank m^* by subtracting out $S^{*2}U^{*-2}$, where S^{*2} is the diagonal matrix defined by the main diagonal of R^{*-1}. Thus, if all diagonal-free submatrices of R have rank less than $n/2$, so must those of R^*, and conversely. Regardless, (30) holds.

In effect, then, (30) implies that to every R for which $\sigma_j^2 \neq u_j^2$ or 1 for all j and where $m < n/2$, there corresponds an R^* which is a generalized "Heywood" case (cf. **4**, 159f). Although all diagonal-free matrices have small rank in R^*, no communalities can be found to make $R - U^2$ of equally small rank and yet be Gramian. It must be that $m^* \geqq n - m$. This again emphasizes that the case $m < n/2$ may be the exception, rather than the rule, for correlation matrices. And it is interesting that this paradox arises precisely for those cases where no σ_j^2 equals the corresponding u_j^2.

To establish (30), we first recall the theorem (**4**, 157f) that if $S^2 - U^2$ is nonsingular, and if s is the non-negative index of $R - S^2$, then

$$s \leqq m \qquad (| S^2 - U^2 | > 0). \tag{31}$$

Now, the proof of (31) in (**4**) can be modified to take care of the case where $S^2 - U^2$ is possibly singular, to establish the weaker but more universal inequality $p \leqq m$, where p is the *positive* index of $R - S^2$. We shall not take space to prove this modification here, but shall merely state it in terms of our needs for R^*:

$$p^* \leqq m^*, \tag{32}$$

where p^* is the positive index of $R^* - S^{*2}$, and we do *not* necessarily assume $S^{*2} - U^{*2}$ to be nonsingular.

Now, from (29), $R^{*-1} = S^{-1}RS^{-1}$, or since the main diagonal elements of R are all unity,

$$S^{*2} = S^2. \tag{33}$$

It is interesting to note that (33) and (29) imply that $(R^*)^* = R$, or R is to R^* as R^* is to R.

Statistically, (33) implies that the relative predictability of the jth anti-image from the $n - 1$ remaining anti-images is the same as for the jth original variable from the $n - 1$ remaining original variables. From (29) and (33) we can write the identity

$$R^* - S^{*2} = S(R^{-1} - I)S. \tag{34}$$

Sylvester's "law of inertia" (cf. **4**, p. 152) applied to (34) shows that p^* equals the positive index of $R^{-1} - I$, which in turn clearly equals the number of latent roots of R^{-1} greater than unity. Hence p^* equals the number of latent roots of R itself which are *less* than unity. But it has been shown in (**4**) that s is not less than the number of latent roots of R which are greater than or equal to unity whenever $I - S^2$ is nonsingular. Since R has n latent roots all told, it follows that

$$s + p^* \geqq n \qquad (| I - S^2 | > 0). \tag{35}$$

Inequality (30) follows from (31), (32), and (35).

To prove that (30) is a "best possible" inequality, we must show that matrices R exist for which the equality sign holds. It suffices to consider an R which has only two distinct latent roots, say $\lambda_1 > 1$ with multiplicity f and $\lambda_2 < 1$ with multiplicity $f^* = n - f$. Then it must be that

$$m = f, \qquad m^* = f^*. \tag{36}$$

For $m \geqq f$ by inequality (39) of (**4**, 159), and hence $m = f$ by considering that $R - \lambda_2 I$ is Gramian and of rank f; $m^* = f^*$ by analogous reasoning on R^{-1}. Since $f + f^* = n$, (36) provides a special case where the equality in (30) holds.

Inequality (31) by itself is similarly a "best possible" one. Consider the case where R^* has two distinct latent roots, say $\lambda_1 < 1$ with multiplicity p^* and $\lambda_2 < 1$ with multiplicity $p = n - p^*$. Since $R^{*-1} - I = S^{-1}RS^{-1} - I = S^{-1}(R - S^2)S^{-1}$, p is the positive index of $R - S^2$ while p^* is that of $R^* - S^{*2}$. Also, since no root vanishes, $p = s$ or the positive and non-negative indices coincide. Since $R^{*-1} - \lambda_2^{-1}I = S^{-1}RS^{-1} - \lambda_2^{-1}I$ is Gramian and of rank $p = s$, so must $R - \lambda_2^{-1}S^2$ be, or the equality in (31) must hold for this case.

VIII. *Relation to Image Analysis*

The ratio of u_j^2 to σ_j^2 indicates the relative predictability of the jth unique-factor scores from the n observed variables of R, according to (5). Closely related is another parameter developed in image theory and denoted by δ_j^2, namely, the variance of the difference between the respective scores on the jth anti-image and the jth unique factor. It turns out (**3**, 293) that δ_j^2 can be computed as the simple difference

$$\delta_j^2 = \sigma_j^2 - u_j^2 \qquad (j = 1, 2, \cdots, n). \tag{37}$$

Hence, a necessary and sufficient condition that $\sigma_j^2 \to u_j^2$ as $n \to \infty$ is that $\delta_j^2 \to 0$. This implies that the unique-factor scores must be essentially the total anti-image scores from the universe of content. Here we have the individual anti-images themselves as increasingly better estimates of the unique-factor scores as $n \to \infty$. This problem of estimating scores is perhaps even more basic than that of estimating only over-all parameters, such as uniquenesses, which are based on the scores. Estimating U^2 by S^2 has the important property of tying in directly with the score estimation problem via image analysis.

REFERENCES

1. Golub, G. H. On the number of significant factors as determined by the method of maximum likelihood. (Stencilled) Digital Computer Laboratory, Univ. of Illinois, 1954.
2. Guttman, L. Multiple rectilinear prediction and the resolution into components. *Psychometrika*, 1940, **5**, 75-99.
3. Guttman, L. Image theory for the structure of quantitative variates. *Psychometrika*, 1953, **18**, 277-296.
4. Guttman, L. Some necessary conditions for common-factor analysis. *Psychometrika*, 1954, **19**, 149-161.
5. Guttman, L. A new approach to factor analysis: the radex. In Paul F. Lazarsfeld (ed.), Mathematical thinking in the social sciences, Glencoe, Ill.: Free Press, 1954.
6. Guttman, L. The determinacy of factor score matrices, with implications for five other basic problems of common-factor theory. *Brit. J. stat. Psychol*, 1955, **8**, 65-81.
7. Rao, C. R. Estimation and tests of significance in factor analysis. *Psychometrika*, 1955, **20**, 93-111.
8. Thurstone, L. L. Multiple-factor analysis. Chicago: Univ. Chicago Press, 1947.

Manuscript received 5/17/55

Revised manuscript received 9/13/55