

An Adaptation of Correspondence Analysis for Square Tables

Michael Greenacre

Facultat de Ciències Econòmiques i Empresariales
Universitat Pompeu Fabra
Ramon Trias Fargas, 23–27, E-08005 Barcelona, Spain
E-mail: michael@upf.es

Abstract

The application of correspondence analysis to square asymmetric tables is often unsuccessful because of the strong role played by the diagonal entries of the matrix, obscuring the data off the diagonal. A simple modification of the centering of the matrix, coupled with the corresponding change in row and column masses and row and column metrics, allows the table to be decomposed into symmetric and skew-symmetric components, which can then be analyzed separately. The symmetric and skew-symmetric analyses can be performed using a simple correspondence analysis program if the data are set up in a special block format.

1 Introduction

Correspondence analysis (CA) is a technique for visualizing tables of frequencies as well as nonnegative data on a commensurate set of ratio-scale variables. There are a number of different ways of defining correspondence analysis. For our purposes we consider two ways, both based on the singular value decomposition (SVD).

The first way can be called the *matrix approximation method*. Suppose that the data matrix \mathbf{N} has been divided by its grand total n to obtain the so-called *correspondence matrix* $\mathbf{P} = \mathbf{N}/n$. Suppose \mathbf{P} has row and column sums \mathbf{r} and \mathbf{c} respectively, and that \mathbf{D}_r and \mathbf{D}_c are diagonal matrices with the elements of \mathbf{r} and \mathbf{c} on the diagonal. For frequency data, \mathbf{P} is an observed discrete bivariate distribution and \mathbf{r} and \mathbf{c} are the marginal distributions. Then CA can be defined as the reduced-rank matrix approximation of \mathbf{P} by weighted least squares, minimizing the following expression:

$$\text{trace}[\mathbf{D}_r^{-1}(\mathbf{P} - \hat{\mathbf{P}})\mathbf{D}_c^{-1}(\mathbf{P} - \hat{\mathbf{P}})^{\top}] = \sum_i \sum_j \frac{(p_{ij} - \hat{p}_{ij})^2}{r_i c_j} \quad (1)$$

for a matrix $\hat{\mathbf{P}}$ of given reduced rank. We know that the best rank 1 approximation is given by $\hat{\mathbf{P}} = \mathbf{r}\mathbf{c}^{\top}$, called the *trivial solution*, so that we can equivalently consider the approximation of the centered matrix $\mathbf{P} - \mathbf{r}\mathbf{c}^{\top}$. The solution for any low rank is given by the *generalized singular value decomposition* (GSVD) of $\mathbf{P} - \mathbf{r}\mathbf{c}^{\top}$ in the metrics \mathbf{D}_r^{-1} and \mathbf{D}_c^{-1} respectively:

$$\mathbf{P} - \mathbf{r}\mathbf{c}^{\top} = \mathbf{U}\mathbf{D}_\alpha\mathbf{V}^{\top} \quad \text{where } \mathbf{U}^{\top}\mathbf{D}_r^{-1}\mathbf{U} = \mathbf{V}^{\top}\mathbf{D}_c^{-1}\mathbf{V} = \mathbf{I} \quad (2)$$

(this differs from the usual SVD where the metrics are simply identity matrices – see Greenacre, 1984, Appendix A). For constructing CA maps, the *principal coordinates* of the row and column points are given by $\mathbf{F} = \mathbf{D}_r^{-1}\mathbf{U}\mathbf{D}_\alpha$ and $\mathbf{G} = \mathbf{D}_c^{-1}\mathbf{V}\mathbf{D}_\alpha$ respectively. For example, to plot the rows and columns in two dimensions, the rank 2 solution given by the first two columns of \mathbf{F} and \mathbf{G} is used. The resulting plot is called the *symmetric map*, as opposed to other so-called *asymmetric maps* (see Greenacre 1984, 1993).

The second way of defining CA may be called the *profile approximation method*. This is the way CA was originally defined by Benzécri (1973). A profile is a row or column of the matrix divided by its corresponding sum. For example, the row profiles are the rows of the matrix $\mathbf{D}_r^{-1}\mathbf{P}$, in which case CA can be defined as the approximation of the row profiles by points in a low-dimensional subspace. Distances and scalar products in the space are computed using the *chi-square metric*, a weighted Euclidean metric using \mathbf{D}_c^{-1} as the weighting matrix. Furthermore, the row profiles are weighted by the respective

elements of \mathbf{r} , called the *row masses*. The objective function in this case is:

$$\text{trace}[\mathbf{D}_r(\mathbf{D}_r^{-1}\mathbf{P} - \hat{\mathbf{Q}})\mathbf{D}_c^{-1}(\mathbf{D}_r^{-1}\mathbf{P} - \hat{\mathbf{Q}})^\top] = \sum_i r_i \sum_j \frac{(p_{ij}/r_i - \hat{q}_{ij})^2}{c_j} \quad (3)$$

