DETECTION OF PATTERNS IN MULTIVARIATE FINANCIAL DATA

M. Kanevski[a], **V. Timonin**[a], **A. Pozdnoukhov**[a], and **M. Maignan**[b]

[a]Institute of Geomatics and Analysis of Risk, University of Lausanne
Amphipôle building, 1015 Lausanne, Switzerland
mikhail.kanevski@unil.ch

[b]Banque Cantonale de Genève
1201 Geneva, Switzerland
michel.maignan@bcge.ch

The research deals with an analysis and modeling of multivariate financial time series using unsupervised machine learning algorithms. The main attention is paid to problems of data clustering in high dimensional feature spaces and nonlinear dimensionality reduction. Feature spaces can be composed of values of several/many financial instruments belonging to the same portfolio, their volatilities and or pair wise correlation structures, etc. Fundamental tasks are 1) to find structures/patterns/clusters of cooperative behavior in these spaces using no a priori information and nonlinear algorithms; 2) to detect low dimensional manifolds where portfolio evolves by using nonlinear dimensionality reduction algorithms. Both solutions can contribute to the optimization of portfolios and financial data risk management. An important results of the study is a visualization of high dimensional multivariate financial data which can contribute to the contemporary exploratory data analysis.

In order to solve the first problem an unsupervised (no a priori information about the number and structure of classes) machine learning algorithms – Self-Organizing Kohonen Map (SOM) [1], is used. SOM is a well-known type of neural network using competitive (unsupervised) learning in order to produce a low-dimensional (one or two dimensional) representation of the input space of the training samples. It preserves the topological properties of the input space. SOM can be used to reveal hidden relationships in a complex set of data. Self-organizing Maps have found many applications in different fields where problem of high dimensional data visualization, clustering and classification is important [1].

One of the several case studies considered deals with the analysis of temporal cooperative behavior (clustering) of several financial indexes (SMI, Dow Jones Industrial (DJI), CAC, FTSE, Nikkei, DAX, EURO STOXX), see Figure 1. The task was to detect typical classes composed of these indexes during the period 1999-2008. Both daily prices and corresponding returns were considered as features of high dimensional spaces (in this particular case – seven dimensional space).

It was found that the indexes considered demonstrate temporal well defined cooperative behavior (see Figure 1). The results for seven classes are presented as dark dots in the upper part of Figure 1. Switching between classes corresponds to the changing in market conditions (bearish, bullish, level of prices, and corresponding transition zones). U-matrix of the trained SOM is shown in Figure 1 (right). Seven classes were extracted using k-means post-processing.

The same analysis was carried out using corresponding returns. It was found that pair wise correlations are basically stable in time (changes are not significant). At the same time three well-separated clusters of returns behavior were detected: DJI, Nikkei and all European indexes including Swiss SMI.

Another group of methods applied to the dataset belongs to the graph-based nonlinear dimensionality reduction techniques. These are the spectral clustering and Laplacian eigenmaps [2]. Obtained low-dimensional embedding and clustering are analyzed and compared to the results of SOM. The issues related to the practical use of the methods are discussed as well.

In conclusion it should be noted that application of powerful exploratory data techniques based on machine learning algorithms like SOM and nonlinear dimensionality reduction can be important tools in the analysis of complex patterns of cooperative in high dimensional feature spaces composed of multivariate financial data.

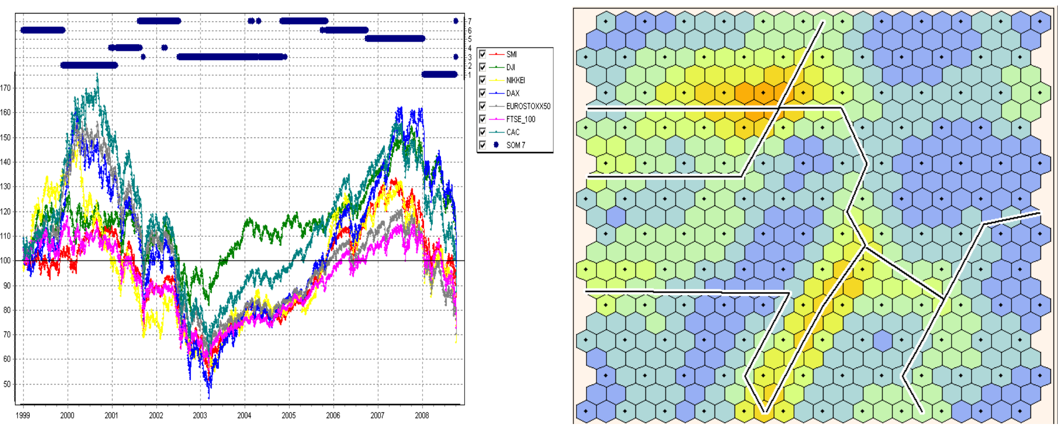

Figure 1. Temporal evolution of indexes and corresponding classes of cooperative behavior (left figure). Data were normalized to the initial point considered as 100%. U-matrix of trained SOM (right figure).

Keywords
multivariate financial data clustering and classification, nonlinear dimensionality reduction, self-organizing maps

References

[1] T. Kohonen, ``Self-Organizing Maps" Springer, Heidelberg, 2000.
[2] A. Izenman, ``Modern Multivariate Statistical Techniques. Regression, Classification, and Manifold Learning" Springer, New York, 2008.