ESANN'2004 proceedings - European Symposium on Artificial Neural Networks
Bruges (Belgium), 28-30 April 2004, d-side publi., ISBN 2-930307-04-8, pp. 63-68

Description of the Group Dynamic of Funds' Managers using Kohonen's Map

AARON Catherine[1], TADJEDDINE Yamina[2].

[1]Samos-Matisse université Paris I ; [2]Forum université Paris X

Abstract : the aim of this paper is to observe the group dynamic of funds' managers during an interesting period : January 99- July 01, which includes many financial events as financial cracks, high speculation on new technologies. We define a strategy as the sensitivity on French stock market indexes. We project strategies on a Kohonen Map. We propose a new approach to analyze the group dynamic by studying movements on this map.

1- Introduction

The aim of this paper is to construct a method that can help to analyse the dynamic of funds manager using a projection on a Kohonen map. We chose a projection on a Kohonen map because this method is really well adapted for our data base (Cf. Verleysen, Kohonen, Maillet). We used two principals approaches of analysing dynamic in Kohonen maps. The first is projecting on as many maps as there are dates and observing the changes in these maps. In our case this may not be efficient because of the great changes in structure of the maps during our time period. For instance, it is well known by financial experts that in crises periods all agents have quite the same behaviour, so maps for crisis periods should have less cells that maps for quiet periods. So deformation analysis may be very complex. The second approach is projecting of all individuals at all dates on the same map and drawing trajectories for each individual. In our case this may be difficult to apply because of the large size of the time period that gives hardly interpretable graphs. Another problem is that we cannot describe 83 graphs. So we chose the following method : an unique projection of all individuals at every date to avoid map deformation problems. Then, instead of dynamic analysis for each individual we prefer to analyse the group dynamic of funds managers. In this way we are going to observe the frequentation of the map for at every moment.

2- Database and preliminary

Our database is originally composed of daily values for 83 funds, specialised on French stock market. We choose to work with 15 days returns: $R_t(fund) = \frac{V_t(fund)}{V_{t-15}(fund)} - 1$. This choice is a good compromise between relevant regression results and not too much smoothing of series.

We introduce French stocks indexes to explain funds returns. We used the same return definition on these indexes. The indexes are : CAC40, SBF80, Second Market

(SM) , High Technologies (IT). First, we studied correlation between indexes. The correlation between indexes are strong because of their dependence with French economy. Then we build uncorrelated indexes by using linear regression

$$\text{(S1)}\quad \begin{aligned} R_t(SM) &= \alpha_1 R_t(CAC) + \varepsilon_t(SM) \\ R_t(SBF80) &= \alpha_2 R_t(CAC) + \beta_2 \varepsilon_t(SM) + \varepsilon_t(SBF80) \\ R_t(IT) &= \alpha_3 R_t(CAC) + \beta_3 \varepsilon_t(SM) + \varepsilon_t(IT) \end{aligned}$$

$R_t(CAC), R_t(NR), \varepsilon_t(SM), \varepsilon_t(SBF80), \varepsilon_t(IT)$ are not correlated.

We propose to characterize a fund at a date t as its sensitivity to these four uncorrelated indexes and to the risk-free interest rate(NR). So, we compute for all funds f and for all dates the sensitivity vectors by regressing $R_t(f)$ on $R_t(CAC), R_t(NR), \varepsilon_t(SM), \varepsilon_t(SBF80), \varepsilon_t(IT)$ on 61 dates : $\{t-30, ..., t+30\}$.

$$R_t(f) = a_t(f) R_t(CAC) + b_t(f) R_t(NR) + c_t(f) \varepsilon_t(SM) + d_t(f) \varepsilon_t(SBF80) + e_t(f) \varepsilon_t(IT) + \varepsilon$$

Then, we reconstruct sensitivity vectors by inverting the system (S1):

We select the sensitivity vectors for which the regression is above 0.6 ($R^2 > 0.6$). Afterwards, we consider only the positive coefficients and those who have a Student's statistic greater than 2. Finally, as the empirical distribution of $\sum_{ind} p(ind, f, t)$ is strongly centred around 1, we decide to work on $q(ind, f, t) = p(ind, f, t) / \sum_{ind \in \{CAC, SM, ...\}} p(ind, f, t)$. This way, the weight vectors are interpretable as percentages.

