PSYCHOMETRIKA—VOL. 58, NO. 1, 101–114
MARCH 1993

ANALYSIS OF ASYMMETRY BY A SLIDE-VECTOR

BERRIE ZIELMAN AND WILLEM J. HEISER

DEPARTMENT OF DATA THEORY
UNIVERSITY OF LEIDEN

The slide-vector scaling model attempts to account for the asymmetry of a proximity matrix by a uniform shift in a fixed direction imposed on a symmetric Euclidean representation of the scaled objects. Although no method for fitting the slide-vector model seems available in the literature, the model can be viewed as a constrained version of the unfolding model, which does suggest one possible algorithm. The slide-vector model is generalized to handle three-way data, and two examples from market structure analysis are presented.

Key words: slide-vector model, unfolding, constrained multidimensional scaling, asymmetry.

Introduction

Broadly defined, the term multidimensional scaling (MDS) indicates a class of techniques that represent dissimilarity measures between n objects by distances between points in a possibly low dimensional space. The objects can be anything: cars, personality traits, ethnic categories, and so on. There are various methods to obtain dissimilarity measures (e.g., Coxon, 1982); some methods of data collection yield *asymmetric* dissimilarity measures. It can be argued that a distance model is not appropriate for the analysis of square asymmetric data tables, because by definition a distance function is *symmetric*. Approximating an asymmetric matrix by a symmetric model makes perfect fit of the symmetric model generally impossible. A number of methods that are proposed to represent asymmetries can viewed as special cases of a general model proposed by Holman (1979), which is called a similarity-bias model by Nosofsky (1991). The similarity-bias model has symmetric parameters for the similarity between objects, and additional row and column parameters, frequently interpreted as bias parameters, to accommodate asymmetry. Examples of asymmetric models that can be interpreted as similarity-bias models have been proposed by, for instance, Weeks and Bentler (1982), Okada (1988a, 1988b), Krumhansl (1978), Tversky (1977), and Caussinus (1965). A review of the similarity bias model and a discussion of some special cases can be found in Nosofsky (1991). Multidimensional models for analyzing asymmetry have been proposed by Gower (1977), Constantine and Gower (1978), and Harshman, Green, Wind, and Lundy (1982).

Examples of asymmetric matrices that might lead to interesting results when subjected to a scaling analysis are occupational mobility tables, brand switching data, sociometric interaction data and communication and volume flows. The dissimilarity measure between object i and j will be denoted by δ_{ij}, ($i = 1, \ldots, n; j = 1, \ldots, n$). An object i is referred to as dominating object j if it is observed that $\delta_{ij} > \delta_{ji}$.

A relatively unknown asymmetric adjustment of the distance model, suggested by Kruskal in 1973 (see de Leeuw & Heiser, 1982), will be discussed in this paper. The

The authors wish to thank Ivo van der Lans, John Gower, and the Editor for their comments on an earlier version of this manuscript.

Requests for reprints should be sent to Willem J. Heiser, Department of Data Theory, University of Leiden, PO Box 9555, 2300 RB Leiden, THE NETHERLANDS.

0033-3123/93/0300-91382$00.75/0

model is mentioned in de Leeuw and Heiser as the *slide-vector model* and in Carroll and Wish (1974) as the *drift vector model*. These two names refer to the same model; this paper uses the name slide-vector model. The slide-vector model represents asymmetry by a uniform shift, or translation, of the difference vector between the points in a multidimensional space.

In this paper an algorithm based on the majorization theory of de Leeuw and Heiser (1980) is presented for fitting the slide-vector model. In addition, a three-way variant will be proposed.

Modeling Asymmetry By a Slide-Vector

The slide-vector scaling model (Kruskal, 1973, referenced by de Leeuw & Heiser, 1982) attempts to account for asymmetry in an observed proximity matrix by a uniform shift into one direction imposed on a symmetric Euclidean representation for the scaled objects. Explicitly, the model is:

$$d_{ij}(X; \mathbf{z}) = \left\{ \sum_s (x_{is} - x_{js} + z_s)^2 \right\}^{1/2}. \tag{1}$$

The asymmetric quasi-distance is a function of the configuration matrix X, with coordinates x_{is} $(i = 1, \ldots, n; s = 1, \ldots, p)$ and the slide-vector $\mathbf{z}$, with coordinates z_s. The parameter p denotes the dimensionality of the model. The model contains the simple Euclidean model as a special case if the slide-vector is equal to zero. Note that the diagonal elements of the model are non-zero, unlike a regular distance model. Furthermore, if all coordinates for two objects coincide, the model predicts a nonzero symmetric dissimilarity.

Inserting $x_{js} - z_s = y_{js}$ into (1) yields:

$$d_{ij}(X; Y) = \left\{ \sum_s (x_{is} - y_{js})^2 \right\}^{1/2}. \tag{2}$$

Distance formula (2) is used in multidimensional unfolding (MDU). From this equation it follows that the slide-vector model is a special case of the unfolding model, where the configuration matrix Y for columns is a translation of the configuration matrix X for rows. In general, the unfolding model merely assumes that there is a configuration for the rows and a configuration for the columns and these two configurations need not be a function of each other. In the context of analyzing square asymmetric matrices, the unfolding method can be interpreted as a column-specific slide-vector model where every column point $\mathbf{y}_j$ can be decomposed as the sum of the row point $\mathbf{x}_j$ and a slide-vector $\mathbf{z}_j$.