Again we have a trivial solution because it turns out that the row vector \mathbf{c}^\top comes closest to all the row profiles in terms of weighted least sum-of-squared distances, so that it is equivalent to approximate the centered profiles $\mathbf{D}_r^{-1}\mathbf{P} - \mathbf{1}\mathbf{c}^\top$. Again this problem is solved using the GSVD of $\mathbf{D}_r^{-1}\mathbf{P} - \mathbf{1}\mathbf{c}^\top$, this time in the metrics \mathbf{D}_r (the diagonal matrix of row masses) and \mathbf{D}_c^{-1} (the metric on the row points):

$$\mathbf{D}_r^{-1}\mathbf{P} - \mathbf{1}\mathbf{c}^\top = \mathbf{X}\mathbf{D}_\mu\mathbf{Z}^\top \quad \text{where } \mathbf{X}^\top\mathbf{D}_r\mathbf{X} = \mathbf{Z}^\top\mathbf{D}_c^{-1}\mathbf{Z} = \mathbf{I} \quad (4)$$

Since the functions (1) and (3) to be minimized are equivalent, there is the following relationship between the two GSVDs (2) and (4): $\mathbf{X} = \mathbf{D}_r^{-1}\mathbf{U}$ (left singular vectors), $\mathbf{D}_\mu = \mathbf{D}_\alpha$ (singular values) and $\mathbf{Z} = \mathbf{V}$ (right singular vectors), Hence the principal coordinates matrices turn out to be: $\mathbf{F} = \mathbf{X}\mathbf{D}_\alpha$ and $\mathbf{G} = \mathbf{D}_c^{-1}\mathbf{Z}\mathbf{D}_\alpha$.

In both definitions of CA the *total inertia* of the table, a measure of the table's total variation, is equal to the weighted sum-of-squares of the centered matrix being approximated:

$$\text{total inertia} = \sum_i \sum_j (p_{ij} - r_i c_j)^2 / (r_i c_j) \quad (5)$$

The inertia accounted for by the rank K^* solution (or K^* -dimensional solution) is equal to the weighted sum-of-squares of the matrix approximation, which is equal to $\sum_{k=1}^{K^*} \alpha_k^2$. The minimum value of (1) (or (3)), which is the residual inertia not accounted for, is equal to the remaining sum-of-squared singular values: $\sum_{k=K^*+1}^K \alpha_k^2$.

The special case considered here is the application of CA to square tables where the rows and columns refer to the same set of objects. Mobility tables, export/import data, migration tables, misclassification tables, confusion matrices and transition matrices are examples of square tables regularly encountered in practice. The application of CA to such tables is generally not very successful and this is often due to the strong effect of the diagonal values on the results.

For example, consider the social mobility table given in Table 1, obtained from Best (1990). These data concern the jobs of parliamentarians in the *Frankfurter Nationalversammlung*, considered to be the first democratic government in Germany. The rows refer to the jobs of the parliamentarians when entering the labour market, and the columns to their jobs in May 1848 when the government started its session.

Table 1
Mobility table from Frankfurter Nationalversammlung

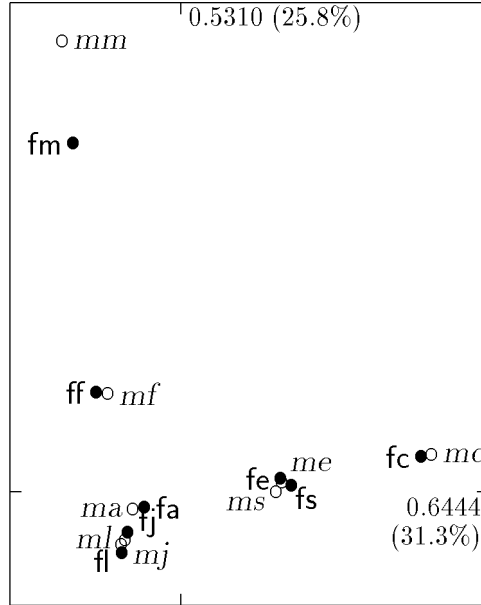
<i>Job when entering labour market</i>		<i>Main occupation in May 1848</i>							
		mj	ma	me	mm	mc	mf	ml	ms
<i>Justice</i>	fj	117	83	19	0	1	17	76	6
<i>Administration</i>	fa	11	37	6	0	0	7	7	6
<i>Education</i>	fe	4	3	67	1	4	3	5	13
<i>Military</i>	fm	0	5	1	16	0	8	1	0
<i>Church</i>	fc	0	1	11	0	26	0	1	0
<i>Farmer</i>	ff	0	4	0	0	0	19	1	0
<i>Lawyer</i>	fl	8	2	1	0	0	0	22	1
<i>Self-employed</i>	fs	0	7	20	0	4	1	0	37

Abbreviations used as labels in Figures 1 to 5 are constructed as follows: first letter is “f” or “m” corresponding to time points “first occupation” and “main occupation” respectively; second letter indicates the occupation.

