ANALYSIS OF THE TIME EVOLUTION OF NON-LINEAR FINANCIAL NETWORKS

Abstract. We treat financial markets as complex networks. It is commonplace to create a filtered graph (usually a Minimally Spanning Tree) based on an empirical correlation matrix. In our previous studies we have extended this standard methodology by exchanging Pearson’s correlation coefficient with information – theoretic measures of mutual information and mutual information rate, which allow for the inclusion of non-linear relationships. In this study we investigate the time evolution of financial networks, by applying a running window approach. Since information-theoretical measures are slow to converge, we base our analysis on the Hirschfeld-Gebelein-Rényi Maximum Correlation Coefficient, estimated by the Randomized Dependence Coefficient (RDC). It is defined in terms of canonical correlation analysis of random non-linear copula projections. On this basis we create Minimally Spanning Trees for each window moving along the studied time series, and analyse the time evolution of various network characteristics, and their market significance. We apply this procedure to a dataset describing logarithmic stock returns from the Warsaw Stock Exchange for the years between 2006 and 2013, and comment on the findings, their applicability and significance.

Keywords: financial networks, non-linear dependence, maximum correlation coefficient, canonical-correlation analysis.

1. INTRODUCTION

Financial markets are not only complex systems, but also complex adaptive systems. Network theory is one of the methods to analyse their complexity, particularly for financial instruments traded on stock markets (Mandelbrot 1963, Mantegna 1991). These studies are useful for both phenomenological advances, and for practical risk and investment assessments. Econophysicists developed a method of network analysis based on single linkage clustering analysis, which is commonly based on the empirical Pearson's correlation coefficient matrix. Such correlation structures have been analysed for time series describing stock returns (Plerou et al. 1999, Mantegna 1999, Laloux et al. 2000), market index returns (Bonanno et al. 2000, Sandoval & Franca 2012) and currency exchange rates

* Ph.D., Cracow University of Economics, Rakowicka 27, 31-510 Kraków. E-mail: Pawel.F.Fiedor@ieee.org

[69]

Most analyses concentrate on the static structure of the network, measuring the similarities between subjects for the whole available time series. For long analyses these may not contain economically homogeneous data, and averaging over these leaves no interesting insight for market participants interested in the current state of the market. Relatively few inquiries look at networks estimated for short time series, and their time evolutions. All of these analyses are based on linear similarity measure (Albert & Barabasi 2000, Dorogovtsev et al. 2008, Fenn et al. 2011, Sienkiewicz et al. 2013, Fiedor 2014d). In this study we concentrate on the time evolution of the Warsaw Stock Exchange in the last few years, including non-linear relationships in the methodology. Using a relatively short running window makes it hard to use mutual information due to it being slow to converge (Cover & Thomas 1991, Paninski 2003). To include non-linearity we use Hirschfeld-Gebelein-Rényi Maximum Correlation Coefficient instead of Pearson’s correlation. It is a theoretical postulate, thus we use the Randomized Dependence Coefficient (RDC), which measures the dependence between random samples as the largest canonical correlation between $k$ randomly chosen non-linear projections of their copula transformations, as its estimator.

This paper is organised as follows. In Section 2 we present the methods used in the analysis. In Section 3 we present the dataset used, obtained results, and the discussion of these. In Section 4 we conclude the study and propose further research.

2. METHODS

In this section we briefly introduce the Randomized Dependence Coefficient, how we base full graphs on it, and how we then filter them into their minimally spanning trees. Measures of non-linear statistical dependence prove hard to derive and analyse. Commonly used measures of dependence such as Pearson's rho, Spearman's rank, or Kendall's tau, deal only with a limited class of association patterns. This is problematic in analysing complex systems with an array of intricate interdependencies. There are a few non-linear dependence measures, including the Alternating Conditional Expectations (Breiman & Friedman 1985), Kernel Canonical Correlation Analysis (Bach & Jordan 2002),
Copula Maximum Mean Discrepancy (Gretton et al. 2005, Poczos et al. 2012), Brownian Correlation (Szekely et al. 2007) and the Maximal Information Coefficient (Reshef et al. 2011), all of which are computationally expensive, and show poor performance under presence of noise. In this study we use the recently proposed Randomized Dependence Coefficient (RDC) (Lopez-Paz et al. 2013), an estimator of the Hirschfeld-Gebelein-Rényi Maximum Correlation Coefficient (Gebelein 1941, Rényi 1959). RDC measures dependence between two random variables as the largest canonical correlation (Hardoon & Shawe-Taylor 2009) between random non-linear projections of their respective empirical copula-transformations. Thus it is invariant to monotonically increasing transformations, and has the computational cost of $O(n \log n)$ with respect to the sample size.

Given two random samples $X \in \mathbb{R}^{p \times n}$ and $Y \in \mathbb{R}^{q \times n}$ together with the parameters $k \in \mathbb{N}$ and $s \in \mathbb{R}$, the Randomized Dependence Coefficient between $X$ and $Y$ is defined as (Lopez-Paz et al. 2013):

$$rdc(X,Y;k,s) = \sup_{\alpha, \beta} \rho(\alpha^T \Phi_{k,s}(X), \beta^T \Phi_{k,s}(Y))$$

(1)

