A simple alternative to Generalized Procrustes Analysis.

Application to sensory profiling data

JOACHIM KUNERT [1] EL MOSTAFA QANNARI [2]

[1] Fachbereich Statistik, Universität Dortmund D 44 221 *DORTMUND, GERMANY*

[2] *ENITIAA/INRA - Unité de Statistique Appliquée à la Caractérisation des Aliments*
Rue de la Géraudière B. P. 82 225, 44322 NANTES cedex 03 - FRANCE

Running title :

A simple alternative to GPA.

Correspondence should be sent to :

EL MOSTAFA QANNARI

ENITIAA/INRA - Unité de Statistique Appliquée à la Caractérisation des Aliments
Rue de la Géraudière B. P. 82 225, 44322 NANTES cedex 03 - FRANCE

email : qannari@enitiaa-nantes.fr

Abstract

A statistical method for analysing sensory profiling data obtained by means of fixed vocabulary or free choice profiling is discussed. The most interesting feature of this method is that it involves only simple statistical treatment and can therefore be performed using standard software packages. The outcomes of this method are compared to those of Generalized Procrustes Analysis on the basis of two data sets obtained respectively by means of fixed vocabulary and free choice profiling. A significance test is also discussed in order to assess whether the overall configuration of the products is meaningful. This significance test is based upon a simulation study involving the permutation procedure.

Keywords : Sensory profiling, Principal Components Analysis, Generalized Procrustes Analysis, Isotropic scaling factors, Permutation test.

Introduction

Generalized Procrustes Analysis (GPA) was introduced and popularized by Gower (1975). It is used for the analysis of sensory profiling data obtained by means of free choice profiling or fixed vocabulary profiling (Arnold and Williams, 1986 ; Dijkterhuis and Gower, 1991). However, a wider use of GPA is impeded by the fact that this statistical method involves sophisticated computations and the practitioner needs specialized software which may not be available for reasons of cost or computing environment. As a by-product, the use of free choice profiling procedure is also hindered.

A simple alternative to GPA is discussed in this paper. A noteworthy feature of this method is that it can be performed using many standard software packages. The method involves simple statistical pre-treatment in order to determine isotropic scaling factors associated with the assessors and, in a second step, Principal Components Analysis (PCA) is carried out on the matrix formed by merging the scaled data sets associated with the assessors. This PCA allows pictorial representations of the products and the interpretation of the principal components is undertaken by considering their correlations with the assessors' attributes.

The method of analysis discussed herein is compared with GPA on the basis of two data sets obtained respectively by means of fixed vocabulary and free choice profiling. It turns out that the outcomes of both methods compare fairly well. This may encourage practitioners in sensory analysis to use this alternative method. Also, it is believed that this alternative method will shed some light on GPA, therefore enhancing its understanding.

We also discuss a significance test which enables the person analysing the data to assess whether the overall configuration of the products is meaningful. This test is based upon the procedure discussed by Wakeling *et al* (1992) within the context of GPA. It should be pointed out, however, that the significance test involves extensive computations.

Sensory profiling data and GPA

The sensory profiling (free choice or fixed vocabulary) of n products by m assessors results in matrices X_1 , X_2 ,... X_m , where the rows refer to the products, and the columns to the attributes. Throughout this paper, it is assumed that the configurations are column centered therefore removing the effect of judges scoring at different levels of the scales. Column centering consists in subtracting from each entry of each data set the average of the corresponding column.

The different stages in GPA make it possible to eliminate some types of variations between assessors (Arnold and Williams, 1986). These stages involve (i) the determination of isotropic scaling factors which adjust for differences in range of scoring and, (ii) the determination of optimal rotations in order to match as closely as possible the assessors' configurations. Eventually, a group average configuration is derived.

An alternative procedure to GPA

The scaling of the configurations entails a standardization of the configurations in order to adjust for variations among assessors in range of scoring. This problem is generally solved by multiplying each assessor's data table by a positive scalar. The following isotropic scaling factors α_i (i = 1, 2,..., m) achieve such objective :

$$\alpha_i = \frac{\sqrt{T}}{\sqrt{t_i}}$$

where t_i is the sum of the variances of the attributes in data matrix X_i and $T = \frac{\sum_{i=1}^{m} t_i}{m}$.

