RR 93-01

TWO METHODS FOR MULTIDIMENSIONAL ANALYSIS OF THREE-WAY SKEW-SYMMETRIC MATRICES

Berrie Zielman

Department of Data Theory
University of Leiden

Two Methods for Multidimensional Analysis of Three-way Skew-symmetric Matrices

Berrie Zielman

Department of Data Theory, University of Leiden, P.O Box 9555

2300 RB Leiden

Abstract

In this paper we propose a method for studying three-way skew-symmetric data. The method is a generalization of the singular value decomposition of a single skew-symmetric matrix. The model consists of a common space that has the same interpretation as a diagram derived from the SVD on a two-way matrix. The differences between the slabs are described by a transformation matrix indicating the departure of a replication from the common space. The skew-symmetry is reflected in the model by a direction in the diagram. An alternating least squares algorithm is presented for estimating the parameters of the model. An example is given to illustrate the method.

Key words: asymmetry, skew-symmetry, singular value decomposition, three-way methods

1. Introduction

The INDSCAL model (Carroll and Chang, 1970) is an individual differences model that describes relations between several proximity-sources. Here, sources refer to experimental conditions, points in time, individuals, and so on. The deviation of a particular source from the common space is modelled by a linear transformation of the common space. In scalar product form the INDSCAL model for symmetric data can be written as:

$$\mathbf{S}_k = \mathbf{T}\mathbf{R}_k\mathbf{T}',$$

where $\mathbf{T}$ is an n by p matrix with coordinates, p denotes the number of dimensions, $\mathbf{R}_k$ is a diagonal matrix with dimension weights, that indicate the relative importance or salience of the dimensions for replication k (k=1,..., m), and m is the number of sources. The IDIOSCAL model has a similar structure as the INDSCAL model. The matrix $\mathbf{T}$ is defined as before, but here the matrix $\mathbf{R}_k$ is a symmetric matrix.

In this paper we consider fitting models to skew-symmetric data that are analogues of the INDSCAL and IDIOSCAL models for symmetric data. Skew-symmetry may arise when we collect asymmetric proximity data Δ_k, and we apply the decomposition

$$\Delta_k = \mathbf{S}_k + \mathbf{A}_k,$$

where the matrix $\mathbf{S}_k$ is symmetric, and the matrix $\mathbf{A}_k$ describes the departures from symmetry and is skew-symmetric: $a_{ijk} = -a_{jik}$. The symmetric matrix is obtained by averaging the corresponding elements across the diagonal, whereas the skew-symmetric matrix is obtained from the equation $\mathbf{A}_k = \Delta_k - \mathbf{S}_k$,

The symmetric part can be analyzed by a symmetric model such as MDS or cluster analysis; the matrix that contains the departures from symmetry can be studied by singular value decomposition (SVD). The matrices $\mathbf{A}_k$ and $\mathbf{S}_k$ are orthogonal, in the sense that the sum of squares of the matrix Δ_k can be decomposed into sum of squares due to symmetry and sum of squares due to skew-symmetry

$$\Sigma_i \Sigma_j \, \delta_{ij}^2 = \Sigma_i \Sigma_j \, s_{ij}^2 + \Sigma_i \Sigma_j \, a_{ij}^2$$

where the subscripts k have been dropped for convenience. Because of this split of sum of squares, the two components can be viewed independently.

A situation where skew-symmetry may arise is data from a psychological choice experiment. In psychological choice modelling π_{ij} denotes the probability that object i is preferred over object j; $\pi_{ji} = 1 - \pi_{ij}$ denotes the complement of that probability. If these probabilities are transformed by the inverse of the standard normal distribution, the transformed probabilities are skew-symmetric. Methods for analyzing these transformed probabilities were originally proposed by Thurstone (1929). Starting with Carroll (1981), Heiser and De Leeuw (1981) and Takane (1980) the number of multidimensional models for skew-symmetric choice tables has been growing; some examples can be found in De Soete, Feger and Klauer (1989)

Gower (1977) and Constantine and Gower (1978) studied the singular value decomposition of a skew-symmetric matrix. No INDSCAL or IDIOSCAL method seems to be available for skew-symmetric data. In this paper we will develop IDIOSCAL and INDSCAL analogues of the method by Gower (1977) and Constantine and Gower (1978) for skew-symmetric matrices. The distance or scalar product interpretation of the diagram in ordinary INDSCAL and IDIOSCAL is replaced by the area interpretation. An algorithm for fitting both models is presented, and an application is given.

2. Singular value decomposition of skew-symmetric matrices

The departures from symmetry can be analyzed by computing the singular value decomposition (SVD) of the skew-symmetric matrix **A**. The properties of the SVD of a skew-symmetric matrix will be discussed after we have the discussed their properties for an arbitrary matrix **B**. The SVD of any matrix **B** can be written as:

$$\mathbf{B} = \mathbf{K}\,\mathbf{\Phi}\,\mathbf{L}',$$

where $\mathbf{K}'\mathbf{K} = \mathbf{L}'\mathbf{L} = \mathbf{I}$, and $\mathbf{\Phi}$ is a diagonal matrix containing the singular values in descending order. The matrices **K** and **L** contain the left and right singular vectors.