3- Projection on a Kohonen Map

We projected the 64773 weight vectors on a Kohonen Map of size 7 times 7. We have chosen this size to be coherent with the future processing. Later, we are going to analyse the frequentation in the map so we have to choose a number of cells not too large in order to have significant percentages. Date by date we only get an average of 75 points that lead us to select small maps. We chose to work on group of three dates to get enough points for 7 by 7 Kohonen maps.

The initial card is presented with the following graph. The first part of it shows the map and its clustering in principal groups. The second part shows more precisely the projection of index on the map.

4- Individuals, variables

As we want to observe the group dynamic of funds managers, we choose to characterize a date t (our new individuals) by a frequentation matrix $G(t) \in M_7(\mathbb{R})$ where $g_{i,j}(t)$ is the number of funds that are projected in the cell (i, j) of the Kohonen

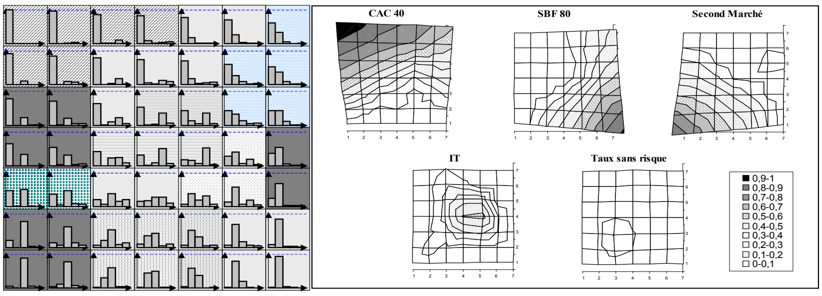

Figure 1 :Kohonen map for weights vectors.
Bars are in the following order : CAC40, SBF80,SM,IT,NR

map at the date t. Then, we construct the normalized frequentation matrix $F(t), f_{i,j}(t) = g_{i,j}(t) / \sum_{i,j} g_{i,j}(t)$. So, we have a set of 272 dates characterized by the $F(t)$ matrix which represents the percentage of funds in each cell of the Kohonen map at the date t and these are our variables. The function $F : t \rightarrow F(t)$ represents the group dynamic in the Kohonen map. Now we want to identify time intervals where F is continuous. In this purpose we are going to define a distance between $F(t)$ and $F(t')$.

5- Distance between frequencies maps

If we use classical distances (as the Euclidian distance) for $F(t)$ and $F(t')$ we do not reflect the topological proximity of Kohonen cells.

Let assume that we have the tree following frequentation matrix

$$A = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}, B = \begin{pmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}, C = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}$$

- When using the Euclidean distance, we get $d(A,B) = d(A,C) = d(B,C) = \sqrt{2}$
- When we deal with frenquencies matrices in a Kohonen map, the topological proximity of the cells $(1,1)$ and $(1,2)$ implies that $d(A,B) < d(A,C)$

Therefore we choose to interpret $F(t)$ as frequencies histograms. This approach allows us to smooth with gaussian kernel :

$$f_t((x,y)) = \sum_{i,j} F_{i,j}(t).K((x,y),(i,j),\sigma) \text{ with : } K((x,y),(i,j),\sigma) = \frac{1}{\sqrt{2\pi\sigma}} \exp\left(-\frac{(x-i)^2 + (y-j)^2}{2\sigma^2} \right).$$

We choose $\sigma = 1/2$ so that $\forall (i,j,i',j')$, $F_{i,j} = F_{i',j'}$ the map $f_t(x,y)$ is almost constant on $[1, i_{\max}] \times [1, j_{\max}]$ (cf figure 2). This choice implies that the value on a cell inflates its closest neighbours with 13%.

The following figure shows the correspondence between punctual frequencies matrix and smoothed frequencies map for two examples with $\sigma = 1/2$.

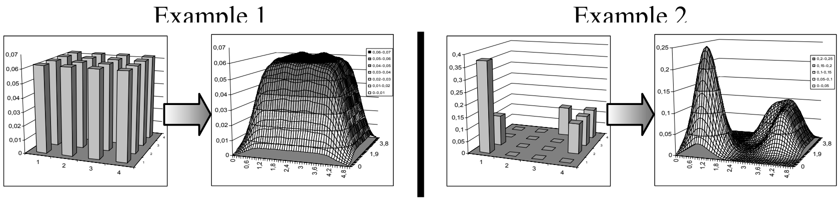

Figure 2 : smoothed map with gaussian kernel and a window size of 0,5

With this function we easily get the distance between two maps :

$$d(f_t, f_{t'}) = \iint (f_t - f_{t'})^2 = \sum_{i,j}\sum_{i',j'} (F_{i,j}(t) - F_{i,j}(t'))(F_{i',j'}(t) - F_{i',j'}(t'))\exp\left(-\frac{(i-i')^2 + (j-j')^2}{4\sigma^2} \right)$$

