

SOME BOUNDARY CONDITIONS FOR A MONOTONE ANALYSIS OF SYMMETRIC MATRICES

JAMES C. LINGOES*

THE UNIVERSITY OF MICHIGAN

This paper gives a rigorous and greatly simplified proof of Guttman's theorem for the least upper-bound dimensionality of arbitrary real symmetric matrices S , where the points embedded in a real Euclidean space subtend distances which are strictly monotone with the off-diagonal elements of S . A comparable and more easily proven theorem for the vector model is also introduced. At most $n-2$ dimensions are required to reproduce the order information for both the distance and vector models and this is true for any choice of real indices, whether they define a metric space or not. If ties exist in the matrices to be analyzed, then greatest lower bounds are specifiable when degenerate solutions are to be avoided. These theorems have relevance to current developments in nonmetric techniques for the monotone analysis of data matrices.

Guttman [1967, p. 78]; [1968, p. 477] formalized a notion having some currency among scaling theorists [*e.g.*, Torgerson, 1958, pp. 270 and 278] to the effect that $n - 2$ dimensions was a possible least upper-bound within the context of nonmetric multidimensional scaling. We intend here to rephrase this theorem in some minor details for the purposes of clarity and rigor and then later offer a simple proof of its validity.

Guttman's $n - 2$ Theorem: The off-diagonal elements $\{s_{ij}\}$ of any real symmetric matrix of order n are strictly monotone with the set of interpoint distances $\{d_{ij}\}$ among some n points in a real Euclidean space having at most $n - 2$ dimensions. Not only are the Euclidean distances monotone with the elements of S , but the $\{d_{ij}^2\}$ are even linear (non-homogeneous) functions of the $\{s_{ij}\}$. For non-degeneracy when $\{s_{ij}\} = c \neq 0$, $n - 1$ dimensions are required for a strict monotone mapping.

By a strict monotone mapping of S into E^m is meant:

$$(1) \quad \text{if } s_{ij} < s_{kt} \text{ then } d_{ij} < d_{kt}$$

and

$$(2) \quad \text{if } s_{ij} = s_{kt} \text{ then } d_{ij} = d_{kt},$$

* This research in nonmetric methods is supported in part by a grant from the National Science Foundation (GS-929 & -2850).

The very helpful comments and encouragement of Louis Guttman and J. Douglas Carroll are greatly appreciated. Finally, to that unknown, but not unsung, reviewer who helped in the clarification of the argument, I express my thanks.

where

$$(3) \quad d_{ij} = \sqrt{\sum_{a=1}^m (x_{ia} - x_{ja})^2},$$

for $i \neq j$ and $k \neq l$ throughout ($i, j, k, l = 1, 2, \dots, n$) and m is the rank of X , a set of real normalized vectors. There is no qualification in the above theorem regarding fulfillment of the triangle inequality for a metric space for the $\{s_{ij}\}$. The elements of S can be anything (positive, negative, or fail the triangle inequality) so long as the conditions of

$$(4) \quad s_{ii} = s_{ii} \text{ (symmetry)}$$

and

$$(5) \quad S \in \text{Re (real)}$$

are satisfied.

In regard to the linearity of the functions relating $\{s_{ij}\}$ to $\{d_{ij}^2\}$, Guttman's theorem entails that

$$(6) \quad d_{ij}^2 = a(s_{ij} + b) \text{ (not all equal to zero)}$$

is the form for an arbitrary multiplicative constant a and a uniquely determined additive constant b .

Rather than directly proceeding to a proof of Guttman's theorem at this point, we wish to introduce a parallel theorem for scalar products, since the latter's proof is far simpler and serves to illustrate the constructive principles involved. Both constructions (for the distance and vector models) follow Guttman's for a $n - 1$ least upper-bound [Guttman, 1965], but go on to provide constant vectors as well to reduce dimensionality to $n - 2$. Following the proofs of these two theorems, we shall deal with the treatment of special cases requiring a least upper-bound of $n - 1$ when degeneracy is to be avoided.