The SVD of a skew symmetric matrix can be written as:

$$\mathbf{A} = \mathbf{K}\,\mathbf{\Phi}\,\mathbf{J}\,\mathbf{K}',$$

where **J** is a block diagonal matrix with block matrices

$$\begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}.$$

The matrix containing the left singular vectors are column wise permutations followed by a reflection of one column of the pair of the right singular vectors. The singular values of a skew symmetric matrix come in pairs: $\phi_1 = \phi_2$; $\phi_3 = \phi_4$; $\phi_{n-1} = \phi_n$. If n is odd the last singular value is zero.

Gower has pointed out that the SVD of a skew-symmetric matrix has to be interpreted in terms of pairs of dimensions, this in contrast to principal components analysis of a correlation matrix where we can also interpret a single dimension. In scalar notation a pair of dimensions can be written as

$$a_{ij} = \phi_1(k_{i1}k_{j2} - k_{i2}k_{j1}),$$

where skew-symmetry is measured by the area between two points and the origin. This is shown in Figure 1 for two points P_i and P_j with coordinates (k_{i1}, k_{i2}) and (k_{j1}, k_{j2}) respectively. The area of the triangle formed by these two points and the origin is $1/2\{k_{j1}k_{i2}-k_{i1}k_{j2}\}$.

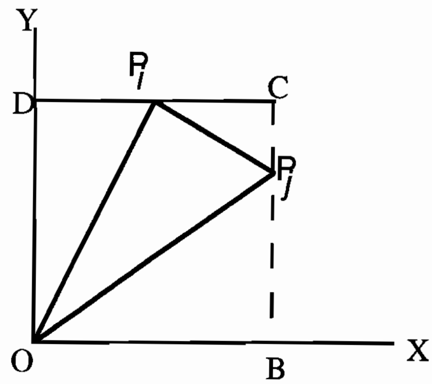

Figure 1: Area as a measure of skew-symmetry

The area of the triangle OP_iP_j can be calculated by subtracting the area of the triangles OBP_j, CP_iP_j and ODP_i from the area of the rectangle OBCD in Figure 1. Specifically, the area of the triangle is given by

$$
\begin{aligned}
OP_iP_j &= k_{j1}k_{i2} - 1/2\{k_{j1}k_{j2} + k_{i1}k_{i2} + (k_{j1} - k_{i1})(k_{i2}-k_{j2})\} \\
&= k_{j1}k_{i2} - 1/2\{k_{j1}k_{i2} + k_{i1}k_{j2}\} \\
&= 1/2\{k_{j1}k_{i2}-k_{i1}k_{j2}\}.
\end{aligned}
$$

A hypothetical example of an interpretation is provided in Figure 2, where four objects A, B, C, D are depicted. The skew-symmetry between A and B equals twice the area of the triangle with vertices OA and OB, this area is dark shaded. If we go counter clockwise, that is from B to A, the area is positive and if we go clockwise from A to B the area is negative, thus modelling skew-symmetry. In this diagram the sign of skew-symmetry is modelled by a direction, hence the name directional plane. The points B and D are collinear on a line with the origin giving zero area, this corresponds to symmetry between objects B and D.

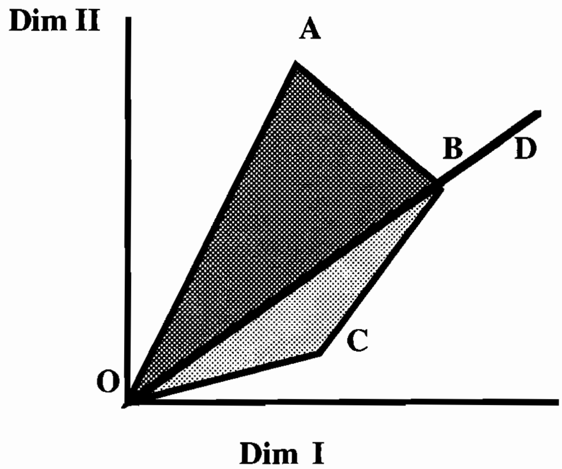

Figure 2: Interpretation of a directional plane in terms of areas

So far, we have restricted our attention to the interpretation of two-dimensional models. The interpretation of models with more dimensions is similar to the two-dimensional case. When we have more than one pair of dimensions, skew-symmetry is measured by additive contributions of different pairs of dimensions.

A special case of this model arises when all points are on a straight line. In this particular case the areas are proportional to the basis of the triangles, giving a linear form of skew-symmetry that can be written as:

$$\mathbf{A} = \mathbf{e}\mathbf{y}' - \mathbf{y}\mathbf{e}',$$

where one singular vector is constrained to a vector of ones. Up to a monotonic transformation this linear model is also known under the name Thurstone Case-V model.

