PLS discriminant analysis applied to conventional sensory profiling data

K. Rossinia,b,c,d, S. Verduna,b,c, V. Carioua,b,c, E.M. Qannari,a,b,c,*, F.S. Fogliattd

* Corresponding author at: ONIRIS, USC “Sensometrics and Chemometrics Laboratory”, Nantes F-44322, France.
E-mail address: elmostafa.qannari@oniris-nantes.fr (E.M. Qannari).

Contents lists available at ScienceDirect

Food Quality and Preference journal homepage: www.elsevier.com/locate/foodqual

1. Introduction

The main goal of the statistical treatment of conventional sensory profiling data is to exhibit inter-product differences while handling the variations among assessors. That is often accomplished by means of factor analytical methods which enable the investigation of similarities between products on the basis of graphical displays. Different approaches have been proposed to address this issue; among them we single out principal components analysis (PCA), generalized procrustes analysis (GPA), and canonical variates analysis (CVA). Each approach has advantages and drawbacks. We focus on conventional sensory data, also referred to as sensory data with a fixed vocabulary. This type of data may be obtained by means of different sensory evaluation procedures (e.g. the quantitative descriptive analysis protocol by Stone & Sidel, 1998), and may be organized as a three way data matrix (products x attributes x assessors). However, depending on the statistical method to be performed on the data, other data arrangements may be more convenient.

PLS discriminant analysis (PLS-DA) is yet another method which is suited for the analysis of conventional sensory profiling, and stands at the intersection of the methods mentioned above. It yields interesting indices, such as the between to total variance ratios, which reflect the agreement among assessors and the discrimination among products. The VIP indices (variable importance in the projection) are also of paramount interest as they highlight the importance of the various attributes. They may be useful in guiding the selection of a subset of relevant attributes from the complete set of attributes. We also show how the graphical displays of the products may be enhanced by using confidence ellipses obtained by means of assessors’ re-sampling (bootstrap). The outcomes of PLS-DA are compared to those of alternative methods through a case study pertaining to the sensory evaluation of varieties of cider.

The rest of the paper is organized as follows. We start by discussing the pre-treatment of the data in order to cope with some known sources of variation among assessors. Next we outline the most popular methods to analyze conventional sensory profiling data and focus on PLS-DA, presenting its advantages in comparison with other methods. In particular, we discuss the VIP indices which are popular within the framework of PLS regression and PLS-DA, and present how they may be a useful guiding criterion to select a subset of relevant attributes from a complete set. Finally,
A comparison of PLS discriminant analysis with other methods is illustrated using a case study dataset. In particular, the various methods' stability is investigated using assessors' re-sampling (bootstrap) and confidence ellipses.

2. Material and methods

2.1. Pre-treatment of the data

Assume that $m$ assessors perform the sensory profiling of $n$ products using a set of $p$ attributes. Data obtained from assessor $k$ ($k=1, \ldots, m$) is organized in a ($n \times p$) matrix $X_k$, with rows referring to products and columns referring to attributes. To cope with some known sources of variation among assessors it is recommended to center each matrix $X_k$ by subtracting from the entries in each column its corresponding average. Centring removes the assessors' main effect (or shift effect) which will be present if assessors use different levels of the scoring scale. Another source of variation among assessors is related to their differences in the ranges of the scoring scale. Isotropic scaling factors are usually introduced to address this problem, as follows. Multiply each dataset $X_k$ by its associated scaling factor $a_k$ to (i) shrink configurations of assessors with tendency to use large ranges of the scoring scale (for that, make $a_k < 1$), or (ii) expand configurations of assessors with tendency to use relatively narrow ranges of the scoring scale (for that, make $a_k > 1$). Appropriate scaling factors may be computed as follows (Kunert & Qannari, 1999):

- Determine $t_k$, the total variance of the dataset $X_k$, by adding the variances of each column of $X_k$.
- Compute $t$ as the average of $t_k$ ($k=1, \ldots, m$).
- Set $a_k = \sqrt{t / t_k}$.

Multiplying each dataset $X_k$ by its associated scaling factor $a_k$ yields new matrices with the same total variance, $t$. In what follows, we denote by $X_k$ the ($n \times p$) matrix obtained by centring the columns of $X_k$ and multiplying the resulting matrix by the isotropic scaling factor $a_k$.

