ELSEVIER

0950-3293(95)00033-X

Food Quality and Preference **6** (1995) 309–314
1995 Elsevier Science Limited
Printed in Great Britain
0950-3293/95 $9.50+.00

A HIERARCHY OF MODELS FOR ANALYSING SENSORY DATA

El Mostafa Qannari,[a] Ian Wakeling[b] & Halliday J. H. MacFie[b]

[a]ENITIAA-INRA, Laboratoire de Statistique appliquée à la caractérisation des aliments, 44072, Nantes, Cedex 03, France
[b]Institute of Food Research, Reading Laboratory, Earley Gate, Whiteknights Road, Reading, RG6 2EF, UK

ABSTRACT

We propose a hierarchy of models for averaging sensory profile data. The models follow from formulating the data from each assessor in terms of association matrices and considering different strategies for weighted averaging of these matrices. It turns out that two forms of weighting contained within the hierarchy are very close to Generalised Procrustes Analysis (GPA) and Individual Differences Scaling (INDSCAL). The advantage of the current approach is that the methods are not iterative. The methods are illustrated using data based on perception of yoghurts.

Keywords: Generalised Procrustes Analysis; Individual Differences Scaling; STATIS; Association Matrices; Sensory Analysis.

INTRODUCTION

In sensory profiling (free choice profiling or conventional profiling) suppose that m assessors score n samples. The results can be displayed in matrices X_1, X_2,... X_m where rows refer to samples and columns refer to attributes. Throughout this paper all data matrices are assumed to be column centred which adjusts for scoring at different levels of the scale by different assessors. Therefore, all configurations are translated to be centred about a common origin.

Generalised Procrustes Analysis (GPA) (Gower, 1975; Dijksterhuis & Gower, 1991; Arnold & Williams, 1986) is aimed at matching the configurations to a group average configuration by an iterative process in the course of which optimal rotations and isotropic scaling factors are computed. STATIS (Lavit, 1988), which is well known among French statisticians and data analysts, provides an alternative to GPA. The computation of isotropic scaling factors and the group average configuration is straightforward (not iterative) and the results are, to some extent, compatible with those given by GPA, as we shall see on the basis of a data set. The computations in STATIS are simplified because this method is based upon association matrices (see next section) instead of the configurations themselves.

INDSCAL (Carroll & Chang, 1970; Krzanowski, 1990 pp.183–193) may be considered as a more general model than GPA and STATIS in the sense that it allows the assessors to weight differentially the several dimensions of a common 'psychological space'. The determination of this common space and the dimension weights are computed by means of an iterative algorithm, but convergence is not always guaranteed.

On the basis of association matrices, we discuss a hierarchy of methods that includes STATIS as a particular case. These methods range from a simple method that may be easily computed to a more general method that is proposed as an alternative to INDSCAL with the advantage that it has no problem of convergence because it is not iterative. In the spirit of Common Principal Components theory (Flury, 1988) which describes a hierarchy of models based upon covariance matrices associated to data tables, we suggest a hierarchy of models using association matrices instead of covariance matrices. However, the approach we adopt is rather exploratory in the sense that no distributional assumptions are made, whereas Common Principal Components is based on assumptions of normality and maximum likelihood estimation. The hierarchy of models postulates three possible relationships between the association matrices. These relationships have direct interpretation in terms of the behaviour of the sensory panellists. Therefore, this interpretation may be at the forefront of the practitioner's mind in judging which model is appropriate. On the other hand, the choice of a particular model among those suggested may be made according to the general principle of parsimony which states that, if a model with few parameters fits the data satisfactorily, then it should be preferred to models that require more parameters.

It is worth noting that a family of models embodied in a hierarchically organised procedure called PINDIS had been suggested by Lingoes and Borg (1978). In this approach the individual differences (between panellists in our context) are assessed using increasingly complex transformations. Like GPA, calculations in PINDIS are based on the configurations rather than either scalar product or distance matrices. This has the advantage to lead to results interpretationally more direct, but the calculations being complex could have the consequence that the properties (stability, significance testing,...) of the computed solution are difficult to study.