3. Directional planes representing individual differences

The skew-symmetric generalization of the INDSCAL model that we wish to propose here is:

$$\mathbf{A}_k = \mathbf{T}\mathbf{U}_k\mathbf{J}\mathbf{T}', \tag{1}$$

where $\mathbf{T}$ is an n by p matrix with coordinates and $\mathbf{U}_k$ is a diagonal matrix with directional plane weights coming in pairs. These directional plane weights indicate the relative importance or salience of the directional plane for replication k. The weights can be negative or positive; unlike the symmetric INDSCAL method where negative weights cannot be interpreted. Here, negativity of the weights is not a problem, because it means that the direction of asymmetry is reversed, thus all positive skew-symmetries become negative and all negative skew-symmetries become positive. The matrix $\mathbf{J}$ takes care of the skew-symmetry of the model and has the same block-diagonal structure as for the two-way model. Obviously, the model is an INDSCAL generalization for skew-symmetric matrices, the only difference between INDSCAL for symmetric data and the skew-symmetric variant is the matrix $\mathbf{J}$.

The IDIOSCAL model that we wish to propose for skew-symmetric data can be written as

$$\mathbf{A}_k = \mathbf{T}\mathbf{U}_k\mathbf{T}', \tag{2}$$

where $\mathbf{T}$ is defined as before, and $\mathbf{U}_k$ is now a p by p skew-symmetric matrix. When only one pair of dimensions is fitted, the two models are identical, because $\mathbf{U}_k$ is of order 2 by 2, consisting of only one diagonal block.

Interpretation of the INDSCAL model can be done in a similar way as the two-way case, with the precaution of keeping the scales of the diagrams, which are recorded in the weight matrix $\mathbf{U}_k$ in mind. For IDIOSCAL, the situation is more complicated because $\mathbf{U}_k$ is square and skew-symmetric, and we have to interpret the $n(n\text{-}1)/2$ elements as well. Interpretation may continue by plotting all possible pairs of dimensions of $\mathbf{T}$, where we may still interpret the areas between points and the origin, because the sum of these pairs of dimensions yield the predicted values by the model. However, when the dimensionality is large we are facing the problem of interpreting a large number of diagrams.

Another possibility for interpreting the IDIOSCAL model, is to analyze the matrix $\mathbf{U}_k$ by SVD. Because $\mathbf{U}_k$ is skew-symmetric, the diagrams obtained by plotting the singular vectors have the same properties as discussed in section 2. Now, the diagram is interpreted in terms of relations between dimensions instead of relations between objects. The success of this type of interpretation depends on the clarity of the dimensions. If we cannot find an interpretation for the dimensions, the interpretation of diagrams with relations between dimensions will not lead to anything useful.

4. Fitting the models

The parameters of the INDSCAL and IDIOSCAL model for skew-symmetric data can be estimated by minimizing the least squares loss function

$$\mathrm{f}(\mathbf{T};\, \mathbf{U}_k) \qquad = \Sigma_k \| \mathbf{A}_k - \mathbf{T}\mathbf{J}\mathbf{U}_k\mathbf{T}' \|^2.$$

This loss function is also written as

$$\mathrm{f}(\mathbf{z}_s;\mathbf{y}_s;u_{ks}) \qquad = \Sigma_k \| \mathbf{A}_k - \sum_{s=1}^{q} u_{ks}(\mathbf{z}_s\mathbf{y}_s' - \mathbf{y}_s\mathbf{z}_s') \|^2, \qquad (3)$$

where the matrix $\mathbf{J}$ is absorbed in the new notation, $\mathbf{z}_s$ denotes an odd column $\mathbf{t}_s$ of $\mathbf{T}$ and $\mathbf{y}_s$ denotes the adjacent column $\mathbf{t}_{s+1}$ of $\mathbf{T}$, and the scalar u_{ks} denotes the sth diagonal element of $\mathbf{U}_k$. Recall that the diagonal elements of $\mathbf{U}_k$ come in pairs thus u_{ks}

equals u_{ks+1}.This notation shows that the elements of $\mathbf{U}_k$ can be interpreted as regression weights indicating the relative importance or salience of the skew-symmetry modelled by the sth directional plane for the kth source.

Loss function (3) can be minimized by using an alternating least squares algorithm, where parameters are divided in subsets and the parameters of one subset are improved keeping the other subsets fixed. Instead of estimating row and column parameters and the weights separately, as is usually done in ALS-algorithms, the present algorithm estimates in one substep the first dimension for rows and the second dimension for columns and in the second substep the first dimension for columns and the second dimension for rows. The algorithm alternates between parameter subsets, and also over directional planes. First, loss function (3) is rewritten in such a way that (3) becomes a function of one pair of dimensions, regarding the other pairs fixed, as