The two-dimensional CA map of this table is given in Figure 1. The total inertia of this table is 2.0571 and 57.1% of this inertia is accounted for in the map. Notice that of all the row-column pairs, only one – military – separates, while the others lie practically at the same position in the display. Since the CA of a symmetric matrix gives coincident pairs of rows and columns, this map might lead us to conclude that the data are close to symmetric, apart from the fourth row and fourth column. In fact, there are many other interesting deviations from symmetry in the table, which could be seen by examining higher-dimensional solutions. However, we do not see these in the best two-dimensional solution, which is the one most users of CA would interpret, since this solution is dominated by the symmetric part of the table, especially the diagonal of the matrix.

In Section 2 we shall show how the analysis of the square table can be separated into two analyses where the symmetric part of the table and the deviations from symmetry are treated separately. In Section 3 we show how to perform both the symmetric and skew-symmetric correspondence analyses using a regular computer program for simple correspondence analysis. Section 4 closes with a discussion of the method.

Figure 1
Correspondence analysis of Table 1



2 Symmetric and skew-symmetric analyses

An alternative approach to visualizing a square asymmetric table, say \mathbf{A} , is to decompose it into two components, a symmetric table and a skew-symmetric table (Constantine & Gower 1978, Gower 1980):

$$\mathbf{A} = \mathbf{Q} + \mathbf{R} \quad (6)$$

where $\mathbf{Q} = \frac{1}{2}(\mathbf{A} + \mathbf{A}^T)$ and $\mathbf{R} = \frac{1}{2}(\mathbf{A} - \mathbf{A}^T)$. The sum-of-squares is similarly partitioned:

$$\sum_i \sum_j a_{ij}^2 = \sum_i \sum_j q_{ij}^2 + \sum_i \sum_j r_{ij}^2 \quad (7)$$

since the sum-of-cross-products vanishes: $\sum_i \sum_j q_{ij} r_{ij} = 0$. In our case the matrix \mathbf{A} is the correspondence table \mathbf{P} , centered with respect to \mathbf{rc}^T , and the sum-of-squares we want to decompose, the total inertia in (5), includes the weighting of each squared term by $1/(r_i c_j)$.

It would be convenient if the CA of \mathbf{P} decomposed into the CA of the symmetric component and the CA of the skew-symmetric component, as in (6):

$$\mathbf{P} = \mathbf{S} + \mathbf{T} \quad (8)$$

For the data of Table 1 the matrices on the right hand side of (8), multiplied by the sample size n , are equal to:

$$n\mathbf{S} = \begin{bmatrix} 117.0 & 47.0 & 11.5 & 0.0 & 0.5 & 8.5 & 42.0 & 3.0 \\ 47.0 & 37.0 & 4.5 & 2.5 & 0.5 & 5.5 & 4.5 & 6.5 \\ 11.5 & 4.5 & 67.0 & 1.0 & 7.5 & 1.5 & 3.0 & 16.5 \\ 0.0 & 2.5 & 1.0 & 16.0 & 0.0 & 4.0 & 0.5 & 0.0 \\ 0.5 & 0.5 & 7.5 & 0.0 & 26.0 & 0.0 & 0.5 & 2.0 \\ 8.5 & 5.5 & 1.5 & 4.0 & 0.0 & 19.0 & 0.5 & 0.5 \\ 42.0 & 4.5 & 3.0 & 0.5 & 0.5 & 0.5 & 22.0 & 0.5 \\ 3.0 & 6.5 & 16.5 & 0.0 & 2.0 & 0.5 & 0.5 & 37.0 \end{bmatrix}$$

$$n\mathbf{T} = \begin{bmatrix} 0.0 & 36.0 & 7.5 & 0.0 & 0.5 & 8.5 & 34.0 & 3.0 \\ -36.0 & 0.0 & 1.5 & -2.5 & -0.5 & 1.5 & 2.5 & -0.5 \\ -7.5 & -1.5 & 0.0 & 0.0 & -3.5 & 1.5 & 2.0 & -3.5 \\ 0.0 & 2.5 & 0.0 & 0.0 & 0.0 & 4.0 & 0.5 & 0.0 \\ -0.5 & 0.5 & 3.5 & 0.0 & 0.0 & 0.0 & 0.5 & -2.0 \\ -8.5 & -1.5 & -1.5 & -4.0 & 0.0 & 0.0 & 0.5 & -0.5 \\ -34.0 & -2.5 & -2.0 & -0.5 & -0.5 & -0.5 & 0.0 & 0.5 \\ -3.0 & 0.5 & 3.5 & 0.0 & 2.0 & 0.5 & -0.5 & 0.0 \end{bmatrix}$$

The only problem is that the CA of \mathbf{P} involves a different centering, and hence also different metrics, from the CA of \mathbf{S} . \mathbf{S} has row and column margins equal to $\mathbf{w} = \frac{1}{2}(\mathbf{r} + \mathbf{c})$, the average of the row and column sums of \mathbf{P} . Notice that \mathbf{T} is already centered in the sense that all its elements sum to zero: $\sum_i \sum_j t_{ij} = 0$.