If multiple tables are available for analysis (for example, data may have been collected for different subjects, under different conditions or at different points in time), a three-way generalization of the slide-vector model that we wish to propose is:

$$d_{ij}(X; U_k; \mathbf{z}) = \left\{ \sum_s u_{ks}(x_{is} - x_{js} + z_s)^2 \right\}^{1/2}, \tag{3}$$

where u_{ks} is a weight indicating the relative importance of dimension s for the k-th data source $(k = 1, \ldots, m)$. The configuration matrix X is usually called the *common space*. The model assumes a set of p dimensions underlying the n objects in the same way as in the simple slide-vector model, and assumed to be common to all sources;

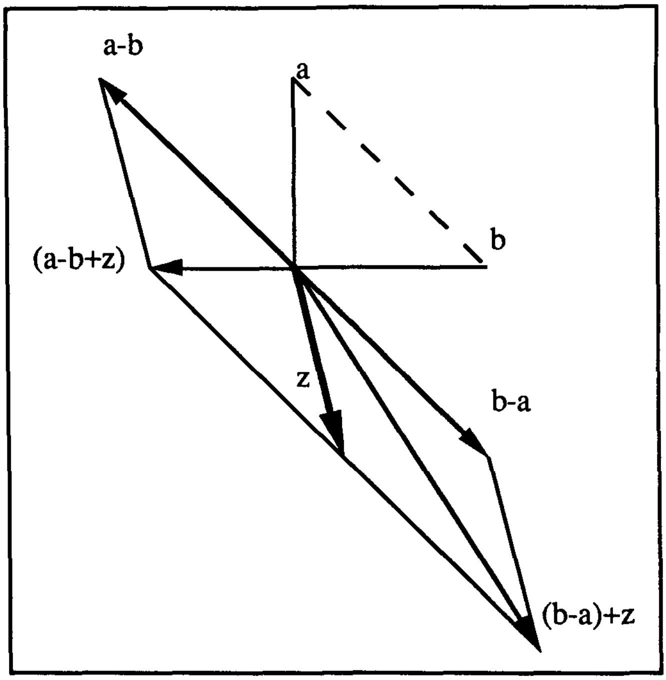


FIGURE 1.
Geometry of the slide-vector model.

hence, the name common space. Formula (3) differs from (1) only in the presence of the dimension weight u_{ks}, which can be interpreted analogously to the weights from the INDSCAL model (Carroll, 1972; Carroll & Chang, 1970) for symmetric dissimilarity data.

Geometry of the Slide-Vector Model

An attractive feature of the slide-vector model is that the objects are depicted in a low dimensional space and the slide-vector is incorporated as an additional point in this space. In contrast with other models (Gower, 1977; Okada, 1988a, 1988b; Weeks & Bentler, 1982) the asymmetry parameters are linked to the *dimensions* of the scaling diagram, and not to the objects themselves. The model predicts large asymmetry for dissimilar objects. The slide-vector model can be viewed as a model of a mountain; walking a mountain uphill will take more effort than walking to the valley, although the distance remains the same. This asymmetry in effort increases with the distance we walk. In the psychological literature, Keren and Baggen (1981) reported a negative correlation (−.733) between asymmetry parameters and average similarity. This correlation is in agreement with the postulated relationship between asymmetry and symmetry by the slide-vector scaling model.

The geometry of the slide-vector model is illustrated in Figure 1, where two objects are depicted as vectors in a two-dimensional space; the displacement between the termini of the vectors is drawn by a dashed line. The distance is computed by first subtracting the vectors (a − b; b − a); these difference vectors are of the same length but with opposite sign. By adding the slide-vector, indicated by the bold vector z in the Figure, to these difference vectors, and then establishing their length, an asymmetric quasi-distance d_{ba} and d_{ab} indicated by (b − a) + z and (a − b) + z is obtained. From Figure 1 it follows that objects located in a direction similar to the slide-vector domi-

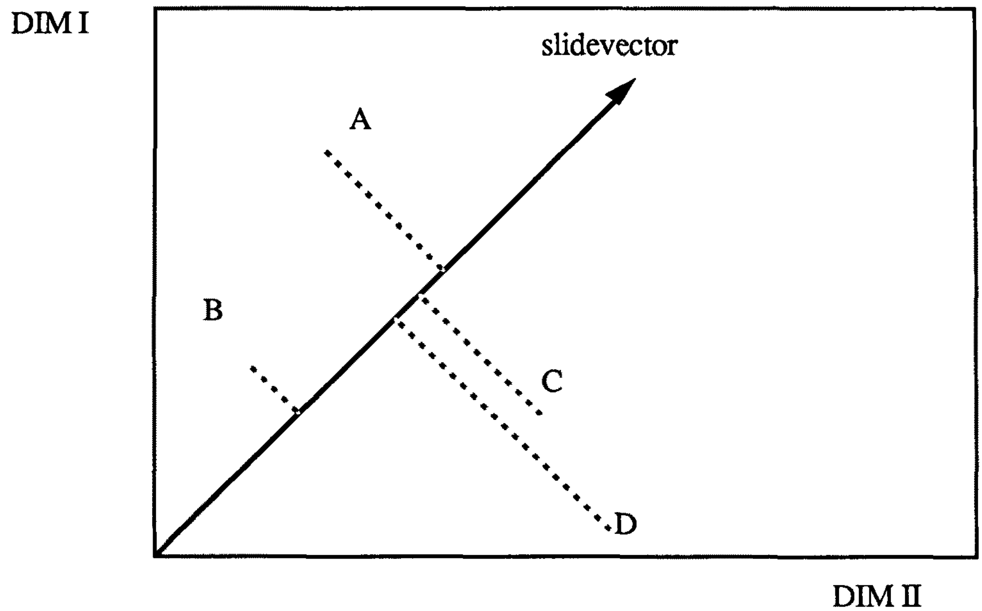


FIGURE 2.
Joint representation of symmetry and asymmetry.

nates the other objects. From the MDU interpretation it follows that the column points are translated in a specific direction.

