

ALTERNATIVE MEASURES OF FIT FOR THE
SCHÖNEMANN-CARROLL MATRIX FITTING ALGORITHM*

JAMES C. LINGOES

THE UNIVERSITY OF MICHIGAN

AND

PETER H. SCHÖNEMANN

PURDUE UNIVERSITY

In connection with a least-squares solution for fitting one matrix, A , to another, B , under optimal choice of a rigid motion and a dilation, Schönemann and Carroll suggested two measures of fit: a raw measure, e , and a refined similarity measure, e_s , which is symmetric. Both measures share the weakness of depending upon the norm of the target matrix, B , e.g., $e(A, kB) \neq e(A, B)$ for $k \neq 1$. Therefore, both measures are useless for answering questions of the type: "Does A fit B better than A fits C ?". In this note two new measures of fit are suggested which do not depend upon the norms of A and B , which are $(0, 1)$ -bounded, and which, therefore, provide meaningful answers for comparative analyses.

Schönemann and Carroll [1970] proposed a generalization of the orthogonal Procrustes problem which would yield a least-squares fit of a given matrix, A , to a target matrix, B , under choice of an orthogonal rotation, a translation, and a central dilation. Such displacements in the context of monotone distance analysis, of course, leave invariant the relative magnitudes of the inter-point distances and do not affect monotone measures of goodness of fit. For monotone vector analyses, however, translations do not preserve angles, thus precluding their use as A matrices. Quite often an investigator would like to compare (or present a single orientation for) several geometric representations over either samples or techniques for which such transformations are permitted either in terms of the constraints of the analysis or in terms of the interpretations that are to be made. The S-C algorithm is ideally suited to these purposes, but the two measures of fit proposed by the authors (*op. cit.*, pp. 248 & 249) have a serious shortcoming in that they are dependent upon the norm of the target matrix, B . For example, as the Euclidean norm $(\text{tr } B'B)^{1/2}$ approaches zero so will the Euclidean norm of the residual matrix, E .

* This research in nonmetric techniques is supported in part by a grant from the National Science Foundation (GS-2850) to the University of Michigan.

In applications where our main interest is in assessing the fit of a given matrix A to several different target matrices, B_1, B_2, \dots, B_n , one may wish to have a measure of fit, $f(A, B)$, which not only is symmetric (so that $f(A, B) = f(B, A)$), but is also invariant under separate stretchings (so that $f(aA, bB) = f(A, B)$ for all non-zero scalars a, b), and which is bounded by zero and one, *i.e.*, is dimension-free.

After refreshing the reader on the notation used by Schönemann and Carroll for both the problem and its solution [1970], we shall advance two measures of fit possessing some or all of the properties listed above.

Notation

Recall that the *problem* was to

$$(1) \quad \min_{c, T, \gamma} e = \text{tr } E'E$$

in

$$(2) \quad B = cAT + J\gamma' + E,$$

where J is a $q \times 1$ column vector of ones, A and B are two known $p \times q$ matrices, T is a $q \times q$ transformation matrix for orthogonal rotation, γ' is the $1 \times q$ translation vector, c is the multiplying scalar for contraction/dilation, and E is the residual $p \times q$ error matrix.

The *solution* for c, T , and γ yields:

$$(3) \quad T = VW'$$

where

$$(4) \quad VDW' = A'QB$$

and

$$(5) \quad Q = I - JJ'/p$$

for the $p \times p$ matrices I (identity matrix) and Q , and where

$$(6) \quad c = \text{tr } T'A'QB / \text{tr } A'QA$$

and

$$(7) \quad \gamma = (B - cAT)'J/p.$$

A measure of *raw* fit is the criterion function itself, the Frobenius norm $|E|$ of the matrix of residuals:

$$(8) \quad E = Q(B - cAT),$$

which is given by

$$(9) \quad e = \text{tr } E'E = \text{tr } B'QB - (\text{tr } T'A'QB)^2 / \text{tr } A'QA.$$

The above criterion function is, however, not symmetric in the sense that a least-squares fit of A to B would give a different value for e than a fit of B to A . To achieve symmetry, the constant

$$(10) \quad u = \text{tr } B'QB / \text{tr } A'QA$$

was introduced to define a *symmetric* measure of fit:

$$(11) \quad e_s = eu^{-1/2} = (\text{tr } A'QA)^{1/2}(\text{tr } B'QB)^{1/2} - (\text{tr } T'A'QB)^2 / [(\text{tr } A'QA)^{1/2}(\text{tr } B'QB)^{1/2}],$$

which does satisfy

$$(12) \quad e_s(A, B) = e_s(B, A),$$

i.e., is invariant under the order of fit, as was shown in *loc. cit.*

However, from (9) and (11), it is clear that both e and e_s depend upon the norm of the target matrix B (although not on A , because of c), so that

$$(13) \quad e(A, kB) = k^2e \quad \text{and} \quad e_s(A, kB) = ke_s.$$

Indeed, both

$$(14) \quad e \rightarrow 0 \quad \text{and} \quad e_s \rightarrow 0 \quad \text{as} \quad \text{tr } B'B \rightarrow 0,$$

as long as we are dealing with real matrices, since

$$(15) \quad \text{tr } B'B = \text{tr } E'E + \text{tr } \hat{B}'\hat{B},$$

where the matrix of best fit, \hat{B} , is defined from (2) as $\hat{B} = B - E$ and satisfies $\text{tr } \hat{B}'E = 0$.