The Vector Model

We now state the parallel

Theorem: The off-diagonal elements $\{s_{ij}\}$ of any real symmetric matrix of order n are strictly monotone with the set of scalar products $\{\theta_{ij}\}$ among some n vectors in a real Euclidean space having maximum rank of $n - 2$. Not only are the scalar products monotone with the elements of S , but the $\{\theta_{ij}\}$ are uniquely determined up to an additive constant of the $\{s_{ij}\}$. For non-degeneracy when $\{s_{ij}\} = c \neq 0$ and $n = 2$, $n - 1$ dimensions are required for a strict monotone mapping.

Similar to (1), (2), (3), and (6) above, we have the definitions:

$$(1a) \quad \text{if } s_{ij} > s_{kl} \text{ then } \theta_{ij} > \theta_{kl}$$

and

$$(2a) \quad \text{if } s_{ij} = s_{kl} \text{ then } \theta_{ij} = \theta_{kl},$$

where

$$(3a) \quad \theta_{ij} = \sum_{a=1}^m x_{ia}x_{ja},$$

for $i \neq j$ and $k \neq l$ throughout $(i, j, k, l = 1, 2, \dots, n)$ and m is the rank of X , a set of real normalized vectors, and, of course, (4) and (5) are fulfilled.

In regard to θ being equivalent to S up to a translation, we have the form

$$(6a) \quad \theta_{ij} = s_{ij} + b,$$

where b is a uniquely determined additive constant.

Proof: Set

$$(7) \quad v_{ij} = \begin{cases} -\sum_{\substack{k=1 \\ k \neq i}}^n s_{ik} & (i = j) \\ s_{ij} & (i \neq j) \end{cases}.$$

V is a singular matrix because

$$(8) \quad \sum_{i=1}^n v_{ij} = 0 \quad (j = 1, 2, \dots, n).$$

From a theorem on real symmetric matrices having row (column) sums equal to a constant c , V has at least one root equal to $c = 0$ and at most one eigen-vector of constant components paired with a root equal to c . The unit length vector defined by

$$(9) \quad y = n^{-1/2}l,$$

where l is a n -element column vector of unities, is such a vector.

Now set

$$(10) \quad W = V - \lambda I,$$

where I is the order n identity matrix and λ is the smallest root algebraically of V not associated with y of (9). Subtracting a scalar λ from the diagonal of V yields roots for W which are those of V minus the scalar, i.e., $\mu(W) = \lambda_1 - \lambda, \lambda_2 - \lambda, \dots, \lambda_n - \lambda$.

If $\lambda = 0$ (or, zero is a multiple root of V), then, of course, $m \leq n - 2$ and we are through. If $\lambda < 0$, then W is Gramian; and, if $\lambda > 0$, then the constant vector y of W has an associated negative but real root. In any event when $\lambda \neq 0$, we can reduce the rank of W by one by constructing a unit rank matrix of constant elements

$$(11) \quad U = \underline{\lambda n^{-1} U'}$$

and subtracting it from W , yielding the matrix

$$(12) \quad Z = W + U,$$

which is Gramian and of rank $m \leq n - 2$. We can, therefore, write (12) in scalar notation as

$$(12a) \quad z_{ij} = s_{ij} + \lambda n^{-1} \quad (i \neq j),$$

since the off-diagonal elements of S have been altered by (12) only.

From a familiar theorem on Gramian matrices, there exists an X of order $n \times m$ such that

$$(13) \quad Z = XX',$$

or, again in scalar notation,

$$(13a) \quad z_{ij} = \sum_{\alpha=1}^m x_{i\alpha} x_{j\alpha},$$

which satisfies the definition of the $\{\theta_{ij}\}$ in (3a). Combining (3a) with (13a), on the one hand, with (12a), on the other, we see immediately that

$$(14) \quad z_{ii} = \sum_{\alpha=1}^m x_{i\alpha} x_{i\alpha} = \theta_{ii} = s_{ij} + \lambda n^{-1} \quad (i \neq j)$$

satisfies (6a) and *a fortiori* (1a) and (2a) in $m \leq n - 2$ dimensions, as we set out to prove.

The theorem for the vector model is relevant to the algorithm developed by Lingoes and Guttman (1967) for the nonmetric factor analysis of arbitrary indices of relationship, in general, and correlations/covariances, in particular.