A solution to this problem is to use the centering of \mathbf{S} and its associated row and column metrics, both equal to \mathbf{D}_w^{-1} , for the CA of \mathbf{P} . In other words, instead of analyzing $\mathbf{P} - \mathbf{r}\mathbf{c}^\top$ in the metrics \mathbf{D}_r^{-1} and \mathbf{D}_c^{-1} , we analyze $\mathbf{P} - \mathbf{w}\mathbf{w}^\top$ in the metrics \mathbf{D}_w^{-1} and \mathbf{D}_w^{-1} . Equation (8) in centered form is thus:

$$\mathbf{P} - \mathbf{w}\mathbf{w}^\top = \mathbf{S} - \mathbf{w}\mathbf{w}^\top + \mathbf{T} \quad (9)$$

and the corresponding decomposition of inertia is

$$\sum_i \sum_j (p_{ij} - w_i w_j)^2 / (w_i w_j) = \sum_i \sum_j (s_{ij} - w_i w_j)^2 / (w_i w_j) + \sum_i \sum_j t_{ij}^2 / (w_i w_j) \quad (10)$$

Because $\mathbf{w}\mathbf{w}^\top$ is not the closest rank 1 matrix to \mathbf{P} we expect the inertia on the left hand side of (10) to be higher than the usual inertia given by (5) – how much higher will depend on how different \mathbf{r} and \mathbf{c} are. In our example, the total inertia of the recentered matrix is 2.2222, that is 0.1651 higher than the matrix with the usual centering analyzed in Figure 1. Incorporating this disparity in the margins of \mathbf{P} into the CA leads to some more of the asymmetry being displayed in the map. Figure 2 shows the map obtained by performing the GSVD on $\mathbf{P} - \mathbf{w}\mathbf{w}^\top$ in the metrics \mathbf{D}_w^{-1} and \mathbf{D}_w^{-1} , and we can see that more pairs of points are separate compared to Figure 1.

Figure 2
 Recentered correspondence analysis of Table 1

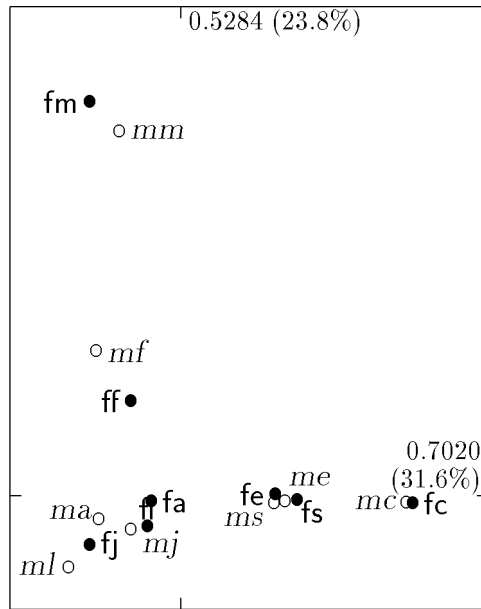


Figure 3
 Correspondence analysis of
 symmetric component **S**

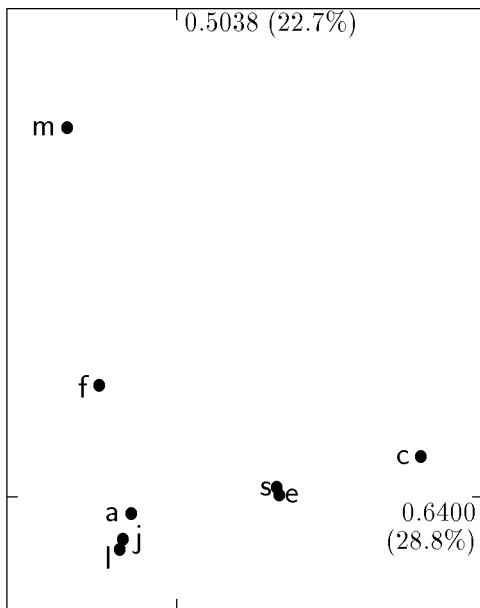
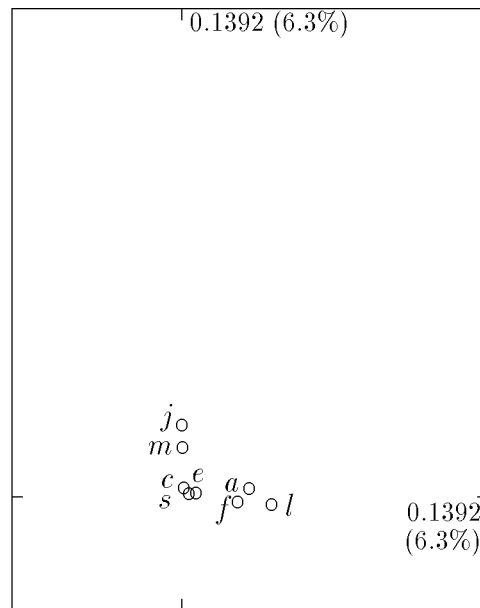


