Journal of Classification 11:79-99 (1994)

A Generalization of GIPSCAL for the Analysis of Nonsymmetric Data

Henk A. L. Kiers

University of Groningen

Yoshio Takane

McGill University

Abstract: Graphical representation of nonsymmetric relationships data has usually proceeded via separate displays for the symmetric and the skew-symmetric parts of a data matrix. DEDICOM avoids splitting the data into symmetric and skew-symmetric parts, but lacks a graphical representation of the results. Chino's GIPSCAL combines features of both models, but may have a poor goodness-of-fit compared to DEDICOM. We simplify and generalize Chino's method in such a way that it fits the data better. We develop an alternating least squares algorithm for the resulting method, called Generalized GIPSCAL, and adjust it to handle GIPSCAL as well. In addition, we show that Generalized GIPSCAL is a constrained variant of DEDICOM and derive necessary and sufficient conditions for equivalence of the two models. Because these conditions are rather mild, we expect that in many practical cases DEDICOM and Generalized GIPSCAL are (nearly) equivalent, and hence that the graphical representation from Generalized GIPSCAL can be used to display the DEDICOM results graphically. Such a representation is given for an illustration. Finally, we show Generalized GIPSCAL to be a generalization of another method for joint representation of the symmetric and skew-symmetric parts of a data matrix.

Keywords: DEDICOM, alternating least squares, multidimensional scaling.

This research has been made possible by a fellowship from the Royal Netherlands Academy of Arts and Sciences to the first author, and by research grant number A6394 to the second author, from the Natural Sciences and Engineering Research Council of Canada. The authors are obliged to Jos ten Berge and Naohito Chino for stimulating comments.

Authors' Addresses: Henk A. L. Kiers, Department of Psychology, University of Groningen, Grote Kruisstraat 2/1, 9712 TS Groningen, The Netherlands. Yoshio Takane, Department of Psychology, McGill University, 1205 Avenue Docteur Penfield, Montreal, Quebec, H3A 1B1, Canada.

1. Introduction

Multidimensional scaling (MDS) is a popular family of techniques for generating graphical representations of data consisting of (dis)similarities among a set of objects. Usually these (dis)similarities are symmetric, in the sense that the relation between object i and object j is the same as that between object j and object i. However, there are many situations in which relations between objects are nonsymmetric. For instance, a person i may like person j very much, while person j dislikes person i. If a data set consists of nonsymmetric relationships among a set of objects, one cannot apply ordinary MDS to obtain a (graphical) representation of the data. Several techniques have been proposed to find representations of nonsymmetric data, some of which first split the data table into a symmetric and a skew-symmetric part, and analyze the symmetric part by, for instance, classical MDS (Torgerson 1958; Gower 1966), and the skew-symmetric part by specially designed methods (e.g., see Gower 1977; Constantine and Gower 1978). That is, if $\mathbf{X}$ is a square matrix with nonsymmetric relationships scores between n objects, then the symmetric part of $\mathbf{X}$, $\mathbf{X}_s = ½\,(\mathbf{X} + \mathbf{X}')$, is analyzed by using an MDS technique, and the skew-symmetric part $\mathbf{X}_k = ½\,(\mathbf{X} - \mathbf{X}')$ is analyzed by the technique proposed by Gower. Thus, one obtains two different representations of the same data, one representing the symmetric part of the relationships and the other representing the remaining skew-symmetric part. If different, unrelated mechanisms underlie the symmetric and skew-symmetric part of the data, the above procedures can be used to study these independently. However, in practice, one usually does not know if one deals with such unrelated mechanisms. Then, it can be useful to find a single representation of the observed nonsymmetric relationships. A method designed for the joint analysis of symmetric and skew-symmetric relationships is DEDICOM (DEcomposition into DIrectional COMponents), proposed by Harshman (1978). This method decomposes the $n \times n$ data matrix $\mathbf{X}$ as

$$\mathbf{X} = \mathbf{ARA}' + \mathbf{E}\,, \tag{1}$$

where $\mathbf{A}$ is an $n \times r$ $(r \le n)$ matrix containing coefficients to relate the objects to 'basic concepts' underlying the objects, the $r \times r$ matrix $\mathbf{R}$ contains measures to represent nonsymmetric relations between those basic concepts, and $\mathbf{E}$ is an $n \times n$ matrix of residuals. DEDICOM finds the $\mathbf{A}$ and $\mathbf{R}$ that fit model (1) in the least squares sense (by minimizing tr $\mathbf{E}'\mathbf{E}$). Although DEDICOM gives a single representation for the stimuli (in $\mathbf{A}$), it does not give a simple graphical representation of the objects to represent the nonsymmetric relationships.

Chino (1978, 1990) proposed a method that both handles the symmetric and skew-symmetric part simultaneously and gives a graphical representation of the nonsymmetric relationships. In matrix notation, his GIPSCAL method (Generalized Inner Product SCALing; Chino 1990, also see Chino 1980, p. 23) is based on the representation

$$\mathbf{X} = a\mathbf{A}\mathbf{A}' + b\mathbf{A}\mathbf{R}_I\mathbf{A}' + c\mathbf{1}\mathbf{1}' + \mathbf{E}\,, \tag{2}$$

where the $n \times r$ matrix $\mathbf{A}$ contains coordinates for the objects on r dimensions, $\mathbf{R}_I$ is a fixed skew-symmetric matrix with off-diagonal elements 1 or -1 in such a pattern that neighboring off-diagonal elements have opposite sign; for example, if

$$r = 4\,, \quad \mathbf{R}_I = \begin{bmatrix} 0 & 1 & -1 & 1 \\ -1 & 0 & 1 & -1 \\ 1 & -1 & 0 & 1 \\ -1 & 1 & -1 & 0 \end{bmatrix}\;;$$

furthermore, $\mathbf{1}$ is an n-vector with unit elements, $\mathbf{E}$ is an $n \times n$ matrix of residuals, and a, b, and c are scalars to express relative importances of the three parts of which the model exists. Although not explicitly done by Chino, we henceforth assume that $a \geq 0$, because for negative a the model would not make sense graphically (unless we deal with dissimilarities, but in such cases we would multiply the data by -1 to obtain similarities). It may be worth noting that the sizes of a, b and $\mathbf{A}$ are undetermined, because multiplying a and b by a positive scalar and dividing $\mathbf{A}$ by the square root of that scalar does not affect the model.

Chino's GIPSCAL model (2) is fitted to the data in the least squares sense by minimizing tr $\mathbf{E}'\mathbf{E}$. The coordinates in $\mathbf{A}$ are used for representing the objects in an r-dimensional space. The configuration represents both the symmetric part of $\mathbf{X}$ and the skew-symmetric part of $\mathbf{X}$. That is, for the symmetric part, the element (i,j) of $\mathbf{X}_s$ is represented by $a\mathbf{a}'_i\mathbf{a}_j + c$, where $\mathbf{a}'_i$ denotes the i-th row of $\mathbf{A}$, and c an additive constant; so the symmetric part is represented, as in classical MDS, by the inner product of the coordinate vectors for i and j. For the skew-symmetric part, each element (i,j) of $\mathbf{X}_k$ is represented in a more complicated way. For each pair of dimensions, we compute the area of the triangle formed by the origin and the projections of i and j on these dimensions. Then element (i,j) is represented by the sum of all these areas, each multiplied by $\pm 2b$, where the sign depends on whether in going from the projection of i to that of j we make a clockwise ($-$) or counterclockwise (+) movement. Clearly, if $r > 2$, it takes considerable effort to deduce the representation of the skew-symmetric part of the data from the graphical displays. In the present paper, we will simplify the model (without affecting the representation and the goodness-of-fit of the model), and next, offer a slightly more general variant of GIPSCAL, which yields a goodness-of-fit which is at least as good as GIPSCAL's and usually better.

Chino (1990) has derived an alternating least squares algorithm for fitting the off-diagonal elements of the GIPSCAL model. For the case where diagonal elements should be fitted as well (denoted as 'fitting the full model'), no alternating least squares algorithm is available. In the present paper, we will derive an alternating least squares algorithm for fitting the generalized GIPSCAL model, which, by some simple modifications can also be used for fitting the full GIPSCAL model.

