ESANN'2004 proceedings - European Symposium on Artificial Neural Networks
Bruges (Belgium), 28-30 April 2004, d-side publi., ISBN 2-930307-04-8, pp. 87-92

Non-linear Analysis of Shocks when Financial Markets are Subject to Changes in Regime

M.Olteanu(1), J.Rynkiewicz(1), B.Maillet(2)

(1)SAMOS - MATISSE
90 Rue de Tolbiac
75013 Paris France
madalina.olteanu,rynkiewi@univ-paris1.fr

(2)TEAM
Maison des Sciences Economiques
106-112 Bd de l'Hopital, 75013 Paris
Bertrand.Maillet@univ-paris1.fr

Abstract

Violent turbulences are often striking the financial markets and an Index of Market Shocks (IMS) was recently introduced in the attempt of quantifying these turbulences. Regime switching linear models have already been used in modelling the conditional volatility of returns. In this paper we propose a description of the IMS with hybrid models integrating multi-layer perceptrons and hidden Markov chains. After sudying the prediction performance of these models, we focus on the series separation and the index behaviour subject to the hidden states.

1 Introduction

The financial "crashes" or "crisis" we hear speaking of in the media, are translated on the markets by large price movements. Whether it is the Great Depression started in 1929, the one day 22.6% drop of US stocks in October 1987 or the World Trade Center attacks in September 2001, the markets have experienced powerful turbulences since their creation. Until recently, there was no way of quantifying these crisis or measuring how large they were. Traditional measures of risk, such as the implied or the historical volatility, have some drawbacks in being considered as suitable tools.

Following the initial approach of Zumbach and al (2000), Maillet and Michel (2003) introduce the Index of Market Shocks (IMS), an easily computable measure that quantifies the importance of market movements. The index was designed in the attempt of having a scale for measuring the extent of financial crisis, and also as a tool for comparing the intensity of strongly agitated periods. Briefly, its construction is a financial translation of the Richter scale for measuring earthquakes intensity. The first step is to compute volatilities, from the highest to the lowest frequency available. Then, a Principal Components Analysis is performed and an increasing function of the "dissipated energy" (the volatility, in this case) is applied to the resulting factors.

The aim of this paper is to seek for a good method of modelling and forecasting the Index of Market Shocks values. The basic idea is to find a hidden Markov chain - neural net hybrid model that would provide, besides good data fitting and forecasts, a state separation that describes the market behaviour in normal times and in crisis periods.

ESANN'2004 proceedings - European Symposium on Artificial Neural Networks
Bruges (Belgium), 28-30 April 2004, d-side publi., ISBN 2-930307-04-8, pp. 87-92

2 Autoregressive Hidden Markov Chain Models

Before giving the results, we make a brief description of the autoregressive models with hidden Markov chains.

Financial series historics, taken over long periods, show important breaks in the series behaviour. These breaks are due to several reasons, such as bankruptcies and panics on the markets, changes in governments policies or wars. One way to model a time series, while taking into account these “regime” changes, is to use hidden Markov chains.

Let us consider $(y_t)_{t\in\mathbb{N}}$ the observed time series and let $(x_t)_{t\in\mathbb{Z}}$ be a homogeneous Markov chain defined by its state space $E = \{e_1, ..., e_N\}$, $N \in \mathbb{N}^\star$ and the $N \times N$ transition matrix A with $a_{ij} = P\left(x_{t+1} = i \mid x_t = j\right)$, $i, j = 1, ..., N$. If we suppose, with no loss of generality, that the chain state space is the canonical basis of $\mathbb{R}^N$ and we note $v_{t+1} = x_{t+1} - E\left(x_{t+1} \mid x_t\right)$, then an autoregressive hidden Markov chain model has the following form:

$$\left\{ \begin{array}{c} x_{t+1} = Ax_t + v_{t+1} \\ y_{t+1} = F_{x_{t+1}}(y_{t-p+1}^t) + \sigma_{x_{t+1}}\varepsilon_{t+1} \end{array} \right.$$

where y_{t-p+1}^t defines the vector $(y_{t-p+1}, ..., y_t)$, $F_{x_{t+1}} \in \{F_{e_1}, ..., F_{e_N}\}$is an autoregressive function of order p, $\sigma_{x_{t+1}} \in \{\sigma_{e_1}, ..., \sigma_{e_N}\}$is a real strictly positive number, $(\varepsilon_t)_{t\in\mathbb{N}}$ are independent, identically distributed $\mathcal{N}(0, 1)$.