Figure 4
 Correspondence analysis of
 skew-symmetric component **T**



Although some more of the deviation from symmetry in the data is displayed in Figure 2, the map is still largely concentrated on the symmetric component. In order to separate these components we can perform two analyses, one using the GSVD, again in the metrics \mathbf{D}_w^{-1} and \mathbf{D}_w^{-1} , on the centered symmetric table $\mathbf{S} - \mathbf{w}\mathbf{w}^T$ and the other a similar GSVD on the already centered skew-symmetric \mathbf{T} . These analyses attempt to account for the respective inertia components on the right hand side of (10) separately. The resultant displays are shown in Figures 3 and 4, both of which are drawn to the same scale. In both of these displays it is necessary only to display the label corresponding to the occupations (see Table 1) because there is no need to indicate the time points.

The interpretation of Figure 3 is essentially the same as Figure 1 (or Figure 2) – each pair of points has been substituted by one point representing the occupation, since none of the deviation from symmetry is included in this analysis. Closeness of two occupations will depict the fact that there is a relatively high level of exchange between them, while distant occupations involve relatively little exchange. There is a gradient of occupations from military at top left to farmer and then to a cluster of jobs at the centre consisting of administration, justice and lawyer, and then across to the right where we first find education and self-employed and then church at the extreme. The positions of the jobs concur with Best’s hypothesis that the occupations fall into groups within which people tend to migrate, with relatively little migration between groups. These groups are (i) military, farmer; (ii) justice, administration, lawyer; and (iii) education, church, self-employed. To verify this grouping in the data we re-order the rows and columns of Table 1, and tabulate the normalized frequencies $s_{ij}/\sqrt{(w_i w_j)}$ of the symmetric matrix, multiplied by the sample size n (to give chi-square components rather than inertia components), given in Table 2. The three sets of rows and columns are delineated and it is clear that the highest values are grouped in the subtables on the diagonal of the table, in accordance with Best’s hypothesis. Notice that the exchange is not always between all pairs of occupations in a group, for example there is a large exchange between justice and administration and justice and lawyer, but not between lawyer and administration.

Table 2

Values of $ns_{ij}/\sqrt{(w_i w_j)}$ of the re-ordered symmetric component

<i>Job when entering labour market</i>		<i>Main occupation in May 1848</i>							
		mm	mf	mj	ma	ml	me	mc	ms
<i>Military</i>	fm	460.0	89.6	0.0	33.8	8.2	13.2	0.0	0.0
<i>Farmer</i>	ff	89.6	331.8	61.5	58.1	6.4	15.5	0.0	6.7
<i>Justice</i>	fj	0.0	61.5	351.7	205.9	223.1	49.3	3.7	16.8
<i>Administration</i>	fa	33.8	58.1	205.9	236.3	34.8	28.1	5.4	53.1
<i>Lawyer</i>	fl	8.2	6.4	223.1	34.8	206.5	22.7	6.6	4.9
<i>Education</i>	fe	13.2	15.5	49.3	28.1	22.7	410.9	80.2	132.1
<i>Church</i>	fc	0.0	0.0	3.7	5.4	6.6	80.2	484.8	27.9
<i>Self-employed</i>	fs	0.0	6.7	16.8	53.1	4.9	132.1	27.9	386.8

Although the inertia displayed in Figure 4 is much smaller than in Figure 3, the advantage for the interpretation is that it concentrates entirely on the deviations from symmetry. As shown by Constantine & Gower (1978), the singular vectors of a skew-symmetric matrix occur in pairs corresponding to pairs of equal singular values. Points close to the centre of the map, clergy, education and self-employed, show that there is very little difference between the inflow and outflow amongst these jobs. The larger changes in flow are between those categories further from the centre. The direction of positive flow can be deduced from the relative orientations of the points – in this case flow is interpreted in a clockwise direction so that categories *j* (justice) and *m* (military) are losing more people to the categories *l* (lawyer), *a* (administration) and *f* (farmer). The actual interpretation is in terms of the triangles formed by each pair of points and the centre (Figure 5): if the area is large then the deviation from symmetry is large (see Constantine & Gower 1978 or Gower 1980). For example, the biggest triangle of this type in Figure 4 is formed by points *j* (justice) and *l* (lawyer), as shown in Figure 5. Thus the highest positive deviation from symmetry (of the normalized value $t_{ij}/\sqrt{w_i w_j}$) is for the job changes from justice to lawyer. A negative change of the same magnitude is from lawyer to justice, i.e. from point *l* to *j*, being a triangle with counter-clockwise orientation. As in Table 2, we can verify this by tabulating the normalized values of $t_{ij}/\sqrt{w_i w_j}$, also multiplied by the sample size *n* and with rows and columns re-ordered as in Table 2 (see Table 3). The large normalized values for the flow from justice and from military can be checked in this table.