The division by $\sqrt{t_i}$ is intended to put all the configurations on the same footing as the sums of squares become equal for all the data sets. The multiplicative constant $\sqrt{T}$ is simply introduced in order to allow direct comparison with the isotropic scaling factors obtained by means of GPA, as the scalar α_i given above are subjected to the same constraint as in GPA, namely :

$$\sum \alpha_i^2 t_i = \sum t_i$$

The isotropic scaling factors α_i (i = 1, 2,..., m) stand as standardization factors. It is worth noting that in GPA, the isotropic scaling factors also take into account the performance of the assessors. As discussed by Collins (1992) and Qannari *et al* (1997), the isotropic scaling factors in GPA enable, in addition to standardizing the configurations, the assessors to be weighted in such a way that those assessors who are not in good agreement with the general point of view are downweighted in the process of computing the average group configuration.

The scaled configurations $Y_i = \alpha_i X_i$ (i = 1, 2,..., m) are considered and merged into a matrix $Y = (Y_1 \mid Y_2 \mid ... \mid Y_m)$ whose columns are formed by all the attributes of all assessors. PCA performed on matrix Y makes it possible to depict relationships among products on the basis of the principal components. This PCA copes with the variations among the assessors in the use of different terms (free choice profiling) or different interpretations of the same attributes. By performing a PCA on matrix Y, those attributes which convey the same meaning will generate the same principal components.

The following property will enhance the understanding of the method as it shows that the analysis outlined herein is based on the same premises as GPA. Assume that X_1 , X_2 ,... X_m are actually scaled and rotated configurations of the same configuration X, that is, $X_i = c_i XR_i$, where c_i is a scalar and R_i is a rotation matrix. Then, the analysis discussed in this paper leads to the same principal components as X (except for a multiplying factor). This means that the analysis exhibits in this case the common dimensions to all the data sets. Details of this feature are given in the appendix.

It is worth noting that a PCA of the merged data sets has been advocated by several authors as a strategy for the simultaneous analysis of multiple data sets (Levin, 1966 ; Jaffrenou, 1978). Kiers (1991) shows how several statistical methods amount to performing a PCA on the merged data set. In particular, the STATIS method (Lavit *et al,* 1994 ; Schlich, 1996 ; Qannari *et al*, 1997) fits within this framework. This procedure of determination of a group average configuration involves heavy computations and, similarly to GPA, it exhibits scaling factors that take account of the differences in the range of scoring as well as of the performance of the assessors. Qannari *et al* (1995) discussed a hierarchy of three models for analyzing sensory data. The first model leads to a PCA on the merged data sets, whereas the second and third model lead respectively to the STATIS model and to an alternative to INDSCAL (Carroll and Chang , 1970). Schlich (1996) suggested without further development the idea of pre-scaling the configurations and performing a PCA on the merged scaled data sets as a means to analyze sensory profiling data. Escofier and Pagès (1984) also developed a method under the acronym AFM (*Analyse Factorielle Multiple*) which bears some similarity to the method outlined herein. This method involves two steps. In the first step, the data sets X_1, X_2,..., X_m are normalized such that the first principal component in each data set explains the same amount of inertia (total variance). This is achieved by dividing each matrix X_i (i = 1, 2,..., m) by the square root of the largest eigenvalue of matrix $X_i^T X_i$ (covariance matrix). The second step in AFM consists in performing a PCA on the merged scaled data sets. This pre-scaling procedure although based on an intuitively sound basis may not be appropriate in situations where an assessor describes the products using a high dimensional space and where these dimensions have almost the same variances. Such a configuration may be obtained by a random generation of numbers. In this case, the isotropic scaling factor associated with the considered assessor is approximately $\sqrt{p}$, where p is the number of dimensions. Therefore, it appears that this isotropic factor tends to favor the considered assessor in the process of computing a group average configuration. Our approach which is also based on a PCA of the merged data sets exhibits isotropic scaling factors which, in addition to being easily computable and easily interpretable, are put on the same footing as the isotropic scaling factors in GPA in order to allow straightforward comparison of the methods. We also point out the connection between our approach and GPA by outlining a property which is discussed above and in detail in the appendix. Furthermore, a significance test to assess the relevance of the overall configuration is provided in the next section.