One of the main problems with the DEDICOM model is its nonuniqueness: Any nonsingular transformation $\mathbf{T}$ of $\mathbf{A}$, yielding $\tilde{\mathbf{A}} = \mathbf{AT}$, can be compensated by replacing $\mathbf{R}$ by $\tilde{\mathbf{R}} = \mathbf{T}^{-1}\mathbf{R}(\mathbf{T}')^{-1}$, without affecting the DEDICOM representation. Such transformations can affect the interpretation of the DEDICOM results considerably. The GIPSCAL model also has some nonuniqueness, but this deficiency does not interfere with the interpretation of the GIPSCAL solution, as will be derived below.

Chino (1980) has mentioned that his model is a constrained version of the DEDICOM model. We will show that our generalization of GIPSCAL also is a constrained variant of DEDICOM. In addition, we will derive necessary and sufficient conditions for equivalence of the two models. These conditions turn out to be rather mild, and hence, in cases of (near) equivalence, the graphical display from generalized GIPSCAL can be seen to provide the DEDICOM results, which until now could only be represented 'numerically', with a graphical representation. This advantage will be illustrated by the analysis of an empirical data set.

A third method for simultaneous representation of the symmetric and skew-symmetric part of a nonsymmetric relationships table has been proposed by Escoufier and Grorud (1980) and independently by Chino (1991). As a final result we show that the Escoufier and Grorud method is a constrained variant of generalized GIPSCAL.

2. A Simplification of GIPSCAL

The interpretation of GIPSCAL results is relatively complicated with respect to the skew-symmetric part of the data. To simplify the GIPSCAL model, we can reparametrize the matrices $\mathbf{A}$ and $\mathbf{R}_I$ as follows. Let $\mathbf{R}_I = \mathbf{U}\Delta\mathbf{U}'$ denote the 'Gower decomposition' of $\mathbf{R}_I$, where Δ is a block-diagonal matrix with 2×2 matrices $\begin{bmatrix} 0 & \lambda_l \\ -\lambda_l & 0 \end{bmatrix}$ along the diagonal and, if n is odd, a zero element in the last diagonal position, and $\mathbf{U}$ is an orthonormal matrix (Gower 1977). Note that the nonzero elements of Δ (being the singular values of $\mathbf{R}_I$) are fixed, because the elements of $\mathbf{R}_I$ are fixed. Using this Gower decomposition, we define $\tilde{\mathbf{A}} \equiv \mathbf{AU}$, and rewrite model (2) as

$$\mathbf{X} = a\tilde{\mathbf{A}}\tilde{\mathbf{A}}' + b\tilde{\mathbf{A}}\Delta\tilde{\mathbf{A}}' + c\mathbf{1}\mathbf{1}' + \mathbf{E}\,. \tag{3}$$

In this representation the skew-symmetric part is represented by

$$\mathbf{X}_k = b\tilde{\mathbf{A}}\Delta\tilde{\mathbf{A}}'\,. \tag{4}$$

Now each element (i,j) of $\mathbf{X}_k$ is represented by the sum of the representation obtained from the first two dimensions, that of the third and fourth, etc. That is,

$$[\mathbf{X}_k]_{ij} = b\lambda_1(\tilde{a}_{i1}\,\tilde{a}_{j2} - \tilde{a}_{i2}\,\tilde{a}_{j1}) + b\lambda_2(\tilde{a}_{i3}\,\tilde{a}_{j4} - \tilde{a}_{i4}\,\tilde{a}_{j3}) + \ldots, \tag{5}$$

hence $[\mathbf{X}_k]_{ij}$ can be seen as the weighted sum of triangle areas lying within the origin, and the projections of $\mathbf{a}'_i$ and $\mathbf{a}'_j$ on the first two dimensions, on the second two dimensions, etc. This description makes it possible to study the contribution of pairs of dimensions (christened "bimensions" by Carroll; see Harshman 1981) separately. For instance, the first pair contributes $b\lambda_1(\tilde{a}_{i1}\,\tilde{a}_{j2} - \tilde{a}_{i2}\,\tilde{a}_{j1})$ to the *skew-symmetric* relation between i and j, and the first and second dimension contribute $a(\tilde{a}_{i1}\,\tilde{a}_{j1})$ and $a(\tilde{a}_{i2}\,\tilde{a}_{j2})$, respectively (or $a(\tilde{a}_{i1}\,\tilde{a}_{j1} + \tilde{a}_{i2}\,\tilde{a}_{j2})$ together), to the *symmetric* relation between i and j. As a result, it is possible to display the total contribution of one pair of dimensions in one plot, where the inner product of two vectors (times a) describes the symmetric portion of the relation, and the related triangle area (times $\pm 2b\lambda_l$) describes the skew-symmetric portion. This representation involves $\frac{1}{2}r$ two-dimensional plots, and hence is considerably simpler than GIPSCAL, which requires $\frac{1}{2}r(r-1)$ two-dimensional plots for a complete representation of the nonsymmetric relations.

3. Generalized GIPSCAL

Above, we have rewritten the GIPSCAL model as (3), or, after dropping overline tildes (~):

$$\mathbf{X} = a\mathbf{A}\mathbf{A}' + b\mathbf{A}\Delta\mathbf{A}' + c\mathbf{1}\mathbf{1}' + \mathbf{E}\,, \tag{6}$$

where Δ is a fixed matrix with singular values of the matrix $\mathbf{R}_I$ in skew-symmetric 2×2 blocks along the diagonal. The elements of Δ serve to indicate the importances of the bimensions in representing the skew-symmetric part of $\mathbf{X}$. The importances are fixed to the singular values of $\mathbf{R}_I$, a matrix which was fixed to have elements equal to 0 and ± 1, which clearly facilitated the interpretation of the $\frac{1}{2}r(r-1)$ GIPSCAL plots. In our description of

GIPSCAL, the fixed values in $\mathbf{R}_l$, and hence in Δ, no longer facilitate interpretation. In fact, there no longer seems to be any reason to keep the nonzero elements of Δ fixed, as long as they indicate the importance of each bimension. Therefore, we propose a "Generalized GIPSCAL" model in which the nonzero elements in Δ are left free, except for the requirement that Δ should have skew-symmetric 2×2 blocks along the diagonal. In fact, this modification comes down to replacing the special skew-symmetric matrix $\mathbf{R}_l$ in Chino's GIPSCAL model by an arbitrary skew-symmetric matrix. Note that, in this relaxed model, we can, without loss of generality, delete the scalars a and b, by subsuming $a^{1/2}$ in $\mathbf{A}$ if $a > 0$, and b in Δ. Obviously, if $a = 0$, the term involving a vanishes, and we would have a simplified model, with b again subsumed in Δ. In that case, the corresponding generalization of GIPSCAL would be equivalent to Gower's (1977) method for decomposing a skew-symmetric matrix. We will only consider the case where $a > 0$. Consequently, we propose Generalized GIPSCAL as the method that minimizes

$$\sigma_g(\mathbf{A},\delta_1,\ldots,\delta_q,c) = \|\mathbf{X} - \mathbf{AA}' - \mathbf{A}\Delta\mathbf{A}' - c\mathbf{11}'\|^2\,, \qquad (7)$$

where Δ is the matrix with 2×2 blocks $\begin{bmatrix} 0 & \delta_l \\ -\delta_l & 0 \end{bmatrix}$, $l = 1,\ldots,q$, along the diagonal, and if r is odd a zero element in the last diagonal position. The interpretation of results of the Generalized GIPSCAL method can be made analogously to the bimensionwise interpretation suggested above for GIPSCAL results.

Generalized GIPSCAL generalizes GIPSCAL to the effect that weights for different bimensions are no longer fixed to prespecified values (computed via the singular values of $\mathbf{R}_l$). For the case with only one bimension ($r = 2$ or $r = 3$), this relaxation can be done without affecting the model's goodness-of-fit, because any rescaling of Δ can be compensated by the free scalar b. In case $r \geq 4$, this strategy is no longer possible, because different elements of Δ are rescaled differently; hence Generalized GIPSCAL differs from GIPSCAL as soon as there are two bimensions or more, that is, in case $r \geq 4$.

