PSYCHOMETRIKA—VOL. 35, NO. 3
SEPTEMBER, 1970

# ON METRIC MULTIDIMENSIONAL UNFOLDING

PETER H. SCHÖNEMANN*

THE OHIO STATE UNIVERSITY

The problem of locating two sets of points in a joint space, given the Euclidean distances between elements from distinct sets, is solved algebraically. For error free data the solution is exact, for fallible data it has least squares properties.

## 1. *Introduction*

In Coombs' own words "the basic assumptions of the theory of preferential choice on which the unfolding technique in one dimension is based are as follows: Each individual and each stimulus may be represented by a point on a common dimension called a "*J*-scale," . . . and each individual's preference ordering of the stimuli from most to least preferred corresponds to the rank order of the absolute distances of the stimulus points from the ideal point, the nearest being the most preferred. The individual's preference ordering is called an "*I*-scale" and may be thought of as the *J*-scale folded at the ideal point. . . . The data consist of a set of *I*-scales from a number of individuals, and the analytical problem is how to unfold these *I*-scales to recover the *J*-scale" [Coombs, 1964, p. 80].

Coombs and his coworkers, notably Bennett [1956], Bennett and Hays [1960], and Hays and Bennett [1961], generalized this basic idea into a theory, and a technique, for multidimensional unfolding, with the objective of locating the points of both sets, given a number of *I*-scales, in a joint space of more than one dimension.

In its original nonmetric form [Coombs, 1958] this model presents considerable technical problems even in the simplest, one-dimensional case. Goode [1957] contributed an ingenious algorithm aimed at easing the effort necessary to unfold a set of *I*-scales in one dimension. The technical difficulties multiply in the multidimensional case, and, although the theoretical work by Bennett and Hays provides a solution in principle, such a solution quickly becomes impractical for "even . . . a moderate number of stimuli in a moderately small dimensionality" [Coombs, 1964, p. 175].

Realizing this, Coombs, very early, looked for a metric version of his unfolding problem which, perhaps, might be easier to solve and which might

*Now at Purdue University.

also be useful as an approximate substitute for the nonmetric case in certain situations [Coombs & Kao, 1960; Coombs, 1964, Ch. 8]. Again we quote Coombs [p. 181f]: "Consider the simple case of a one-dimensional latent attribute generating the preferences of individuals over a set of alternatives. . . . Consider the *I*-scale of an individual at the extreme left end of the scale and that of another individual very close to him. Clearly, their preference orderings will be almost identical and will correlate close to $+1$. Individual *A*'s *I*-scale will correlate progressively less with the *I*-scales of other individuals as they are farther removed from him on the joint scale. In fact, the correlation will be zero between individual *A* and the median individual in the distribution, and will ultimately be $-1$ between him and the individual at the extreme opposite end of the scale. . . . If we factor analyze the correlation matrix for such a configuration by the method of principal components, the space obtained will be two-deminsional, and under rotation one dimension will be the original line that generated the preferential choices and the second dimension will be the vector of the median individual on the line. . . . The higher the projection of an individual's point on this extra dimension the more central the individual in the configuration of individuals. Hence the nearer he is to the others on the average and the better he represents them."

Coombs used metric data (*i.e.*, numerical distances instead of just their orderings) to verify this intuitive reasoning was essentially correct. We might point out, in passing, that his definition of the "median individual" comes very close to that of a centroid (being the point least distant from a set of given points, *i.e.*, the mean individual) so that the problem of having to "rotate out" this extra dimension could have been avoided by first removing the centroid, prior to obtaining the principal axes decomposition. (See also below.)

In [1964] Ross and Cliff took a closer look at the metric version of Coombs' unfolding model. They found "first, (that the) Coombs and Kao [1960] conjecture is nearly correct in that under some conditions the rank of the matrix of correlations between individuals of the distance between an individual and a stimulus will be approximately one greater than the dimensionality of the set of points. This is only approximately true, however. Exact statements are possible if squared distances are used instead of distances themselves. Here we see that the obtained rank will be one greater than the number of common dimensions. The 'extra' factor will give the distance from the centroid of one set of points. Doubly-centering the matrix will eliminate the extra factor" [p. 176]. Ross and Cliff also found that, while "determination of the rank of the space . . . is thus quite straightforward . . . solution for the two sets of coordinates which will reproduce *D* (the distance matrix) is another matter. . . . A relatively simple (*viz.* one-dimensional) case was investigated. . . . In the errorless case, an attempt to solve for the two unknowns resulted in sixth degree equations. . . . The intractableness of even this simple case led us to abandon attempts at a solution" [p. 174]. Notwithstanding, the present paper owes much to their thoughtful exposition of some of the aspects of the metric multidimensional scaling problem.

## 2. An Algebraic Formulation of the Metric Unfolding Problem

Given two sets of points $S_1, \cdots, S_p$ and $P_1, \cdots, P_q$ with coordinate vectors $a_i'$ $(i = 1, \cdots, p)$ and $b_j'$ $(j = 1, \cdots, q)$ relative to some origin in a joint Cartesian space of $m$ $(\geq 1)$ dimensions one can compute all distances $d_{ij}$ between pairs with one member from each set as

$$d_{ij}^2 = (a_i - b_j)'(a_i - b_j) = a_i'a_i - 2a_i'b_j + b_j'b_j \tag{2.1}$$

which could be written

$$\Delta_{12}^{(2)} = (d_{ij}^2) = a^{(2)}J_q' - 2AB' + J_pb^{(2)\prime}, \tag{2.2}$$

where $J_p$ and $J_q$ are two column vectors of ones with $p$ and $q$ components respectively, $a^{(2)} = (a_i'a_i)$, $b^{(2)} = (b_j'b_j)$ and $A$, $B$ are the two coordinate matrices of the $S_i$ and $P_j$, respectively. These two matrices are determined only up to a rotation which will leave the distances and the scalar products $a_i'b_j$ invariant, as is well known.