Significance test

The purpose of the significance test suggested herein is to assess whether the overall configuration obtained by means of PCA on the merged data sets is meaningful. This test is based upon a simulation study. The permutation procedure suggested by Wakeling *et al* (1992) within the context of GPA is considered. A set of m new data tables is obtained by randomly permuting the rows of the original data sets and the analysis outlined above is carried out. Note that the permutation procedure consists in exchanging (for each assessor independently) the whole rows of the data matrix. This leaves the correlation structure between the descriptors used by the same assessor unaltered, it only influences the structure of the products.

This process is repeated N (say N=100) times and at each time a loss function is computed. This loss function is discussed further in the appendix. It expresses how far the overall configuration is removed from the simulated configurations. The empirical distribution of this loss function is drawn and the loss function corresponding to the original data tables is compared to this empirical distribution. This simulation study can be simplified by considering that the isotropic scaling factors are unchanged by permutation of the rows of the original data sets as the variances of the variables (columns of the original data sets) remain the same.

Examples

To illustrate the method outlined in this paper, two sets of sensory profiling data are analyzed. The first data set is obtained by means of the fixed vocabulary profiling procedure with untrained students. The experiment was designed as a circular-balanced, complete block design with 16 assessors, see e.g. Kunert (1998). However, 3 assessors backed out, so there was no circular-balance in the experiment. The experiment had a run-in period. In the run-in period, each assessor received the same beer that he/she were to test as the last one. They were told that this run-in-beer must not be evaluated. They also knew that this beer would come again as the last one. The panelists assessed 5 German beers by scoring each product on a unstructured line scale for 4 attributes : 'bitterness', 'intensity of yeast taste', 'fruity' and 'strength'. The variables were deliberately chosen to have a broad and not very clear meaning. The experiment was intended to illustrate and compare several statistical methods (GPA, ANOVA, …).

Table 1 gives the isotropic scaling factors associated with the assessors. It also gives the isotropic scaling factors obtained by means of GPA. The scaling factors derived from both methods are in good agreement.

Table 1 (about here)

In a second step, each configuration was multiplied by its corresponding scaling factor and a PCA was performed on the table obtained by merging all the data sets. The positions of the samples relative to the first two principal components are depicted in figure 1a. The same configuration obtained by performing GPA is given in figure 1b. The two figures are very similar.

Figure 1a and figure 1b (about here)

A significance test was performed by undertaking a simulation study. At each stage, new data sets are obtained by permuting randomly and independently the original scaled data sets, and a loss function is computed (see appendix). Figure 2 gives the empirical distribution of the loss function after 100 simulations. The loss function for the original data sets is equal to 6.53, which indicates that the performance of the assessors was very poor because the loss function is of the same magnitude as the loss function obtained with „ random “ data. This is not surprising because the assessors in this experiment were not trained and most of them took part in a sensory experiment for the first time. The permutation test (Wakeling *et al*, 1992) was performed in order to assess whether the consensus obtained by GPA was meaningful. It led to the same conclusion.

Figure 2 (about here)

The second set of data was obtained from an experiment involving free choice profiling. The data are discussed by Dijksterhuis and Punter (1990). They are also used by Dijksterhuis and Gower (1991) to illustrate GPA. Seven assessors profiled eight yogurts (labelled 1, 2,..., 8).

The data are analyzed by the method discussed in this paper and the outcomes are compared with those of GPA. The isotropic scaling factors obtained by both methods (table 2) are in good agreement.

table 2 (about here)

The data sets were scaled and merged into one super-matrix on which a PCA was performed. The 8 yogurts are depicted in figure 3a on the basis of the first two principal components. Similarly, figure 3b gives the configuration of the yogurts on the basis of the first two principal components of the group average configurations obtained by means of GPA. It is clear that the graphical displays are similar.

Figure 3a and figure 3b (about here)

A permutation test was performed on the yogurt data. Figure 4 gives the empirical distribution of the loss function after 100 simulations. The observed value of the loss function for the original data sets is equal to 0.99, which indicates that the overall configuration of the products is highly significant. The same conclusion was drawn from a permutation test performed on the consensus obtained by means of GPA (Wakeling *et al*, 1992).

Figure 4 (about here)

Conclusion

The method of analysis discussed in this paper can be recommended as an alternative to GPA. Its rationale is easy to grasp and it does not need specific software packages. Although it is very simple, this statistical method enables the practitioner to cope with the different variations in scoring among the assessors and to analyze data obtained by means of free choice profiling.

References

ARNOLD G. M. and WILLIAMS, A. A. 1986. The use of Generalized Procrustes analysis in sensory analysis. In *statistical Procedures in Food Research,* ed. J. R. Piggott. Elsevier.