ASSOCIATION MATRICES

For a data matrix, X, $(n \times p)$, assumed to be column centred, the association matrix with which we are concerned is defined by $W = XX'$ $(n \times n)$. The diagonal elements of this matrix are the squared distances of the samples from the origin and the off-diagonal elements are scalar products between samples which are quantities proportional to the cosine between samples and, therefore, characterise similarities between them. Thus, the association matrix describes how samples relate to each other, that is, the configuration of the samples.

A very useful property of the association matrix is that, given two configurations X and Y with association matrices denoted respectively by W_X and W_Y, it can be proven (Glaçon, 1981) that W_X and W_Y are equal if and only if, X and Y can be matched by means of a rigid rotation. As a consequence of this property it appears that in dealing with association matrices, the calculations are tremendously simplified because it is not necessary to explicitly determine beforehand the rotations which match the configurations.

Another property is that given a matrix, W $(n \times n)$, that is symmetric and positive semi-definite (i.e. for each vector $(n \times 1)$, x: $x' Wx \geq 0$), it is possible to derive a configuration X that has W as its association matrix: $W = XX'$. The determination of such matrix, X, is given by:

$$X = Q\Lambda;$$

where Q is the orthogonal matrix the columns of which are the normalised eigenvectors of W and Λ is the diagonal matrix whose diagonal elements are the square roots of the eigenvalues of W.

An association index between two configurations X and Y is defined as follows (Robert & Escoufier, 1976):

$$I(X, Y) = \text{trace}(W_x W_y)$$

A normalised version of this index is given by the so-called RV-coefficient:

$$RV(X, Y) = \frac{\text{trace}(W_x W_y)}{\sqrt{\text{trace}(W_x W_x)}\ \sqrt{\text{trace}(W_y W_y)}}$$

RV (X, Y) ranges between 0 and 1. It is equal to 0 if, and only if, configurations X and Y are embedded in orthogonal subspaces (complete disagreement between configurations). RV (X, Y) is equal to 1, if and only if, X and Y can be matched by means of a rotation and multiplication by a scalar.

For any two assessors i and j, the RV coefficient between configurations X_i and X_j provides a measure of similarity that shows to what extent these assessors have the same view of the sample configuration. Consequently, the table of RV coefficients between assessors can be subjected to Multidimensional Scaling, and the outcome will be a graphical display of m points. Each point will represent one assessor and the distance between any two points will represent the extent to which the corresponding assessors have different views of the configuration of the products (Lavit, 1988).

We now consider how to obtain the best group average configuration based on these properties and using increasingly complex assumptions about the differences between the assessors.

LEVEL ONE: EQUALITY OF ASSOCIATION MATRICES

The simplest assumption is that assessors all perceive the inter-relationships between the samples similarly. This is equivalent to assuming that any two of the original sample by sensory attribute configurations X_i and X_j can be matched by a rigid rotation. In terms of association matrices, this is equivalent to assuming that all the association matrices are equal (apart from random errors).

This would be appropriate when assessors have used free choice profiling, or where a fixed choice profile has been used, but assessors are likely to confuse some descriptors or interpret them differently. Note that, at this level, assessors are not assumed to vary in their range of scoring, since individual scaling factors are not given.

Denote by $W_1, W_2, \ldots, W_m$, the association matrices corresponding respectively to configurations $X_1, \ldots X_m$. To compute a common matrix W we search for the matrix that is as close as possible to $W_1, W_2, \ldots, W_m$ in that sense that it minimises the criterion:

$$\sum_{i=1}^{m} \text{trace}((W_i - W)'(W_i - W))$$

$$= \sum_{i=1}^{m} \| W_i - W \|^2$$

where $\|A-B\|^2$ denote the sum of elementwise squared differences between matrices A and B. This criterion leads to the solution:

$$W = \frac{1}{m}\sum_{i=1}^{m} W_i.$$

This matrix, being positive semi-definite, may be written as follows: $W = CC'$. C may be considered as a group average configuration. A graphical display for products may be obtained by performing Principal Component Analysis (PCA) on C and Principal Components may be interpreted by considering their correlations with columns of data tables $X_1, X_2,... X_m$ as is customary with GPA.