Figure 5

Interpretation of skew-symmetric map in terms of area of triangle subtended with origin

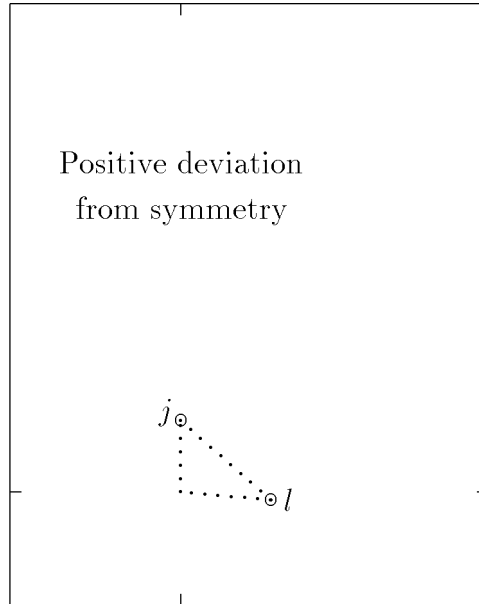


Table 3

Values of $nt_{ij}/\sqrt{(w_i w_j)}$ of the re-ordered skew-symmetric component

<i>Job when entering labour market</i>		<i>Main occupation in May 1848</i>							
		mm	mf	mj	ma	ml	me	mc	ms
<i>Military</i>	fm	0.0	89.6	0.0	33.8	8.2	0.0	0.0	0.0
<i>Farmer</i>	ff	-89.6	0.0	-61.5	-15.8	6.4	-15.5	0.0	-6.7
<i>Justice</i>	fj	0.0	61.5	0.0	157.7	180.6	32.2	3.7	16.8
<i>Administration</i>	fa	-33.8	15.8	-157.7	0.0	19.3	9.3	-5.4	-4.0
<i>Lawyer</i>	fl	-8.2	-6.4	-180.6	-19.3	0.0	-15.1	-6.6	4.9
<i>Education</i>	fe	0.0	15.5	-32.2	-9.3	15.1	0.0	-37.4	-28.0
<i>Church</i>	fc	0.0	0.0	-3.7	5.4	6.6	37.4	0.0	-27.9
<i>Self-employed</i>	fs	0.0	6.7	-16.8	4.0	-4.9	28.0	27.9	0.0

3 Obtaining complete solution using simple CA

The results of Figure 3 can be obtained by applying simple CA to the symmetric matrix **S**. The skew-symmetric analysis, however, needs specific additional programming which is easy if a SVD routine is available, for example in statistical languages such as S-PLUS, Gauss or Matlab. However, in this section we demonstrate that, by setting up the input

data in a special format, it is possible to obtain both the symmetric and skew-symmetric results in one application of simple CA.

If the original square table \mathbf{N} is $p \times p$, then the idea is to set up a $2p \times 2p$ block matrix:

$$\tilde{\mathbf{N}} = \begin{bmatrix} \mathbf{N} & \mathbf{N}^\top \\ \mathbf{N}^\top & \mathbf{N} \end{bmatrix}$$

We now prove that the simple CA of this matrix provides both sets of results in Figures 3 and 4.

The correspondence matrix for $\tilde{\mathbf{N}}$ is

$$\tilde{\mathbf{P}} = \frac{1}{4} \begin{bmatrix} \mathbf{P} & \mathbf{P}^\top \\ \mathbf{P}^\top & \mathbf{P} \end{bmatrix} \quad (11)$$

and has row and column sums equal to

$$\tilde{\mathbf{w}} = \frac{1}{2} \begin{bmatrix} \mathbf{w} \\ \mathbf{w} \end{bmatrix} \quad (12)$$

It is easily shown that each of the four submatrices of $\tilde{\mathbf{P}}$ contributes exactly one quarter of the inertia on the left hand side of (10), so that the total inertia of $\tilde{\mathbf{P}}$ is the same as the proposed recentered form of the CA:

$$\text{trace}[\mathbf{D}_w^{-1}(\tilde{\mathbf{P}} - \tilde{\mathbf{w}}\tilde{\mathbf{w}}^\top)\mathbf{D}_w^{-1}(\tilde{\mathbf{P}} - \tilde{\mathbf{w}}\tilde{\mathbf{w}}^\top)^\top] = \text{trace}[\mathbf{D}_w^{-1}(\mathbf{P} - \mathbf{w}\mathbf{w}^\top)\mathbf{D}_w^{-1}(\mathbf{P} - \mathbf{w}\mathbf{w}^\top)^\top]$$