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.

COLLINS A. J. 1992. Scaling factors in generalized Procrustes analysis. In *Computational statistics, vol. 1. Proceedings of the 10th symposium on computational statistics.* COMPSTAT, ed. Y. Dodge and J. Whittaker.

DIJKSTERHUIS G. B. and GOWER J. C. 1991. The interpretation of generalized Procrustes analysis and allied methods. *Food Quality and Preference,* **3**, 67-87.

DIJKSTERHUIS G. B. and PUNTER P. H. 1990. Interpreting generalized Procrustes analysis 'analysis of variance' tables. *Food Quality and Preference*, **2**, 255-265.

ESCOFIER B. and PAGÈS J. 1984. L'analyse factorielle multiple. *Cahiers du Bureau universitaire de recherche opérationnelle*, 42, Université Pierre et Marie Curie, Paris.

GOWER J. C. 1975. Generalized Procrustes analysis. *Psychometrika*, **40**, 33-51.

KIERS H. A. L. 1991. Hierarchical relations among three-way methods. *Psychometrika*, **56,** 449-470.

KUNERT J. 1998. Sensory experiments as crossover studies. *Food Quality and Preference. in press*

LAVIT CH. , ESCOUFIER Y. , SABATIER R. and TRAISSAC P. 1994. The ACT STATIS method. *Computational Statistics and Data Analysis*, 18, 97-119.

LEVIN, J. 1966. Simultaneous factor analysis of several gramian matrices. *Psychometrika*, 31, 413 - 419

JAFFRENOU, P. A. 1978. Sur l'analyse des familles finies de variables vectorielles. Thesis, Université de Saint - Etienne. France

QANNARI E. M. , WAKELING I. and MACFIE H. J. H. 1995. A hierarchy of models for analysing sensory data. *Food Quality and Preference*, 6, 309-314.

QANNARI E. M., COURCOUX PH., LEJEUNE M. and MAYSTRE O. 1997. Comparaison de trois stratégies de détermination d'un compromis en évaluation sensorielle. *Revue de Statistique Appliquée.*

SCHLICH P. 1996. Defining and validating assessor compromises about product distances and attribute correlations. In *Multivariate analysis of data in sensory science*. ed. Naes and Risvik. Elsevier.

ROBERT P. and ESCOUFIER Y. 1976. A unifying tool for linear multivariate statistical methods : the RV coefficient. *Applied Statistics*, **25**, 257-267.

WAKELING, I. N. , RAATS, M. N. and MACFIE , H. J. H. 1992. A new significance test for consensus in generalized Procrustes analysis. *Journal of sensory studies*, 7, 91-96.

Appendices

A theoretical property of the consensus

Given a centered matrix X (the true structure of the data) and m centered matrices X_1, X_2 ,... X_m, (the assessments of the judges), which are influenced by rotation and individual scaling factors, such that :

$X_i = c_i XR_i$, where c_i is a scalar and R_i is a rotation matrix e. g. $R_iR_i^T = I$ (I being the identity matrix and R_i^T the transpose matrix of R_i).

We prove that the analysis outlined in the paper leads to the same consensus as the matrix X (except for a multiplicative constant).

The standardization of the matrix X_i leads to the matrix $Y_i = \sqrt{\frac{\sum_j c_j^2}{m}} X_i R_i$. Let $Y = (Y_1 \mid Y_2 \mid ... \mid Y_m)$ be the matrix obtained by merging all the scaled configurations. The normalized principal components of X are given by the eigenvectors of matrix YY^T :

$$YY^T = \sum_i Y_i \; Y_i^T = \frac{\sum_j c_j^2}{m} \sum_i X \;\; R_i R_i^T X^T = \left(\sum_j c_j^2 \right) XX^T .$$

Consequently, Y and X have the same normalized principal components and all the eigenvalues (variances of the principal components) of the matrix YY^T are obtained from those of $X X^T$ by multiplying by the constant $\sum_j c_j^2$.