In Schönemann, [in press] use was made of the fact that any translation of the coordinates $a_i'$ to some new origin with coordinates $a_0'$ can be achieved by use of an oblique projection operator $Q$ which annihilates vectors of type $J$ which appear in (2.2). For example, if $a_0' = c'A$ and $c$ satisfies $c'J = 1$ (which it can always be made to satisfy, see *ibid.*) then the matrix

$$Q_1 = I - J_pc_1' \tag{2.3}$$

will be idempotent and will also annihilate all (column-) vectors in the space of $J$. Applying such a matrix to $A$ one finds

$$A^* = Q_1A = A - J_p(c_1'A) = A - J_pa_0' \, . \tag{2.4}$$

$A^*$ is a matrix with rows $a_i^{*\prime} = a_i' - a_0'$, which are the coordinates of the $i$'th point of the first set relative to some new origin $a_0'$ with coordinate representation $a_0' = c_1'A$ prior to the translation. One similarly obtains a matrix $B^*$ with rows $b_j^{*\prime} = b_j' - c_2'B$ if one premultiplies $B$ with the idempotent matrix

$$Q_2 = I - J_qc_2' \qquad (c_2'J_q = 1).$$

However, it is important to note that even where $p = q$ and $Q_1 = Q_2$, $a_0' \neq b_0'$, in general, *i.e.*, the origins to which $A$ and $B$ are referred to will not be the same.

Applying two such projectors, $Q_1$ and $Q_2$, to eq. (2.2) one obtains

$$C_{12} = A^*B^{*\prime} = Q_1AB'Q_2' = Q_1(-\Delta_{12}^{(2)}/2)Q_2' \tag{2.5}$$

which is a matrix with elements $c_{ij}$ which formally resemble scalar products between pairs of elements from distinct sets, except that the coordinates of the left member $(a_i^{*\prime})$ are expressed relative to a different origin than those of the right member $(b_j^{*\prime})$. To avoid confusion let us call these elements

$$c_{ij} = (a_i - a_0)'(b_j - b_0) = a_i^{*\prime}b_j^* \tag{2.6}$$

"quasi-scalar products."

An off-diagonal matrix of such quasi-scalar products can be computed from the off-diagonal matrix of the observed distances $\Delta_{12} = (d_{ij})$ where the choice of $Q_1$ and $Q_2$ is quite arbitrary. A case could be made, perhaps, for preferring the centroids of both subsets as new origins, in which case

$$Q_1 = (I - J_pJ_p'/p) \text{ and } Q_2 = (I - J_qJ_q'/q). \tag{2.7}$$

The matrix $C_{12}$ can be decomposed into a product of two matrices $G, H$ of full column rank in a number of ways, *e.g.*, by an Eckart–Young (1936) decomposition. Any two such factorizations must be related by a non-singular transformation. In particular, let

$$C_{12} = GH' = A^*T^{-1}TB^{*\prime} \tag{2.8}$$

relate $A^*$ with $G$ $(GT = A^*)$ and $B^*$ with $H$ $(HT^{-1\prime} = B^*)$ where $G$ (with rows $g_i'$) and $H$ (with rows $h_j'$) are the factors of a given full rank factorization of $C_{12}$.

One then has

$$a_i' = a_i^{*\prime} + a_0' = g_i'T + a_0' \tag{2.9}$$

and

$$b_j' = b_j^{*\prime} + b_0' = h_j'T^{-1\prime} + b_0' .$$

Returning to (2.1) one now finds for the squares of the observed distances

$$d_{ij}^2 = g_i'Mg_i + h_j'M^{-1}h_j + 2g_i'T(a_0 - b_0) - 2h_j'T^{-1\prime}(a_0 - b_0) + (a_0 - b_0)'(a_0 - b_0) - 2g_i'h_j \tag{2.10}$$

where

$$M = TT'. \tag{2.11}$$

Given $\Delta_{12}^{(2)} = (d_{ij}^2)$, $g_i'$ and $h_j'$ (which is computable from (2.5)) the problem is now to solve for the unknowns $m_{rs}$ in $M = (m_{rs}) = M'$ and the vector of unknowns $a_0 - b_0$. Having obtained $M$, it could be factored for $T$ as in (2.11) which then could be used to find $A_0$, $B_0$ from (2.8) as

$$GT = A - J_pa_0' = A_0 \qquad HT^{-1\prime} = B - J_qb_0' . \tag{2.12}$$

$HT^{-1\prime}$, in turn could be converted, with the help of the vector $a_0 - b_0$ into

$$HT^{-1\prime} - J_q(a_0 - b_0)' = B - J_qa_0' = B_0 . \tag{2.13}$$

$A_0$, $B_0$ would give the coordinates of both sets of points in a joint space relative to an origin at $a_0'$ up to a rotation.

### 3. *An Algebraic Solution of the Metric Unfolding Problem*

Eqs. (2.10) can be rearranged to read

$$f_{ij} = d_{ij}^2 + 2g_i'h_j = g_i'Mg_i + 2g_i'T(a_0 - b_0) \qquad \begin{array}{l} i = 1, \cdots, p \\ j = 1, \cdots, q \end{array} \tag{3.1}$$

+ terms not involving the subscript $i$.

Therefore, differencing on $i$ will eliminate all terms which are either constants or which depend on $j$ alone, *e.g.*,

(3.2)

$$f_{ij} - f_{pj} = g_i'Mg_i - g_p'Mg_p + 2(g_i' - g_p')T(a_0 - b_0) \qquad \begin{array}{l} i = 1, \cdots, p - 1 \\ j = 1, \cdots, q. \end{array}$$

There now are $(p - 1)q$ equations in $m(m + 1)/2 + m = m(m + 3)/2$ unknowns. Moreover, eqs. (3.2) can be written as a linear system of $(p - 1)q$ inhomogeneous equations in $m(m + 3)/2$ unknowns.