LEVEL TWO: PROPORTIONALITY OF ASSOCIATION MATRICES

The second level of comparison between configurations proposes that the associations matrices are proportional. That is, for two association matrices W_i and $W_j (i,j = 1, 2,..., m)$ there exist positive scalars α_i and α_j such that $\alpha_i W_i = \alpha_j W_j$.

In terms of the original configurations, this level implies that any two configurations may be matched by means of a rotation and multiplication by a scalar. Consequently, this level may be used with free choice profiling and conventional profiling as an alternative to GPA. In addition to possible rotations between configurations, it allows for an isotropic factor for each configuration, adjusting, therefore, for differences in range of scaling for different assessors.

In order to compute the isotropic scaling factors we maximise the criterion:

$$F = \sum_{i=1}^{m}\sum_{j=1}^{m} \alpha_i \alpha_j \mathrm{trace}(W_i W_j)$$

under the constraint

$$\sum_{i=1}^{m} \alpha_i^2 \mathrm{trace}(W_i W_i) = \sum_{i=1}^{m} \mathrm{trace}(W_i W_i)$$

The rationale behind this strategy is that we search for isotropic scaling factors that maximise the agreement between assessors. The constraint is chosen in the spirit of GPA which uses the same constraint with X_i instead of W_i and, as in GPA, other constraints may be considered but these are likely to lead to more awkward computation (Gower, 1975).

Let R be the matrix of RV coefficients between configurations (Robert and Escoufier, 1976):

$$R_{ij} = \frac{\mathrm{trace}(W_i W_j)}{\sqrt{\mathrm{trace}(W_i W_i)}\sqrt{\mathrm{trace}(W_j W_j)}}$$

Let $\beta = (\beta_1, \beta_2,...,\beta_m)'$ be the first normalised eigenvector of the matrix R. As all the elements of R are positive, the components of vector β are positive according to Perron's theorem (see for instance Horn and Johnson, 1990). We can prove that the isotropic factor scalars α_i that maximise the criterion F subject to the constraint considered above are given by:

$$\alpha_i = \frac{\sqrt{T}\beta_i}{\sqrt{\mathrm{trace}(W_i W_i)}}$$

where $T = \sum_{i=1}^{m} \mathrm{trace}(W_i W_i)$.

These scalars include pre-scaling factors given by $\alpha_i = \frac{\sqrt{T}}{\sqrt{\mathrm{trace}(W_i W_i)}}$ and coefficients, β_i, that range between 0 and 1. A coefficient, β_i, close to zero indicates that the corresponding assessor, i, is not in good agreement with the others. Thus, greater weight is given to assessors in general agreement with the others and less to an assessor who has a different view of the configuration. A similar interpretation for isotropic scaling factors was given by Collins (1992) in the framework of GPA.

Once the scaling factors are determined, a common association matrix for the configuration may be obtained by simply averaging the scaled association matrices:

$$W = \frac{1}{m}\sum_{i=1}^{m} \alpha_i W_i.$$

As above, a group average configuration, C, may be obtained by writing W in the form: $W = CC'$. Such an expression is possible because W is positive semi-definite. Depicting product maps is achieved by performing PCA on matrix C.

As a matter of fact, this level leads to STATIS which had been widely performed on several kinds of data (Lavit, 1988) and particularly on sensory analysis data (Schlich, 1993).