6- Clustering algorithm

Now we want to identify time intervals when F is continuous. There is an equivalence between $F : I \to F(I)$ continuous and $I \times F(I)$ a connected space if I is a one-dimension set (our present case).
The hierarchical clustering algorithm with minimum distance between two sets : $d(A,B) = \min(d(a,b), a \in A, b \in B)$ leads to "connected" classes. Classically, we use an intra-class distance diagram to choose the "good" number of classes. But intra-class distance is built to be relevant with the connexity concept (see Aaron [2003]).
We choose the first great gap of intra class distance : for 29 classes.

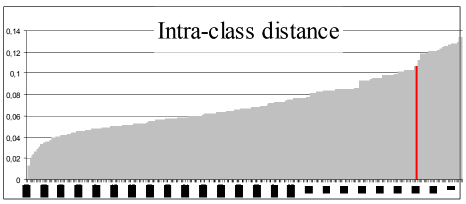

Figure 3 : Intra-class distance for choosing the number of cluster

7- Results

In figure 4 we have the different clusters. In its first part we can see CAC40 and IT evolution during the time period with the changes of classes indicated by vertical lines. We can observe that in the beginning of the period, changes of classes are not frequent. But after the new economy crash (may-august 2000) and more generally after a significant recession there are fast changes.
In the second part of the graph we have the average frequentation in Kohonen map. That gives us really resumed information for large classes but is relevant for short ones. Thus, we can notice some results such as the large dispersion during periods of economically good results (eg the second columns of graphics) and low dispersion

during crises periods. But for more detailed interpretation we have to observe the "date after date" picture. We chose to detail the third class because of its financial interest (high speculation on the new market and crash) see figure 5.