$$\begin{aligned} \mathrm{f}(\mathbf{z}_r;\mathbf{y}_r;u_{kr}) &= \Sigma_k \|\mathbf{A}_k - \sum_{s\neq r}^{q} u_{ks}(\mathbf{z}_s\mathbf{y}_s' - \mathbf{y}_s\mathbf{z}_s') - u_{kr}(\mathbf{z}_r\mathbf{y}_r' - \mathbf{y}_r\mathbf{z}_r')\|^2 \\ &= \Sigma_k \| \mathbf{A}_{kr} - u_{kr}(\mathbf{z}_r\mathbf{y}_r' - \mathbf{y}_r\mathbf{z}_r')\|^2 \end{aligned}$$

with $\mathbf{A}_{kr}$ defined as $\mathbf{A}_k - \sum_{s\neq r}^{q} u_{ks}(\mathbf{z}_s\mathbf{y}_s' - \mathbf{y}_s\mathbf{z}_s')$. Using the Kronecker product and Vec notation $\mathrm{f}(\mathbf{z}_r;\mathbf{y}_r;u_{kr})$ can be written as

$$\begin{aligned} \mathrm{f}(\mathbf{z}_r;\mathbf{y}_r;u_{kr}) &= \Sigma_k \| \mathrm{Vec}\,(\mathbf{A}_{kr}) - \mathrm{Vec}\,(u_{kr}(\mathbf{z}_r\mathbf{y}_r' - \mathbf{y}_r\mathbf{z}_r'))\|^2 \\ &= \Sigma_k \| \mathrm{Vec}\,(\mathbf{A}_{kr}) - u_{kr}\{(\mathbf{y}_r \otimes \mathbf{I}) - (\mathbf{I} \otimes \mathbf{y}_r)\}\mathbf{z}_r\|^2 \\ &= \Sigma_k \| \mathrm{Vec}\,(\mathbf{A}_{kr}) - u_{kr}\mathbf{P}(\mathbf{y}_r)\mathbf{z}_r\|^2, \end{aligned} \tag{4}$$

where $\mathbf{P}(\mathbf{y}_r)$ is defined as $(\mathbf{y}_r \otimes \mathbf{I}) - (\mathbf{I} \otimes \mathbf{y}_r)$, $\mathbf{I}$ is the identity matrix of order n, and the symbol $\otimes$ denotes the Kronecker product of matrices. For fixed $\mathbf{y}_r$ and $\mathbf{u}_{kr}$, the global minimum of (4) is attained by

$$\mathbf{z}_r = \{\mathrm{P}'(\mathbf{y}_r)\Sigma_k u_{kr}^2\, \mathrm{P}(\mathbf{y}_r)\}^{-}\mathrm{P}'(\mathbf{y}_r)\mathrm{Vec}\,(\mathbf{A}_{kr}u_{kr}). \tag{5}$$

The matrix $\{\mathbf{P}'(\mathbf{y}_r)u\backslash O(^{2},_{kr}\ )\mathbf{P}(\mathbf{y}_r)\}^-$ denotes the generalized inverse of the matrix $\{\mathbf{P}'(\mathbf{y}_r)u_{kr}^2\,\mathbf{P}(\mathbf{y}_r)\}$, that can be calculated by

$$\{\mathbf{P}'(\mathbf{y}_r)\Sigma_k\, u_{kr}^2\,\mathbf{P}(\mathbf{y}_r) + \mathbf{y}_r(\mathbf{y}'_r\mathbf{y}_r)^{-1}\mathbf{y}'_r\ \}^{-1} - \mathbf{y}_r(\mathbf{y}'_r\mathbf{y}_r)^{-1}\mathbf{y}'_r,$$

because the vector $\mathbf{y}_r$ spans the null-space of the matrix. Given the current estimates of $\mathbf{y}_r$ and $\mathbf{z}_r$ and assuming the inverse exists, the importance weights u_{kr} are given by:

$$u_{kr} = \{\mathbf{z}'_r\,\mathbf{P}'(\mathbf{y}_r)\mathbf{P}(\mathbf{y}_r)\mathbf{z}_r\ \}^{-1}\mathbf{z}'_r\,\mathbf{P}'(\mathbf{y}_r)\mathrm{Vec}\ \mathbf{A}_{kr}. \qquad (6)$$

Loss function (4) can also be written as

$$\begin{aligned} f(\mathbf{z}_r;\mathbf{y}_r;u_{kr}) \quad &= \Sigma_k \|\ \mathrm{Vec}\ (\mathbf{A}_{kr}) - \mathrm{Vec}\ (u_{kr}(\mathbf{z}_r\mathbf{y}_r' - \mathbf{y}_r\mathbf{z}_r'))\|^2 \\ &= \Sigma_k \|\ \mathrm{Vec}\ (\mathbf{A}_{kr}) - u_{kr}\{(\mathbf{I}\otimes\mathbf{z}_r) - (\mathbf{z}_r\otimes\mathbf{I})\}\mathbf{y}_r\|^2 \\ &= \Sigma_k \|\ \mathrm{Vec}\ (\mathbf{A}_{kr}) - u_{kr}\mathbf{Q}(\mathbf{z}_r)\mathbf{y}_r\|^2, \end{aligned}$$