The Distance Model

For the proof of Guttman's theorem, let us define

$$(15) \quad \bar{s}_i = \frac{1}{n} \sum_{j=1}^n s_{ij} \quad (i = 1, 2, \dots, n)$$

and

$$(16) \quad \bar{\bar{s}} = \frac{1}{n} \sum_{i=1}^n \bar{s}_i,$$

from which we construct

$$(17) \quad t_{ij} = \bar{s}_i + \bar{s}_j - \bar{\bar{s}} - s_{ij} \quad (i, j = 1, 2, \dots, n).$$

T^* is a double-centering of $-S$ and the rank of $T \leq n - 1$ by reason of

* The reader will recognize the similarity of T to Torgerson's (1958, Eq. 16, p. 258) B^* matrix. Our proof, however, does not rely in any way upon Torgerson's treatment of B^* .

the constant row sums of zero. Associated with a zero root is a constant vector y as was the case for the vector model.

Define the scalar λ to be the smallest root algebraically of T not associated with y , the constant vector, and set

$$(18) \quad R = T - \lambda I.$$

The construction in (18) ensures that R will have at most one negative root. Using the present definition of λ we can use U , as in (11), to construct

$$(19) \quad G = R + U,$$

where G is Gramian and of rank $m \leq n - 2$, thus eliminating the constant vector. Expressing G in terms of T , we have in scalar notation

$$(20) \quad g_{ii} = t_{ii} - \lambda + \lambda n^{-1},$$

because of (18) and (19), while

$$(20a) \quad g_{ij} = t_{ij} - \lambda \delta_{ij} + \lambda n^{-1} \quad (i \neq j),$$

because of (19) and symmetry with (20), where δ_{ij} is Kronecker's *delta*.

Invoking our theorem on the decomposition of Gramian matrices and substituting the right member of (17) in (20) and (20a) yields

$$(21) \quad \sum_{a=1}^m x_{ia}^2 = 2\bar{s}_i - (\bar{s} + s_{ii}) - \lambda + \lambda n^{-1} \quad (i = j)$$

and

$$(21a) \quad \sum_{a=1}^m x_{ia}x_{ja} = \bar{s}_i + \bar{s}_j - (\bar{s} + s_{ij}) - \lambda \delta_{ij} + \lambda n^{-1} \quad (i \neq j).$$

On the other hand, squaring and expanding (3) we have

$$(22) \quad d_{ij}^2 = \sum_{a=1}^m x_{ia}^2 - 2 \sum_{a=1}^m x_{ia}x_{ja} + \sum_{a=1}^m x_{ja}^2.$$

Substituting (21) and (21a) into (22) results in

$$(23) \quad \begin{aligned} d_{ij}^2 &= [2\bar{s}_i - (\bar{s} + s_{ii}) - \lambda + \lambda n^{-1}] \\ &\quad - 2[\bar{s}_i + \bar{s}_j - (\bar{s} + s_{ij}) - \lambda \delta_{ij} + \lambda n^{-1}] \\ &\quad + [2\bar{s}_j - (\bar{s} + s_{jj}) - \lambda + \lambda n^{-1}]. \end{aligned}$$

The right member of (23) reduces to

$$(24) \quad d_{ij}^2 = 2s_{ij} - (s_{ii} + s_{jj}) - 2\lambda(1 - \delta_{ij}).$$

If the diagonal elements $\{s_{ii}\}$ are constant (and we can set them so, since our theorem is concerned solely with the off-diagonal elements of S), then the right member of (24) differs from the semi-metric on the left by an

additive constant and a multiplicative factor of 2^* only, *i.e.*,

$$(24a) \quad d_{ii}^2 = 2[s_{ii} - (c + \lambda)] \quad (i \neq j \text{ and diagonal constant}),$$

which has the form of (6) and, therefore, satisfies (1) and (2) as well in $m \leq n - 2$ dimensions. Taking the square roots of (24a) yields the Euclidean metric (where (1) and (2) are also satisfied), which was to be proven.