Hamilton’s study [3] on the GNP (Gross National Product) growth is one of the first applications of the HMC models, where the series is supposed to behave as a mixture of linear autoregressive models. But one can also consider nonlinear autoregressive functions, such as the multilayer perceptrons and thus, we get the so called hybrid MLP-HMC models (multilayer perceptron - hidden Markov chain models).

3 Research of a MLP-HMC Model

The data used in the analysis are daily values of the IMS, computed for the MSCI (Morgan Stanley Capital International) World Equities Index, from June 25th, 1998 until July 25th, 2002 (1066 observations).

We are investigating whether the series is “regime changing” and whether we can distinguish between different market behaviours by using a hybrid MLP-HMC model. The linear model and the one hidden-layer perceptron that were onsidered in a preliminary study showed that the significant lags were 1, 2, 4, 5 and 6. We fixed this input vector and we let the number of experts and hidden units vary up to three.

In the end, two configurations were selected and further investigated, on the basis of two criteria : the architecture with the smallest mean squared error and the one with has the “best” transition matrix. The “best” transition matrix is chosen such that its trace divided by its dimension is the closest to one and thus, for the corresponding configuration, we have the most powerful segmentation of the series. These "empirical" criteria were chosen because of the impossibility of using the usual likelihood ratio test (one of the regularity conditions is not fulfilled, that is the nonsingularity of the information matrix).

ESANN'2004 proceedings - European Symposium on Artificial Neural Networks
Bruges (Belgium), 28-30 April 2004, d-side publi., ISBN 2-930307-04-8, pp. 87-92

3.1 Prediction Performance

The first step of this study was to compare the estimated MLP-HMC model with the results of a linear ARMA and of a multi-layer perceptron.

After investigating all ARMA(p,q) models, for p and q in a specified range, an AR(6) minimizing the BIC criterion was selected. The research of the multi-layer perceptron is done using the REGRESS software [5], which performs an automatic research through all possible configurations with a hidden layer. The "best" model is chosen to minimize the BIC criterion and the "Statistical Stepwise" [1] algorithm is used to eliminate the non significant connexions, once the "dominating" perceptron was found. The final selected model has two hidden units, twelve parameters and the input vector goes up to the sixth lag of time.

The summary in the next table shows that, for this series, there is no interest in considering MLP-HMC in forecasting purposes as there is no improvement when compared to the linear model or to the perceptron. The global mean squared error is computed on the whole series (1060 observations), while the training and test errors are computed when splitting the series (800 values and, respectively, 260).

Model	Global MSE	Training MSE	Test MSE
AR(6)	0.35714	0.358	0.34879
MLP(1,2,4,5,6)	0.34191	0.33989	0.34989
MLP-HMC (min MSE)	0.32799	0.32796	0.37429
MLP-HMC (trans. mat.)	0.34789	0.35228	0.36626

Next, we will focus on studying the state separation of the series. Since the transition matrix does not suggest an interesting separation and since the table above implies overfitting, we will not further investigate the model which minimizes the mean squared error and we will only consider the models which maximize the trace of the transition matrix. We chose to study a two-state model and a three-state model.

3.2 A Two State Model

For the two-state model, we have the following estimated transition matrix :

$$\hat{A} = \begin{pmatrix} 0.90808 & 0.14065 \\ 0.09192 & 0.85935 \end{pmatrix}$$

The first expert has three hidden units with an associated noise variance of 0.43045, while the second is a linear model (0 unit) with an associated variance of 0.05637.

In Figure 1, we plotted the conditional probabilities of the first expert for Russia, Brazil and United States. We note that they are close to one when the market passes through a crisis period.

The first example concerns the russian crisis in August 1998 for which the main causes were the strong ruble devaluation and the country debt default, combined with the echos of the Eastern Asia crisis in 1997. On the graphic, one can see that the peaks correspond to the key moments of the crisis : August 11th, the stock market is at its lowest level, while the short-term bond market crashes, August 17th, the Central Bank