To interpret the Generalized GIPSCAL solution, we may use the above simplified interpretation derived for GIPSCAL. This simplification facilitated the interpretation by decreasing the number of plots to be interpreted, but the interpretation of one plot is still quite complicated since it involves inferring the contribution of a stimulus pair to symmetry and to skew-symmetry via entirely different processes (viz., via taking inner products, and via computing triangle areas, respectively). To represent the total nonsymmetric relations we have to sum these portions. The total nonsymmetric relation between two stimuli cannot directly be read from the plot. This drawback seriously detracts from the usefulness of the plot for inferring contributions of stimulus pairs to the whole matrix of nonsymmetric relations. Therefore, we propose an alternative procedure, which is closely related to a suggestion by Gower and Zielman (1992), and made in a slightly different context.

The model's representations for all stimulus pairs are given by the matrix $\hat{\mathbf{X}} = \mathbf{AA}' + \mathbf{A\Delta A}'$, where we ignored the additive constant which contributes equally to all relationships. For the l-th bimension, we can write the associated vectors of $\mathbf{A}$ in an $(n \times 2)$ matrix $\mathbf{A}_l$, and the associated portion of Δ can be written as $\Delta_l = \begin{bmatrix} 0 & \delta_l \\ -\delta_l & 0 \end{bmatrix}$. Then, the representation by this bimension is given by

$$\hat{\mathbf{X}}_l = \mathbf{A}_l\mathbf{A}'_l + \mathbf{A}_l\Delta_l\mathbf{A}'_l = \mathbf{A}_l(\mathbf{I} + \Delta_l)\mathbf{A}'_l\,. \tag{8}$$

Now, we can write $(\mathbf{I} + \Delta_l) = \beta_l\mathbf{T}'_l$, where $\beta_l \equiv (1 + \delta_l^2)^{1/2}$, and $\mathbf{T}_l \equiv \beta_l^{-1}\begin{bmatrix} 1 & -\delta_l \\ \delta_l & 1 \end{bmatrix}$, and verify that $\mathbf{T}_l$ is orthonormal. Specifically, let $\phi_l \equiv \mathit{arctan}\,(\delta_l)$, then $\mathbf{T}_l$ is a clockwise rotation over the angle ϕ_l. Substituting $(\mathbf{I} + \Delta_l) = \beta_l\mathbf{T}'_l$ for $(\mathbf{I} + \Delta_l)$ in (8) we find

$$\hat{\mathbf{X}}_l = \beta_l\mathbf{A}_l\mathbf{T}'_l\mathbf{A}'_l = \tilde{\mathbf{A}}_l\tilde{\mathbf{B}}'_l\,, \tag{9}$$

with $\tilde{\mathbf{A}}_l \equiv \beta_l^{1/2}\mathbf{A}_l$, and $\hat{\mathbf{B}}_l \equiv \tilde{\mathbf{A}}_l\mathbf{T}_l$. By means of (9) the nonsymmetric relations between the stimuli are represented by inner products between rows of $\tilde{\mathbf{A}}_l$ and of $\tilde{\mathbf{B}}_l$. Thus, a scalar-product representation (also known as biplot, see Gabriel 1971) is obtained for the nonsymmetric relations. Specifically, the relation of stimulus i to stimulus j is represented by the inner product of the i-th row of $\mathbf{A}_l$ and the j-th row of $\mathbf{B}_l$, whereas the relation of stimulus j to stimulus i is represented by the inner product of the j-th row of $\mathbf{A}_l$ and the i-th row of $\mathbf{B}_l$, as in ordinary biplots. This biplot is special, however, in that the configuration for the column stimuli is equal to that for the row stimuli, except for a rotation over the angle ϕ_l (just as in Gower and Zielman 1992). As a result, stimuli close to each other will have more or less the same nonsymmetric relationships with other stimuli.

An advantage of this biplot type of interpretation is that the importances of the skew-symmetric and the symmetric part of a particular relation can be compared more easily than in the original GIPSCAL interpretation. In the latter, the contributions to symmetry and skew-symmetry were displayed in incomparable units (inner products and triangle areas, respectively). In our new interpretation procedure, we can 'read' the inner products between i and j, and between j and i from the plot, and the difference between these indicates the amount of skew-symmetry, whereas the sum of these indicates the amount of symmetry. In practice, it may suffice to compare size and sign of these two inner products: Having $\hat{x}_{ij}$ and $\hat{x}_{ji}$ both positive (or both negative) and of more or less the same size indicates little asymmetry and much symmetry; $\hat{x}_{ij}$ large and $\hat{x}_{ji}$ small (both positive or both negative) indicates considerable amounts of both symmetry and asymmetry; $\hat{x}_{ij}$ and $\hat{x}_{ji}$ of opposite sign, and at least one of these being large indicates much asymmetry and little symmetry. To aid 'reading' the inner products from the plot, one may use the projections of the vectors for i (one for row i and one for column i) on the vectors for j from the other configurations, having lengths that are proportional to the required inner products.

With the above interpretation procedure, Generalized GIPSCAL becomes a method which yields bimensional biplots between row and column configurations that are constrained to be rotations of each other. The rotation angle ϕ_l is determined for each bimension separately and may thus vary considerably over bimensions. The size of the angle indicates the amount of skew-symmetry represented by bimension l: if ϕ_l is close to zero, hardly any skew-symmetry is represented by the bimension. If ϕ_l is close to 90°, the representation is dominated by skew-symmetry. The interpretation procedure sketched here will be illustrated in an analysis at the end of the paper.

4. An Alternating Least Squares Algorithm for GIPSCAL

Chino (1990) has derived an alternating least squares (ALS) algorithm for fitting the GIPSCAL model to off-diagonal elements only. All diagonal elements are considered missing, and his derivation does not yield an ALS algorithm for cases where the diagonal is to be fitted as well. That is, the steps used in the algorithm for off-diagonal fitting are not the least squares optimal steps for fitting the full (including diagonal) model, and hence need not decrease the function value monotonically. In the present section, we will derive an ALS algorithm for fitting the (full) Generalized GIPSCAL model, with an option for handling missing data (thereby including the off-diagonal case). At the end of the present section, we indicate how the algorithm can be modified to fit the original GIPSCAL model.

We will first derive an ALS algorithm for the case without missing data. The problem is to minimize (7) alternately over $\mathbf{A}$, Δ, and c. For fixed $\mathbf{A}$ and Δ, minimizing (7) over c amounts to minimizing a quadratic function in c, which is solved by taking

$$c = 1'(\mathbf{X} - \mathbf{AA}' - \mathbf{A}\Delta\mathbf{A}')1/n^2 . \tag{10}$$

For fixed Δ and c, function (7) can be minimized over $\mathbf{A}$ row by row as follows. Consider the problem of minimizing σ_g over the i-th row $\mathbf{a}'_i$ (where the prime is used to emphasize that we are dealing with row rather than column vectors) of $\mathbf{A}$, with the other rows of $\mathbf{A}$ considered fixed. Let $\mathbf{A}_{-i}$ denote the matrix $\mathbf{A}$ with the i-th row deleted, let the element x_{ij} denote the element (i,j) of $(\mathbf{X} - c\mathbf{11}')$, and let $\mathbf{x}_{r(i)}$ and $\mathbf{x}_{c(i)}$ denote the i-th (transposed) row and column, respectively, of $(\mathbf{X} - c\mathbf{11}')$ with the element (i,i) deleted. Then, with $\mathbf{a}'_i\Delta\mathbf{a}_i = 0$ because Δ is skew-symmetric, it remains to minimize

$$g(\mathbf{a}'_i) = (x_{ii} - \mathbf{a}'_i\mathbf{a}_i)^2 + \left\| \begin{bmatrix} \mathbf{x}_{r(i)} \\ \mathbf{x}_{c(i)} \end{bmatrix} - \begin{bmatrix} \mathbf{A}_{-i}(\mathbf{I} + \Delta') \\ \mathbf{A}_{-i}(\mathbf{I} + \Delta) \end{bmatrix} \mathbf{a}_i \right\|^2$$

$$= (x_{ii} - \mathbf{a}'_i\mathbf{a}_i)^2 + \|\phi - \mathbf{F}\mathbf{a}_i\|^2 , \tag{11}$$

where we define $\phi \equiv \begin{bmatrix} \mathbf{x}_{r(i)} \\ \mathbf{x}_{c(i)} \end{bmatrix}$, and $\mathbf{F} \equiv \begin{bmatrix} \mathbf{A}_{-i}(\mathbf{I} + \Delta') \\ \mathbf{A}_{-i}(\mathbf{I} + \Delta) \end{bmatrix}$. This problem is of the form discussed and solved by Ten Berge (1991). Thus, we have derived a solution for minimizing (7) over the i-th row of $\mathbf{A}$ given the other rows of $\mathbf{A}$. By updating each row of $\mathbf{A}$ in the way described above, we decrease (7) monotonically.