Although we were free to set the $\{s_{ii}\} = c$ in our construction, actually a weaker condition suffices for strict monotonicity for $i \neq j$. If $s_{ii} < s_{kk}$ ($i \neq j$ and $k \neq \ell$), then it is sufficient that

$$(25) \quad \frac{1}{2}[(s_{kk} + s_{\ell\ell}) - (s_{ii} + s_{jj})] < s_{kk} - s_{ii}$$

satisfies (1); or, if $s_{ii} = s_{kk}$ ($i \neq j$ and $k \neq \ell$), then

$$(26) \quad s_{ii} + s_{jj} = s_{kk} + s_{\ell\ell}$$

must satisfy (2), for all i, j, k , and ℓ . Setting $\{s_{ii}\}$ to c (and zero has the virtue of simplicity) insures that the semi-metric be a linear transformation of $S(i \neq j)$, whereas permitting the weaker condition above of unequal diagonal elements, fulfilling conditions (25) and (26), makes $\{d_{ij}^2\}$ a monotonic (but, nevertheless, strict) transformation of $S(i \neq j)$.

The above theorem for the distance model is pertinent to those techniques which perform a monotone distance analysis on symmetric matrices ($i \neq j$), *e.g.*, those by de Leeuw, Guttman & Lingo, Kruskal, McGee, Roskam & Lingo, Shepard, Torgerson, and Young. Some of the more familiar "brand" names of computer programs for such analyses are: M-D-SCAL, TORSCA, SSA-I, to name but a few.

Properly speaking, of course, our theorem is not concerned with the diagonal elements of S at all and, therefore, they can be set to any convenient value in the construction. Indeed, in many experimental settings estimates of self-distances are not even determined. The $\{s_{ii}\}$ are, however, relevant to the qualification of a given index as a measure of distance or dissimilarity. Since some programs (most notably Kruskal's M-D-SCAL) permit inclusion of the diagonal of S , it is germane to consider the limits of our theorem in such cases.

If the diagonal is to be included, then it is necessary that

$$(27) \quad \{s_{ii}\} = c,$$

otherwise (1) would be violated, *i.e.*, we would have a degenerate solution, since untied $\{s_{ii}\}$ would map into tied $\{d_{ii}\}$. Furthermore, the constant must satisfy

* The $\frac{1}{2}$ in defining Torgerson's B^* matrix would absorb the 2 associated with s_{ij} and divide the sum of the diagonals, but, for our purposes, this fact is of no consequence (other than motivating the particular form that (6) assumes), since the rank of T and B^* are identical and they have the same unit length vectors.

$$(28) \quad c \leq s_{i i_{m i n}} \quad (i \neq j),$$

because $d_{m i n} = 0$ and all $d_{i i} = 0$.

Obviously, when the inequality is valid in (28) we have satisfied (1) and (2) for strict monotonicity. When, however, the equality holds in (28), then it must also be true that

$$(29) \quad s_{i k} = s_{j k} \quad (k = 1, 2, \dots, n)$$

be satisfied (yielding zero off-diagonal distances for all $s_{i j} = c$) for strict monotonicity between all elements of S and D . If (29) is not met for all k for equality in (28), then we have what Guttman [1968] refers to as "semi-strong" monotonicity, *i.e.*, (1) is fulfilled, but we are free to map unequal distances into equal dissimilarities in (2).

It remains to treat the degenerate or special cases for our two models as well as to indicate greatest lower bounds.

Degenerate Cases

Certainly, when $n = 2$ our constructions for both models insure that the two implied vectors of V and T vanish, yielding $m = 0$ and $\theta_{i j} = d_{i j} = 0$. If this result is to be disallowed for $s_{i j} \neq 0$, then $n - 1$ dimensions are required for strict monotonicity. The construction of both V and T guarantees at most $n - 1$ dimensions by virtue of zero row (column) sums. The null matrix, of course, is not a degenerate case because its rank is properly zero.

For $n > 2$ and all off-diagonal elements of $S = c \neq 0$, the construction of V and T will yield one zero root and $n - 1$ equal, non-zero roots (*i.e.*, $\lambda(V) = -nc$ and $\lambda(T) = c$ when $\{s_{i i}\} = 0$), which, when subjected to the transformation implicit in (10) and (18), representing a translation of the origin of the eigen-solution, will vanish. As a consequence, the constant vector of W will perfectly reproduce the off-diagonal elements of S if we do not deflate X by removing the constant vector as in (12). Not so, however, for the distance model. The constant vector of R in (18) will result in zero distances, which is equivalent to $m = 0$ (a degenerate case). Thus, if we do not permit such a solution we will require $n - 1$ dimensions as the least upper-bound for S in the case of the distance model when the off-diagonal elements are constant (but not zero).