The interpretation of a joint plot of the configuration and the slide-vector can be simplified by distinguishing a *preference* part indicating the dominant member of a pair of objects and a *symmetric* part indicating similarity. Squaring (1) and expanding yields:

$$d_{ij}^2(X;\ \underline{\mathbf{z}}) = \sum_s (x_{is} - x_{js})^2 + \sum_s z_s^2 + 2 \sum_s z_s(x_{is} - x_{js}). \tag{4}$$

The first and second terms on the right-hand side of (4) are symmetric; the first term corresponds to the squared Euclidean distance indicating the similarity between the objects. The second term indicates the length of the slide-vector, corresponding to the general amount of symmetry explained by the model. The third term on the right-hand-side of (4) is the preference part, and has the property of *skew-symmetry*:

$$\sum_s z_s(x_{is} - x_{js}) = - \sum_s z_s(x_{js} - x_{is}).$$

This skew-symmetric term corresponds to the vector model of preference data (Carroll, 1972). So the relative popularity or dominance of an object with respect to the other objects can be found by the projection of the object on the slide-vector. The vector model interpretation is illustrated in Figure 2, where four objects A, B, C, D are depicted as points in a two-dimensional space. The dashed lines in Figure 2 correspond to the projections of the objects on the slide-vector. Objects with high projections dominate objects with low projections. In this example, object A dominates all others. The distances among the points can be interpreted as the similarity or resemblance of the objects; object C is more similar to object D than to object A. In this small example the quasi-distance from point A to point C is larger than the quasi-distance from point C to point A.

A Restricted Unfolding Algorithm

In this section it will be shown how the coordinates and the slide-vector of the various models can be obtained by using an unfolding algorithm that can handle restrictions on the configuration. A detailed description of this algorithm for unfolding can be found in Heiser (1987). How well a particular configuration matches the data will be measured by the STRESS badness-of-fit index (Kruskal, 1964a, 1964b):

$$\sigma(X; Y) = \sum_i \sum_j w_{ij}(\delta_{ij} - d_{ij}(X; Y))^2. \tag{5}$$

The iterative majorization approach to MDS of de Leeuw and Heiser (1980), leading to the so-called SMACOF algorithm, basically consists of a repetition of two steps: First, improve the locations of the points in the multidimensional space by computing the Guttman transform; and second, take care that these improved locations satisfy the desired model restrictions by solving a metric projection problem. A third step can be added to this two-step algorithm by introducing optimal scaling of the dissimilarities, but this generality will not be discussed in the sequel. The weights w_{ij} in (5) can be used to code missing data, or to simulate the behavior of other loss functions. In our application the weights will be used to suppress the effects of the diagonal elements of the data matrix.

In the unfolding case the Guttman transform can be split up into a Guttman transform for the row points and a Guttman transform for the column points. We first define the auxiliary matrix A with elements:

$$a_{ij} = \frac{w_{ij}\delta_{ij}}{d_{ij}(X; Y)}, \qquad \text{if } d_{ij}(X; Y) > 0,$$
$$a_{ij} = 0, \qquad \text{if } d_{ij}(X; Y) = 0.$$

The weights are collected in a matrix $W = \{w_{ij}\}$, and four diagonal matrices are defined as follows:

$$P = \text{diag}\,(Ae),$$

$$Q = \text{diag}\,(e'A),$$

$$R = \text{diag}\,(We),$$

$$C = \text{diag}\,(e'W),$$

where **e** denotes the n-vector of ones, and diag ($\cdot$) denotes a diagonal matrix having diagonal elements equal to its vector argument. The preliminary matrices X^* and Y^* are computed from previous estimates $\mathbb{X}$ and $\mathbb{Y}$ as follows:

$$X^* = P\mathbb{X} - A\mathbb{Y},$$

$$Y^* = Q\mathbb{Y} - A'\mathbb{X}.$$

The unconstrained updates X^+ and Y^+, of the row and column configurations (the Guttman transforms), can be found by solving the system of equations (Heiser, 1987):

$$RX^+ - WY^+ = X^*,$$

$$CY^+W'X^+ = Y^*.$$

The SMACOF algorithm is quite simple now: as long as the improvement of STRESS exceeds a predetermined small positive constant, set $\mathbb{X}$ and $\mathbb{Y}$ equal to X^+ and Y^+, respectively, and compute new updates. A proof of convergence can be found in de Leeuw and Heiser (1980) and de Leeuw (1988). To incorporate the restrictions imposed by the slide-vector model, a metric projection problem has to be solved. This can be done by introducing the auxiliary matrix E, of order $2n$ by $n + 1$, having the structure:

