HORIBA Scientific

Technical Note

Software

SO-TN03

LabSpec 6 Multivariate Analysis Module

HORIBA Scientific's LabSpec 6 software platform offers advanced data processing for characterization of increasingly complex sample systems. It includes a comprehensive multivariate analysis module which is seamlessly powered by Eigenvector Research Inc. (EVRI), a world leader in chemometric and multivariate technology.

HORIBA Scientific

EIGENVECTOR RESEARCH INCORPORATED

HORIBA's cutting edge instrumentation delivers the flexibility, adapatability and expandability that are critical to perform diverse yet optimized experiments over a wide range of samples throughout a researcher's successful career. Consequently, data acquired (typically single spectra or spectral arrays) are multidimensional and multilayered, requiring robust mathematical/statistical processing and intelligent physical/chemical interpretation.

The LabSpec 6 *Multivariate Analysis Module* is designed to provide just such functionality, and includes a range of analysis algorithms.

Classical Least Squares (CLS) Fitting

Applicable to both individual spectra and spectral arrays, Classical Least Squares (CLS) fitting is an easy and fast analysis tool for mixtures. Based on the simple axiom that a spectrum from a mixture of chemical ingredients is a mixture of the spectra from the pure ingredients, the target spectrum or spectral array is least squares fitted with the user selected reference (pure ingredients) spectra. Such a method is ideally suited for samples where the composition is known in advance, and spectra of the pure ingredients are available.

Decomposition

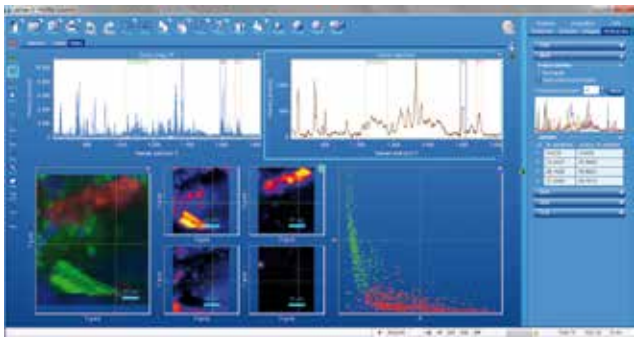
Featuring Principal Component Analysis (PCA) and Multivariate Curve Resolution (MCR), decomposition methods are good for exploring data from unknown or little known samples.

Applied to a spectrum array, decomposition methods resolve the mixture spectra into underlying patterns (called latent variables or loadings) that are statisically unique. They are likely to be representative of pure ingredients.

Each spectrum, I, of the spectrum array is expressed as a linear combination of loadings (a vector, $\vec{l}$) and their scores (a scalar, s_i). A high score of a spectrum with respect to a loading indicates that more of the spectral features represented by the principal component are present in the spectrum than would be in the case of a low score.

$$\mathrm{I} = \Sigma\, s_i \vec{l}_i + \mathrm{E}$$

LabSpec 6 offers the option to plot scores of two loadings. The scores systematically emphasizes the heterogeneity in the sample, helping to classify spectra into groups of similar chemistry. For a sample that is a complex mixture, comparing scores of two loadings may be necessary for the classification. The scatter plot of scores allows the classification to be easily visualized.



JOBIN YVON Technology

A full range of preprocessing options are available within LabSpec 6 including baseline correction, 1st and 2nd derivative, and standard normal variate (SNV). Preprocessing options that are specifically recommended for decomposition methods are included directly within the *Multivariate Analysis Module* for the ease of use.

Decomposition methods are unsupervised methods, requiring only the target data (a spectral array) to operate.

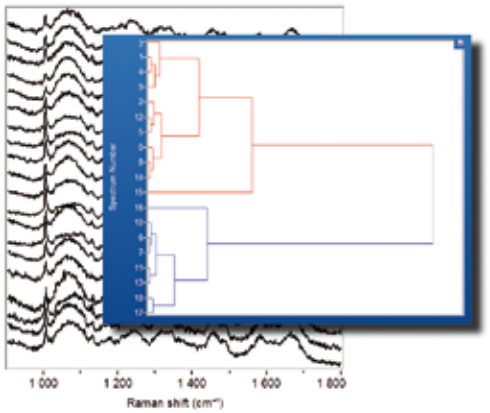
- **Principal Component Analysis (PCA)**

  Principal Component Analysis is often the first method to try on a new data to explore its behavior. Analysis is faster than MCR, and its loadings (conventionally called Principal Components) are mutually orthogonal, if not chemically pure.

- **Multivariate Curve Resolution (MCR)**

  Multivariate Curve Resolution is an automated mixture analysis method used to resolve mixtures of unknown responses, and provide more physically-interpretable results than Principal Component Analysis. Analysis often takes longer than PCA, but the resulting loadings tend to be chemically purer than those from PCA.

Clustering



There are two types of clustering analyses – agglomerative and partitional. LabSpec 6 offers hierarchical clustering analysis (agglomerative) using Ward's method, and divisive clustering analysis (partitional) using K-means.

Hierarchical clustering analysis (HCA) initially assumes that each spectrum is a class of its own. It successively compares two spectra or groups of spectra, and combines those that are similar, ultimately ending up with all spectra in one class. The process is demonstrated in a dendrogram, and a threshold is applied to produce the specified number of classes. The results are class memberships for each spectrum.

Divisive clustering analysis (DCA) initially assumes that all spectra belong to a class. Through an iterative procedure, spectra are divided into the specified number of groups, so that spectra belonging to the same group are more similar to each other than the spectra belonging to different groups. The results are class memberships for each spectrum.

Clustering methods are unsupervised methods, requiring only the target data (a spectral array) to operate.

JOBIN YVON Technology

HORIBA Scientific

info.sci@horiba.com www.horiba.com/scientific

USA: +1 732 494 8660
UK: +44 (0)20 8204 8142
China: +86 (0)21 6289 6060
France: +33 (0)1 69 74 72 00
Italy: +39 2 5760 3050
Brazil: +55 (0)11 5545 1500
Germany: +49 (0)89 4623 17-0
Japan: +81 (0)3 6206 4721
Other: +33 (0)1 69 74 72 00