Finally, for fixed $\mathbf{A}$ and c, we have to minimize (7) over Δ, or rather over $\delta_1, \ldots, \delta_q$. To solve this problem, we first write Δ as $\mathbf{KD}$, where $\mathbf{K}$ is a fixed matrix of the same form as Δ, with the elements $\delta_1, \ldots, \delta_q$ replaced by unit elements, and, if n is odd, a unit element on the final diagonal position, and $\mathbf{D}$ is the diagonal matrix with diagonal elements $\delta_1, \delta_1, \delta_2, \delta_2, \ldots, \delta_q, \delta_q$, and a final 0 if r is odd. Then, the problem of minimizing σ_g over Δ reduces to minimizing

$$\sigma_g(\mathbf{D}) = \|(\mathbf{X} - \mathbf{AA}') - \mathbf{AKDA}'\|^2$$

$$= \|Vec(\mathbf{X} - \mathbf{AA}') - Vec(\mathbf{AKDA}')\|^2$$

$$= \|Vec(\mathbf{X} - \mathbf{AA}') - (\mathbf{A} \otimes \mathbf{AK})\, Vec(\mathbf{D})\|^2$$

$$= \|Vec(\mathbf{X} - \mathbf{AA}') - (\mathbf{A} \times (\mathbf{AK}))\mathbf{d}\|^2 , \tag{12}$$

where Vec denotes a matrix strung out columnwise into a vector, $\otimes$ denotes the Kronecker product, $\mathbf{A} \times (\mathbf{AK})$ denotes the $n^2 \times r$ matrix containing Kronecker products of corresponding columns of $\mathbf{A}$ and $\mathbf{AK}$ (see Carroll & Chang, 1970, p.286, for an earlier use of this product), and the vector $\mathbf{d}$ contains the diagonal elements of $\mathbf{D}$. Now constructing the $n \times q$ matrix $\mathbf{B}$ with sums of consecutive columns of $(\mathbf{A} \times (\mathbf{AK}))$, ignoring the last column of $(\mathbf{A} \times (\mathbf{AK}))$ if r is odd, and defining $\mathbf{w} \equiv (\delta_1, \ldots, \delta_q)'$, the problem reduces to the regression problem of minimizing $\| Vec(\mathbf{X} - \mathbf{AA}') - \mathbf{Bw} \|^2$, for which the solution is given by

$$\mathbf{w} = (\mathbf{B}'\mathbf{B})^{-}\mathbf{B}' \, Vec(\mathbf{X} - \mathbf{AA}') , \tag{13}$$

where $(.)^-$ denotes a generalized inverse. We thus find a solution for the parameters $\delta_1, \ldots, \delta_q$, and hence for Δ. Alternately updating c, $\mathbf{A}$, and Δ, we have a monotonically converging algorithm for Generalized GIPSCAL.

The algorithm derived above can straightforwardly be modified to allow for missing data. Let $\mathbf{W}$ be an $n \times n$ binary indicator matrix with unit elements for nonmissing data and zeros for missing data. Then the problem is to minimize

$$\sigma_g(\mathbf{A}, \delta_1, \ldots, \delta_q, c) = \sum_{i=1}^{n} \sum_{j=1}^{n} w_{ij}(x_{ij} - \mathbf{a}'_i\mathbf{a}_j - \mathbf{a}'_i\Delta\mathbf{a}_j - c)^2 . \tag{14}$$

It is readily verified that this problem is equivalent to the minimization of

$$\sigma_g^*(\mathbf{A}, \delta_1, \ldots, \delta_q, c, \mathbf{X}) = \| \mathbf{X} - \mathbf{AA}' - \mathbf{A}\Delta\mathbf{A}' - c\mathbf{11}' \|^2 \tag{15}$$

over c, $\mathbf{A}$, Δ, and the *missing elements* of $\mathbf{X}$. To see why these methods are equivalent, note that the missing elements of $\mathbf{X}$ are updated by setting them equal to the corresponding elements in the model representation, thus setting the loss for these elements to zero. Using the second description, an ALS algorithm is obtained at once by alternatingly minimizing σ_g^* over $\mathbf{X}$, c, $\mathbf{A}$, and Δ, which differs from the original algorithm only in that an additional step for updating the missing values in $\mathbf{X}$ is needed.

The algorithm above can be used for missing data at any position. However, in the special case where all diagonal elements are considered missing (the off-diagonal case), we prefer to use a different procedure for updating $\mathbf{A}$, as follows. In the off-diagonal case, the problem of minimizing σ_g^* over $\mathbf{a}'_i$ reduces to minimizing

$$g^*(\mathbf{a}'_i) = \left\| \begin{bmatrix} \mathbf{x}_{r(i)} \\ \mathbf{x}_{c(i)} \end{bmatrix} - \begin{bmatrix} \mathbf{A}_{-i}(\mathbf{I} + \Delta') \\ \mathbf{A}_{-i}(\mathbf{I} + \Delta) \end{bmatrix} \mathbf{a}_i \right\|^2 = \| \phi - \mathbf{F}\mathbf{a}_i \|^2 \tag{16}$$

where we dropped the term $(x_{ii} - \mathbf{a}'_i\mathbf{a}_i)^2$ from (11) since this term can, regardless of the value of $\mathbf{a}_i$, be set equal to zero by immediately updating the missing value x_{ii}. The problem of minimizing (16) is a simple regression problem with solution $\mathbf{a}_i = (\mathbf{F}'\mathbf{F})^{-}\mathbf{F}'\phi$.

The Generalized GIPSCAL algorithm derived above can be used for fitting the original GIPSCAL model by setting Δ equal to $\mathbf{K}\Sigma$, where Σ is the diagonal matrix with singular values of $\mathbf{R}_J$, and allowing for a scalar b in the skew-symmetric model part. The scalar a can still be considered subsumed in $\mathbf{A}$. Then the problem of minimizing (16) is modified into that of minimizing

$$\sigma_g^*(\mathbf{A},b,c,\mathbf{X}) = \|\mathbf{X} - \mathbf{AA}' - b\mathbf{A}\Delta\mathbf{A}' - c\mathbf{11}'\|^2 . \qquad (17)$$

An ALS algorithm can be constructed by updating the missing values of $\mathbf{X}$, the rows of $\mathbf{A}$ and c as before, and, instead of updating Δ, we now have to update b as

$$b = (tr\mathbf{A}'(\mathbf{X} - c\mathbf{11}')\mathbf{A}\Delta') / (tr\mathbf{A}\Delta\mathbf{A}'\mathbf{A}\Delta'\mathbf{A}') , \qquad (18)$$

which follows from minimizing a quadratic function in b. In the off-diagonal case, this algorithm is basically equivalent to Chino's (1990) algorithm.