It is worth noting that the pre-scaling of the data is optional depending on whether the sensory analyst considers that the differences among the assessors in the range of scoring are merely an experimental artifact or reflect genuine differences among the products (Macfie & Hedderley, 1993). The pre-scaling is proposed as an option in methods such as GPA (Dijksterhuis & Gower, 1991; Gower, 1975) or STATIS (Schlich, 1996) but there is no reason why it should not be proposed as such prior to more common methods such as PCA on the average dataset.

2.2. Statistical treatments of sensory profiling data

There are different ways to organize the data obtained from a conventional sensory profiling procedure. Certain statistical analyses follow naturally from each choice of arrangement, as depicted in Fig. 1.

As stated previously, sensory profiling data may be presented as a three way array in which entries refer to products, attributes and assessors, and methods devoted to this type of arrangement (e.g. PARAFAC) may be performed on the data (Bro, Qannari, Kiers, Næs, & Frøst, 2008; Brockhoff, Hirst, & Næs, 1996; Cocchi et al., 2006). Such methods are not very popular in sensory evaluation probably because their rationale is not fully grasped by practitioners. Notwithstanding, we believe that they should be further investigated by sensory analysts.

2.2.1. Average dataset

The most popular practice in sensory analysis is the averaging of datasets over assessors, resulting in a two way dataset (prod-
ucts × attributes). Obviously, the rationale behind such practice is to consider evaluations given by assessors as replicates which differ from each other by random noise. PCA is usually performed on the average dataset in order to depict the relationships among products. However, this strategy of analysis does not explore the within-products variation structure, and unless further investigation is carried out it does not provide tools, such as indices and graphical displays, to assess differences among the assessors.

2.2.3. Vertically unfolded data

In this case each dataset $X_k$ is placed sideways horizontally, and the resulting supermatrix $X^v$ is suitable to be analyzed using methods developed within the framework of multiblock datasets. The assumption is that although using the same attributes, assessors might interpret them differently, and analytical methods that adjust for this source of variation should be considered. Undoubtedly, the most popular among such methods is GPA (Dijksterhuis & Gower, 1991; Gower, 1975). Alternatively, a family of methods based on performing PCA on $X^v$ may also be used; MFA (Husson, Le, & Pagès, 2005) and STATIS (Schlich, 1996) are examples of such methods. However, in the sections to follow, these methods are not considered since we are concerned with genuine situations of conventional sensory profiling, where the assumption is that assessors interpret attributes similarly.

2.2.3. Vertically unfolded data

In this case datasets $X_k$ are stacked up vertically; the resulting supermatrix $X^v$ has $(n \times m)$ rows and $p$ columns. PCA may be performed on $X^v$ and we refer to Luciano and Naes (2009) for an interesting discussion on how the PCA outcomes could be used in conjunction with ANOVA to investigate both the sensory data structure and the similarities among products. For the same analytical purpose, Monrozier and Danzart (2001) among others propose the use of canonical variate analysis (CVA). In CVA the number of groups is equal to the number of products, and we seek components (also called canonical variates) that best discriminate the products. As a result, we want products as far removed from each other as possible (maximizing the variation between groups) and observations within groups (i.e. products) as much clustered around their centroids as possible (minimizing the variation within groups). CVA is clearly more aligned with the objectives of sensory profiling. However, it is well known that, similar to multiple linear regression, CVA may lead to unstable results in the presence of high colinearity among attributes; PLS-DA is recommended to overcome such drawback (Barker & Rayens, 2003; Naes & Indahl, 1998; Nocairi, Qannari, Vigneau, & Bertrand, 2005). Martens and Martens (2001) were the first to propose the use of PLS-DA on sensory profiling data but, to the best of our knowledge, this paper is the first to elaborate on the topic.

