Weighted PLS-Discriminant analysis with application to conventional sensory profiling

Stéphane Verdun  Véronique Cariou  El Mostafa Qannari

Sensometrics and Chemometrics Laboratory
Oniris, Nantes

10th Sensometrics
Overview

1. Discrimination of the products in sensory profiling
2. Weight assignment
3. Illustration
4. Conclusion
Conventional sensory profiling

The assessors score the products for various descriptors, leading to the 3-way matrix $X$:
Popular methods

- **PCA**
  - **Average dataset**
  - **Assessors x Products**

- **PCA**

- **CVA**
  - **Product 1**
  - **Product n**

- **PLS-DA**
  - **Product 1**
  - **Product n**
Focus on PLS-DA in this presentation
Seek components which maximize the between product variation
See poster by Rossini et al.
Assessors’ performance

- A good performance ensures a good discrimination
- Which actions should be taken in case of bad performance?
  - Discard the assessors
  - Downweight the assessors (Statis (Schlich, 1996), GPA (Qannari et al., 1999))
Overall downweighting of assessors

General strategy involves in GPA or Statis:

- Objective: find weights for the assessors, according to their agreement
- Compute the similarity matrix between the assessors
- Extract the first eigenvector
- Assign the components of the eigenvector as assessors’ weights
It may happen that an assessor has a good agreement with the panel except for one specific product

→ Downweight each case (assessor x product)
Within each product

- Compute a similarity matrix $S$
- Normalize to a stochastic matrix $P$
- Extract the dominant eigenvector
- Assign the components of the eigenvector as weights

Weighted PLS-DA

- The weights can be used to compute robust means, variances...
- The algorithm of weighted PLS-DA is the same as PLS-DA except that the means and the between products covariance matrix is computed using the weights
Example of similarities between two cases $i$ and $j$

- Gaussian similarity:
  \[ s_{ij} = \exp\left(-\frac{d_{ij}^2}{2\sigma^2}\right) \text{ where } \sigma \text{ is a tuning parameter} \]

- Proportion of common neighbours within the $k$ nearest neighbours, $k$ is a tuning parameter
How to tune the parameter $\sigma$ or $k$ (number of nearest neighbours) ?

Jacknife procedure (leave-one-out) on the assessors and choose a parameter $\sigma$ or $k$ that ensures the highest stability of a two (or three...) dimensional representation of the products (by means of PLS-DA).
Data

- QDA experiment
- 10 varieties (ciders) evaluated according to 10 descriptors by 7 assessors
Discrimination of the products in sensory profiling

Factorial plane

**Figure:** First factorial plane - Red squares are weighted means and the green triangle is a classic mean

**Figure:** First factorial plane - Map of the variables
Example of weights

Weights for some products and some assessors:

<table>
<thead>
<tr>
<th></th>
<th>P1</th>
<th>...</th>
<th>P3</th>
<th>...</th>
<th>P5</th>
<th>P7</th>
<th>P9</th>
<th>Mean weight</th>
</tr>
</thead>
<tbody>
<tr>
<td>A1</td>
<td>0.063</td>
<td>...</td>
<td>0.086</td>
<td>...</td>
<td>0.183</td>
<td>0.056</td>
<td>0.143</td>
<td>0.132</td>
</tr>
<tr>
<td>A2</td>
<td>0.055</td>
<td>...</td>
<td>0.195</td>
<td>...</td>
<td>0.100</td>
<td>0.204</td>
<td>0.206</td>
<td>0.157</td>
</tr>
<tr>
<td>A3</td>
<td>0.162</td>
<td>...</td>
<td>0.116</td>
<td>...</td>
<td>0.100</td>
<td>0.167</td>
<td>0.206</td>
<td>0.165</td>
</tr>
<tr>
<td></td>
<td>...</td>
<td>...</td>
<td>...</td>
<td>...</td>
<td>...</td>
<td>...</td>
<td>...</td>
<td></td>
</tr>
<tr>
<td>A5</td>
<td>0.169</td>
<td>...</td>
<td>0.143</td>
<td>...</td>
<td>0.176</td>
<td>0.000</td>
<td>0.048</td>
<td>0.093</td>
</tr>
</tbody>
</table>

**TABLE:** Example of weights and mean weights for some assessors and some products

The assessors’ capability over all products can be measured by the mean of the weights, but there is a loss of information.
A perturbation is introduced in the data by permutating the answers of an assessor.

We have instances of (local) disagreement involving product 3.

The same with the product 9.
Confidence ellipses

Perturbated products: products 3 and 9

**Figure:** First factorial map, without weights

**Figure:** First factorial map, neighbourhood similarity
Conclusion

- Better insight into the assessors’ performance
- The weighting strategy improves the stability of the factorial plane, leading to more robust representations of the products
- The weighting strategy is flexible (use of different similarities) and versatile (use within different factorial methods)
- The parameters of the similarities can be tuned according to different objectives: stability, discrimination...
