Lecture for CMA4CH 2010, Mediterraneum Meeting

Application of Multivariate Analysis and Chemometry to Cultural Heritage and Environment
3rd ed., Taormina, Sicily island, Italy, Europe, 26-29 September 2010

Three-way methods for the analysis of hyphenated chromatographic data

F. Marini, A. D'Aloise, R. Nescatelli, R. Bucci, A.D. Magrì, A.L. Magrì
Dept. Chemistry, University of Rome "La Sapienza", P.le Aldo Moro 5, 00185 Rome, Italy

Abstract

Hyphenated chromatographic instruments, coupling the elution dimension with a multivariate detection method (usually a spectroscopic device) provide so-called second order data, i.e. each analyzed sample is described by a signal that takes the form of a 2-dimensional landscape. When more than one sample is analyzed, the data-set becomes what is defined as a three-way array (or a data cube).

In this communication, the main chemometric techniques used to deal with three-way data will be presented, focusing on their application to the solution of chromatographic problems.

Introduction

Chromatographic analytical systems are being increasingly used for the analysis of many complex samples, for instance in the fields of drug and food chemistry, but also in proteomics and metabolomics. In particular, hyphenated chromatographic systems, such GC-MS, HPLC-DAD and HPLC-MS, are extensively used to obtain detailed qualitative and quantitative information about the analyzed samples. Indeed, the high sensitivity, the low limit of detection, the possibility of analyzing a great number of analytes and identifying these using the spectral dimension, makes hyphenated chromatographic systems among the most widespread analytical techniques in many scientific fields.

Accordingly, many approaches have been proposed during the years to deal with multidimensional data obtained from these techniques in the situation when there is perfect resolution of the eluting peaks. However, in more complex cases, analysis of hyphenated chromatographic data is sometimes hampered by different problems, mainly derived from the chromatographic separation and/or spectral measurements. Sometimes, it is impossible to achieve perfect separation, because the samples are complex or faster chromatographic runs are preferred.

Additionally, when the samples show a similar spectral profile, even moderate overlap between the peaks can result in an unreliable calibration. Moreover, other fundamental problems such as low signal-to-noise ratio of peaks, resulting in background interference in the spectral dimension, or varying baseline and peak shifts may decrease the quality of the final result of the analysis.

In this communication, the main chemometric tools to deal with those kind of problems will be presented by means of specific examples.

A brief presentation of the methods

Different techniques will be presented but the main attention will be focused to the multi-set implementation of Multivariate Curve Resoultion (MCR) and on PARAFAC/PARAFAC2. The hypothesis of MCR [1] is that the overall chromatographic landscape for a sample can be decomposed into the contribution of the elution profile and the spectral fingerprint of individual components. This means that, ideally, if two analytes are coeluting and their spectral profile is sufficiently different, the corresponding chromatographic landscape resulting in overlapping signals can be separated into the individual contribution for the two chemical species. When more than one sample is measured, since the method works on matrix decomposition, an unfolding step is necessary: operationally this means that the individual landscapes corresponding to the different

samples are aligned one after the other, providing that they share the same spectral dimension: this corresponds to the hypothesis that the same component are present in all the samples but allows for the elution profiles to be different not only in terms of analyte concentration but also of peak positions (therefore allowing to deal with the presence of shifts) [1].

Alternatively, tensor methods for the direct decomposition of multidimensional data array, which don't require an unfolding (augmentation) step, exists, the most suitable to deal with chromatographic data being PARAFAC [2]. The PARAFAC concept, as applied to chromatography, implies that every contribution to the overall signal (analytes, interferents, backrground) can be modeled by an individual component showing the same elution and spectral profiles along all the analyzed samples (trilinearity). If this is the case, then the three-dimensional data array can be uniquely decomposed into these individual components on a data-driven basis.

However, when chromatographic data are involved, the presence of shifts and shape changes along the elution profiles can destroy the trilinear structure and make the PARAFAC model inappropriate. Three-mode modeling is still possible, using PARAFAC2 [3], an advanced variant of the original PARAFAC algorithm, which handles shifts and changes in the peak shapes by introducing a different set of elution mode loadings for each sample.

Results

The utility of these methods to resolve overlapping peaks and deal with instrumental and baseline effects in chromatographic experiments will be exemplified using data from the HPLC-DAD analysis of some polyphenol fractions in olive oil. An example of the outcomes of the analysis is reported in Figure 1, where the resolution of 4 phenolic acids that were coeluting, giving rise to a single peak is reported.

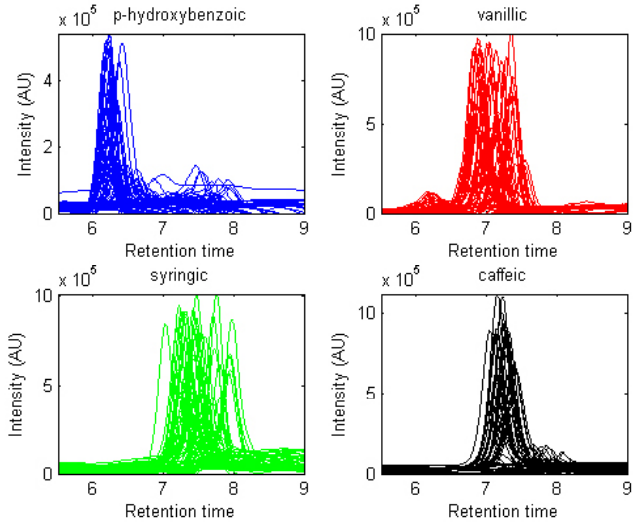

Fig.1 , Example of resolution of overlapping peaks

Conclusions

Three-way chemometric methods are useful tools to analyze data from hyphenated chromatography. Their active use of the so-called second-order advantage, i.e. the wealth of information contained in the two-dimensional experimental landscape allow to achieve accurate calibrations even in the presence of severe baseline effects or peak shifts. Moreover, the availability of different techniques allows versatility in the data analytical step.

References

1) Romà Tauler, Multivariate curve resolution applied to second order data, *Chemom. Intell. Lab. Syst.*, 30(1), (1995) 133-146.
2) Rasmus Bro, PARAFAC: Tutorial & applications, *Chemom. Intell. Lab. Syst.,* 38 (2), (1997) 149-171.
3) Rasmus Bro, Henk A. L. Kiers, Claus A. Andersson, PARAFAC2 - Part II. Modeling chromatographic data with retention time shifts, *J. Chemom.,* 13 (3-4), (1999) 295-309.