2.3. PLS-DA applied to conventional sensory profiling

In short, PLS-DA seeks to determine components, also called latent variables, which maximize the variation between groups (products); see Barker and Rayens (2003), and Nocairi, Qannari, Vigneau, and Bertrand (2005). From this standpoint, PLS-DA appears to be at the intersection of the following three analytical approaches:

(i) PCA applied on the average dataset – the objective is to recover the total variance in the sensory data averaged over all assessors (i.e. the between-products variation), but assessors’ individual data are not taken into account.
(ii) PCA applied on matrix $X^v$ – the objective is to recover the total variance from all datasets, but the presence of groups (products) in the data is not explicitly taken into account in the computation of the principal components.
(iii) CVA – the objective is to recover the between-products variation while minimizing the within-products variation; however, as stated above, this method is sensitive to the presence of colinearity among attributes.

PLS-DA yields graphical displays and offers interpretation tools that enable investigating the structure of the sensory data. PLS-DA components may be used to represent the rows in matrix $X^v$. We may also represent on the basis of the same components any given product as the centroid (or average) point of the rows corresponding to this product; that is, the point which stands at the barycenter of the assessors’ observations corresponding to this product. The number of significant components to be retained is usually determined by a cross-validation procedure (Martens & Naes, 1989). Moreover, ANOVAs may be performed on PLS-DA components rather than on principal components extracted from matrix $X^v$, as proposed by Luciano and Naes (2009). PLS-DA components are likely to offer a better discrimination of products if compared to PCA components. We may also compute a discrimination index for each component, given by the ratio of the between-products variance (i.e. the variance of the averaged scores over all assessors) to the total variance. The same calculation procedure may be extended to components derived through PCA performed on $X^v$ and through CVA. In the case of principal components derived from PCA on the average dataset, the discrimination indices correspond to the same ratio above, with variances obtained as follows. The between-products variance is equivalent to the variance of the principal component scores. To obtain the total variance, individual data from each assessor are superimposed on the principal component under consideration (i.e. the same vector of loadings associated with this principal component is applied to the individual datasets); the variance of the superimposed scores is used as an estimate of the total variance.

From a technical standpoint, PLS-DA seeks, step by step, latent variables (or components) which are linear combinations of the columns in $X^v$. Let us denote by $t = X^v a$ the first latent variable, where $a$ is the vector of loadings constrained to unit length. As mentioned above, $t$ is sought such that the between groups (or products) variance is as large as possible. The solution to this problem leads to set $a$ as the eigenvector of the between-groups variance–covariance matrix associated with the largest eigenvalue. Thereafter, the so-called deflation procedure is applied in order to determine a second latent variable. This consists in regressing all the variables in $X^v$ upon $t$ and considering the dataset formed by the residuals. Then, taking this latter dataset instead of $X^v$ the same procedure aiming at maximizing the between groups variance is again performed, thus leading to a new latent variable which is by construction orthogonal to $t$. The same strategy can be reiterated in order to determine subsequent latent variables.

Table 1

<table>
<thead>
<tr>
<th>Assessor</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
</tr>
</thead>
<tbody>
<tr>
<td>Isotropic scaling factor</td>
<td>1.207</td>
<td>1.117</td>
<td>1.020</td>
<td>0.737</td>
<td>0.819</td>
<td>10.854</td>
<td>1.087</td>
</tr>
</tbody>
</table>
Within the context of PLS regression and PLS-DA, the VIP (variable importance in the projection) indices are of paramount interest (Chong & Jun, 2005). VIPs are associated with attributes, and reflect their contribution in discriminating the products. They may be computed for each component separately or they may be computed for a PLS-DA model which includes several components. As a rule of thumb proposed by Wold, Johansson, and Cocchi (1994), we may discard from the model attributes with small VIP values (smaller than 0.8). Therefore, as a strategy to select a subset of attributes from a larger set one may set up a model with an appropriate number of components, and discard attributes with small VIP values (Chong & Jun, 2005).