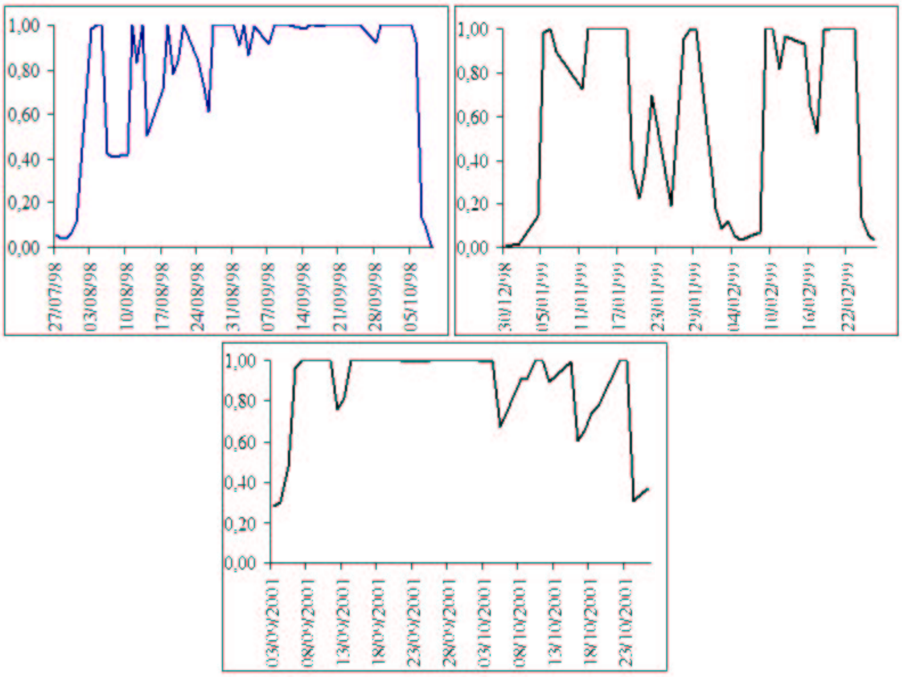

Figure 1: Conditional probabilities of the first expert (Russia, Brazil, U.S.)

widens the ruble/dollar exchange rate corridor and a 90-day moratorium on foreign debt payments is announced, August 27th, Russia's main foreign exchange market, the MICEX, shuts down every currency trade.

The next example concerns the Brazilian crisis in January 1999, mainly due to the important fiscal deficit and the high inflation rate. Here too, we see the decisive moments : January 7th, the governor of Brazil's third biggest state announces a 90-day moratorium on the debt payement to the federal government, January 13th, the finance minister resignes and the real/dollar currency band is lowered, January 15th, the Central Bank abandons its defense of the real, despite the holding of important foreign reserves.

Finally, the last example we propose here is the September 2001 crisis. Even before the terrorist attacks, the american economy showed signes of recession. On the graphic, we see the two moments : September 6th, the stock prices drop due to the bad news on the economy health, September 13th, the delayed opening of the markets after September 11th.

3.3 A Three State Model

The research of a three state model having the most powerful segmentation of the series selected an architecture with the following estimated transition matrix :

ESANN'2004 proceedings - European Symposium on Artificial Neural Networks
Bruges (Belgium), 28-30 April 2004, d-side publi., ISBN 2-930307-04-8, pp. 87-92

$$\hat{A} = \begin{pmatrix} 0.88536 & 0.00457 & 0.03387 \\ 0.00755 & 0.82081 & 0.11433 \\ 0.10701 & 0.17461 & 0.85181 \end{pmatrix}$$

The first expert is a three hidden units perceptron, while the two others have one hidden unit. The associated estimated variances are 0.42687, 0.04422 and 0.275537.

Studying the results for this model is less obvious than for the two-state one. We chose to describe the behaviour of the shock index subject to the three experts by a Kohonen map classification. We considered as input variables the IMS and the conditional probabilities of the three perceptrons.

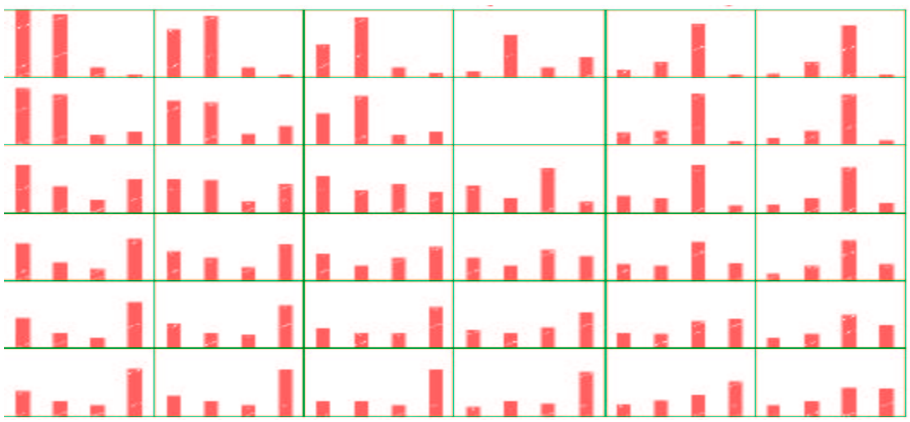

Figure 2: Kohonen map for the three state model (IMS,e1,e2,e3)