where $\mathbf{Q}(\mathbf{z}_r)$ is defined as $(\mathbf{I}\otimes\mathbf{z}_r)$ - $(\mathbf{z}_r\otimes\mathbf{I})$. Because $\mathbf{Q}(\mathbf{z}_r)$=-$\mathbf{P}(\mathbf{z}_r)$, $\mathbf{y}_r$ can be obtained for fixed $\mathbf{z}_r$ and $\mathbf{u}_{kr}$, by

$$\mathbf{y}_r = -\{\mathrm{P}'(\mathbf{z}_r)\Sigma_k u_{kr}^2\,\mathrm{P}(\mathbf{z}_r)\}^-\mathrm{P}'(\mathbf{z}_r)\mathrm{Vec}\ (\mathbf{A}_{kr}u_{kr}). \qquad (7)$$

The algorithm can now be described as

0. initialize u_{ks}, $\mathbf{y}_s$, and $\mathbf{z}_s$,
1. for r=1 compute $\mathbf{A}_{kr}$
2. update $\mathbf{z}_r$ using (5)
3. update u_{kr} using (6)
4. update $\mathbf{y}_r$ using (7)
5. set r=r+1 and repeat steps 2, 3 and 4 until r=q.

6. check for convergence; if the process converged stop, otherwise reset r to one and do step 2, 3, 4 and 5 again.

By every sub-step of the algorithm the function value of the loss function is decreased and the algorithm converges to at least a local minimum. The algorithm alternates between parameter subsets, and over directional planes.

For IDIOSCAL we may choose the same loss function, and now we have to take care that the matrix $\mathbf{U}_k$ is a skew-symmetric matrix of full rank. The algorithm for IDIOSCAL uses the same dimension wise approach, although the specific steps differ. We have to minimize

$$L(\mathbf{T};\mathbf{U}_k) = \Sigma_k \| \mathbf{A}_k - \mathbf{T}\mathbf{U}_k\mathbf{T}' \|^2,$$

where for fixed $\mathbf{T}$ the global minimum of $\mathbf{U}_k$ is given by

$$\mathbf{U}_k = (\mathbf{T}'\mathbf{T})^{-1}\mathbf{T}'\mathbf{A}_k\mathbf{T}(\mathbf{T}'\mathbf{T})^{-1}. \quad (8)$$

In (8), $\mathbf{U}_k$ is skew-symmetric because $\mathbf{A}_k$ is skew-symmetric. The parameters of the matrix $\mathbf{T}$ can be estimated dimension after dimension if we partition $\mathbf{T}$ into $\mathbf{t}_r$, the dimension we are updating, and $\mathbf{T}_{-r}$, the dimensions that are kept fixed. Similarly the skew-symmetric matrix $\mathbf{U}_k$ can be partitioned into

$$\mathbf{U}_k = \begin{pmatrix} u_{rrk} & \mathbf{u}_{-rk} \\ \mathbf{u}'_{-rk} & \mathbf{U}_{-rk} \end{pmatrix},$$

and writing the model in partitioned form, gives

$$\mathbf{t}_r u_{rrk}\, \mathbf{t}_r' + \mathbf{t}_r \mathbf{u}_{-rk}\, \mathbf{T}'_{-r} + \mathbf{T}_{-r}\mathbf{u}'_{-rk}\, \mathbf{t}_r + \mathbf{T}_{-r}\mathbf{U}_{-rk}\, \mathbf{T}'_{-r}$$

where the first term vanishes because a skew-symmetric matrix has zero diagonal, thus u_{rrk}=0. Similarly, we define matrices $\mathbf{A}_{kr} = \mathbf{A}_k - \mathbf{T}_{-r}\mathbf{U}_{-rk}\mathbf{T}'_{-r}$, and proceed by minimizing

$$L(\mathbf{t}_r;\mathbf{U}_k) = \Sigma_k \|\mathbf{A}_{kr} - \mathbf{t}_r\mathbf{u}_{-rk}\mathbf{T}'_{-r} + \mathbf{T}_{-r}\mathbf{u}'_{-rk}\mathbf{t}_r\|^2$$

for every dimension r. Here, a similar notation can be used for converting the model matrix into a vector, first define

$$\mathbf{P}(\mathbf{T}_{-i}\mathbf{U}) = (\mathbf{T}_{-i}\mathbf{u}_{-ik} \otimes \mathbf{I}) - (\mathbf{I} \otimes \mathbf{T}_{-i}\mathbf{u}_{-ik}),$$

since $\mathbf{u}_{-ik} = -\mathbf{u}'_{-ik}$. The update $\mathbf{t}_r$ can now be calculated by

$$\mathbf{t}_r = \{\Sigma_k\mathbf{P}'(\mathbf{T}_{-i}\mathbf{u}_{-ik})\mathbf{P}(\mathbf{T}_{-i}\mathbf{u}_{-ik})\}^{-}\Sigma_k\mathbf{P}'(\mathbf{T}_{-i}\mathbf{u}_{-ik})\text{Vec}(\mathbf{A}_k) \quad (9)$$

where the generalized inverse of $\{\Sigma_k\mathbf{P}'(\mathbf{T}_{-i}\mathbf{u}_{-ik})\mathbf{P}(\mathbf{T}_{-i}\mathbf{u}_{-ik})\}^{-}$ can be calculated as before using the null vector $\Sigma_k(\mathbf{T}_{-i}\mathbf{u}_{-ik})$.