### Table 2

<table>
<thead>
<tr>
<th>Method</th>
<th>Axis 1</th>
<th>Axis 2</th>
<th>Axis 3</th>
</tr>
</thead>
<tbody>
<tr>
<td>PLS-DA</td>
<td>0.887</td>
<td>0.701</td>
<td>0.417</td>
</tr>
<tr>
<td>CVA</td>
<td>0.929</td>
<td>0.781</td>
<td>0.430</td>
</tr>
<tr>
<td>PCA on the average set</td>
<td>0.886</td>
<td>0.698</td>
<td>0.410</td>
</tr>
<tr>
<td>PCA on the concatenated datasets</td>
<td>0.868</td>
<td>0.600</td>
<td>0.274</td>
</tr>
</tbody>
</table>

2.4. Confidence ellipses

In sensory profiling analysis, Husson et al. (2005) have stressed the benefits of setting up confidence ellipses on the graphical displays that depict the similarity between products. That enables identifying products that are significantly different from the group in a multivariate setting. The authors propose a bootstrapping approach to set up the confidence ellipses. The approach relies on intensive computation since it generates a large number of (virtual) panels, which are in turn submitted to PLS-DA. As a result, it is possible to assess fluctuations in the position of the products on the graphical displays. Ideally, these fluctuations should be small, reflecting a good stability of the model under consideration.

The resampling (bootstrap) strategy may be implemented in four steps:

**Step 1.** Perform a PLS-DA on the sensory data obtained from \( m \) assessors. Assume that a PLS-DA model with \( r \) components is retained. As stated above, these components may be used to depict relationships between products.

**Step 2.** Perform a resampling from the \( m \) assessors. In other words, create a new panel comprised of \( m \) assessors by randomly selecting assessors from the initial panel, with replacement.

**Fig. 2.** Configuration of products on the first factorial plan obtained using (a) PLS-DA, (b) CVA, (c) PCA on the average dataset and (d) PCA on the concatenated dataset.
Consequently, in a bootstrapped panel a given assessor may be selected several times whereas others may never be selected. Step 3. Perform a PLS-DA on data from the bootstrapped panel and retain a model with \( r \) components. Next, the new positions of the products are adjusted to the original positions obtained in Step 1 by means of a procrustean rotation (see, for instance, Krzanowski, 2000).

Step 4. Reiterate steps 2 and 3 a large number of times (say, 5000 times). Eventually, for each product, a confidence ellipse containing a desired percentage (e.g. 95%) of the resampled points associated with this product may be drawn.

We use the bootstrap resampling technique described above to compare several methods for the analysis of conventional sensory profiling data; namely: PCA on the average dataset, PCA on the supermatrix \( X^v \), CVA, and PLS-DA.

2.5. Case study

To illustrate the use of PLS-DA in conventional sensory profiling and compare its outcomes with those obtained using other methods of analysis, we consider a case study in which a quantitative descriptive analysis is performed on ten varieties of cider. The sensory panel is formed by seven trained assessors who were asked to score products using a list of ten sensory attributes: sweet, intensity of odor, acid, bitter, astringency, strength, pungent, alcohol, perfume and fruity.

3. Results and discussion

3.1. Pre-treatment of sensory data

Table 1 shows the isotropic scaling factors which were applied to the datasets associated with the seven assessors. Isotropic scaling factors smaller than 1 indicate that configurations of the associated assessors were shrunk to correct their tendency to use a relatively large range of the scale; that corresponds to assessors 4 and 5 in Table 1. In opposition, configurations of assessors 1, 2, 6, and to a lesser extent assessors 3 and 7, were expanded since their associated scaling factors are larger than 1.

3.2. Discrimination indices

PLS-DA was performed on the cider data. In a cross-validation procedure only the first three components turned out significant. To assess the extent to which products are discriminated by the three retained PLS-DA components, we propose computing for each component a discrimination index given by the between-products variance to total variance ratio.

Fig. 3. Configuration of products on the first factorial plan obtained using (a) PLS-DA, (b) CVA, (c) PCA on the average dataset, and (d) PCA on the concatenated dataset and their corresponding 95% confidence ellipses.
Table 2 gives the discrimination indices associated with the first three PLS-DA components. To allow comparisons, we also present the discrimination indices obtained through CVA, PCA performed on X' and PCA performed on the average dataset. Not surprisingly, CVA gives the largest discrimination indices. In fact, as stated above, CVA aims at maximizing the between-products variance and minimizing the within-products variance, and we can show that this is equivalent to maximizing the discrimination index (between-products variance to total variance ratio).

For the first component, PLS-DA and PCA on the average dataset have the same performance in terms of discrimination. In general, PLS-DA leads to better results than the other methods, except for CVA.