LEVEL 3: COMMON UNDERLYING DIMENSIONS, DIFFERENTIALLY WEIGHTED

Level 2 allows for an isotropic scaling for each assessor. This may be restrictive in situations where assessors have little or no training. Level 3 enables us to take into account differences in range of scoring for several dimensions and not merely an overall isotropic scaling. Suppose that S_i is the space in which the products can be represented by points in such a way that inter-point distances match as closely as possible the inter-products distances as perceived by the ith assessor. The rationale behind level 3 is that the m spaces $S_1, S_2,... S_m$ have a common set of dimensions, but these dimensions may be differentially important for specific assessors. In other words, there exists a common space, S, with a set of orthogonal dimensions generated by vectors $q_1, q_2,... q_n$, and the space for each assessor can be obtained from S by weighting each of these dimensions by an appropriate amount. One may recognise here the basic premise of the INDSCAL model (Carroll & Chang, 1970; Krzanowski, 1990 pp.183–193).

A formulation of this model in terms of the associations matrices is as follows:

$$W_i = Q\Lambda_i Q';$$

Q being the matrix whose columns are vectors q_1, q_2,... q_n (common underlying dimensions) and Λ_i being a diagonal matrix whose diagonal elements are denoted $\lambda_1^{(i)}$, $\lambda_2^{(i)}$,... $\lambda_n^{(i)}$. The element $\lambda_j^{(i)}$ is the weight assigned by the ith assessor to the jth (common) axis; this reflects the relative salience attached by that assessor to this particular dimension.

An equivalent form for this level is:

$$W_i = \sum_{j=1}^{n} \lambda_j^{(i)} q_j q_j';$$

this form suggests that once the common vectors q_1, q_2,... q_n are determined, the weights $\lambda_1^{(i)}$, $\lambda_2^{(i)}$,... $\lambda_n^{(i)}$ for the ith assessor may be computed in such a way that the following quantity is minimised:

$$\| W_i - \sum_{j=1}^{n} \lambda_j^{(i)} q_j q_j' \|.$$

This leads to a linear regression problem, the solution of which is obvious because vectors q_j are orthogonal:

$$\lambda_j^{(i)} = \frac{\text{trace}(W_i q_j q_j')}{(q_j' q_j)^2}$$

In order to determine vectors q_1, q_2 ,...,q_n, we suggest a step by step procedure. Firstly, we determine a vector q_1 of unit length that has the greatest association with matrices W_1, W_2,... W_m as measured by:

$$\frac{1}{m}\sum_{i=1}^{m} \text{trace}(W_i q_1 q_1').$$

Equivalently vector q_1 may be determined as the minimiser of the quantity:

$$\sum_{i=1}^{m} \| W_i - \lambda_1^{(i)} q_1 q_1' \|^2,$$

it may be shown that vector q_1 is given by the first normalised eigenvector of $\frac{1}{m}\sum_{i=1}^{m} W_i$. In a second step, the vector q_2 may be determined as a vector with norm 1, orthogonal to q_1 and such that the quantity $\frac{1}{m}\sum_{i=1}^{m}$ trace$(W_i q_2 q_2')$ is maximised. In a similar manner, we may derive vectors q_3,..., q_n. These vectors are the eigenvectors associated respectively to eigenvalues of matrix $\frac{1}{m}\sum_{i=1}^{m}$ trace W_i arranged in a decreasing order.

This solution seems intuitively appealing because common dimensions are provided by the average association matrix as in level 1, but each assessor is accommodated further by fitting weights (or saliences) to each dimension.

EXAMPLE

Dijksterhuis and Punter (1990) describe an experiment conducted on yoghurts using the free choice profiling method of assessment. Seven assessors (labelled A, B,..., G) with little or no training profiled eight different yoghurts (labelled 1, 2,..., 8). These data were analysed using GPA and related methods by Dijksterhuis and

TABLE 1. Results of Fitting Levels 1–3: Lack of Fit Showing Deviation of Each Assessor Association Matrix from the Model; Isotropic from Level 2 and Assessor Saliences from Level 3.