Loss function for the permutation test

Let Y_1, Y_2 , ... , Y_m be the scaled data sets, a group average configuration C is sought such that the following loss function is minimized :

$$L = \sum_{i=1}^{m} trace\left(\left(Y_i Y_i^T - CC^T\right)\left(Y_i Y_i^T - CC^T\right)\right) = \sum_{i=1}^{m} \left\| Y_i Y_i^T - CC^T \right\|^2 .$$

The matrix $Y_i Y_i^T$ reflects the structure of the data set Y_i as it contains the cross-products between the rows of Y_i which indicate similarities between these rows. The use of the cross-products matrices instead of matrices Y_i leads to a simplification of the calculations as it obviates the determination of rotations that match the configurations because cross-products between individuals (products in our context) do not depend upon the orientation of the configurations (Robert and Escoufier, 1976). The cross-product matrices play a central role in STATIS method.

It can be shown that a solution to the previous optimization problem is given by : $C = \frac{1}{\sqrt{m}} Y$, where Y is obtained by merging the scaled data tables. It follows therefore that a PCA can be carried out on Y in order to depict the relative positions of the products and the relationships among assessors' attributes. Moreover, the loss function L can be used as a means to assess how far removed the assessors configurations are from the consensus $C = \frac{1}{\sqrt{m}} Y$. This loss function is used in the simulation study and the significance test outlined in the paper is based upon the empirical distribution of the loss function L.

TABLE 1.

ISOTROPIC SCALING FACTORS OBTAINED FROM THE ALTERNATIVE ALGORITHM TO GPA AND FROM GPA

(FIRST EXPERIMENT)

Assessor	Alternative algorithm	GPA
A1	0.963	0.8745
A2	0.996	1.000
A3	2.082	2.128
A4	1.122	1.012
A5	1.093	1.167
A6	0.838	0.784
A7	1.062	1.088
A8	0.936	0.941
A9	1.248	1.350
A10	0.890	0.914
A11	0.932	0.989
A12	1.017	0.966
A13	0.796	0.788

TABLE 2.

ISOTROPIC SCALING FACTORS OBTAINED FROM THE ALTERNATIVE ALGORITHM TO GPA AND FROM GPA

(SECOND EXPERIMENT).

Assessor	Alternative algorithm	GPA
A1	1.384	1.412
A2	0.987	1.036
A3	0.754	0.776
A4	1.202	1.225
A5	0.981	0.992
A6	0.987	0.926
A7	1.035	0.987

