Food Quality and Preference 23 (2012) 25–29

Contents lists available at ScienceDirect

Food Quality and Preference

journal homepage: www.elsevier.com/locate/foodqual

Weighted PLS-discriminant analysis with application to conventional sensory profiling

Stéphane Verdun *, Véronique Cariou, El Mostafa Qannari

College of Veterinary Medicine, Food Science and Engineering, Nantes Atlantic, ONIRIS, Sensometrics and Chemometrics Laboratory, rue de la Géraudière, BP 82225, 44322 Nantes Cedex 3, France

ARTICLE INFO

Article history:
Received 30 November 2010
Received in revised form 8 July 2011
Accepted 9 July 2011
Available online 27 July 2011

Keywords:
Sensory profiling
Fixed vocabulary
Weight assignment
PLS-discriminant analysis

ABSTRACT

Within the framework of sensory conventional profiling, we propose a general strategy of weight assignment to the various statistical units (products × assessors). This means that for each product, an assessor who does not agree with the other assessors on the evaluation of the product under consideration will be down-weighted thus limiting their impact on the outcome of the analysis. The weights are derived from similarity measures between assessors which reflect the extent to which the assessors agree on the position of each product in the perceptual space. The weights are used to compute the usual statistical parameters such as means and variance–covariance matrices. They are also used to compute the PLS-discriminant components which reflect directions of discrimination among the products. The usefulness of the strategy of weight assignment is demonstrated on the basis of a case study where it is shown that it improves the stability of the representation of the products to small perturbations.

1. Introduction

The assessment of the performance of the panel is of paramount interest in sensory profiling studies and not surprisingly this topic has caused much ink to flow. As pointed out by Latreille et al. (2006), some authors advocate the use of the analysis of variance to assess the reliability and the discrimination ability of the panel. In other studies, indices are computed to evaluate the repeatability and reproducibility of the assessors' evaluations (Rossi, 2001). The assessment of the performance of the panel provides guidelines to the sensory analysts in order to identify which assessors need more training and which particular difficulties they are facing. Other strategies that can be adopted to take account of the outcomes of the assessment of the panel performance could be to discard altogether the evaluations of those assessors who turned out to be in high disagreement with the rest of the panel. Alternatively, several authors recommended weighting the assessors taking account of their performance. Thus, assessors with a bad performance are downweighted and their evaluations have a limited impact on the subsequent analyses (Ledauphin, Hanafi, & Qannari, 2006). This weighting strategy is central to methods such as STATIS (Schlich, 1996) or Generalised Procrustes Analysis (GPA) (Collins, 1992; Qannari, MacFie, & Courcoux, 1999). The rationale behind these methods is to assign a unique weight to each assessor that reflects their overall agreement with the rest of the panel. However, it may occur that the performance of a given assessor is highly affected by the disagreement of this assessor with the rest of the panel with respect to one particular product. For instance, this assessor may be given by mistake a wrong product or there might be a high variability for this specific product. The rationale behind the strategy of analysis that we advocate herein is to down-weight specifically the evaluation of the assessor under consideration for this specific product. In other words, we propose a general strategy of assigning weights to the various statistical units or cases formed by (assessors and products). This means that we will attach to each assessor as many weights as there are products with the understanding that a relatively high weight associated with a given product reflects a high agreement of the assessor with the rest of the panel on the evaluation of the product under consideration and vice versa. Thereafter, the system of weights associated with each assessor is used in the subsequent analyses to compute weighted means, weighted covariance matrices, etc. The implication is that these analyses are more robust because they are less influenced by marginal evaluations. We focus hereinafter on the use of PLS-discriminant analysis (PLS-DA), taking the products as groups. Indeed, this method of analysis stands at the cross-roads of most of the usual strategies of analysis of conventional sensory profiling data. On the basis of a case study, we will illustrate the weight assignment strategy and show that it improves the stability of the final products configuration obtained by means of PLS-DA.

It is worth mentioning that several statistical methods aim at identifying some variations among the assessors and, in order to minimise these variations, some authors introduced *ad hoc* param-

* Corresponding author.
E-mail address: stephane.verdun@oniris-nantes.fr (S. Verdun).

0950-3293/$ - see front matter © 2011 Elsevier Ltd. All rights reserved.
doi:10.1016/j.foodqual.2011.07.002

eters in the models. A typical example of such a strategy of analysis was proposed within the context of the analysis of variance by Brockhoff and Skovgaard (1994). However, this strategy considers one sensory attribute at a time whereas our strategy of analysis operates in a multivariate setting.

The rest of the paper is organised as follows. In Section 2, the strategy of assigning weights to the assessors is outlined. The weights thus obtained are used within PLS-DA and more generally within factorial methods. Thereafter, a case study pertaining to sensory evaluation of ciders is discussed and the results are detailed in Section 3. We end the paper by drawing general conclusions and future developments.