To see this, consider the simplified case of $p$ equations of the type

$$k_i = c_i'Xc_i , \qquad X = (x_{rs}) = X' \qquad \begin{array}{l} i = 1, \cdots, p \\ r, s = 1, \cdots, n \end{array} \tag{3.3}$$

where the $p$ vectors $c_i'$ and the $p$ constants $k_i$ are known and where the problem is to solve for the elements $x_{rs}$ in the symmetric matrix $X$. One could write such a system in summation notation as

$$\sum_r^n \sum_s^n c_{ir}c_{is}x_{rs} = k_i \qquad \begin{array}{l} i = 1, \cdots, p \\ r, s = 1, \cdots, n \end{array} \tag{3.4}$$

or, equivalently, since $x_{rs} = x_{sr}$ , as

$$\sum_{r \le s} \sum c_{ir}c_{is}y_{rs} = k_i \qquad i = 1, \cdots, p \tag{3.5}$$

where

$$\begin{aligned} y_{rs} &= x_{rs} \quad \text{if} \quad r = s \\ &= 2x_{rs} \quad \text{if} \quad r < s. \end{aligned}$$

It is clear that (3.4) is an inhomogeneous system of $p$ linear equations in $n^2$ unknows, some of which appear in pairs. The matrix of coefficients, with entries $c_{ir}c_{is}$ , has repeated columns, rendering it of deficient column rank. The change from the $x_{rs}$ to the $y_{rs}$ eliminates these redundant columns, rendering the system, hopefully, of full column rank (provided $p \geq (n^2 + n)/2$ and no other dependencies exist).

One may therefore rewrite (3.2) as

$$(3.6) \qquad f_{ij} - f_{pj} = d_{ij}^2 - d_{pj}^2 + 2(g_i - g_p)'h_j$$
$$= \sum_{r \leq s}^{m} \sum^{m} (g_{ir}g_{is} - g_{pr}g_{ps})n_{rs} + \sum_{r}^{m} (g_{ir} - g_{pr})x_r$$

where $x_r$ $(r = 1, \cdots, m)$ are the components of the vector $2T(a_0 - b_0)$ and

$$(3.7) \qquad n_{rs} = m_{rs} \quad \text{if} \quad r = s$$
$$= 2m_{rs} \quad \text{if} \quad r < s.$$

Not all of the resulting $(p - 1)q$ linear equations are independent. Rather, there are at most $p - 1$ linearly independent equations in (3.6) since each row of the coefficient matrix is repeated $q$ times.

Hence, in the exact case, where all these equations are identities, it would suffice to solve the system (3.6) for any given $j$. The solution should be unique provided $p - 1 \geq m(m + 3)/2$ since there is no obvious reason why the $m(m + 3)/2$ columns should not be linearly independent. We thus have a necessary, and, in general, also sufficient condition for a unique solution of the metric multidimensional unfolding problem. It should also be clear that there is, of course, perfect symmetry between the two sets of points, *e.g.*, between subjects and stimuli, so that our condition for a unique solution applies to either $p$ or $q$, whichever is larger. If $q > p$ one simply interchanges the role of both subsets in all of the preceding equations to arrive at the same conclusions.

In the fallible case the rows of the coefficient matrix would still repeat $q$ times, but there is no assurance that the constants on the left of (3.6) also will. Indeed, the system will probably be inconsistent as soon as the number of linearly independent rows exceeds the number of columns, *i.e.*, when it is "overdetermined." But in this case one can obtain a least squares solution by multiplying (3.6) on the left with the Moore inverse of the coefficient matrix since (3.6) is formally identical with a set of regression equations if the system is overdetermined. To be explicit, let $K$ be the matrix of coefficients in (3.6) with $q$ identical submatrices $K_0$ of order $(p - 1)\mathrm{x}m(m + 3)/2$ of the form

$$(3.8) \quad K_0 = (g_{ir}g_{is} - g_{pr}g_{ps} \,;\, g_{ir} - g_{pr}), \qquad K' = (K_0', K_0', \cdots, K_0')$$
$$i = 1, \cdots, p - 1, \qquad r = 1, \cdots, m \qquad s = r, \cdots, m$$

and let the vector of known constants on the left of (3.6) be denoted $f$ with $q$ subvectors

$$(3.9) \quad f_j = (d_{ij}^2 - d_{pj}^2 + 2(g_i - g_p)'h_j), \qquad f' = (f_1', f_2', \cdots, f_q'),$$

let the $m(m + 1)/2$ unknowns $n_{rs}$ be assembled in a vector $\nu$ and the remaining $m$ unknowns $x_r$ in a vector $\xi$. Then (3.6) becomes

$$f = K\binom{\nu}{\xi} \tag{3.10}$$

whence

$$\binom{\hat{\nu}}{\xi} = (K'K)^{-1}K'f \tag{3.11}$$

$$= (K_0'K_0)^{-1}K_0'\bar{f}$$

(where $\bar{f} = (1/q) \sum_i^q f_i$)

is either an exact solution or a least squares solution, depending on whether (3.10) is consistent or not. For error free data it must be consistent, so that the solution will be exact. For fallible data, provided one has enough to overdetermine the system, it will probably be inconsistent, so that in this case one obtains a least squares solution for the unknowns $m_{rs}$ and $x_r$ .

Having solved (3.10), one can assemble the components of the partition $\nu$ of the solution vector into a symmetric matrix $M = (m_{rs}) = M'$ where

$$m_{rs} = n_{rs} \quad \text{if} \quad r = s \tag{3.12}$$

$$= n_{rs}/2 \quad \text{otherwise.}$$

In the exact case this matrix $M$ must be positive definite and can therefore be factored for $T$ (2.11) which is determined up to a rotation. In the fallible case $M$ may not necessarily be positive definite. If only one root is non-positive, we suggest† to construct from $M$ a new matrix $M^*$ by replacing the non-negative root by some ETA ($> 0$) which is less than the smallest positive root of $M$, and then to factor $M^*$. Not only must $M^*$, so constructed, be positive definite, but it also is a least squares approximation to $M$ subject to the constraint that no roots be less than ETA. This is so because the sum of squares of the residuals $(M_{rs} - M_{rs}^*)$ is given by the sum of squared differences between the two sets of latent roots in $M$ and $M^*$.

Once $T$ has been found by factoring $M$ or $M^*$ the difference vector $a_0 - b_0$ can be computed from (3.6) as

$$a_0 - b_0 = T^{-1}\xi/2. \tag{3.13}$$

$T$ and this difference vector can then be used as in (2.12), (2.13) to obtain the coordinates of both sets of elements up to a rigid motion.

†This remedy did not work as well in practice as had been hoped when these lines were written. If $M$ has nonpositive roots the program should be terminated. One of my students, Miss Wang, is presently working on a more robust least squares solution to handle the fallible case.