3.3. Graphical displays

Fig. 2 displays the biplots where both variables and products are plotted on the same plane. For simplicity, we have retained only configurations on the first factorial plan, comprised of the first two retained components. To allow comparisons, we also present the biplots obtained through CVA, and through PCA on the average dataset and on the dataset obtained by stacking the assessors’ datasets on top of the others. It is clear that the four configurations lead to similar results. The first PLS-DA component (axis 1) opposes ciders 4, 8 and 10 to ciders 7, 5, 2, and 6. The former group of ciders is sweeter, fruitier, more perfumed, less bitter, less strong and less pungent than the second group. The second component (axis 2) mainly singles out cider 9 which has a more intense odor.

3.4. Confidence ellipses around the products

Fig. 3 presents the 95% confidence ellipses around the products on the basis of the two first components. These ellipses were obtained using the bootstrapping procedure in Section 2.4. They reflect the variability of the sensory evaluations and the discrimination of products in a multivariate setting. Four sets of ellipses are displayed, one for each analytical approach. Regardless of the method used to analyze the sensory data, products 1, 3 and 9 are clearly separated from the others. Product 10 is separated from the other products, except in the graphical display obtained through CVA. PLS-DA provides a good separation of products 2 and 6, while for the other methods of analysis these products’ confidence ellipses overlap with those of other products. In general, it is clear that PLS-DA best discriminates products since their corresponding confidence ellipses are smaller than those obtained by alternative methods. In addition, it is noteworthy that CVA which is by principle focused on the discrimination of products generates confidence ellipses that overlap the most. That is probably due to the instability of the method when colinearity among attributes is present in the dataset.

As stated above, in order to save space the configurations of the products on the basis of the third component are not shown herein. However, it is worth mentioning that they follow exactly the same pattern as in Fig. 3 (i.e. similar positioning of the products but relatively smaller confidence intervals for PLS-DA than for the other methods).

3.5. VIP indices and selection of a subset of attributes

Fig. 4 presents the VIP indices associated with a PLS-DA model comprised of three components, and their 95% confidence intervals obtained by means of the bootstrap procedure in Section 2.4. It highlights the importance of the attributes fruity, perfume, intensity of odor and sweet. The least important attributes are acid, bitter astringent and strength of odor which have VIP indices smaller than 0.8 or, considering the confidence intervals, not significantly larger than 0.8; in subsequent studies these attributes could be discarded thus saving time and fatigue to assessors. For completeness we have discarded these attributes and performed a PLS-DA on the remaining variables. Fig. 5 presents the configuration of products on the first factorial plane corresponding to the first two PLS-DA components, as well as the 95% confidence ellipses obtained by means of bootstrap resampling. As expected, the overall structure of product configuration is preserved; i.e. disregarding attributes with low VIP values does not compromise the characterization of products.

4. Conclusion

In this paper we propose the use of PLS-DA for the analysis of conventional sensory profiling. PLS-DA stands at the crossroads of popular methods of statistical treatment of conventional sensory profiling data, namely PCA on the concatenated dataset, PCA on average dataset and CVA. It provides statistical tools which yield a better interpretation of the analytical outcomes; among them we single out graphical displays, discrimination indices and VIP indices. The last two indices reflect the importance of variables
in the discrimination and may be used within a strategy to select a subset of attributes from a larger set. In addition, assessors re-sampling (bootstrap) techniques provide tools which enable checking the stability of outcomes and the discrimination of products in a multivariate setting. In the light of a case study presented in the paper it turned out that PLS-DA provides better stability and achieves better discrimination of products than competing methods.

Future research will be deployed in the following directions: (i) deeper analysis on the merits of PLS-DA over alternative methods; (ii) broader investigation on the proposed strategy to select a subset of attributes, and comparison with other variable selection methods (as reviewed in Sahmer & Qannari, 2008); and (iii) proposition of a method to directly relate PLS-DA outcomes to preference data, when such data are available. Research deployment (iii) is particularly challenging since it consists of relating preference data to sensory data, while taking account of the individual assessors' evaluations instead of merely considering averaged scores, as it is usually done.

---

**References**