$$E=\begin{pmatrix} I & e \\ I & \mathbf{0} \end{pmatrix},$$

and defining the matrix S of order $2n$ by p in which the matrices X^+ and Y^+ are stacked on top of each other. The matrix E can be regarded as a design matrix where the first n columns code the equality constraints on the coordinates for the row and column points. The weights are recorded in a new matrix V that is built up from the old matrices; the V matrix is partitioned as:

$$V=\begin{pmatrix} R & -W \\ -W' & C \end{pmatrix}.$$

The metric projection problem can now be formulated as follows: estimate the matrix B that minimizes:

$$L(B)=\text{tr } (S-EB)'V(S-EB). \tag{6}$$

The solution of the metric projection problem (6) is:

$$B=(E'VE)^{-}E'VS,$$

where the matrix $(E'VE)^-$ denotes the generalized inverse of the matrix (E'VE). This inverse can be computed as $\{(E'VE) + \alpha(\alpha'\alpha)^{-1}\alpha'\}^{-1} - \alpha(\alpha'\alpha)^{-1}\alpha'$, where α spans the null-space of the matrix (E'VE); this vector α is an $n + 1$ vector with unities and in the last position a zero. The remaining part of the algorithm is to compute EB, set X equal to the first n rows of this matrix, and Y equal to the last n rows of this matrix. With these constrained X and Y matrices we can compute new Guttman transforms. The slide-vector is given as the last row of the matrix B.

Next it will be shown for the three-way slide-vector model how the INDSCAL and IDIOSCAL model can be estimated with the SMACOF algorithm by solving another metric projection problem. The INDSCAL model allows a differential weighting of the dimensions by different data sources. The IDIOSCAL model allows a differential rotation of the dimensions as well. The model is extended with an additional matrix U_k incorporating the various restrictions required by the individual differences model. The metric projection problem can now be formulated as:

$$L(B; U_k) = \frac{1}{m} \text{tr} \sum_k (S_k - EBU_k)'V_k(S_k - EBU_k), \tag{7}$$

where S_k contains the stacked Guttman transforms for the k-th data source, and V_k denotes the weight matrix for the k-th set of residuals. The matrix product EB is called the common space. This common space satisfies the slide-vector restrictions. The matrix U_k contains the linear transformation of the common space for source k. The matrix U_k can be constrained to be a diagonal matrix in which case we obtain an INDSCAL representation. If this matrix is free, an IDIOSCAL representation is obtained (see Heiser & Stoop, 1986). The loss function in (7) can be minimized by alternating least squares (ALS); for fixed B the solution of U_k is, assuming the inverse exists:

$$U_k = (B'E'V_kEB)^{-1}B'E'V_kS_k.$$

For fixed U_k, the matrix B can be found by solving the system of equations:

$$\sum_k \{U_kU'_k \otimes E'V_kE\} \text{ vec } (B) = \text{vec}\left(\sum_k E'V_kSU_k\right), \quad (8)$$

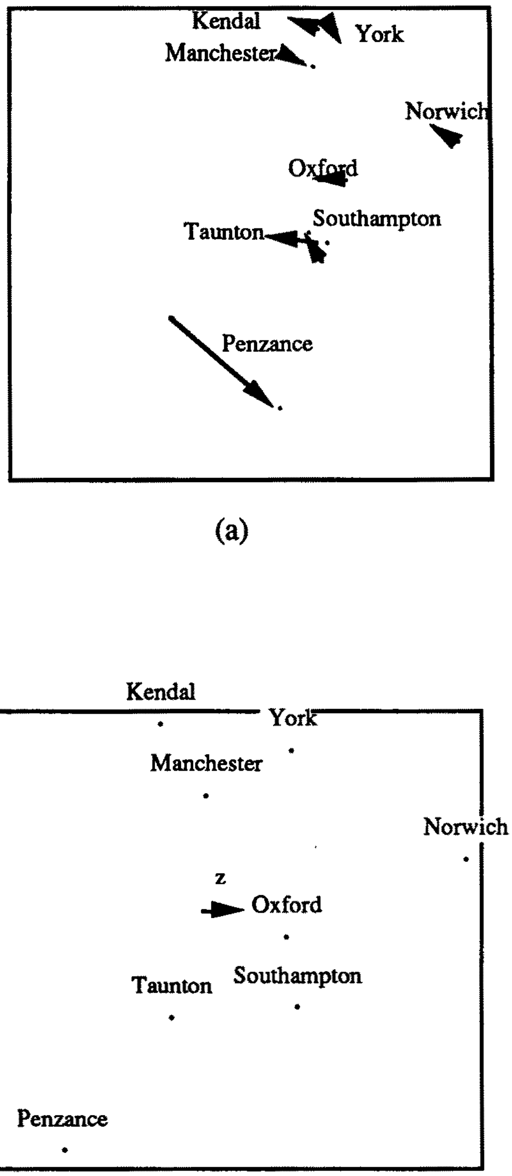
where the symbol $\otimes$ denotes the right Kronecker product of matrices; for instance, $A \otimes B = [a_{ij}B]$, and vec ($\cdot$) denotes the operator that transforms a matrix into a vector by stacking the columns of a matrix one underneath the other. If the individual differences are modeled by weighting the dimensions, the estimation problem of the common space reduces to p smaller subproblems; because of the blockwise structure of the matrix on the left-hand-side of (8), the common space coordinates can be estimated dimension after dimension.

Applications

In this section we will discuss two applications, one that illustrates the potential value of the model over simple unfolding, and one with real data from market structure analysis.