Assessor	**A**	**B**	**C**	**D**	**E**	**F**	**G**	**Total**
Lack of fit level 1	0.129	0.113	0.077	0.149	0.140	0.200	0.190	0.998
Lack of fit level 2	0.129	0.103	0.072	0.152	0.125	0.193	0.167	70.941
Lack of fit level 3	0.105	0.058	0.064	0.140	0.075	0.130	0.118	0.690
Isotropic scaling factor for level 2	1.023	1.028	1.184	0.954	0.947	0.900	0.962	
Assessor saliences from level 3	A	B	C	D	E	F	G	Total
Dim 1	0.417	0.610	0.479	0.477	0.660	0.240	0.221	3.104
Dim 2	0.342	0.054	0.143	0.284	0.116	0.365	0.220	1.524
Dim 3	0.028	0.102	0.177	0.084	0.037	0.109	0.236	0.773
Dim 4	0.049	0.092	0.074	0.041	0.035	0.149	0.109	0.548
Dim 5	0.079	0.071	0.026	0.054	0.081	0.101	0.057	70.469
Dim 6	0.068	0.060	0.059	0.045	0.062	0.016	0.050	0.362

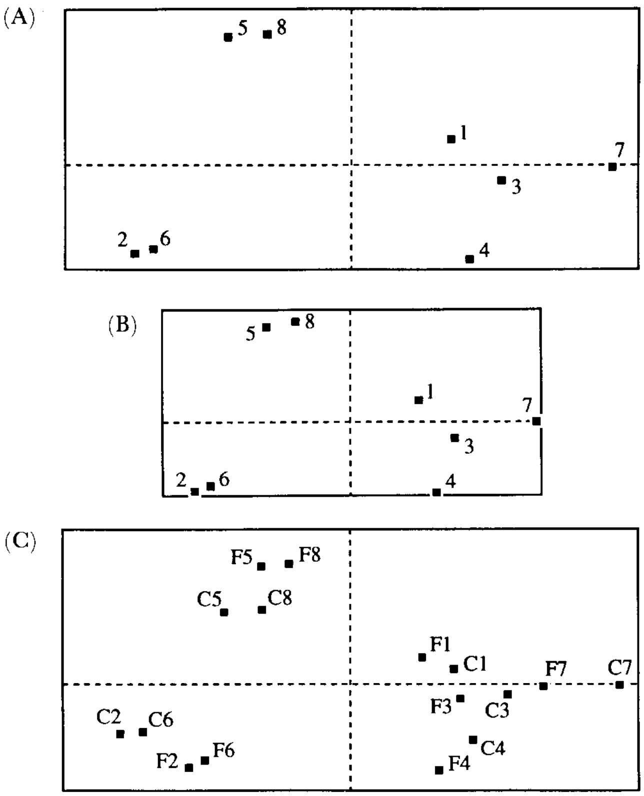

FIG. 1. (A) First two principal components of a PCA of the group average derived from level one, showing the eight yoghurts. (B) First two principal components of a PCA of the group average derived from level two, showing the eight yoghurts. (C) Common dimensions q_1 and q_2 and private spaces for assessor C and assessor F.

Gower (1991), and we refer to this publication in order to compare our results with those of GPA. The positions of the samples relative to the first two principal components of the group average configurations are given in Figure 1A–C. These are all very similar to each other and to the solution given by Dijksterhuis and Punter (1990) and Dijksterhuis and Gower (1991)

The detailed results of fitting levels 1–3 are given in Table 1. Examining the lack of fit statistics for each assessor indicates only a very marginal improvement in going from level 1 to level 2.

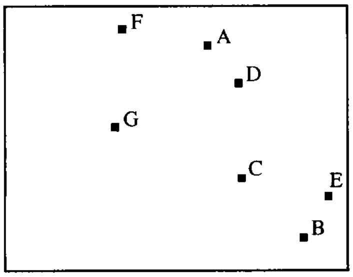

FIG. 2. Subject space showing assessor saliences for dimension one and dimension two.