5. Uniqueness of the Generalized GIPSCAL Representation

In general, the Generalized GIPSCAL model does not provide a unique representation for $\mathbf{A}$, c, and Δ. However, if $\mathbf{A}$ has full rank and does not contain $\mathbf{1}$ in its column space, and if the elements of Δ that belong to different bimensions are distinct, then Δ and c are determined uniquely given the total Generalized GIPSCAL representation, and $\mathbf{A}$ is unique up to bimensionwise rotations. To prove this claim, suppose we have a different set of parameters $\mathbf{A}^*$, Δ^*, and c^*, satisfying the same assumptions and giving exactly the same representation for $\mathbf{X}$. That is, suppose that

$$\mathbf{AA}' + \mathbf{A}\Delta\mathbf{A}' + c\mathbf{11}' = \mathbf{A}^*\mathbf{A}^{*\prime} + \mathbf{A}^*\Delta^*\mathbf{A}^{*\prime} + c^*\mathbf{11}' . \qquad (19)$$

Then, for the symmetric part of (19) we have

$$\mathbf{AA}' + c\mathbf{11}' = \mathbf{A}^*\mathbf{A}^{*\prime} + c^*\mathbf{11}' . \qquad (20)$$

Let $\mathbf{J} \equiv (\mathbf{I} - n^{-1}\mathbf{11}')$, then from (20) we have $\mathbf{AA}'\mathbf{J} = \mathbf{A}^*\mathbf{A}^{*\prime}\mathbf{J}$. From the assumption that $\mathbf{A}^*$ does not contain $\mathbf{1}$ in its column space it follows that $\mathbf{JA}^*$ has full rank; hence $\mathbf{A}^* = \mathbf{AT}$, for a certain matrix $\mathbf{T}$, which is nonsingular because $\mathbf{A}$ and $\mathbf{A}^*$ are assumed to have full rank. It follows that

$$\mathbf{AA}'\mathbf{J} = \mathbf{ATT}'\mathbf{A}'\mathbf{J} , \qquad (21)$$

which upon premultiplication by $(\mathbf{A}'\mathbf{A})^{-1}\mathbf{A}'$ (using the full rank of $\mathbf{A}$) and postmultiplication by $\mathbf{JA}(\mathbf{A}'\mathbf{JA})^{-1}$ yields $\mathbf{TT}' = \mathbf{I}$; that is, $\mathbf{T}$ is orthonormal. As a by-product, we now have $\mathbf{AA}' = \mathbf{A}^*\mathbf{A}^{*\prime}$; hence $c^* = c$. For the skew-symmetric part of (19), we have

$$\mathbf{A\Delta A'} = \mathbf{A^*\Delta^*A^{*'}} \ . \tag{22}$$

Now we can derive from (22) that $\mathbf{A\Delta A'} = \mathbf{AT\Delta^*T'A'}$ and, hence, exploiting the full rank of $\mathbf{A}$, that $\Delta = \mathbf{T}\Delta^*\mathbf{T}'$. If we substitute $\Delta = \mathbf{KD}$ and $\Delta^* = \mathbf{KD}^*$ in $\Delta = \mathbf{T}\Delta^*\mathbf{T}'$ and define the orthonormal matrix $\mathbf{U}$ as $\mathbf{U} \equiv \mathbf{TK}$, we find $\mathbf{KDI} = \mathbf{UD^*T'}$, of which left- and right-hand side both define a singular value decomposition (SVD). From the uniqueness properties of the SVD with partly distinct singular values, it follows that $\mathbf{D} = \mathbf{D}^*$ (hence $\Delta = \Delta^*$) and, for the right-hand singular vectors, $\mathbf{T} = \mathbf{IN} = \mathbf{N}$, where $\mathbf{N}$ is an orthonormal matrix that commutes with $\mathbf{D}$, which hence contains nonzero elements only in the 2×2 blocks along the diagonal (because $\mathbf{D}$ has distinct singular values for different bimensions). To conclude, we find that, under the mild conditions specified above, c and Δ are unique given the Generalized GIPSCAL representation, and $\mathbf{A}$ is determined up to a bimensionwise rotation.

For the interpretation, the above derived uniqueness is quite useful. Indeed, the only indeterminacy is the orientation of the axes of a bimension. However, this orientation is of no importance in the bimensionwise interpretation proposed above.

6. Conditions for Equivalence of Generalized GIPSCAL and DEDICOM

The basic difference between DEDICOM and Generalized GIPSCAL is that in the former we have $\mathbf{R}$, where we have $(\mathbf{I} + \Delta)$ in the latter. It follows that Generalized GIPSCAL is a constrained variant of DEDICOM, where the constraint is that $\mathbf{R}$ can be written as $(\mathbf{I} + \Delta)$. This constraint is not as stringent as it seems, because we can replace $\mathbf{R}$ by $\mathbf{TRT}'$ for any nonsingular matrix $\mathbf{T}$, if we replace $\mathbf{A}$ by $\mathbf{AT}^{-1}$. In this way, it may be possible to find a matrix $\mathbf{R}$ that does satisfy the constraint, even though that was not apparent at first. In the following theorem, we describe a necessary and sufficient condition under which the DEDICOM solution can be made to satisfy the constraint implicitly used in Generalized GIPSCAL.

Theorem. *Given a DEDICOM solution with full rank matrices* $\mathbf{A}$ *and* $\mathbf{R}$, *the DEDICOM representation* $\mathbf{ARA}'$ *can be written in the form of the Generalized GIPSCAL model with* $c = 0$, *if and only if the symmetric part of matrix* $\mathbf{R}$ *in the DEDICOM solution is positive definite (p.d.).*

Proof. Let $\mathbf{R}_s$ and $\mathbf{R}_k$ denote the symmetric and skew-symmetric parts of $\mathbf{R}$, respectively. If $\mathbf{R}_s$ is p.d., then we can decompose it as $\mathbf{R}_s = \mathbf{TT}'$ (e.g., using the Cholesky decomposition). Defining $\hat{\mathbf{X}} \equiv \mathbf{ARA}'$, we find $\hat{\mathbf{X}} = \mathbf{AR}_s\mathbf{A}' + \mathbf{AR}_k\mathbf{A}' = \mathbf{ATT'A}' + \mathbf{ATT}^{-1}\mathbf{R}_k(\mathbf{T}')^{-1}\mathbf{T}'\mathbf{A}'$. Let $\mathbf{T}^{-1}\mathbf{R}_k(\mathbf{T}')^{-1} = \mathbf{U\Gamma U}'$ denote the Gower decomposition of the skew-symmetric matrix

$\mathbf{T}^{-1}\mathbf{R}_k(T')^{-1}$, with $\mathbf{U}$ orthonormal and Γ block-diagonal. Then, with the definition $\tilde{\mathbf{A}} \equiv \mathbf{ATU}$, we obtain $\hat{\mathbf{X}} = \tilde{\mathbf{A}}\tilde{\mathbf{A}}' + \tilde{\mathbf{A}}\Gamma\tilde{\mathbf{A}}'$, which is of the form of the Generalized GIPSCAL model with $c = 0$.

Conversely, if the DEDICOM representation can be written in the form of the Generalized GIPSCAL model with $c = 0$, then, $\mathbf{ARA}' = \underline{\mathbf{A}}\underline{\mathbf{A}}' + \underline{\mathbf{A}}\Delta\underline{\mathbf{A}}'$, for certain matrices $\underline{\mathbf{A}}$ and Δ, where Δ is skew-symmetric. Then it follows that $\underline{\mathbf{A}}$ is in the column space of $\mathbf{A}$, and hence $\underline{\mathbf{A}} = \mathbf{AT}$ for a certain matrix $\mathbf{T}$. Since $\mathbf{A}$ and $\mathbf{R}$ have full rank, $\underline{\mathbf{A}}$ also has full rank; hence $\mathbf{T}$ is nonsingular. It follows that $\mathbf{ARA}' = \mathbf{ATT}'\mathbf{A}' + \mathbf{AT}\Delta\mathbf{T}'\mathbf{A}'$, and hence $\mathbf{R} = \mathbf{TT}' + \mathbf{T}\Delta\mathbf{T}'$. As a result, $\mathbf{R}_s = \mathbf{TT}'$, which is p.d. ■

The above theorem implies that for every DEDICOM solution with $\mathbf{R}_s$ p.d., it is possible to give a graphical representation in exactly the same way as in Generalized GIPSCAL. A similar result has been found by Zielman and Heiser (1991) for the case where $r = 2$. For this case they mentioned the (near) equivalence of DEDICOM and GIPSCAL after the $\mathbf{A}$ in DEDICOM has been rotated such that $\mathbf{R}_s$ is diagonalized. Here, we have generalized their result to the case where $r \geq 2$.