The reader may argue that a slide-vector representation could also be obtained by an unfolding program without restrictions. An example that this can be difficult is given in Constantine and Gower (1978), who used unfolding to analyze a table giving distances between eight English towns. These distances were corrupted by adding a skew-symmetric term $x_i - x_j$ to these true distances, where the values of x were the coordinates of the towns on the east-west axis. Thus a distance becomes asymmetric, and the asymmetry is related to the east-west dimension. The diagram obtained from the unfolding analysis, reported in Constantine and Gower, shows that although the model fitted well in terms of STRESS, the translations of the individual points varied not only in magnitude, but points were also translated in different directions. This effect is shown in Figure 3a, where the translations of column points to row points have been marked by arrows. This result is rather surprising because we know in advance that the skew-symmetric component is related to the dimensions. Analysis by the slide-vector model (See Figure 3b) revealed the correct direction of asymmetry and the true configuration of points. This masking of the true direction of asymmetry in the unconstrained unfolding analysis is probably due to the large number of parameters.

The second example uses brandswitching data. This type of data can be used to study competition between brands and may say something about the structure of the market. The aim of market structure analysis is to partition brands into submarkets with close competition. The data of this example are taken from a study by Harshman et al. (1982), and deal with the phenomenon of brand switching between 16 car segments. The entries in the table indicate the number of people who bought brand i in the first period and currently buy brand j in the second. Brand switching data are likely to be nonsymmetric; usually a different proportion of brand A users switch to brand B than vice versa within a given time period. A detailed description of the data can be found in Harshman. Under the present model the brands are positioned in a multidimensional space where brands located close together compete more with each other than with brands far apart. The slide-vector represents the direction in the space where the most



(a)

(b)

FIGURE 3.
(a) Unfolding configuration of England; (b) slide-vector configuration of England.

popular brands, in the sense of more "switched to" than "switched from" can be found.

The raw brand switching matrix is not appropriate as input for the scaling program. The matrix has to be adjusted for differences in market share and the data have to be converted from similarities to dissimilarities. These two steps can be performed by applying the gravity model (Tobler & Wineburg, 1971), which amounts to first dividing the raw frequencies n_{ij} by their row and column sum; and secondly, inverting the standardized frequencies. These inverted numbers yield *squared* distances according to the gravity model, so as a last step the square root of the quantities is taken. Another problem with these sets of data is that the diagonal (containing the loyal non-switchers) tends to dominate the analysis; in our example these numbers will influence the length of the slide-vector. This effect is eliminated from the analysis by giving these diagonal elements zero weights; in other words, we ignore brand loyalty and focus on the switching segment.

The transformed car switching matrix was analyzed in two dimensions. The coordinates are reported in Table 1 and graphically displayed in Figure 4a. The first three letters of the labels indicated the size of the cars: sub = subcompact; sma = small; com = compact; mid = midsize; std = standard; lux = luxury. The last letter indicates a distinguishing feature within a size category: c = captive import; d = domestic; i = import; s = specialty; m = medium price; l = low price. The first dimension corresponds to an import-domestic distinction and the second dimension discriminates the small cars from the large cars. The fit of the two-dimensional model is quite good (STRESS = 0.066).

A general result from linear algebra is that we may decompose any asymmetric matrix $\underline{\Delta}$ into a symmetric part, by averaging across the diagonal, and a skew-symmetric part, which can be obtained by subtracting the corresponding elements across the diagonal. Thus one obtains

$$\Delta = S + A,$$

where S is a symmetric matrix of averages $s_{ij} = \{\delta_{ij} + \delta_{ji}\}/2$, and A a skew-symmetric matrix with elements $a_{ij} = \{\delta_{ij} - \delta_{ji}\}/2$. The matrix A describes the departures from symmetry, and can be viewed as the preference or dominance part of an asymmetric matrix, since if $\delta_{ij} > \delta_{ji}$, then $a_{ij} > 0$. The matrices A and S are orthogonal; therefore, the sum of squares of the matrix Δ can be decomposed into a sum of squares due to symmetry and a sum of squares due to skew-symmetry. If one applies this decomposition to the matrix of residuals, it is possible to decompose the STRESS into a symmetric and a skew-symmetric part. First define:

$$f_{ij} = (d_{ij}(X, Y) - d_{ji}(X; Y))/2;$$

$$g_{ij} = (d_{ij}(X; Y) + d_{ji}(X; Y))/2,$$

$$h_{ij} = \delta_{ij} - f_{ij}.$$

The STRESS can now be written in terms of these new quantities as

$$\sigma(X; Y) = \underline{\sum_i \sum_j} w_{ij}(\delta_{ij} - g_{ij} - f_{ij})^2 = \underline{\sum_i} \sum_j w_{ij}(h_{ij} - g_{ij})^2,$$

which can be decomposed into

$$\sigma(X; Y) = \sum_i \sum_j w_{ij}(h_{ij} - \underline{\hat{h}_{ij})^2 + \sum_i} \sum_j \underline{w}_{ij}(\hat{h}_{ij} - g_{ij})^2, \quad (9)$$

where $\underline{w_{ij}}\hat{h}_{ij} = (w_{ij}h_{ij} + w_{ji}h_{ji})$ and $\underline{w}_{ij} = w_{ij} + w_{ji}$. The first term at right-hand-side of (9) denotes the STRESS due to skew-symmetry and the last term denotes the STRESS due to symmetry. To arrive at (9) we could also have started by defining $h_{ij} = \delta_{ij} - g_{ij}$, and this yields the *same* decomposition of the STRESS if the weight matrix is *symmetric*; in the case of asymmetry in the weight matrix, the decompositions differ.