Such dependency on the norm of B is clearly undesirable when one wants to compare the fits for different target matrices. We shall, therefore, propose two alternative measures which, since they are so defined as to be invariant under the norm of B , render such comparisons meaningful.

Alternative Measures of Fit

Perhaps the simplest solution to achieve the desired scale invariance is to divide e in (9) by $\text{tr } B'B$, *i.e.*, to define

$$(16) \quad L = \text{tr } E'E / \text{tr } B'B.$$

Since

$$(17) \quad 0 \leq \text{tr } E'E \leq \text{tr } B'B$$

from (14), it follows that this coefficient is not only invariant under separate stretchings of A and B (in view of (13)), but is also bounded by 0 and 1.

The measure, L , is perfectly adequate for comparison designs where a given matrix A is fitted to several target matrices: B_1, B_2, \dots , because it

permits the selection of that B which fits A best. Such designs arise in comparisons of multiple scaling solutions or they might involve multiple hypotheses about the target matrix of a configuration.

In contrast to e_s , which is norm-dependent, but symmetric, L , although norm-invariant, is no longer symmetric, *i.e.*,

$$(18) \quad L(A, kB) = L(A, B), \text{ but } L(A, B) \neq L(B, A),$$

except when $u^{1/2} = \text{tr } B'B$.

Our second proposal for a measure of fit combines the advantages of both e_s and L . If we divide the raw measure, e , in (9) by $\text{tr } B'QB$ instead of by $\text{tr } B'B$ as in (16), we obtain

$$(19) \quad S = 1 - (\text{tr } T'A'QB)^2 / (\text{tr } A'QA \cdot \text{tr } B'QB).$$

S results from L defined in (16) as a special case when the matrices $A = (a_{ij})$ and $B = (b_{ij})$ are preprocessed to satisfy

$$(20) \quad \sum_i a_{ij} = \sum_j b_{ij} = 0 \quad \text{or} \quad J'A = J'B = \phi'$$

and

$$(21) \quad \sum_i \sum_j a_{ij}^2 = \sum_i \sum_j b_{ij}^2 = 1 \quad \text{or} \quad \text{tr } A'A = \text{tr } B'B = 1.$$

With these restrictions the translation vector γ vanishes and one finds that

$$(22) \quad e = e_s = L = S,$$

so that all four coefficients coincide. All will be invariant under separate stretchings and order of fitting and all will be (0, 1)-bounded. It then becomes meaningful to ask whether A fits B better than C fits D , regardless of the number of points, the number of dimensions, the various norms, and regardless of the fitting order.

The coefficient, S , defined in (19) makes preprocessing unnecessary, *i.e.*, it has the virtues of both e_s (symmetry) and L (scale-invariance). S is symmetric because it differs from e_s only by the symmetric multiplier, $(\text{tr } A'QA \cdot \text{tr } B'QB)^{-1/2}$, and it is norm-invariant and bounded by 0 and 1 because it is a special case of L , which is norm-invariant and (0, 1)-bounded for all matrices, and S results from L when A and B satisfy the constraints (20) and (21).

That $S^{1/2}$, in fact, is the matrix analogue of a coefficient of alienation can be most easily seen by considering the limiting case where A and B are p -element column vectors. Then QA , QB are vectors of deviation scores and T is merely a reflection (± 1). The $\text{tr } A'QA$ and $\text{tr } B'QB$ are then proportional to sample variances and $c = \text{tr } T'A'QB / \text{tr } A'QA$ is a regression coefficient, which becomes a correlation coefficient, r , if A and B are normed by $(\text{tr } A'QA)^{1/2}$ and $(\text{tr } B'QB)^{1/2}$, respectively. Thus, S is the matrix analogue of $1 - r^2$.

Illustration

Using the 4×2 and 16×5 matrices given by Schönemann and Carroll [1970] in their Tables 2 and 3 (but renaming the matrices of Table 3 as C and D), we determined $S^{1/2}(A, B) = .109$ and $S^{1/2}(C, D) = .499$. The corresponding values for $e_s(A, B)$ and $e_s(C, D)$ were .042 and .058, respectively. As can be seen from the coefficients of alienation, C fits D about five times poorer than A fits B , whereas the values for e_s would suggest not too much difference in the respective fits.*

* A computer program is available from Lingo [1973], which substitutes $S-C$'s e and e_s measures by L and $S^{1/2}$, respectively.

REFERENCES

- Lingo, J. C. *The Guttman-Lingo Nonmetric Program Series*. Ann Arbor, Michigan: Mathesis Press, 1973.
- Schönemann, P. H. & Carroll, R. M. Fitting one matrix to another under choice of a central dilation and a rigid motion. *Psychometrika*, 1970, **35**, 245-255.

Manuscript received 6/1/73

Revised manuscript received 4/25/74