If DEDICOM and Generalized GIPSCAL are equivalent, we have thus furnished DEDICOM with a plotting procedure which, moreover, gives a unique sequence of two-dimensional plots (except for rotation of the plots themselves). If $\mathbf{R}_s$ in DEDICOM is not p.d., and hence Generalized GIPSCAL (with $c = 0$) and DEDICOM are not equivalent, we can apply Generalized GIPSCAL to the same data, and compare the goodness-of-fit value to that obtained by DEDICOM. If these differ only slightly, one may prefer the Generalized GIPSCAL representation because it gives almost the same representation of the data but adds a graphical display to the output.

7. Escoufier and Grorud's Method as a Special Case of Generalized GIPSCAL

A third method for the simultaneous representation of the symmetric and the skew-symmetric part of a nonsymmetric data table is given by Escoufier and Grorud (1980; also see Chino 1991). We will now show how this method is related to DEDICOM and Generalized GIPSCAL. The method is based on the eigendecomposition of the Hermitean (usually complex) matrix $\mathbf{X}_s + i\mathbf{X}_k$. As noted by Escoufier and Grorud, the eigenvectors of $\mathbf{X}_s + i\mathbf{X}_k$ are directly related to the eigenvectors of the matrix $\mathbf{H} = \begin{bmatrix} \mathbf{X}_s & -\mathbf{X}_k \\ \mathbf{X}_k & \mathbf{X}_s \end{bmatrix}$. Escoufier and Grorud have shown that the eigenvalues of this matrix have multiplicity 2, and that to an eigenvalue λ_l correspond eigenvectors $\begin{bmatrix} \mathbf{u}_l \\ \mathbf{v}_l \end{bmatrix}$ and $\begin{bmatrix} -\mathbf{v}_l \\ \mathbf{u}_l \end{bmatrix}$. It can be verified that the $2q$-dimensional Eckart-

Young (1936) approximation of $\mathbf{H}$ reduces to the following representations for $\mathbf{X}_s$ and $\mathbf{X}_k$:

$$\hat{\mathbf{X}}_s = \sum_{l=1}^{q} \lambda_l(\mathbf{u}_l\mathbf{u}'_l + \mathbf{v}_l\mathbf{v}'_l)\,, \tag{23}$$

and

$$\hat{\mathbf{X}}_k = \sum_{l=1}^{q} \lambda_l(\mathbf{v}_l\mathbf{u}'_l - \mathbf{u}_l\mathbf{v}'_l)\,. \tag{24}$$

In fact, Escoufier and Grorud's method minimizes $\|\hat{\mathbf{X}}_s - \mathbf{X}_s\|^2 + \|\hat{\mathbf{X}}_k - \mathbf{X}_k\|^2$ over λ_l, $\mathbf{u}_l$ and $\mathbf{v}_l$, $l = 1,\ldots,q$. We will now show how this method is related to DEDICOM and Generalized GIPSCAL.

Collecting $\mathbf{u}_l$ and $\mathbf{v}_l$ into a matrix $\mathbf{A} \equiv (\lambda_1^{1/2}\,\mathbf{v}_1 \mid \lambda_1^{1/2}\,\mathbf{u}_1 \mid \cdots \mid \lambda_q^{1/2}\,\mathbf{v}_q \mid \lambda_q^{1/2}\mathbf{u}_q)$, we find

$$\mathbf{X}_s = \mathbf{AA}' + \mathbf{E}_s\,, \tag{25}$$

and

$$\mathbf{X}_k = \mathbf{AKA}' + \mathbf{E}_k\,, \tag{26}$$

where $\mathbf{K}$ is the $2q \times 2q$ matrix with 2×2 blocks $\begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix}$ along the diagonal and zeros elsewhere. Minimizing $\|\hat{\mathbf{X}}_s - \mathbf{X}_s\|^2 + \|\hat{\mathbf{X}}_k - \mathbf{X}_k\|^2$ then reduces to minimizing

$$\begin{aligned}\sigma_h(\mathbf{A}) &= \|\mathbf{X}_s - \mathbf{AA}'\|^2 + \|\mathbf{X}_k - \mathbf{AKA}'\|^2 \\ &= \|\mathbf{X}_s + \mathbf{X}_k - (\mathbf{AA}' + \mathbf{AKA}')\|^2 \\ &= \|\mathbf{X}_s + \mathbf{X}_k - (\mathbf{A}(\mathbf{I} + \mathbf{K})\mathbf{A}')\|^2\,,\end{aligned} \tag{27}$$

using the fact that $tr(\mathbf{X}_s - \mathbf{AA}')(\mathbf{X}_k - \mathbf{AKA}') = 0$, because it is the trace of a product of a symmetric and a skew-symmetric matrix. If we denote $(\mathbf{I} + \mathbf{K})$ as $\mathbf{R}$, it follows that minimizing σ_h is equivalent to DEDICOM, subject to the constraint that $\mathbf{R} = \mathbf{I} + \mathbf{K}$; hence the Escoufier and Grorud method is a constrained variant of DEDICOM. It is also a constrained variant of Generalized GIPSCAL with c set equal to zero, since the latter fits the model $\mathbf{X} = \mathbf{AA}' + \mathbf{A\Delta A}' + \mathbf{E} = \mathbf{A}(\mathbf{I} + \Delta)\mathbf{A}' + \mathbf{E}$, with Δ containing 2×2 blocks $\begin{bmatrix} 0 & \delta_l \\ -\delta_l & 0 \end{bmatrix}$ along the diagonal, for arbitrary δ_l. Clearly, the Escoufier and Grorud method is the constrained variant of Generalized GIPSCAL with the scalars $\delta_1,\ldots,\delta_q$ fixed to unity. In conclusion, we have shown that the Escoufier and Grorud method is a constrained variant of Generalized GIPSCAL, which in turn is a constrained variant of DEDICOM.

Incidentally, it should be mentioned that Chino has also established certain relations between Escoufier and Grorud's method and GIPSCAL. Chino (1991) has recently pointed out that Escoufier and Grorud's method has a finite-dimensional complex Hilbert space structure if all (nonzero) eigenvalues of $\mathbf{X}_s + i\mathbf{X}_k$ are positive, and Chino and Shiraiwa (1993) established the corresponding metric properties for this method as well as for GIPSCAL and DEDICOM.

We have shown that the Escoufier and Grorud (1980) method is a constrained version of Generalized GIPSCAL. However, there seems to be no data analytic reason for constraining the Generalized GIPSCAL model in this particular way. The main use of the Escoufier and Grorud method then seems to rest in the fact that it gives a closed-form solution, whereas both GIPSCAL and DEDICOM require iterative procedures to find a solution. The Escoufier and Grorud method may serve as a useful starting configuration for the iterative algorithms used in Generalized GIPSCAL and DEDICOM.

8. Example

As an example, we reanalyzed the car switching data for 16 car types reported by Harshman, Green, Wind, and Lundy (1982). Each cell (i,j) in this 16×16 data matrix pertains to the frequency with which type i car owners switch to a new type j car. The car types are Subcompact/Domestic (SUBD), Subcompact/Captive Imports (SUBC), Subcompact/Imports (SUBI), Small Specialty/Domestic (SMAD), Small Specialty/Captive Imports (SMAC), Specialty/Imports (SMAI), Low Price Compact (COML), Medium Price Compact (COMM), Import Compact (COMI), Midsize Domestic (MIDD), Midsize Imports (MIDI), Midsize Specialty (MIDS), Low Price Standard (STDL), Medium Price Standard (STDM), Luxury Domestic (LUXD), and Luxury Import (LUXI). For more details the reader is referred to Harshman et al. (1982).