Using this decomposition, the STRESS due to symmetry was 0.035, and the STRESS due to skew-symmetry was 0.031. The skew symmetric variation in the data was 5%; thus, the model predicts the symmetry relatively better than the skew-sym-

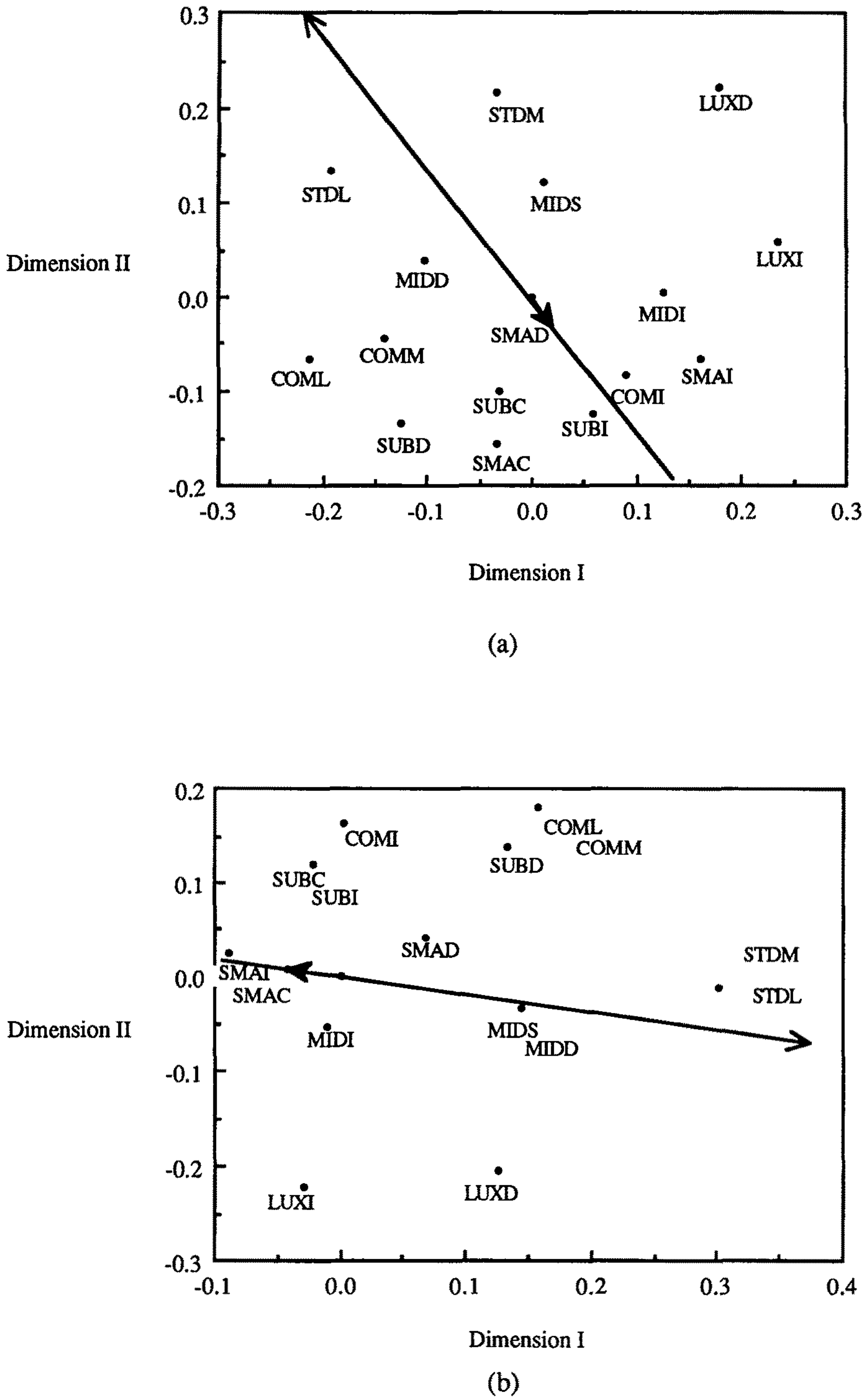


FIGURE 4.
(a) Two-dimensional solution for the car switching data; (b) constrained solution for the car switching data.

metry but this effect is probably due to the relatively small amount of skew-symmetric variation compared to the symmetric part.

From Figure 3 it is seen that the imported cars compete more with each other than

Table 1
Configuration Matrices for the Car Switching Data

	Unconstrained		Constrained	
Segment	Dim I	Dim II	Dim I	Dim II
SUBD	-0.125	-0.135	0.134	0.138
SUBC	-0.031	-0.101	-0.021	0.120
SUBI	0.058	-0.125	-0.021	0.120
SMAD	0.012	-0.022	0.067	0.042
SMAC	-0.034	-0.156	-0.088	0.024
SMAI	0.162	-0.065	-0.088	0.024
COML	-0.212	-0.067	0.158	0.181
COMM	-0.140	-0.045	0.158	0.181
COMI	0.090	-0.082	0.003	0.163
MIDD	-0.103	0.039	0.144	-0.035
MIDI	0.125	0.005	-0.011	-0.053
MIDS	0.011	0.123	0.144	-0.035
STDL	-0.192	0.134	0.301	-0.012
STDM	-0.033	0.216	0.301	-0.012
LUXD	0.178	0.221	0.125	-0.204
LUXI	0.235	0.059	-0.030	-0.223
slidevector	0.015	-0.023	-0.042	0.007