Whenever there are at least two distinct off-diagonal elements of S , there must be at least two distinct eigenvalues of V or T (excluding the zero root associated with the constant vector). By selecting the algebraically smallest root (whether positive, negative, or zero) not associated with the constant vector as a basis for translation, we establish the possibility of reducing the rank of V or T to $n - 2$ always.

From the above argument we have demonstrated that $n - 1$ is a least upper-bound for $n = 2(s_{i j} \neq 0)$ for both models, while for $n > 2$ only the

distance model requires $n - 1$ dimensions for a non-degenerate solution when the off-diagonal elements of $S = c \neq 0$. Based on these results we can establish greatest lower bounds for our two models.

If p is the order of the largest principal minor containing identical values in $S(i \neq j)$, then the greatest lower-bound for the vector model is one dimension and $p - 1$ dimensions for the distance model when strict monotonicity is insisted upon.

For the *proof*: the constant vector alone will metrically reproduce the off-diagonal elements of V for the vector model when $p = n$ (v. s.) and, therefore, for $p < n$ a constant sub-vector of p components will suffice. In the case of the distance model, any Euclidean representation of the p points alone is mathematically equivalent to a $p - 1$ dimensional regular simplex with each point a vertex. This subspace, as a consequence, sets a lower limit for embedding the remaining points. The Guttman-Lingoes SSA-I algorithm prefers to break such ties in an optimal fashion for reducing dimensionality below this minimum (yielding a lower-bound of one dimension) by requiring semi-strong rather than strong (strict) monotonicity.

Concluding Remarks

Our constructions in the above proofs are always specifically of a Euclidean (vector or point) space (and semi-metrics which are monotone functions thereof) and underline the importance of the lack of uniqueness of a solution in a large space and the necessity of empirically having a small space for scientifically meaningful results. Additionally, these constructions also offer a basis for comparison of initial configurations used in these two models, e.g., the Guttman-Lingoes initial configuration used in SSA-I is quite similar to our construction of V for the vector model. A more consistent initial configuration for the distance model is that proposed by Lingoes and Roskam [1970] as embodied in

$$(30) \quad t_{ij}^* = \frac{1}{n} \sum_{i=1}^n \rho_{ij} + \frac{1}{n} \sum_{j=1}^n \rho_{ij} - \frac{(n-1) \left[\binom{n}{2} + 1 \right]}{2n} - \rho_{ij},$$

where ρ_{ij} is the rank-order value of s_{ij} (i.e., $\rho_{ij} = 1$ for $s_{ij_{max}}$ when S is a similarity matrix, otherwise for $s_{ij_{min}}$ when S is a dissimilarity matrix) and the fractional term involving n in both the numerator and denominator represents the average of the first $\binom{n}{2}$ integers taken over the n^2 elements of the symmetric matrix of ranks having a null diagonal. The use of ranks rather than the given $\{s_{ij}\}$ is essentially neutral regarding the shape of the Shepard diagram relating the input coefficients to the distances and provides a standard way of converting similarity/dissimilarity coefficients into distance values.

REFERENCES

- Guttman, L. A definition of dimensionality and distance for graphs. (Unpubl. mimeograph), University of Michigan, 1965.
- Guttman, L. The development of nonmetric space analysis: A letter to John Ross. *Multivariate Behavioral Research*, 1967, 2, 71-82.
- Guttman, L. A general nonmetric technique for finding the smallest coordinate space for a configuration of points. *Psychometrika*, 1968, 33, 469-506.
- Lingoes, J. C. & Guttman, L. Nonmetric factor analysis: A rank reducing alternative to linear factor analysis. *Multivariate Behavioral Research*, 1967, 2, 485-505.
- Lingoes, J. C. & Roskam, E. An empirical study of two multidimensional scaling algorithms. *Multivariate Behavioral Research*, 1970, 5, (in press).
- Torgerson, W. S. *Theory and Methods of Scaling*. John Wiley & Sons, Inc., New York, 1958.

Manuscript received 8/5/69

Revised manuscript received 8/31/70