The GSVD of the matrix $\tilde{\mathbf{P}} - \tilde{\mathbf{w}}\tilde{\mathbf{w}}^\top$ can be assembled from the respective GSVDs of its submatrices. Suppose that $\mathbf{S} - \mathbf{w}\mathbf{w}^\top$ and \mathbf{T} have GSVDs:

$$\mathbf{S} - \mathbf{w}\mathbf{w}^\top = \mathbf{X}\mathbf{D}_\lambda\mathbf{X}^\top \quad \text{where } \mathbf{X}^\top\mathbf{D}_w^{-1}\mathbf{X} = \mathbf{I} \quad (13)$$

$$\mathbf{T} = \mathbf{Y}\mathbf{D}_\mu\mathbf{J}\mathbf{Y}^\top \quad \text{where } \mathbf{Y}^\top\mathbf{D}_w^{-1}\mathbf{Y} = \mathbf{I} \quad (14)$$

where \mathbf{J} is a block diagonal matrix made up of 2×2 blocks $\begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix}$. (We can always assume that \mathbf{J} has an even number of blocks because if p is odd then \mathbf{T} has only $p - 1$ nonzero singular values.) Note the following properties of \mathbf{J} :

$$\mathbf{J}^\top = -\mathbf{J} \quad \text{and} \quad \mathbf{J}^\top\mathbf{D}_\mu = -\mathbf{D}_\mu\mathbf{J} \quad (15)$$

From (9), (11–14), we have:

$$\tilde{\mathbf{P}} - \tilde{\mathbf{w}}\tilde{\mathbf{w}}^\top = \frac{1}{4} \begin{bmatrix} \mathbf{P} - \mathbf{w}\mathbf{w}^\top & \mathbf{P}^\top - \mathbf{w}\mathbf{w}^\top \\ \mathbf{P}^\top - \mathbf{w}\mathbf{w}^\top & \mathbf{P} - \mathbf{w}\mathbf{w}^\top \end{bmatrix}$$

$$\begin{aligned}
&= \frac{1}{4} \begin{bmatrix} \mathbf{S} - \mathbf{w}\mathbf{w}^\top & \mathbf{S} - \mathbf{w}\mathbf{w}^\top \\ \mathbf{S} - \mathbf{w}\mathbf{w}^\top & \mathbf{S} - \mathbf{w}\mathbf{w}^\top \end{bmatrix} + \frac{1}{4} \begin{bmatrix} \mathbf{T} & \mathbf{T}^\top \\ \mathbf{T}^\top & \mathbf{T} \end{bmatrix} \\
&= \frac{1}{4} \begin{bmatrix} \mathbf{X}\mathbf{D}_\lambda\mathbf{X}^\top & \mathbf{X}\mathbf{D}_\lambda\mathbf{X}^\top \\ \mathbf{X}\mathbf{D}_\lambda\mathbf{X}^\top & \mathbf{X}\mathbf{D}_\lambda\mathbf{X}^\top \end{bmatrix} + \frac{1}{4} \begin{bmatrix} \mathbf{Y}\mathbf{D}_\mu\mathbf{J}\mathbf{Y}^\top & \mathbf{Y}\mathbf{J}^\top\mathbf{D}_\mu\mathbf{Y}^\top \\ \mathbf{Y}\mathbf{J}^\top\mathbf{D}_\mu\mathbf{Y}^\top & \mathbf{Y}\mathbf{D}_\mu\mathbf{J}\mathbf{Y}^\top \end{bmatrix} \\
&= \frac{1}{4} \begin{bmatrix} \mathbf{X}\mathbf{D}_\lambda\mathbf{X}^\top & \mathbf{X}\mathbf{D}_\lambda\mathbf{X}^\top \\ \mathbf{X}\mathbf{D}_\lambda\mathbf{X}^\top & \mathbf{X}\mathbf{D}_\lambda\mathbf{X}^\top \end{bmatrix} + \frac{1}{4} \begin{bmatrix} \mathbf{Y}\mathbf{D}_\mu\mathbf{J}\mathbf{Y}^\top & -\mathbf{Y}\mathbf{D}_\mu\mathbf{J}\mathbf{Y}^\top \\ -\mathbf{Y}\mathbf{D}_\mu\mathbf{J}\mathbf{Y}^\top & \mathbf{Y}\mathbf{D}_\mu\mathbf{J}\mathbf{Y}^\top \end{bmatrix} \quad (\text{from (14)}) \\
&= \begin{bmatrix} \frac{1}{2}\mathbf{X} & \frac{1}{2}\mathbf{Y} \\ \frac{1}{2}\mathbf{X} & -\frac{1}{2}\mathbf{Y} \end{bmatrix} \begin{bmatrix} \mathbf{D}_\lambda & \mathbf{0} \\ \mathbf{0} & \mathbf{D}_\mu \end{bmatrix} \begin{bmatrix} \frac{1}{2}\mathbf{X} & \frac{1}{2}\mathbf{Y}\mathbf{J}^\top \\ \frac{1}{2}\mathbf{X} & -\frac{1}{2}\mathbf{Y}\mathbf{J}^\top \end{bmatrix}^\top \quad (16)
\end{aligned}$$