2. Material and methods

2.1. Pre-treatment of the data

Let us assume that m assessors have scored n products according to a set of p attributes. The data associated with each assessor can be presented as an $(n \times p)$ matrix denoted X_k^* ($k = 1, \ldots, m$). The rows of this matrix refer to the products and the columns to the attributes. In order to cope with some known variations among the assessors, each matrix X_k^* is centred by subtracting from the entries of each column the average of the column under consideration. This makes it possible to remove the assessors' main effect (or shift effect) which consists in the fact that assessors may use different levels of the scoring scale. Another source of variation lies in the different ranges of the scoring scales. Isotropic scaling factors are generally applied to circumvent this problem. This consists in multiplying each dataset X_k^* ($k = 1, \ldots, m$) by a scaling factor α_k in order to shrink the configurations of those assessors who have a tendency to use large ranges of the scales and, contrariwise, expand the configurations of those assessors who have a tendency to use relatively narrow ranges of the scales. Appropriate scaling factors can be computed as follows (Kunert & Qannari, 1999).

- Compute t_k as the total variance of dataset X_k^*. Formally, t_k corresponds to the sum of the variances of the columns of X_k^*;
- Compute t as the average of t_k ($k = 1, \ldots, m$);
- Set $\alpha_k = \sqrt{\frac{t}{t_k}}$.

By multiplying each dataset X_k^* by its associated scaling factor α_k,we obtain new datasets which have the same total variance that is, t.

In the following, we denote by X_k the $(n \times p)$ matrix which is obtained from X_k^* by centring the columns and multiplying by the isotropic scaling factor α_k.

2.2. Weight assignment

Ledauphin et al. (2006) proposed a strategy of weight assignment to the assessors whereby a weight is attached to each assessor depending on his or her overall agreement with the rest of the panel. However, an overall agreement may hide a large disparity ranging from the situation where an assessor is in complete disagreement for all the products, to the situation where an assessor disagrees with the rest of the panel on the evaluation of only few cases that is, one or few products. In order to take account of this situation, different strategies can be adopted. For instance, one may identify those particular cases corresponding to high disagreement and discard them altogether. Anzanello, Fogliatto, and Rossini (2011) proposed to retain only those assessors and attributes that achieve a good discrimination of the products. For this purpose, they defined overall indices associated with the assessors (Ledauphin et al., 2006). Obviously, such a strategy of analysis poses the problem of finding a dividing line between those data which should be discarded and data which should be kept. We propose a strategy of analysis which consists in assigning weights to the various cases (assessors $\times$ products) taking account of the agreement of the assessors on the evaluation of each product. In other words, instead of assigning a global weight to each assessor, we propose to assign a weight to each case (assessor $\times$ product) taking account of how each product is assessed by the various assessors.

The weight assignment strategy is based on the determination for each product of a similarity measure between assessors. Such a similarity measure reflects the extent to which the assessors agree on the position of the product under consideration in the perceptual space. Thereafter, the rows of the similarity matrix are scaled to unit sums thus leading to a so-called stochastic matrix. The dominant left eigenvector corresponding to the stationary probability vector of the stochastic matrix is extracted. Finally, the components of this vector are assigned to the assessors as weights, reflecting the performance of each assessor for this particular product. It should be noted that the entries of the similarity measure should be positive but the similarity matrix is not necessarily symmetric. Examples of dissimilarity measures will be discussed in the following section.

The procedure of weight assignment is an adaptation of the general strategy of weight assignment proposed by DeGroot (1974). Starting from a stochastic matrix $P = (p_{ij})$; ($i, j = 1, \ldots$, m; $p_{ij} \geqslant 0$; $\sum_{j=1}^{m} p_{ij} = 1$), DeGroot considered that each row is associated with a judge and reflects degrees of confidence assigned by this judge to the members of the panel, including himself. In an attempt to reach a consensus regarding a parameter of interest, θ (say), each member of the group is assumed to revise his or her own assessment of parameter θ to accommodate the information of the rest of the group. More precisely, it is also assumed that each revised assessment of θ by a judge is a weighted linear combination of the individual assessments $\theta_1, \theta_2, \ldots, \theta_m$ given by the members of the panel. Namely:

$$\theta_i^{(new)} = \sum_{j=1}^{m} p_{ij} \theta_j^{(old)}$$

Thus, the revision process is assumed to be iterative with a constant matrix of weights. Under very weak assumptions on matrix P similar to those assumed for the transition matrix of a Markovian process, Degroot proved that the individual assessments of parameter θ converge to the same limit: $\theta = \sum_{i=1}^{m} w_i \theta_i$; where $w = (w_1, w_2, \ldots, w_m)$ is such that $wP = w$ and $\sum_i w_i = 1$. That is w is the first eigenvector (normalised such that the sum of its components is equal to 1) associated with the eigenvalue 1.

The rationale behind Degroot's strategy of weight assignment precisely fits our purpose if we consider that the assessors are trying to reach a consensus about the position of each product in the perceptual space.

2.3. Examples of similarity measures

As stated above, the first step consists in defining, for each product, a similarity measure among the assessors. Two examples of similarity measures are given herein. However, one should bear in mind that the strategy of analysis is flexible enough to allow other choices of (positive) similarity measures.

The first similarity measure, based on the Gaussian function, is popular within the framework of data analysis and data mining. For a given product, let us denote by x_i and x_j the two profiles of this product given by two assessors i and j (say). These two profiles consist of the (pre-treated) scores given by assessors i and j with respect to the p (common) attributes. The similarity between the