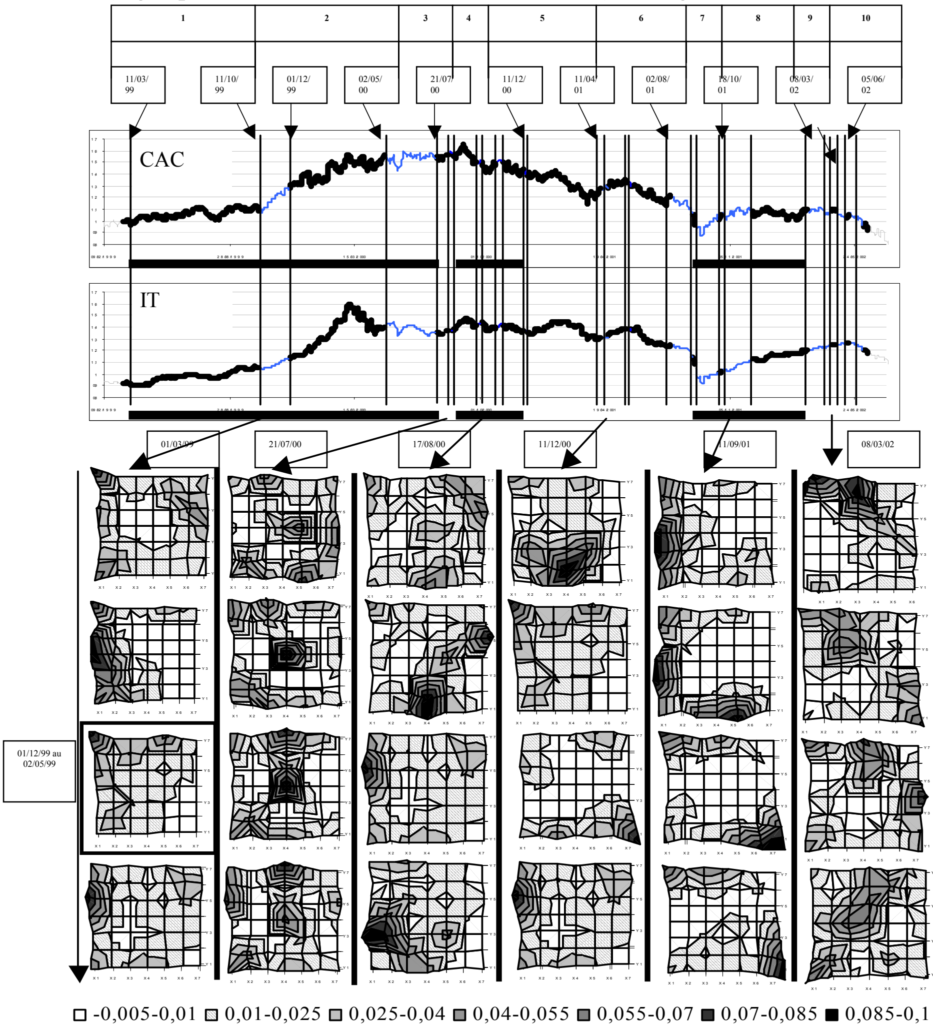

Figure 4: clusters of the time period and mean frequency maps

Figure 5 shows the date after date analysis of the third period of figure 4. We can clearly observe the increasing and the decreasing trends of the IT attraction. The first reflex after the crash were a fly on the CAC40 index but quickly there was a huge attraction of SBF80. At the end of the period, when the economy became more quiet we can observe a great dispersion on all indexes.

8- Conclusion

Our method seems to give encouraging results for analysing group dynamic with Kohonen maps. Once the individuals, the variables and the distance are defined we can enlarge this study applying other data analyses methods. In the case of our application some questions have to be solved. Is the dynamic correlated with performance of index ? can we identify leaders and followers in funds management ?

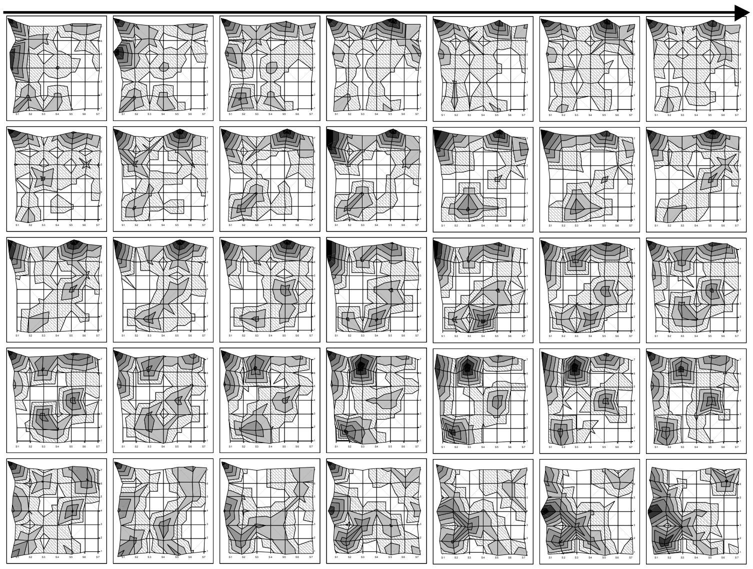

Figure 5 detail of a period

9- References

Aaron C. [2004] : « Connected Classification », *Student*, forthcoming.
Brown S, Goetzmann N. [1997] : «Mutual fund styles», *Journal of Financial Economics*, 43, p. 373-399.
Cardon P., Lendasse A., Wertz V., De Bodt E., Verleysen M. [2002] : «Classification de fonds communs d'investissement par cartes auto-organisées», *ACSEG 2002*.
Chan, Chen, Lakonishok [2002]: «On Mutual Fund Investment Styles", *Review of financial studies*, 15, p. 1407-1437.
Cottrell M, De Bodt E., Pagès G. [1997] : «Theoretical aspects of the Kohonen Algorithm », *WSOM'97*, Helsinki.
DiBartolomeo D., Witkowski E. [1997] : « Mutual fund misclassification : Evidence based on style analysis », *Financial Analysts Journal*; 53 (5), p. 32-43
Falkenstein E. G. [1996] : «Preference s for stock characteristics as revealed by mutual fund portfolio holdings», *Journal of finance*, 51 (1), mars, p. 111-135
Kim M., Shukla R., Tomas M. [2000] : «Mutual fund objective misclassification», *Journal of Economics and Business*, 52, p. 309-323
Kohonen T. [1995] : *Self-Organizating maps*, Springer Series in Information Sciences, 30, Springer.
Maillet B., Rousset P. [2003] : «Classifying Hedge Funds using Kohonen Map», Forthcoming in *Computational Economic*s, Series in Adavnces in Computational Economics and Management Sciences.
Oja E., Kaski S. [1999] : *Kohonen Maps*, Elsevier.