We first obtained (using the Kiers, Ten Berge, Takane, and De Leeuw 1990 DEDICOM algorithm) the two-, three- and four-dimensional DEDICOM solutions for these data, and found that the solutions accounted for 77.2%, 86.4%, and 92.0%, respectively, of the total sum of squares of the data matrix. In all three cases the symmetric part of $\mathbf{R}$ turned out to be p.d., so in all these cases DEDICOM is equivalent to Generalized GIPSCAL. Using this equivalence, we can plot the first and second bimensions as in Generalized GIPSCAL, and interpret the results on the basis of these plots. We therefore transformed the DEDICOM solution into 'GIPSCAL form' and plotted the coordinates ($\tilde{\mathbf{A}}_1$ in (9)) for the most contributing car types on the first two dimensions in Figure 1, and on the third and fourth dimensions ($\tilde{\mathbf{A}}_2$ in (9)) in Figure 2.

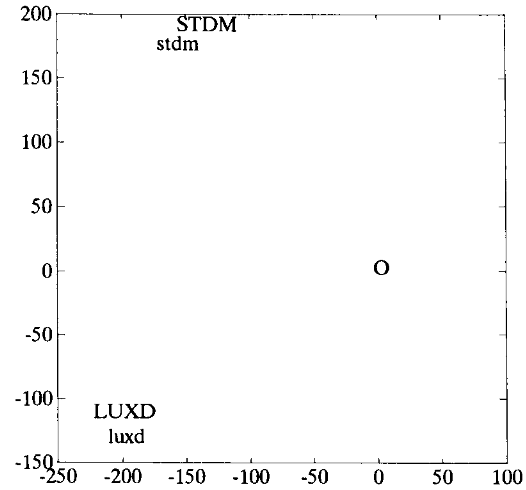

Figure 1. Plot for the first bimension of the car switching data. Explanation of the labels: SUBD: Subcompact/Domestic; SUBC: Subcampact/Captive Imports; SUBI: Subcompact/Imports; SMAD: Small Specialty/Domestic; SMAC: Small Specialty/Captive Imports; SMAI: Specialty/Imports; COML: Low Price Compact; COMM: Medium Price Compact; COMI: Import Compact; MIDD: Midsize Domestic; MIDI: Midsize Imports; MIDS: Midsize Specialty; STDL: Low Price Standard; STDM: Medium Price Standard; LUXD: Luxury Domestic; LUXI: Luxury Import. Lower case labels refer to rows of the data table, upper case labels refer to columns.

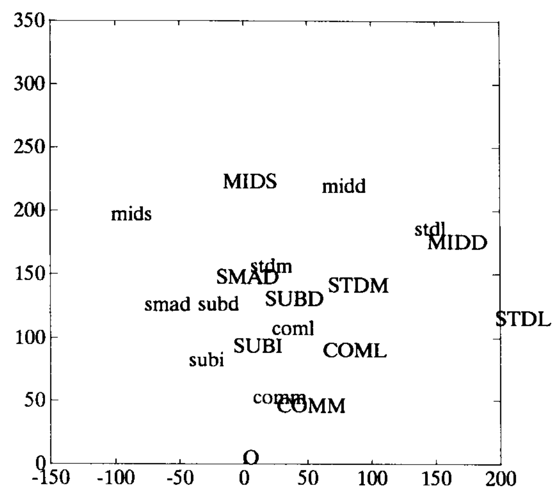

Figure 2. Plot for the second bimension of the car switching data. For the explanation of the labels, see Figure 1.

From the bimension weights δ_1 (.4479) and δ_2 (.0904), we computed the rotation angles $\phi_1 = 24.1°$ and $\phi_2 = 5.2°$. Clearly, the second bimension does not make an important contribution to the skew-symmetric part of the data. The first bimension's contribution to skew-symmetry is considerably larger, although this bimension also contributes more to symmetry than to skew-symmetry. This result is an implication of the fact that the sum of squares of the skew-symmetric part of the data is much smaller than the sum of squares of the symmetric part.

To interpret results for the first bimension, we rotated the configuration over 24.1° and superimposed the resulting configuration ($\tilde{\mathbf{B}}_1$ in (9); with labels capitalized) on the original configuration (with lower case labels). As represented by this bimension, the largest (approximations of) car switching frequencies (i.e., the largest inner products between vectors emanating from the origin to the plotted points) are found for pairs of car types that are far from the origin, and close to each other, like the pairs (stdl,MIDD), (stdl,STDL), (stdl,MIDS), (mids,MIDS), (midd,MIDS), and (midd,MIDD), which involve the car types with the largest market shares: MIDD, MIDS, and STDL; car switchings of medium frequency involve SUBD, SMAD, and STDM. So the first bimension mainly reflects car switchings between these car types. Of these, the car types 'midd' and 'stdl' are displayed far to the right, near many capitalized versions of other car types. Apparently, there is a considerable amount of car switchings from these car types to others, for instance, from 'midd' to MIDS, SMAD, SUBD, STDM, and MIDD, and from 'stdl' to MIDD, STDM, SUBD, COML, and MIDS. These switchings reflect a general tendency (also reported by Harshman et al.) of owners of medium sized or standard cars to switch to small and/or specialty cars. In line with this finding is that the main 'receivers' (the car types whose capitalized versions are far to the left) turn out to be the MIDS specialty cars and to some extent the small SMAD cars, to which many switches occur, for instance, from 'mids', 'midd', 'smad', 'subd', and 'stdm'. It is thus possible to read the (approximated) car switching frequencies from the plot.

To find the largest asymmetric relations (i.e., with largest difference between cells (i,j) and (j,i)), it is easier to use the original GIPSCAL interpretation, based on triangles' areas in the original (or the superimposed) configuration. It can be seen then that the largest triangle is formed by the stimulus pair (STDL,MIDS), and other large triangles are found for (STDL,MIDD), (STDL,SMAD), and (MIDD,MIDS), thus indicating relatively large differences in 'car switchings to' and 'car switchings from'. These large asymmetries reflect real asymmetries in the data, except the asymmetry between STDL and MIDD, which occurred as a result of a sizable modeling error for this stimulus pair. It should be mentioned that, apart from this large modeling error, we found only one similarly large residual, namely for the (symmetric) relation of STDL to itself. The (real) largest asymmetries again display the tendency to switch to smaller and/or specialty cars.

To interpret the results from the second bimension, we rotated the configuration for the second bimension over 5.2° and superimposed the resulting configuration ($\tilde{\mathbf{B}}_2$ in (9)) on the original configuration for the second bimension. The second bimension mainly represents the relations between car types STDM and LUXD; all other car types were located too close to the origin and to each other to display them separately. The inner products between 'stdm' and LUXD and between 'luxd' and STDM are both fairly large, indicating that the relation between these stimuli is mainly symmetric. As far as there is any asymmetry, we can see that 'stdm' is closer to LUXD than 'luxd' is to STDM, and hence that there are more car switchings from 'stdm' to LUXD than from 'luxd' to STDM, but as mentioned, this asymmetry is small compared to the symmetric relations implied by car switchings within types STDM and LUXD. This interpretation is in accordance with the data table telling that there were 21,974 switches from type STDM to type LUXD, and only 9,187 in the other direction, but compared to switches from STDM to STDM (81,808) and from LUXD to LUXD (63,509), this amount of asymmetric car switchings is almost negligible.

If we had interpreted the plot for the second bimension on the basis of triangle areas, we would have inferred that there is a sizable asymmetry between the car switchings from 'luxd' to STDM and those from 'stdm' to LUXD. This observation may seem to contradict the above (biplot based) interpretation, holding that the car switchings between LUXD and STDM are mainly symmetric. However, it should be noted that the size of the triangle area approximates (up to a scalar multiplication) the difference (12,787) between the amount of switches from STDM to LUXD (21,974) and from LUXD to STMD (9,187), which is indeed a large difference compared to what is found for other pairs of car types. Nevertheless, this sizable difference is small compared to the (very large) total amount of switches from LUXD and STDM to LUXD or STDM (176,478), and it should be concluded that, globally, the majority of switches *from* 'luxd' (to LUXD or STDM) is accompanied by switches *to* LUXD (from 'luxd' or 'stdm'), and a similar result holds for STDM. Hence, the car switchings between LUXD and/or STDM take place mainly symmetrically, as was already concluded from the biplot type of interpretation of Figure 2. We have thus demonstrated that interpretation on the basis of triangle areas alone is somewhat hazardous. Such an interpretation should be accompanied by an interpretation of the symmetric relations between the car types, and the contributions of these should be weighted in the ratio $2\delta_l$:1, where δ_l is from the solution of minimizing (7). The fact that for the present bimension $\delta_l = .09$ confirms that the skew-symmetric part has very little importance indeed. The advantage of the biplot approach is that such an a posteriori weighting of contributions is not necessary, because the necessary information is already reflected in the plot.