After initialization of the matrices **T** and $\mathbf{U}_k$, the algorithm proceeds by

1. for r=1 compute $\mathbf{A}_{kr}$
2. update $\mathbf{t}_r$ using (9)
3. update $\mathbf{U}_k$ using (8)
4. set r=r+1 and repeat steps 1 and 2 until r=q.
5. check for convergence if the process converged stop, otherwise reset r at one and do step 1, 2, 3 and 4 again.

The reader may argue that existing methods such as CANDECOMP or PARAFAC for obtaining parameters can also be used. However, these algorithms estimate row and column parameters separately, which in the present case are required to be identical. There is no guarantee that the row parameters and column parameters are identical and that the diagrams obtained via these methods can be interpreted in terms of areas. This is a similar difficulty as may occur in applying an asymmetric method for obtaining a symmetric INDSCAL model (Ten Berge & Kiers, 1991).

6. Example: Differences in Social Mobility

To illustrate the three-way method, we re-analyzed data from a study by Kerckhoff, Campbell, and Winfield-Laird (1985). The data concern differences in social mobility between the United States and Great Britain. The American data came from the second Occupational changes in a Generation Survey (OCG II) and the British data were based on the Oxford Social Mobility survey. The analysis was restricted to men who were between the age of 25 and 64 years. Kerckhoff, Campbell, and Winfield-Laird (1985) adjusted the occupational categories to make the two tables comparable. The rows of the three-way array correspond to occupational category of the father; the columns to occupational category of the sons; the slices correspond to countries. A detailed description of the data can be found in Kerckhoff et al. (1985).

These tables, were the rows classify the occupations of the fathers and the columns the occupations of the sons, are likely to be asymmetric; usually a different number of doctors have sons that are farmers than that farmers have sons that are doctors. Differences in mobility are usually ascribed to different value orientations in the two countries, the Americans are more achievement oriented whereas British are more ascription oriented. Furthermore, America experienced relatively more inflow of immigrants at the turn of the century, which started at the bottom of the social prestige ladder. Finally the two school systems differ, where the British system is said to produce more intergenerational continuity compared to the American system.

From the tables reported by Kerckhoff et al. (1985) the skew-symmetric tables are easily constructed. By decomposing a social mobility table into a symmetric and skew-symmetric part, we can focus the analysis to the skew-symmetric part, which can be interpreted as net mobility, this in contrast to decomposing such a table into structural and circular mobility as is usually done. Structural mobility corresponds to mobility due to different margins. Rather than controlling for differences in marginal distributions, the present method studies how these differences are built up. After all, these differences in marginal distribution affect structure in society, which is precisely the reason for studying mobility tables. Changes in a society can be better seen by analyzing the skew-symmetric part of the table, because this part consists of pure changes.

The twelve occupational categories were: Professional (self-employed), Professional (salaried), Manager, Proprietor, Sales, Clerical, Craftsmen (foreman), Operative, Service, Labourer, Farmer and Farm labourer.

Prior to the analysis the tables were normalized to unit sum of squares for each country to make them comparable. We only report results from the analysis of the father-sons-first occupation table; results from father to sons second occupation and sons first to second occupation will be reported elsewhere. The analysis was done by specifying 1, 2, 3, and 4 planes, and the fit did not improve by specifying more than two planes. Two planes provided an adequate representation of the data. The first plane accounting for 67 percent of the variation in the skew-symmetric part of the table is shown in Figure 3.