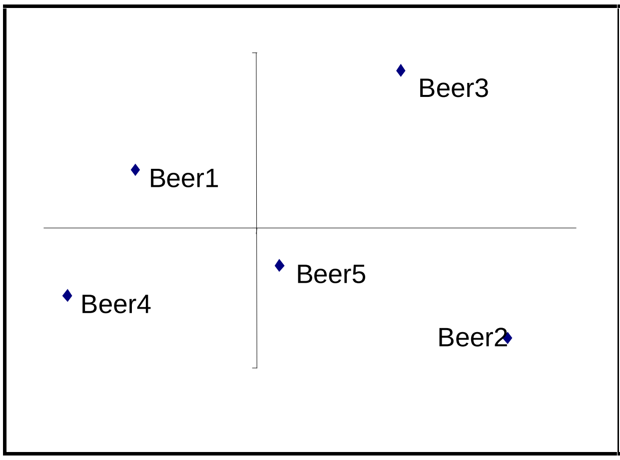

Fig. 1a

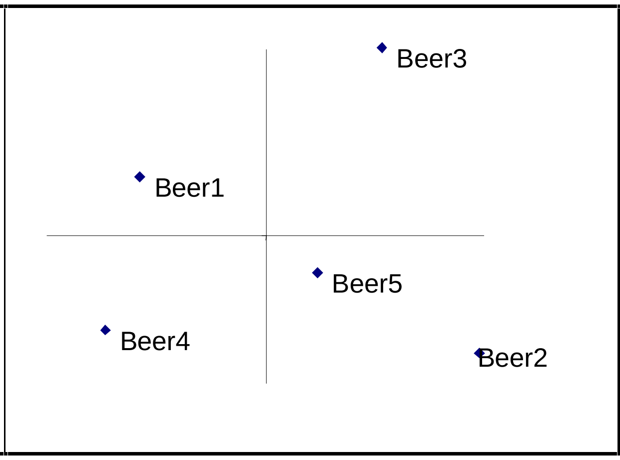

Fig. 1b

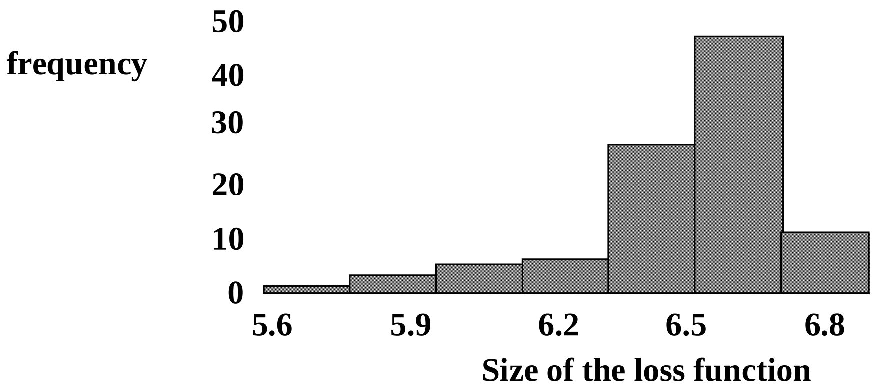

Fig. 2

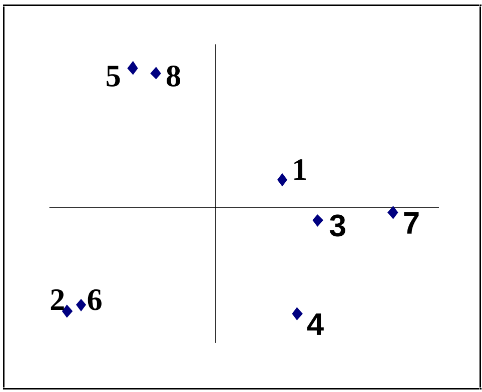

Fig. 3.a

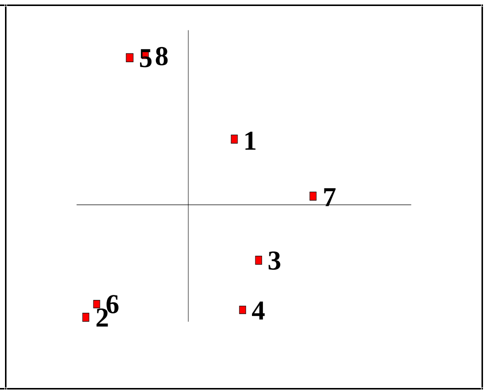

Fig. 3.b

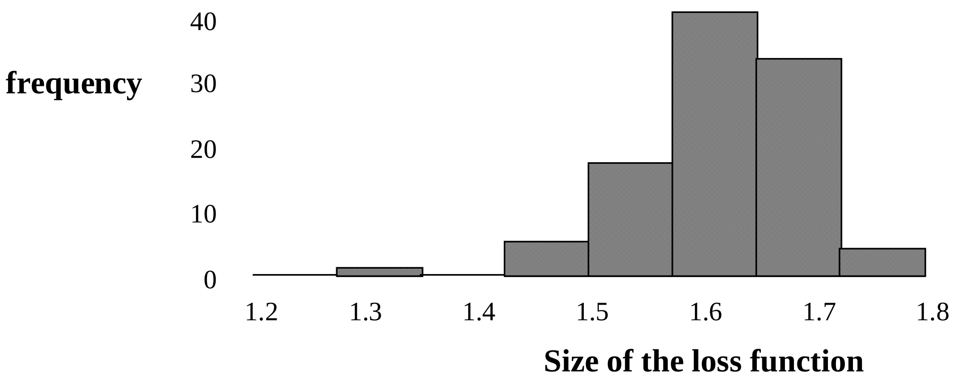

Fig. 4

Figure 1a. First two principal components of the merged data sets, showing the five beers. (Experiment 1)

Figure 1b. First two principal components of GPA group average configuration, showing the five beers. (Experiment 1)

Figure 2. Empirical distribution of the loss function from the simulated data. (The loss function for the original beer data is equal to 6.53). (Experiment 1)

Figure 3a. First two principal components of the merged data sets, showing the eight yogurts (Experiment 2)

Figure 3b. First two principal components of GPA group average configuration, showing the eight yogurts (Experiment 2)

Figure 4. Empirical distribution of the loss function from the simulated data. (The loss function for the original yogurt data is equal to 0.99). (Experiment 2)

Acknowledgment

This work was funded by a Gambrinus Fellowship. The second author wishes to thank warmly the board of trustees of the Dortmund Gambrinus Fellowships and the University of Dortmund for their support.

Part of the work of the first author was supported by the Volkswagen Foundation.