## 4. Computational Notes

While the foregoing algebra may appear involved, the classical, one-dimensional case of the unfolding model can actually be computed by hand as we shall demonstrate in the next section. However, as the number of dimensions $m$ exceeds 1 the computations tend to become unwieldy, especially in view of the need to solve the linear system (3.10), (3.11), which involves five unknowns if $m = 2$, nine unknowns if $m = 3$ and, in general, $m(m + 3)/2$ unknowns for a solution in $m$ dimensions. For this reason a subroutine was written, in FORTRAN IV, which is available from the author upon request. For the benefit of potential users who prefer to write their own, perhaps more efficient program, we include some computational suggestions.

The overall flow of such a program should not present any difficulties. To summarize the major steps:

[1] Read the distance matrix $\Delta_{12}$, $p$, $q$, $m$ (optional), ETA (optional).

[2] Compute the matrix of quasi-scalar products $C_{12} = Q_1(-\Delta_{12}^{(2)}/2)Q_2'$, according to (2.5). This can be done by squaring the elements $d_{ij}$ in $\Delta_{12}$, doubly centering the resulting matrix, dividing all elements by 2 and reversing all signs, in which case the elements $c_{ij}$ in $C_{12}$ are expressed relative to the two different centroids of both sets.

[3] Decide on $m$, the dimensionality of the joint space and decompose $C_{12}$ into a product $C_{12} = GH'$ where both $G$, $H$ are of full column rank $m$. This is most conveniently done in terms of an Eckart–Young decomposition of $C_{12}$ in view of the well-known least squares properties of such a decomposition. Moreover, the Eckart–Young roots in $D_c$(in $C_{12} = XD_cY'$, $X'X = Y'Y = I_{\min(p,q)}$, $D_c$ = diagonal) can be used to estimate $m$ as the number of roots in $D_c$ which are larger in magnitude than some preset ETA. However, care should be taken that $m$, whichever way arrived at, satisfies the admissibility condition $m(m + 3)/2 \leq \max(p - 1, q - 1)$, as a necessary condition for a unique solution of the linear system (3.10).

[4] Construct the coefficient matrix $K_0$ in eq. (3.8) using the rows of the larger matrix $G$, or $H$. Let $G$ be the larger matrix (*i.e.*, $p \geq q$) and let $G = XD_c$ have elements $c_{ir}$ $(i = 1, \cdots, p, r = 1, \cdots, m)$. A reasonable procedure would be to construct $K_0$ in two steps:

(i) construct the $p$ rows $\kappa_i^{*\prime}$ of an intermediate matrix $K^*$ as

$$\kappa_i^{*\prime} = (c_{i1}^2, c_{i1}c_{i2}, \cdots, c_{i1}c_{im}, c_{i2}^2, \cdots, c_{im}^2, c_{i1}, c_{i2}, \cdots, c_{im}) \qquad i = 1, \cdots, p$$

and

(ii) obtain the $(p - 1) \times (m(m + 3)/2)$ matrix $K_0$ by subtracting the $p$'th row of $K^*$ from all preceding $p - 1$ rows.

[5] Construct the vector of constants $\bar{f} = (1/q \sum_i^q (d_{ij}^2 + 2c_{ij} - d_{pj}^2 - 2c_{pj}))$ (eqs. (3.1) and (3.6))

[6] Find the vector of unknowns $\begin{pmatrix}\hat{\nu}\\ \xi\end{pmatrix}$ by solving the linear system $\bar{f} = K_0\begin{pmatrix}\nu\\ \xi\end{pmatrix} + e$ in a least squares sense, *i.e.*, compute

$$\begin{pmatrix}\hat{\nu}\\ \xi\end{pmatrix} = (K_0'K_0)^{-1}K_0'\bar{f}.$$

This step may require some care to avoid indeterminate solutions, see below.

[7] Expand the components of the partition $\nu$ of the solution vector into a symmetric $m \times m$ matrix $M$ using eq. (3.12).

[8] Check $M$ for positive-definiteness by obtaining its eigendecomposition $M = VD_mV'$. If necessary replace the non-positive root in $D_m$ by some ETA ($> 0$) to obtain a positive definite least squares approximation $M^*$ as discussed at the end of Section 3.

[9] Factor $M$ (or $M^*$) into $TT'$, preferably using the eigendecomposition obtained at step [8], *i.e.*, $T = VD_m^{1/2}$, where $D_m$ contains the latent roots of $M$ (or $M^*$).

[10] Carry $G$ into $A_0 = GT$, which expresses the coordinates of the larger set relative to their own centroid as origin.

[11] Compute the difference vector $\alpha_0 - \beta_0 = T^{-1}\xi/2 = D_m^{-1/2}V'\xi/2$, avoiding explicit inversion by use of the eignedecomposition obtained at steps [8], [9].

[12] Compute $B_0 = HT^{-1\prime} - J(\alpha_0 - \beta_0)'$ by first computing $HT^{-1\prime} = HVD_m^{-1/2}$ (avoiding explicit inversion as before) and then subtracting the difference vector $\alpha_0' - \beta_0'$ obtained at step [11] from each row of $HT^{-1\prime}$. This gives the coordinates of the second set relative to the centroid of the first set as an origin.

[13] (Optional) Translate all $p + q$ coordinate vectors to the joint centroid as origin by subtracting the mean coordinate vector (computed over both $A_0$, $B_0$) from $A_0$, $B_0$ to obtain $A$, $B$.

These computations are fairly straightforward once an eigenroutine is available. However, special care should be exercised at steps [3] and [6] to ensure a unique solution for $\begin{pmatrix}\hat{\nu}\\ \xi\end{pmatrix}$. Upon solving the admissibility constraint for $m_{\max}$, one obtains $m_{\max} \leq \sqrt{8 \max (p, q) + 1} - 3)/2$, *i.e.*, $m$ should not exceed this number as a necessary condition for a unique solution of $\begin{pmatrix}\hat{\nu}\\ \xi\end{pmatrix}$. Thus, if either a preset $m$ or an $m$ obtained by counting the Eckart–Young roots larger than ETA exceeds this number then $m$ should be overwritten by $m_{max}$ before the factorization at step [3] is carried out. But even this precaution may not be enough for certain data. Although in theory, as was pointed out in Section 3, "there is no obvious reason why the $m(m + 3)/2$ columns (of $K_0$) should not be linearly independent," they may not be in practice, for no obvious reason. To be sure one should probably check the matrix $K_0'K_0$ for near-singularity by obtaining its eigendecomposition, before taking the inverse, and print out its latent roots. If any of these roots are smaller than some preset ETA the user should be warned that $K_0'K_0$ may be ill-conditioned and all subsequent calculations may be in error. Alternately one might consider stepping down $m$ by 1 before resuming at step [8] for a better determined solution in a smaller space. In practical applications it is probably wise to plan the experiment ahead of time so as to include a few more elements in the larger set than the absolutely necessary minimum for a solution of specified dimension $m$.