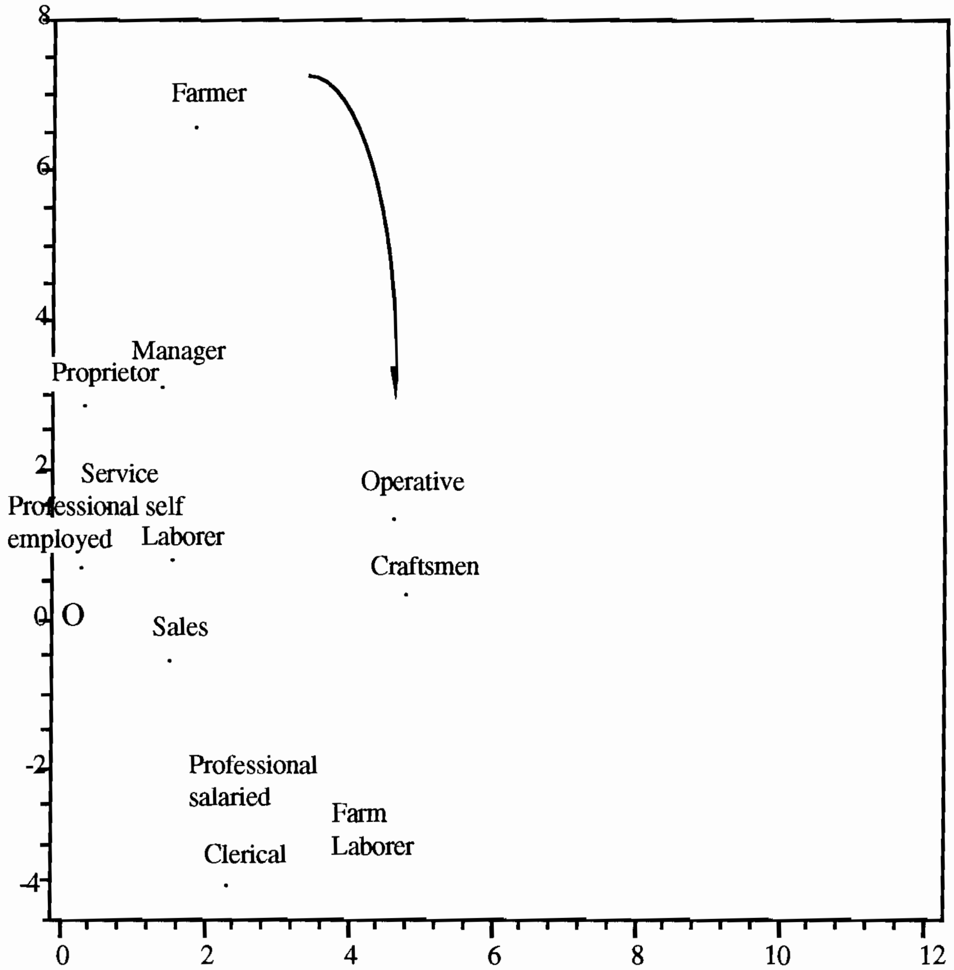

Figure 3: First Directional Plane for the Social Mobility Data

In Figure 3 we see the common changes in occupations in the two societies. The weights for the first plane were almost equally important, the weight for Great Britain was .49 and the weight for the United States was .48. The arrow indicates the direction of positive skew-symmetries. The occupation farmer is on top of the figure, indicating that the farmers are very mobile and their sons are most likely to be employed as operative, craftsmen, clerical, professional (salaried) and farm labourer, but they fly out to the other professions as well. This may correspond to a general industrialization of the two societies or to larger farms. In this net mobility pattern, the occupations proprietor, service and professional (self employed) are positioned near the origin. This indicates absence of mobility for these professions in the first plane. A second chain is manager, labourer, sales, professional salaried and clerical indicating mobility to more skilled work. Symmetries can be found between professional (self employed), service and manager indicating that these professions are in balance.

The second plane accounted for 30 percent of the variation in the skew-symmetric part, giving a total fit of the two planes of 97 percent. The second plane is shown in Figure 4. The pattern of net mobility in Figure 4 shows the differences between the two societies. The arrow indicates positive skew-symmetries for the US data (weight = .42) and negative skew-symmetries for the GB data (weight = -.36), thus the two countries differ in direction of mobility.