However, the lack of fit indices for assessors B, E, F and G are sharply reduced by moving to level 3 and the assessors saliences plotted in Figure 2 indicate that these assessors are the most extreme, with B and E weighting heavily on dimension 1 and no other (see Table 1). Assessors F and G weight least heavily on dimension 1 and reasonably heavily on dimension 2. These dimensions may be interpreted by considering their correlations with columns of data tables $X_1, X_2, \ldots X_m$. Figure 1C depicts the differences in perception between these two groups of assessors as it shows the samples plotted along the first two dimensions of the 'private' spaces of assessors F and E.

Thus, although a formal significance testing framework is not available, the lack of fit statistics lead us to conclude that there is evidence that assessors were differentially weighting the underlying dimensions. This information can enable a number of further analyses to proceed. The common space obtained by equal saliences can be used to relate the sensory to physical and chemical measurements. The existence of differential weighting may stimulate the instigation of further trials to establish whether the subgroupings are evident in a wider population of sensory assessors. Alternatively, further training of quality assurance panels in the use of particular scales may be required.

CONCLUSION

The hierarchy of levels proposed here gives a superior analysis to the use of GPA because it gives the choice of simpler and more sophisticated models and does not require an iterative solution. The models are expressed in terms of the association matrices as follows (ranging from the more complex model to the simplest one):*model 3*: $Wi = \underset{\sim}{Q}\,\Lambda_i\,\underset{\sim}{Q}'$ (common dimensions but each assessor is accommodated further by fitting saliences to each dimension); *model 2*: $Wi = \lambda_i\,\underset{\sim}{Q}\,\underset{\sim}{Q}'$ (common dimensions but each assessor is accommodated further by fitting an isotropic scaling); *model 1*: $Wi = \underset{\sim}{Q}\,\underset{\sim}{Q}'$ (common dimensions for all assessors).

Further research to compare more precisely the results of the suggested strategy with those of GPA, INDSCAL and PINDIS and to provide a soundly based significance testing framework is indicated.

REFERENCES

Arnold, G. M. & Williams, A. A. (1986). The use of Generalised Procrustes Analysis in sensory analysis. In *Statistical Procedures in Food Research*, ed. J. R. Piggott. North Holland, Amsterdam.

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

Collins, A. J. (1992). Scaling factors in generalised procrustes analysis. In *Computational Statistics, Vol. 1. Proceedings of the 10th Symposium on Computational Statistics,* eds Y. Dodge and J. Whittaker.

Dijksterhuis, G. B. & Gower, J. C. (1991). The interpretation of generalised Procrustes analysis and allied methods. *Food Quality and Preference,* **3,** 67–87.

Dijksterhuis, G. B. & Punter, P. H. (1990). Interpreting generalised Procrustes analysis 'analysis of variance' tables. *Food Quality and Preference,* **2,** 255–65.

Flury, B. (1988). *Common Principal Components and Related Multivariate Models.* Wiley, Chichester.

Glaçon, F. (1981). Analyse conjointe de plusieurs matrices de données; comparaisons de plusieurs méthodes. PhD Thesis, Université des Sciences de Grenoble.

Gower, J. C. (1975). Generalised Procrustes Analysis. *Psychometrika,* **40,** 33–51.

Horn, R. A. & Johnson, C. R. (1990). *Matrix Analysis.* Cambridge University Press, Cambridge.

Krzanowski, W. J. (1990). *Principles of Multivariate Analysis, a User's Perspective.* Oxford Statistical Science Series.

Lavit, C. (1988). *Analyse conjointe de tableaux quantitatifs.* Masson, Paris.

Lingoes, J. B. & Borg, I. (1978). A direct approach to individual differences using increasingly complex transformations. *Psychometrika,* **43,** 491–519.

Robert, P. & Escoufier, Y. (1976). A unifying tool for linear multivariate statistical methods: the RV coefficient. *Appl. Stat.,* **25,** 257–67.

Schlich, P. (1993). Contribution à la sensométrie. PhD Thesis, Université Paris Sud.