Table 1

A Detailed Computational Example

Errorfree Data, m=1, p=4, q=3

Coordinates used to generate $\Delta_{12}$:

A' = (1.0 3.0 6.0 9.0) B' = (0.0 4.0 7.0)

Distance Matrix $\Delta_{12}$ generated with A, B:

$$\Delta_{12} = \begin{bmatrix} 1.0 & 3.0 & 6.0 \\ 3.0 & 1.0 & 4.0 \\ 6.0 & 2.0 & 1.0 \\ 9.0 & 5.0 & 2.0 \end{bmatrix} = (d_{ij})$$

Computations:

Matrix of Squares $\Delta_{12}^{(2)}=(d_{ij}^2)$ with row- and column means

(Step [2])

| | | | | | $\Sigma_j$ | $\bar{d}_{i.}$ |
|---|---|---|---|---|---|---|
| | 1.0 | 9.0 | 36.0 | 81.0 | 127 | 31.75 |
| | 9.0 | 1.0 | 4.0 | 25.0 | 39 | 9.75 |
| | 36.0 | 16.0 | 1.0 | 4.0 | 57 | 14.25 |
| $\Sigma_i$ | 46.0 | 26.0 | 41.0 | 110.0 | 223 | |
| $\bar{d}_{.j}$ | 15.33 | 8.67 | 13.67 | 36.67 | | 18.58 = $\bar{d}_{..}$ |

Overall Mean

Doubly Centered Matrix $Q_1 \Delta_{12}^{(2)} Q_2' = (d_{ij}^2 - \bar{d}_{i.} - \bar{d}_{.j} + \bar{d}_{..})$

(Step [2])

| | | | | | $\Sigma_j$ |
|---|---|---|---|---|---|
| | -27.50 | -12.84 | 9.17 | 31.17 | .00 |
| | 2.50 | 1.17 | -.83 | -2.84 | .00 |
| | 25.00 | 11.67 | -8.34 | -28.33 | .00 |
| $\Sigma_i$ | .00 | .00 | .00 | .00 | |

**Table 1 Continued**

Matrix of Quasi-Scalar Products: and its Factorization:

$$C_{12} = Q_1 (-\Delta_{12}^{(2)}/2) Q_2' = G \quad H'$$

(Step [2]) (Step [3])

$$\begin{bmatrix} 13.75 & -1.25 & -12.5 \\ 6.42 & -.58 & -5.83 \\ -4.58 & .42 & 4.17 \\ -15.58 & 1.42 & 14.17 \end{bmatrix} = \begin{bmatrix} .619 \\ .289 \\ -.206 \\ -.701 \end{bmatrix} [22.23 \quad -2.02 \quad -20.21]$$

Intermediate Coefficient Matrix $K^* = (g_i^2, g_i)$:

(Step [$4_i$])

$$\begin{bmatrix} .383 & .619 \\ .083 & .289 \\ .043 & -.206 \\ .492 & -.701 \end{bmatrix}$$

Coefficient Matrix $K_0 = (g_i^2 - g_4^2, g_i - g_4)$

(Step [$4_{ii}$])

$$\begin{bmatrix} -.109 & 1.320 \\ -.408 & .990 \\ -.449 & .495 \end{bmatrix}$$

Intermediate Vector of Constants $f_j^* = (d_{ij}^2 + 2c_{ij})$:

(Step [5])

$f_1^{*\prime} = (28.50 \quad 21.84 \quad 26.83 \quad 49.83)$

$f_2^{*\prime} = (\ 6.50 \quad -.17 \quad 4.83 \quad 27.84)$

$f_3^{*} = (11.00 \quad 4.33 \quad 9.34 \quad 32.33)$

Vector of Constants $\bar{f}$ (= $f_j$ here, since Data are Errorfree)

(Step [5])

$\bar{f}' = (-21.33 \quad -28.00 \quad -23.00)$

## *5. Numerical Illustrations*

Three numerical illustrations will be discussed in this section. Two of these are one-dimensional, one is two-dimensional. Two are based on fallible data, one on exact data. Two are metric applications while one was chosen to illustrate the feasibility of using the present procedure for analyzing non-metric data.

The first example (Table 1) is based on exact data and illustrates in complete numerical detail the computational steps in the one-dimensional case. A 4 $\times$ 3 distance matrix $\Delta_{12}$ was constructed from two coordinate vectors $A$, $B$ for two types of points, say four stimuli $S_i$ and three people $P_j$. Since all computational steps are labeled in accordance with the computational sequence outlined in Section 4, Table 1, if read in conjuction with Section 4, should be self-explanatory.

Table 1 Continued

Linear System $K_0 \binom{\nu}{\xi} = \bar{f}$: (Step [6])

$$\begin{bmatrix} -.109 & 1.320 \\ -.408 & .990 \\ -.449 & .495 \end{bmatrix} \begin{bmatrix} m_{11} \\ x_1 \end{bmatrix} = \begin{bmatrix} -21.33 \\ -28.00 \\ -23.00 \end{bmatrix}$$

$$K_0 \qquad \binom{\nu}{\xi} = \qquad = \qquad \bar{f}$$

Inverse $(K_0' K_0)^{-1}$:

(Step [6])

$$\begin{bmatrix} 5.55 & 1.44 \\ 1.44 & .71 \end{bmatrix}$$

Solution Vector $\binom{\nu}{\xi} = (K_0' K_0)^{-1} K_0' \bar{f}$:

(Step [6])

$$\begin{bmatrix} m_{11} \\ x_1 \end{bmatrix} = \begin{bmatrix} 36.75 \\ -13.13 \end{bmatrix}$$

(Steps [7], [8], [9])

$T = \sqrt{m_{11}} = 6.062$

(Step [11])

$T^{-1} = .1650 \qquad a_0 - b_0 = T^{-1} x_1/2 = -1.08$

Coordinates Relative to Centroid of Larger Set:

(Step [10])

$A_0 = GT$

$= (3.75 \quad 1.75 \quad -1.25 \quad -4.25)$

(Step [12])

$B_0 = HT^{-1} - J(a_0 - b_0)$

$= (4.75 \quad .75 \quad -2.25)$

These coordinates relate to the input coordinates by the rigid motion:

Input Coordinates = - Output Coordinates + 4.75

and hence reproduce all input distances exactly.