9. Discussion

The idea of plotting two configurations for the stimuli in which one is a rotation of the other is rather similar to the slide vector model suggested by Kruskal (1973, personal communication to De Leeuw; see De Leeuw and Heiser 1982, Gower and Zielman 1992). In this model, nonsymmetric data are modeled by distances between points of two configurations, one for the rows and one for the columns. As in our interpretation of Generalized GIPSCAL, the configurations are the same but located differently. Whereas in Generalized GIPSCAL the two configurations differ by a rotation from each other, in the slide vector model, the two configurations differ by a translation (slide vector) from each other. After having been ignored for a long time, the slide vector model has recently been revisited and provided with an algorithm by Zielman and Heiser (1993).

We have shown that Generalized GIPSCAL is equivalent to DEDICOM if the symmetric part of **R** (in DEDICOM) is p.d. This condition was met in the example analyzed here, for all three dimensionalities. Of course, it may happen that this condition is not met. Then, we may still have 'near equivalence' of DEDICOM and Generalized GIPSCAL. To check for near equivalence, we suggest applying both methods and comparing the goodness-of-fit values. If they differ only slightly, the solutions can be called nearly equivalent; if the values differ considerably, the solutions can be concluded to be 'clearly nonequivalent'.

When DEDICOM and Generalized GIPSCAL are equivalent, the DEDICOM representation is of a very special kind: $\hat{\mathbf{X}} = \Sigma_l \beta_l \mathbf{A}_l \mathbf{T}'_l \mathbf{A}'_l = \mathbf{ARA}'$, where **R** is the block-diagonal matrix with 2×2 blocks $\beta_l \mathbf{T}'_l$ along the diagonal. The Generalized GIPSCAL representation is, apparently, based on transforming **A** and **R** such that **R** becomes a block-diagonal matrix that is row- and columnwise orthogonal.

In the data analysis reported above, the representation was dominated by the car types with large market shares. This result stems from interpreting 'relations' by the amounts of car switchings, and these amounts are obviously largest for the most prevalent car types. DEDICOM aims at minimizing the sum of squared modeling errors; hence the model will focus on an optimal representation of the car types with large market shares and will model car switchings involving underrepresented cars only as far as doing so fits in with the mainstream of car switchings. Therefore, in the interpretation, car switchings involving car types like SUBC, SMAC, SMAI, COMM, COMI, MIDI, and LUXI were ignored. When more information is desired on such car types, it seems necessary to decrease the large size differences in the reported car switching frequencies. One might, for instance, preprocess the data so that the rows sum to 1, or take logarithms, to mention two possible transformations. However, in doing so, one actually analyzes derived measures, and in interpreting the results one should take this fact into account. For further discussion of this problem, see, for instance, DeSarbo and De Soete (1984, p.602).

Chino (1990) suggested preprocessing **X** by centering **X** row- and columnwise. If this preprocessing is done, the symmetric part of **X** is represented exactly according to the classical MDS approach, and may hence be seen as a procedure for representing the elements of $(-2\mathbf{X})$ by squared distances between the associated points in the plot. It is, however, doubtful if the symmetric part of $(-2\mathbf{X})$ can indeed often be considered as a matrix of squared distances between stimuli.

References

CARROLL, J. D., and CHANG, J.-J. (1970), "Analysis of Individual Differences in Multidimensional Scaling Via an n-way Generalization of 'Eckart-Young' Decomposition," *Psychometrika, 35*, 283-319.

CHINO, N. (1978), "A Graphical Technique for Representing Asymmetric Relationships Between n Objects," *Behaviormetrika, 5*, 23-40.

CHINO, N. (1980), "A Unified Geometrical Interpretation of the MDS Techniques for the Analysis of Asymmetry and Related Techniques," Paper presented at the Symposium on Asymmetric Multidimensional Scaling at the Spring Meeting of the Psychometric Society, Iowa, May, 1980.

CHINO, N. (1990), "A Generalized Inner Product Model for the Analysis of Asymmetry," *Behaviormetrika, 27*, 25-46.

CHINO, N. (1991), "A Critical Review for the Analysis of Asymmetric Relational Data," *Bulletin of the Faculty of Letters of Aichi Gakuin University, 21*, 31-52.

CHINO, N., and SHIRAIWA, K. (1993), "Geometrical Structures of Some Non-distance Models for Asymmetric MDS," *Behaviormetrika, 20*, 35-47.

CONSTANTINE, A.G., and GOWER, J.C. (1978), "Graphical Representations of Asymmetric Matrices," *Applied Statistics, 27*, 297-304.

DE LEEUW, J., and HEISER, W.J. (1982), "Theory of Multidimensional Scaling with Restrictions on the Configuration," in *Handbook of Statistics*, Vol.2, Eds., P.R. Krishnaiah, and L.N. Kanal, Amsterdam: North Holland, 285-316.

DESARBO, W.S., and DE SOETE, G. (1984), "On the Use of Hierarchical Clustering for the Analysis of Nonsymmetric Proximities," *Journal of Consumer Research, 11*, 601-610.

ECKART, C., and YOUNG, G. (1936), "The Approximation of One Matrix by Another of Lower Rank," *Psychometrika, 1*, 211-218.

ESCOUFIER, Y., and GRORUD, A. (1980), "Analyse factorielle des matrices carrées non symmetriques [Factor analysis of asymmetric square matrices]," in *Data Analysis and Informatics*, Eds., E. Diday, L. Lebart, J.P. Pagès, and R. Tomassone, Amsterdam: North Holland, 263-276.

GABRIEL, K.R. (1971), "The Biplot Graphic Display of Matrices with Application to Principal Components Analysis," *Biometrika, 58*, 453-467.

GOWER, J.C. (1966), "Some Distance Properties of Latent Root and Vector Methods Used in Multivariate Analysis," *Biometrika, 53*, 325-338.

GOWER, J.C. (1977), "The Analysis of Asymmetry and Orthogonality," in *Recent Developments in Statistics*, Eds., J.R. Barra, F. Brodeau, G. Romier, and B. van Cutsem, Amsterdam: North Holland, 109-123.

GOWER, J.C. and ZIELMAN, B. (1992), "Some Remarks on Orthogonality in the Analysis of Asymmetry," Research Report RR-92-08, University of Leiden, Department of Data Theory.

HARSHMAN, R.A. (1978), "Models for Analysis of Asymmetrical Relationships Among N Objects or Stimuli," Paper presented at the First Joint Meeting of the Psychometric Society and the Society for Mathematical Psychology, Hamilton, Ontario, August.

HARSHMAN, R.A. (1981), "DEDICOM Multidimensional Analysis of Skew-symmetric Data. Part 1: Theory," Unpublished technical memorandum, Bell Laboratories, Murray Hill, NJ.

HARSHMAN, R.A., GREEN, P.E., WIND, Y., and LUNDY, M.E. (1982), "A Model for the Analysis of Asymmetric Data in Marketing Research," *Marketing Science, 1*, 205-242.

KIERS, H.A.L., TEN BERGE, J.M.F., TAKANE, Y., and DE LEEUW, J. (1990), "A Generalization of Takane's Algorithm for DEDICOM," *Psychometrika, 55*, 151-158.

TEN BERGE, J.M.F. (1991), "A General Solution for a Class of Weakly Constrained Linear Regression Problems," *Psychometrika, 56*, 601-609.

TORGERSON, W.S. (1958), *Theory and Methods of Scaling*, New York: Wiley.

ZIELMAN, B., and HEISER, W.J. (1991), "Spatial Representations of Asymmetric Proximities," Research Report RR-91-10, University of Leiden, Department of Data Theory.

ZIELMAN, B., and HEISER, W.J. (1993), "Analysis of Asymmetry by a Slide Vector," *Psychometrika, 58*, 101-114.