with the domestic cars. There is also strong competition among small cars. The slide-vector is depicted as an arrow with a bold tail. The long vector with the plain tail is the "switch from" vector. This arrow is drawn larger to facilitate the interpretation. For instance, there are more switches from Luxury Domestic to Luxury Import than the other way round. The trend is to buy a luxury imported car instead of a luxury domestic car. There are also more switches from the standard and midsize cars to the small and subcompact cars than the other way round; the trend is that the small cars are winning in the market. These results are in close agreement with the results found by Harshman et al. (1982) who analyzed these data with the DEDICOM model.

Given the squared version of the model in (4) and the above described decomposition of the data into a symmetric matrix and a skew-symmetric matrix, it is possible to fit the symmetric and skew-symmetric component to the data separately and to investigate afterwards how well the skew-symmetric part can be modeled as a function of the dimensions in a scaling configuration. Using the SPSS-X release of ALSCAL, the symmetric part of the data was scaled in two dimensions. A linear skew-symmetry model was fitted to the skew-symmetric part of the data by calculating the means of the rows of the skew-symmetric matrix. The linear model explained 43% of the variation in the skew-symmetric part of the data; thus, the model did not fit very well. By regressing the row-means on the dimensions, an approximate slide-vector model was obtained; the squared multiple correlation between the dimensions and the row-means was .25, giving only 10% of the skew-symmetric variation that is related to the dimensions. The slide-vector model explained 38% of the variation in the skew-symmetric part of the data, which comes very close to the unconstrained linear model and is considerably better than the post-analysis regression method.

As a second step the data will be re-analyzed by a slightly more constrained version of the slide-vector model. For a discussion of constrained MDS, see Heiser and Meulman (1983a, 1983b). The coordinates of the objects in the space are now required to be a linear combination of external categorical variables, or product-related attributes. Within the SMACOF theory this is a relatively simple task; the first q columns, where q is the total number of categories of the variables, of the matrix E are replaced by vectors with unities and zero's. A one indicates the presence of the attribute and a zero indicates the absence of the attribute; see Meulman and Heiser (1984) for details. In the car switching data the external variables are the size of the cars, with categories subcompact, small, compact, midsize, standard, and luxury. The second variable is import, with two categories of import and domestic. Captive imports are treated as imported cars. The variables are coded as an ANOVA-type design matrix.

The STRESS of the two dimensional constrained solution is 0.11; this is only slightly larger than the unconstrained solution. The STRESS due to symmetry was 0.079; the STRESS due to asymmetry was 0.031. The constrained solution is displayed in Figure 3b. The horizontal dimension is related to the size of the cars; the vertical dimension can be interpreted as a luxury dimension. The imported cars are located in the left quadrants of the space. The trend is to buy small cars in 1979, and within a class of cars of the same size, there are more switches from domestic cars to imported cars than conversely.

Discussion

The perspective of constrained unfolding allows us to think of other generalizations of the slide-vector model; for instance, one in which there are multiple slide-vectors and some points share the same slide-vector. The last column of E codes the slide-vector; if multiple slide-vectors would be desired, the matrix E has to be augmented with additional columns. This model will be called the *multiple slide-vector model*, and written as

$$d_{ij}(X; Z) = \left\{ \sum_s (x_{is} - x_{js} + z_{ks})^2 \right\}^{1/2},$$

where k denotes a set of objects sharing the same slide-vector and Z is the matrix with coordinate values z_{ks}, with $(k = 1, \ldots, K)$. These sets of objects can be defined by the user, or estimated from the data. As examples of what these sets or classes might be, we would have a distinction between diet and non-diet soft drinks, import and domestic cars in the analysis of brand switching data, or a distinction between "hard" and "soft" psychology journals in the analysis of scientometric transaction matrices. If these sets are defined by the users, no new complications arise; if they have to be estimated from the data, one has to solve a combinatorial problem. In general the computational burden can be formidable, but a practical solution is possible by using a K-means type of re-allocation strategy. The procedure could be initialized by means of a K-group clustering procedure on the individual slide-vectors $\mathbf{z}_j$ that can be constructed from an unconstrained unfolding solution. Since this presumably yields a good starting partition, the re-allocation strategy is expected to be efficient (Milligan, 1980).

Another generalization of the slide-vector would be the *row-weighted slide-vector model*, which can be written as

$$d_{ij}(X; \mathbf{z}; \mathbf{a}) = \left\{ \sum_s (x_{is} - x_{js} + z_s \, a_j)^2 \right\}^{1/2} .$$

Here, a_j is a weight indicating the importance of the slide-vector for modeling the asymmetry in dissimilarity between object i to object j. A reason for using this general model would be to model the diagonal entries, rather than ignoring them as we have done here. Within the SMACOF theory the estimation of such models does not lead to new complications. These generalizations of the slide-vector model seems to be worth investigating further.

A related model is the Jet-stream model proposed by Gower (1977), which represents the asymmetry in the data by a direction in the space. The model differs from the slide-vector model in the sense that it divides the distance by an asymmetric term. Also, the slide-vector model resembles the wandering vector model of Carroll (1981) for paired comparison data, although our model is additive instead of multiplicative and not appropriate for paired comparison data. The wandering vector model is designed for skew-symmetric data and models dominance or preference as a direction in a multidimensional space.