The second example (Table 2) is intended to illustrate the possibility of using the present, metric, procedure for analyzing non-metric data. Starting with the same coordinates as in Table 1, and the corresponding distances, it is now assumed that this metric distance information is filtered through three respondents $P_1$, $P_2$, $P_3$ who return only ordinal distance information in the form of three permutations of the stimuli $S_i$. These permutations correspond to the rank order of the distances of the four stimuli from each of the three "ideal points" $P_j$. Given these three permutations as the basic input information, they are then encoded into numerical distance information on an ordinal scale by (arbitrarily) assigning "1" to the stimulus closest to $P_j$ (the first element of the $j$'th permutation), "2" to the second stimulus in the $j$'th permutation, and so on.

The resulting distance matrix $\Delta_{12}$ in Table 2 differs from that in Table 1 numerically, and it also contains an order reversal between $d_{31}$ and $d_{42}$. The corresponding matrix $C_{12}$ of quasi-scalar products is no longer exactly of rank one, as it was in Table 1.

The results obtained upon unfolding these distorted distances in one dimension are shown in Table 2. If one compares the coordinate vectors $A$, $B$ in Table 2 with those in Table 1 one finds that, while the order relations within each set have been preserved, there are two transpositions between sets, viz., $P_1$ with $S_1$ and $P_3$ with $S_3$. The rank order correlation between the metric ($U$) and the non-metric ($V$) coordinates is Rho $(U, V) = .93$. Similarly, a comparison of (i) the metric distances used for input in example 1 ($X$), (ii) the distorted distances used for input in example 2 ($Y$), and (iii) the distances reproduced from the output coordinates of example 2 ($Z$), shows that the rank order has been essentially preserved for all three sets: there is one error from $X$ to $Y$ (Rho $(X, Y) = .93$), another from $Y$ to $Z$ (Rho $(Y, Z) = .91$), and two from $X$ to $Z$ (Rho $(X, Z) = .84$).

While these results, obtained for a relatively small example, are of

Table 2

A Non-Metric Application

Fallible Data, $m=1$, $p=4$, $q=3$

Permutations of four stimuli, $S_1$, $S_2$, $S_3$, $S_4$ returned by three subjects, $P_1$, $P_2$, $P_3$:

$$\begin{array}{ccc} P_1 & P_2 & P_3 \\ \end{array}$$
$$\begin{bmatrix} S_1 & S_2 & S_3 \\ S_2 & S_3 & S_4 \\ S_3 & S_1 & S_2 \\ S_4 & S_4 & S_1 \end{bmatrix}$$

Permutations encoded into distances on an ordered scale:

$$\begin{array}{c|ccc} & P_1 & P_2 & P_3 \\ \hline S_1 & 1 & 3 & 4 \\ S_2 & 2 & 1 & 3 \\ S_3 & 3 & 2 & 1 \\ S_4 & 4 & 4 & 2 \end{array} = \Delta_{12}$$

Matrix of Quasi-Scalar Products

$$C_{12} = Q_1 (-\Delta_{12}^{(2)}/2)\, Q_2'$$

$$\begin{bmatrix} 3.83 & -.17 & -3.67 \\ .33 & 1.83 & -2.17 \\ -2.17 & .33 & 1.83 \\ -2.00 & -2.00 & 4.00 \end{bmatrix}$$

Eckart-Young Roots $D_c$ of $C_{12}$

$$D_c = (7.76 \quad 2.87 \quad .00)$$

Latent Roots of $K_0' K_0$

$$(2.50 \quad .12)$$

Table 2 Continued

Coordinates A
Relative to Joint Centroid

| | $S_1$ | $S_2$ | $S_3$ | $S_4$ |
|---|---|---|---|---|
| A' = | (2.78, | 1.14, | -1.68, | -2.86) |

Coordinates B
Relative to Joint Centroid

| | $P_1$ | $P_2$ | $P_3$ |
|---|---|---|---|
| B' = | (1.23 | .55 | -1.16) |

Reproduced Distance
Matrix $\hat{\Delta}_{12}$

| | $P_1$ | $P_2$ | $P_3$ |
|---|---|---|---|
| $S_1$ | 1.54 | 2.23 | 3.93 |
| $S_2$ | .09 | .59 | 2.30 |
| $S_3$ | 2.91 | 2.22 | .52 |
| $S_4$ | 4.09 | 3.41 | 1.70 |

Residual Matrix
$E = \Delta_{12} - \hat{\Delta}_{12}$

| | $P_1$ | $P_2$ | $P_3$ |
|---|---|---|---|
| $S_1$ | -.54 | .77 | .07 |
| $S_2$ | 1.91 | .41 | .70 |
| $S_3$ | .09 | -.22 | .48 |
| $S_4$ | -.09 | .59 | .30 |

Comparison Between Metric Input and Non-Metric Output

| (Centroid) | Coordinates | |
|---|---|---|
| | U | V |
| | Metric Input | Non-Metric Output |
| $P_1$ | 4.29 | 1.23 |
| $S_1$ | 3.29 | 2.78 |
| $S_2$ | 1.29 | 1.14 |
| $P_2$ | .29 | .55 |
| $S_3$ | -1.71 | -1.68 |
| $P_3$ | -2.71 | -1.16 |
| $S_4$ | -4.71 | -2.89 |