In Figure 2, the first two graphics contain the variables means (normalized values) in the clusters, which are homogeneous. We see that the high values of the index (strongly agitated periods on the markets) are associated with the first expert, while the small ones correspond to the second one, and this opposition is obvious on the map. As for the intermediary values of the index, we may expect them to go with the third experts. This is only partially true : we have, indeed, mean values of the index for which the conditional probability of the third expert is the largest, but also cases of mixture between the first and the third or between the second and the third expert.

In Figure 3 and 4 we have the results of an hierarchical classification performed on the map. If we keep three clusters, we note that the index values evoluate through clusters, from small to high values, the first and the last cluster are homogeneous and correspond, respectively, to calm periods and to crisis moments. The middle cluster is less homogeneous because it integrates the intermediate values that are associated to the third expert, but to the mixtures of experts mentioned above as well. This is even more obvious if we cut the classification tree at five clusters : we do not change the first and the last cluster, but the middle one is splitted according to different expert combination.

An immediate application of the Kohonen classification would be constructing a portfolio selection strategy, by identifying the calm periods and the turbulences on the values of the computed IMS.

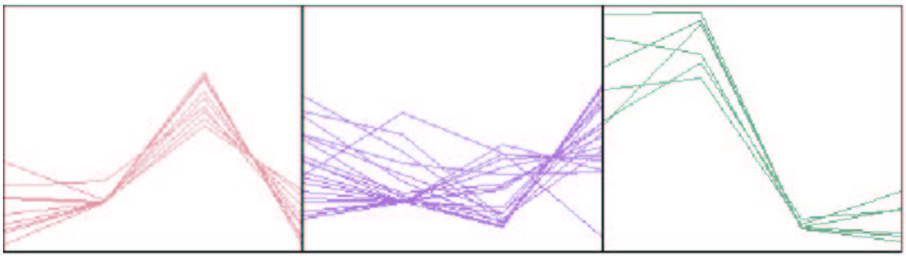

Figure 3: Clustering of the Kohonen map in three clusters

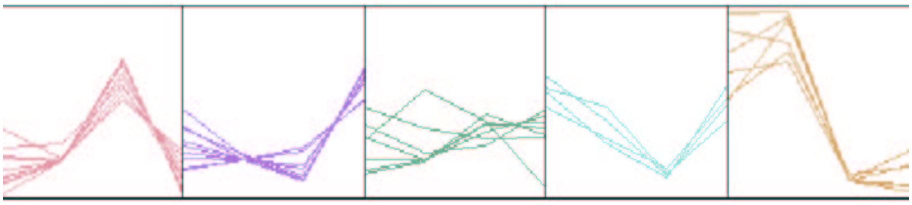

Figure 4: Clustering of the Kohonen map in five cluster

4 Conclusion

Using a hidden Markov chain model for this kind of financial series does not improve the prediction results obtained with classical linear or non-linear models. The interest in using them is related to the state separation they provide. Thus, one may, on one hand, identify key dates in market evolution and, on the other hand, derive strategies for portfolio selection using the experts conditional probabilities and the index values.

References

[1] Cottrell M. and al. (1995), "Neural Modeling for Time Series : a Statistical Stepwise Method for Weight Elimination", IEEE Transaction on Neural Networks, 6, 1355-1364

[2] Hamilton J.D. (1994) "Time series analysis", Chapter 22 "Modeling Time Series with Changes in Regime", Princeton University Press, 677-703

[3] Maillet B., Michel T. (2003) "An Index of Market Shocks Based on Multiscale Analysis", Quantitative Finance 3, 88-97

[4] Kohonen T. (2000) "Self-Organizing Maps", Springer

[5] Rynkiewicz J. (2000) "Modèles hybrides intégrant des réseaux de neurones artificiels à des modèles de chaînes de Markov cachées : applications à la prédiction des séries temporelles", Thèse de doctorat