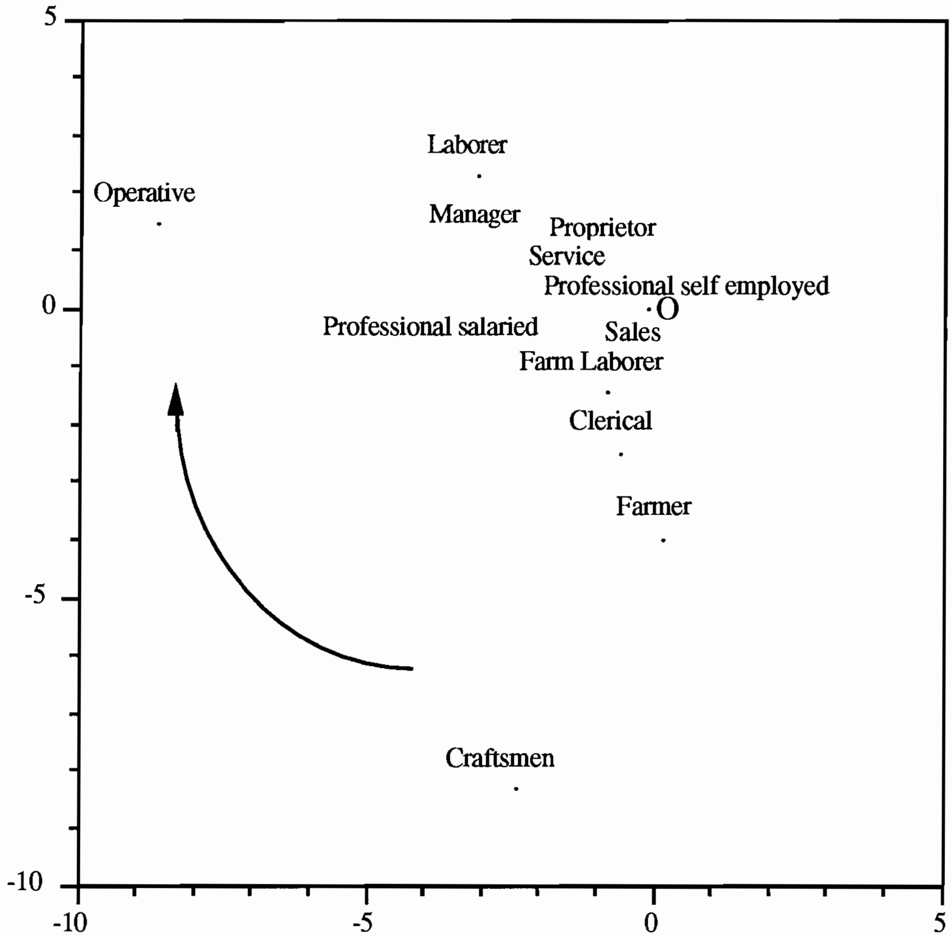

Figure 4: Second Plane for the Social Mobility Data

In Figure 4 we see that the occupations sales, farm labourer, clerical, farmer and craftsmen are on a line through the origin indicating absence of skew-symmetry between these occupations. Labourer, manager, proprietor and service are also on a line through the origin indicating again absence of net mobility in the second plane. The most pronounced skew-symmetry is from craftsmen to operative for the United States and from operative to craftsmen for Great Britain. A less important mobility pattern is from craftsmen and operative to labourer, manager, proprietor and service for the United States and the other way round for Great Britain. In Great Britain, operatives have sons in occupational categories sales, farm labourer, clerical and farmer.

6. Conclusion

In the present paper an INDSCAL and IDIOSCAL analogue for analyzing three-way skew-symmetric data have been developed. In the INDSCAL model the relation between a pair of objects is represented as the area of a triangle formed by the two corresponding points and the origin. For both the INDSCAL and IDIOSCAL model an alternating least squares algorithm for columnwise updating of the parameters was proposed.

The present algorithm has been developed with an explicit decomposition of the data in mind. The analysis of skew-symmetry and symmetry for three-way arrays can now be treated on an equal footing, in a similar way as in the methodology for two-way data proposed by Gower (1977). After the analysis it can be desirable to compare the results of a symmetry analysis to the results of the skew-symmetry analysis. If the symmetric component is analyzed by a distance model, this can be done by rotating the configurations towards each other. In this way, it is sometimes possible to come up with a single model for the complete data. If the two configurations from the symmetric and skew-symmetric part show little differences, it may be possible to fit model for the complete data. A possible method for joint analysis of symmetry and skew-symmetry is called DEDICOM (Harshman, Green, Wind and Lundy, 1982), where an asymmetric table is decomposed into dimensions that are asymmetrically related .

Interpretation of the diagrams obtained from an IDIOSCAL analysis proved to be more difficult than in the INDSCAL model, because of a high number of parameters. When the number of dimensions is large, say larger than 4, than the number of diagrams becomes too large. This is not very surprising, because a similar difficulty arises in the interpretation of symmetric IDIOSCAL results.

Acknowledgement

I thank Willem Heiser and John Gower for stimulating comments on previous drafts of the manuscript.

References

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. (1972). Individual differences and multidimensional scaling. In: R. N. Shepard et al (Eds). *Multidimensional Scaling: Theory and Applications in the Behavioral Sciences.* Seminar press, New York, pp 115-155.

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.* Huber, Bern, 234-289.

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

Constantine, A. G., & Gower, J. C. (1982). Models for the analysis of interregional migration. *Environment and Planning A, 14*, 477-497.

De Soete, G., & Carroll, J. D. (1986). A maximum likelihood method for fitting the wandering vector model. *Psychometrika, 48*, 553-566.

De Soete, G., Feger, H., & Klauer, K. C. (1989). *New developments in psychological choice modeling.* North Holland, Amsterdam.

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.* North Holland, Amsterdam, 109-123.

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., & De Leeuw, J. (1981). Multidimensional mapping of preference data. *Mathematiques et Sciences humaines, 19*, 39-96.

Kerckhoff, A. C., Campbell, R. T., & Winfield-Laird, I. (1985). Social mobility in Great Britain and the United States. *American journal of sociology,91*,281-308.

Takane, Y. (1980). Maximum likelihood estimation in the generalized case of Thurstone's model of comparitive judgment. *Japanese Psychological Research*, *22*, 188-196.

Ten Berge, J. M. F. & Kiers, H. A. L. (1991). Some clarifications of the candecomp algorithm applied to INDSCAL. *Psychometrika*, *2*, 317-326.

Thurstone, L. L. (1929). Theory of attitude measurement. *Psychological Review*, *36*, 222-241.