Rho (U,V) = .93

| | Distances | | |
|---|---|---|---|
| | X | Y | Z |
| $d_{ij}$ | Metric Input | Non-Metric Input | Non-Metric Output |
| $d_{41}$ | 9 | 4 | 4.09 |
| $d_{13}$ | 6 | 4 | 3.93 |
| $d_{31}$ | 6 | 3 | 2.91 |
| $d_{42}$ | 5 | 4 | 3.41 |
| $d_{23}$ | 4 | 3 | 2.30 |
| $d_{12}$ | 3 | 3 | 2.23 |
| $d_{21}$ | 3 | 2 | .09 |
| $d_{32}$ | 2 | 2 | 2.22 |
| $d_{43}$ | 2 | 2 | 1.70 |
| $d_{11}$ | 1 | 1 | 1.54 |
| $d_{22}$ | 1 | 1 | .59 |
| $d_{33}$ | 1 | 1 | .52 |

Rho(X,Y)=.93 Rho(X,Z)=.84 Rho(Y,Z)=.91

course not conclusive, they do seem to suggest that further study of the possible non-metric use of the present metric unfolding method may be worthwhile.

Our final example (Table 3) illustrates the general, multidimensional case for fallible data. There are eight elements of one set and five of another. The magnitude of the Eckart–Young roots $D_c$ indicates that a fit in two dimensions may be feasible. The coefficient matrix $K_0$ of the linear system (not given) is now $7 \times 5$, *i.e.*, there are two more stimuli in the larger set than the necessary minimum five. Despite this redundancy, the smallest latent root of $K_0'K_0$ (.0234) is already fairly close to zero. In another two-dimensional example (not given) with only six elements in the larger set the smallest root dropped to .0009 for exact data, and to less than .0005 for fallible data which illustrates the need for caution at this step to ensure a unique solution for the unknowns. In the present instance the solution for the five unknowns $m_{11}$, $2m_{12}$, $m_{22}$, and $x_1$, $x_2$ is unique and the resulting $2 \times 2$ matrix $M$ positive definite, so that no least squares adjustments of its roots became necessary. The two matrices of coordinates $A$ $(8 \times 2)$ and $B$ $(5 \times 2)$ are expressed relative to their joint centroid (so that their columns, if summed over all 13 rows, sum to zero). These coordinate matrices are defined up to a common rotation by an orthogonal $2 \times 2$ matrix. The rows of $A$ and $B$ were then used to recompute an $8 \times 5$ matrix of distances $\hat{\Delta}_{12}$ which is compared with the matrix of input distances $\Delta_{12}$ in $E = \Delta_{12} - \hat{\Delta}_{12}$. As can be seen from Table 3 these residuals are fairly small so that the

Table 3

General Case. Fallible Data, m=2, p=8, q=5

Distance Matrix $\Delta_{12}$ Used for Input

| | $P_1$ | $P_2$ | $P_3$ | $P_4$ | $P_5$ |
|---|---|---|---|---|---|
| $S_1$ | 4.0 | 3.0 | 3.0 | 4.0 | 6.0 |
| $S_2$ | 3.0 | 9.0 | 9.0 | 2.0 | 4.0 |
| $S_3$ | 3.0 | 5.0 | 6.0 | 4.0 | 2.0 |
| $S_4$ | 2.0 | 4.0 | 4.0 | 3.0 | 3.0 |
| $S_5$ | 4.0 | 2.0 | 3.0 | 5.0 | 5.0 |
| $S_6$ | 4.0 | 1.0 | 2.0 | 5.0 | 5.0 |
| $S_7$ | 3.0 | 9.0 | 9.0 | 3.0 | 2.0 |
| $S_8$ | 4.0 | 6.0 | 7.0 | 4.0 | 2.0 |

Matrix of Quasi-Scalar Products $C_{12}$

| | $P_1$ | $P_2$ | $P_3$ | $P_4$ | $P_5$ |
|---|---|---|---|---|---|
| $S_1$ | -4.41 | 8.96 | 10.96 | -2.85 | -12.66 |
| $S_2$ | 9.59 | -16.54 | -14.54 | 13.65 | 7.84 |
| $S_3$ | -.51 | 1.36 | -2.14 | -2.45 | 3.74 |
| $S_4$ | -1.61 | 2.26 | 4.26 | -2.55 | -2.36 |
| $S_5$ | -5.11 | 10.76 | 10.26 | -8.05 | -7.86 |
| $S_6$ | -5.91 | 11.46 | 11.96 | -8.85 | -8.66 |
| $S_7$ | 8.89 | -17.24 | -15.24 | 10.45 | 13.14 |
| $S_8$ | -.91 | -1.04 | -5.54 | .65 | 6.84 |

Eckart-Young roots $D_c$ of $C_{12}$

$D_c$ = (54.22 11.41 3.35 1.56 .00)

Latent roots of $K_0'$ $K_0$

(3.41 1.07 .23 .06 .02)

Coordinates A Relative to Joint Centroid

| | $S_1$ | $S_2$ | $S_3$ | $S_4$ | $S_5$ | $S_6$ | $S_7$ | $S_8$ |
|---|---|---|---|---|---|---|---|---|
| $A'=$ | 2.00 | -4.47 | -.31 | .46 | 2.35 | 2.65 | -4.50 | -1.05 |
| | 2.70 | 1.97 | -2.52 | -.12 | -.45 | -.37 | -.56 | -2.91 |

Coordinates B Relative to Joint Centroid

| | $P_1$ | $P_2$ | $P_3$ | $P_4$ | $P_5$ |
|---|---|---|---|---|---|
| $B'=$ | -1.51 | 4.54 | 4.42 | -2.18 | -2.41 |
| | .86 | .12 | 1.27 | 1.41 | -1.41 |