This shows how the GSVD of $\tilde{\mathbf{P}} - \tilde{\mathbf{w}}\tilde{\mathbf{w}}^\top$ is constructed from the symmetric and skew-symmetric components of $\mathbf{P} - \mathbf{w}\mathbf{w}^\top$. Since the singular values of $\tilde{\mathbf{P}} - \tilde{\mathbf{w}}\tilde{\mathbf{w}}^\top$ are in descending order, there will be a re-ordering of the singular values in \mathbf{D}_λ and \mathbf{D}_μ as well as their corresponding singular vectors in the solution (16). The normalization can be checked, remembering that $\mathbf{D}_{\tilde{w}}$ has the elements of $\frac{1}{2}\mathbf{w}$ twice down its diagonal, so that for example for the rows:

$$\begin{aligned}
\begin{bmatrix} \frac{1}{2}\mathbf{X} \\ \frac{1}{2}\mathbf{X} \end{bmatrix}^\top \begin{bmatrix} \frac{1}{2}\mathbf{D}_w & \mathbf{0} \\ \mathbf{0} & \frac{1}{2}\mathbf{D}_w \end{bmatrix}^{-1} \begin{bmatrix} \frac{1}{2}\mathbf{X} \\ \frac{1}{2}\mathbf{X} \end{bmatrix} &= \frac{1}{2}\mathbf{X}^\top\mathbf{D}_w^{-1}\mathbf{X} + \frac{1}{2}\mathbf{X}^\top\mathbf{D}_w^{-1}\mathbf{X} \\
&= \mathbf{I}
\end{aligned}$$

Finally we have to show that the principal coordinates we obtain for $\tilde{\mathbf{P}} - \tilde{\mathbf{w}}\tilde{\mathbf{w}}^\top$ are the same as those for $\mathbf{P} - \mathbf{w}\mathbf{w}^\top$. From the formulae following equation (2) for the principal coordinates \mathbf{F} , the principal coordinates in the latter case are $\mathbf{F}_S = \mathbf{G}_S = \mathbf{D}_w^{-1}\mathbf{X}\mathbf{D}_\lambda$ for the symmetric part and $\mathbf{F}_T = \mathbf{D}_w^{-1}\mathbf{Y}\mathbf{D}_\mu$ and $\mathbf{G}_T = \mathbf{D}_w^{-1}\mathbf{Y}\mathbf{J}^\top\mathbf{D}_\mu = -\mathbf{D}_w^{-1}\mathbf{Y}\mathbf{D}_\mu\mathbf{J}$ for the skew-symmetric part. These coordinates are recovered exactly in the principal coordinates of the block matrix $\tilde{\mathbf{P}} - \tilde{\mathbf{w}}\tilde{\mathbf{w}}^\top$ corresponding to its first set of rows and columns. All the other coordinates are repetitions or 90 degree rotations of these coordinates.

As an example, we set up the 16×16 matrix $\tilde{\mathbf{N}}$ using Table 1 and performed a simple CA, using program *SimCA* (Greenacre 1986). The matrix has a dimensionality of 15, as always in CA this is one less than the number of rows or number of columns, whichever is the smaller. Seven of these dimensions correspond to the symmetric CA and eight to the skew-symmetric CA. The decomposition of the total inertia of 2.2222 (exactly the same total as the recentered matrix displayed in Figure 2), is as follows:

Table 4
Decomposition of inertia along principal axes

<i>Principal axis</i>	<i>Principal inertia</i>	<i>Percentage inertia</i>
1	0.63995	28.80
2	0.50375	22.67
3	0.38072	17.13
4	0.16836	7.58
5	0.14738	6.63
6	0.13920	6.26
7	0.13920	6.26
8	0.06167	2.78
9	0.01423	0.64
10	0.01423	0.64
11	0.00667	0.30
12	0.00667	0.30
13	0.00012	0.01
14	0.00006	0.00
15	0.00006	0.00

Thus the dimensions of the symmetric analysis are numbers 1 to 5, 8 and 13, while the other dimensions whose inertias occur in pairs are those of the skew-symmetric analysis. Figures 3 and 4 are thus the maps of dimensions 1 & 2 and 6 & 7 respectively.

4 Discussion

We have illustrated how the CA of a square asymmetric table can fail by trying to accommodate two different aspects of the table simultaneously. In our example, the symmetric part of the table was seen to dominate the CA, with the skew-symmetric part hardly revealing itself. This is the most typical situation in practice. By splitting the table into symmetric and skew-symmetric components and analyzing these separately, we can visualize the phenomena separately, with a gain in interpretability and representation of information. In order to achieve this, we need to consider a slightly different centering and metric for the row and column points, so that the symmetric and skew-symmetric components have the same geometry.

Notice that there is no difference in the number of free parameters between the simple CA and the symmetric/skew-symmetric analysis. In the former case we arrive at a two-dimensional map showing the rows and the columns in a joint display. In the latter case we arrive at separate two-dimensional maps involving one set of points each. It is valuable to construct the latter pair of maps on the same scale as this gives an idea of the