References

Carroll, J. D. (1972). Individual differences and multidimensional scaling. In R. N. Shepard, A. K. Romney, & S. B. Nerlove (Eds.), *Multidimensional scaling: Theory and applications in the behavioral sciences* (pp. 105–155). New York: Seminar Press.

Carroll, J. D. (1981). Models and methods for multidimensional analysis of preferential choice (or other dominance) data. In E. D. Lantermann & H. Feger (Eds.), *Similarity and choice* (pp. 234–289). Bern: Huber.

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

Carroll, J. D., & Wish, M. (1974). Multidimensional perceptual models and measurement methods. In E. C. Carterette & M. P. Friedman (Eds.), *Handbook of perception* (pp. 391–447). New York: Academic Press.

Caussinus, H. (1965). Contribution a l'analyse statistique des tableaux de correlation. [Contributions to the statistical analysis of correlation matrices]. *Annals of the Faculty of Science, University of Toulouse, 29*, 77–182.

Constantine, A. G., & Gower, J. C. (1978). Graphical representation of asymmetric matrices. *Journal of the Royal Statistical Society, Series C, 27*, 297–304.

Coxon, A. P. M. (1982). *The user's guide to multidimensional scaling.* London: Heinemann Educational Books.

de Leeuw, J. (1988). Convergence of the majorization method for multidimensional scaling. *Journal of Classification, 5*, 163–180.

de Leeuw, J., & Heiser, W. (1980). Multidimensional scaling with restrictions on the configuration. In P. R. Krishnaiah (Ed.), *Multivariate analysis-V* (pp. 501–522). Amsterdam: North Holland.

de Leeuw, J., & Heiser, W. J. (1982). Theory of multidimensional scaling. In P. R. Krishnaiah & L. N. Kanal (Eds.), *Handbook of statistics, Vol. 2* (pp. 285–316). Amsterdam: North Holland.

Gower, J. C. (1977). The analysis of asymmetry and orthogonality. In J. R. Barra, F. Brodeau, G. Romer, & B. van Cutsem (Eds.), *Recent developments in statistics* (pp. 109–123). Amsterdam: North Holland.

Harshman, R. A., Green, P. E., Wind, Y., & Lundy, M. E. (1982). A model for the analysis of asymmetric data in marketing research. *Marketing Science, 1*, 204–242.

Heiser, W. J. (1987). Joint ordination of species and sites: the unfolding technique. In P. Legendre & L. Legendre (Eds.), *Developments in numerical ecology* (pp. 189–221). Berlin: Springer Verlag.

Heiser, W. J., & Meulman, J. (1983a). Analyzing rectangular tables by joint and constrained multidimensional scaling. *Journal of Econometrics, 22*, 139–167.

Heiser, W. J., & Meulman, J. (1983b). Constrained multidimensional scaling, including confirmation. *Applied Psychological Measurement, 7*, 381–404.

Heiser, W. J., & Stoop, I. (1986). *Explicit SMACOF algorithms for individual differences scaling* (Research Report RR-86-14). Leiden: University of Leiden, Department of Data Theory.

Holman, E. W. (1979). Monotonic models for asymmetric proximities. *Journal of Mathematical Psychology*, *20*, 1–15.

Keeren, G., & Baggen, S. (1981). Recognition models of alphanumeric characters. *Perception and Psychophysics*, *29*, 234–246.

Krumhansl, C. L. (1978). Concerning the applicability of geometric models to similarity data: The interrelationship between similarity and spatial density. *Psychological Review*, *84*, 445–463.

Kruskal, J. B. (1964a). Multidimensional scaling by optimizing goodness of fit to a nonmetric hypothesis. *Psychometrika*, *29*, 1–27.

Kruskal, J. B. (1964b). Nonmetric multidimensional scaling: A numerical method. *Psychometrika*, *29*, 115–129.

Meulman, J., & Heiser, W. J. (1984). Constrained multidimensional scaling: More directions than dimensions. In T. Havránek, Z. Šidák, & M. Novák (Eds.), *Compstat 1984, Proceedings in computational statistics* (pp. 51–56). Vienna: Physica Verlag.

Milligan, G. W. (1980). An examination of the effect of six types of error perturbation on fifteen clustering algorithms. *Psychometrika*, *45*, 325–342.

Nosofsky, R. M. (1991). Stimulus bias, asymmetric similarity, and classification. *Cognitive Psychology*, *23*, 94–140.

Okada, A. (1988a). Asymmetric multidimensional scaling of car switching data. In W. Gaul & M. Schader (Eds.), *Data, expert knowledge and decisions* (pp. 279–290). Berlin: Springer Verlag.

Okada, A. (1988b). An analysis of intergenerational occupational mobility by asymmetric multidimensional scaling. *Proceedings of the SMABS 88 Conference* (pp. 1–15). Groningen: University of Groningen.

Tobler, W. R., & Wineburg, S. (1971). A Cappadocian speculation. *Nature*, *231*, 39–42.

Weeks, D. G., & Bentler, P. M. (1982). Restricted multidimensional scaling models for asymmetric proximities. *Psychometrika*, *47*, 201–208.

Tversky, A. (1977). Features of similarity. *Psychological Review*, *84*, 327–352.

Manuscript received 9/24/91
Final version received 4/25/92