Table 3 Continued

| | Reproduced Distance Matrix $\hat{\Delta}_{12}$ | | | | | Residual Matrix $E = \Delta_{12} - \hat{\Delta}_{12}$ | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| | $P_1$ | $P_2$ | $P_3$ | $P_4$ | $P_5$ | $P_1$ | $P_2$ | $P_3$ | $P_4$ | $P_5$ |
| $S_1$ | 3.96 | 3.62 | 2.81 | 4.37 | 6.02 | .04 | -.62 | .19 | -.37 | -.02 |
| $S_2$ | 3.17 | 9.20 | 8.92 | 2.36 | 3.96 | -.17 | -.20 | .18 | -.36 | .04 |
| $S_3$ | 3.59 | 5.52 | 6.06 | 4.35 | 2.37 | -.59 | -.52 | -.06 | -.35 | -.37 |
| $S_4$ | 2.20 | 4.09 | 4.19 | 3.05 | 3.14 | -.20 | -.09 | -.19 | -.05 | -.14 |
| $S_5$ | 4.07 | 2.26 | 2.69 | 4.89 | 4.85 | -.07 | -.26 | .34 | .11 | .15 |
| $S_6$ | 4.34 | 1.96 | 2.41 | 5.14 | 5.16 | -.34 | -.96 | -.41 | -.14 | -.16 |
| $S_7$ | 3.31 | 9.06 | 9.10 | 3.04 | 2.26 | -.31 | -.06 | -.10 | -.04 | -.26 |
| $S_8$ | 3.80 | 6.36 | 6.88 | 4.46 | 2.02 | -.20 | -.36 | .88 | -.46 | -.02 |

fit in two dimensions can be judged satisfactory. There is a noticeable bias towards negative discrepancies in $E$ for which we cannot offer an explanation.

In conclusion, it should be pointed out that the present method fails if one set of elements is contained entirely within a proper subspace of the other, *e.g.*, if the locus of one set of points is a straight line contained in a plane defined by the other set of points, and if the data are exact. Such configurations are not likely to arise with real data but the reader should be aware of this limitation of the present method.

Our conjecture is that such configurations call for an analysis within the larger space (*i.e.*, in two dimensions, in the example) while the rank of $C_{12}$ is bounded by the dimensionality of the smaller space (*i.e.*, one, in the example). Therefore, only a subset of the necessary Eckart–Young vectors (one such vector pair, in the example) would be uniquely defined, while the rest would be defined only up to a rotation. This indeterminacy (of the second columns in $G$, $H$, in the example) would carry into a corresponding indeterminacy of the columns of the coefficient matrix $K_0$ , and render the present approach inapplicable.

Reviewers wondered about the relation of the present solution to other individual difference models in multidimensional scaling, especially the metric procedure proposed by Bloxom [1968] and Carroll and Chang [1970] and the non-metric procedures described by Lingoes [in press], Kruskal [1968], and others. It should be clear that our present solution is nothing more or less than an algebraic solution of a geometric problem: to locate two sets of points in a joint space (by assigning coordinates to them) given the Euclidean distances between elements from distinct sets. As such it bears no direct relationship to either the non-metric unfolding techniques (which, by definition, do not require knowledge of Euclidean distances) nor to the model proposed by Bloxom [1968] and Carroll and Chang [1970] (which postulates a different space for each individual and which does require knowledge of all distances within one set—usually the stimuli).

Earlier work by Carroll and Chang [1967] does seem to include the metric unfolding paradigm as a special case. The difference is that knowledge about the stimulus space is obtained by conventional multidimensional scaling techniques (which require all distances within one set) which then is used to determine subject specific scale values. In the present solution these assignments are made simultaneously in the spirit of Coombs' original formulation of the unfolding paradigm.

As to the choice between metric versus non-metric techniques, we feel that the present emphasis on non-metric techniques, on the presumed strength of their greater generality, may well be ill-advised in the long run, at least as long as it is not realized that the input information has to be justified as meaningful in both cases. The need for such a justification is perhaps more obvious with metric techniques. This may well prove a blessing in disguise, precisely because it limits their use. For an excellent discussion of this problem see Krantz [1967].

## REFERENCES

Bennett, J. F. Determination of the number of independent parameters of a score matrix from the examination of rank orders. *Psychometrika*, 1956, **21**, 383–393.

Bennett, J. F., & Hays, W. L. Multidimensional unfolding: Determining the dimensionality of ranked preference data. *Psychometrika*, 1960, **25**, 27–43.

Bloxom, B. Individual differences in multidimensional scaling. ETS Research Bulletin 68–45. Princeton: Educational Testing Service, 1968.

Carroll, J. D., & Chang, Jih-Jie. Relating preference data to multidimensional scaling solutions via a generalization of Coombs' unfolding model. Paper read at the Spring Meeting of the Psychometric Society, Madison, Wisconsin, 1967.

Carroll, J. D., & Chang, Jih-Jie. Analysis of individual differences in multidimensional scaling via an N-way generalization of "Eckart–Young" decomposition. *Psychometrika*, 1970, **35**, 283–319.

Coombs, C. H. Inconsistencies of preferences in psychological measurement. *Journal of Experimental Psychology*, 1958, **55**, 1–7.

Coombs, C. H. *A theory of data*. New York: Wiley, 1964.

Coombs, C. H., & Kao, R. C. Nonmetric factor analysis. Department of Engineering Research Bulletin No. 38. Ann Arbor: University of Michigan, 1955.

Eckart, C., & Young, G. The approximation of one matrix by another of lower rank. *Psychometrika*, 1936, **1**, 211–218.

Goode, F. M. Interval scale representation of ordered metric scales. Dittoed manuscript, University of Michigan, 1957.

Hays, W. L., & Bennett, J. F. Multidimensional unfolding: Determining configuration from complete rank order preference data. *Psychometrika*, 1961, **26**, 221–238.

Krantz, D. H. Rational distances functions for multidimensional scaling. *Journal of Mathematical Psychology*, 1967, **4**, 226–244.

Kruskal, J. B. How to use MDSCAL, a program to do multidimensional scaling and multidimensional unfolding. Unpublished report, Bell Telephone Laboratories, 1968.

Lingoes, J. C. A general nonparametric model for representing objects and attributes in a joint space. In J. C. Gardin (Ed.), *Les Comptes-rendus de Colloque International sur L'emploi des Calculateurs en Archeologie: Problèmes Sémiologiques et Mathématiques.* Marseilles: Centre Nationale de la Recherche Scientifiques, in press.

Ross, J., & Cliff, N. A generalization of the interpoint distance model. *Psychometrika,* 1964, **29**, 167–176.

Schönemann, P. H. Fitting a simplex symmetrically. *Psychometrika,* 1970, **35**, 1–21.

*Manuscript received 6/25/69*
*Revised manuscript received 11/18/69*