The above expression means that RDC is calculated in three distinct phases: copula transformations of the empirical data, random non-linear projections on these copulas, and finally a canonical correlation analysis to find the maximal Pearson’s correlation coefficient between randomly chosen non-linear projections of the copula transformations of the empirical data. We present these phases briefly, step by step. RDC operates on the empirical copula transformation of the data (Nelsen 2006, Poczos et al. 2012), so that the results are invariant to any monotonically increasing transformations. Considering a random vector $X = (X_1, ..., X_d)$ with continuous marginal cumulative distribution functions $P_i$, $1 \leq i \leq d$, we may express the copula transformation of this random vector $X$ as a vector $U = (U_1, ..., U_d) = P(X) = (P_i(X_1), ..., P_i(X_d))$, which has uniform marginals. In the next step, such copula transformations are augmented with non-linear projections, so that we are able to use linear methods to capture non-linear dependencies. It is a widely used approach (Rahimi & Recht 2008). The choice of the non-linear dependencies $\phi : \mathbb{R} \to \mathbb{R}$ out of the infinite number of possibilities is the main, unavoidable, assumption. It is common to use sigmoids, parabolas, radial basis functions, complex sinusoids, sines and cosines. A further, more detailed study, of the types of non-linear behaviour found within financial data is to follow, but is outside the scope of this preliminary research. Thus, in this study we follow the authors of this method and use sine and cosine...
projections: \( \phi(w^T x + b) = (\cos(w^T x + b), \sin(w^T x + b)) \). Thus, shift-invariant kernels are approximated with these features when using the appropriate random parameter sampling distribution. The choice of \( w \sim N(0, sI) \) being Normal is equivalent to the use of the Gaussian kernel for MMD or KCCA, with \( s \) being the kernel width parameter (Rahimi & Recht 2008). We also choose \( b \sim U[-\pi, \pi] \).

Given a dataset \( X = (x_1, \ldots, x_n) \) we can write the the \( k \)-th order random non-linear projection from \( X \in \mathbb{R}^{d \times n} \) to \( \Phi_{X;k} = \Phi(X; k, s) \in \mathbb{R}^{2k \times n} \) as:

\[
\Phi(X; k, s) = \begin{pmatrix}
\phi(w_1^T x_1 + b_1) & \cdots & \phi(w_1^T x_k + b_k) \\
\vdots & \ddots & \vdots \\
\phi(w_n^T x_1 + b_1) & \cdots & \phi(w_n^T x_k + b_k)
\end{pmatrix}
\]  

(2)

The final step consists in finding the linear combinations of these augmented copula transformations which have maximal correlation. We use the Canonical Correlation Analysis (Haerdle & Simar 2007, Hardoon & Shawe-Taylor 2009) as the calculation of pairs of basis vectors \( \alpha, \beta \) such that the projections \( \alpha^T X \) and \( \beta^T Y \) of two random samples \( X \in \mathbb{R}^{p \times n} \) and \( Y \in \mathbb{R}^{q \times n} \) are maximally correlated (Lopez-Paz et al. 2013). We follow the authors of the RDC with respect to the parameter selection \( k = 10, s \) as the squared Euclidean distance).

We use standard financial data in our study, that is the logarithmic price changes. Let us denote the most recent price for stock \( c \) occurring at the end of day \( t \) during the studied period by \( p_c(t) \). Then for each stock the logarithmic returns are sampled:

\[
r_{c,t} = \log(p_c(t)) - \log(p_c(t-1))
\]  

(3)

throughout the studied period. These time series constitute columns in a matrix \( R \). From these matrices an empirical correlation matrix \( C \) is constructed using the Randomized Dependence Coefficient of columns of matrix \( R \):

\[
C_{f,c} = rdc(R_f, R_c; k, s)
\]  

(4)

From the correlation matrix \( R \) we create a matrix of Euclidean distances between the studied currencies \( D \) in the following way (Mantegna 1999):

\[
D_{f,c} = \sqrt{2(1-C_{f,c})}
\]  

(5)
Finally, we describe the algorithm used for the construction of the minimally spanning trees. Hierarchical networks based on adjacency or distance matrices may be constructed in two ways. The first is to force topological restraints (a threshold) on the similarity measure. Threshold networks are robust with regards to the statistical uncertainty in the estimation of similarity measures, but it is difficult to find a single threshold to appropriately display the nested structure of the similarity matrix. The other method is to force topological restraints, which creates intrinsically hierarchical networks (but less stable with respect to the statistical uncertainty in the data). To create a minimally spanning tree (MST) we use the distance matrix $D$ (Aste et al. 2010) connecting $N$ financial instruments. We transform the distance matrix $D$ into an ordered list $S$, which contains the distances listed in decreasing order. Then we go through the list sequentially, and add the corresponding link to the network if and only if the resulting graph is still a forest or a tree (Tumminello et al. 2005). After all appropriate links are added the graph is guaranteed to reduce to a tree. For detailed description of these methods (see presented references: Tumminello et al. 2005, Tumminello et al. 2007a, Tumminello et al. 2007b).

3. RESULTS AND DISCUSSION

We use a dataset containing end of day prices for 125 securities listed on the Warsaw Stock Exchange between August of 2006 and July of 2013. We wish to include the same set of securities in every constructed network, thus we are not able to construct either a very large set of securities or a very long time horizon, due to the securities entering, leaving and merging on the stock market. We believe that the above choice is optimal for this analysis. These prices are then transformed into logarithmic returns. We create minimally spanning trees based on running window approach with width of 100, 200, and 400 days. Since the differences are not huge we report the results obtained with the running window of width of 200 days. In the results we always report the network as belonging to the last day of the used window, so that we do not include any information not available on this day. For each of the graphs we calculate its characteristics, such as degree distribution or radius. We mostly analyse the degree distribution as it is thought to be the most important feature of a network, and particularly important to the analysis and controllability of socio-economic networks.

First, we show whether MST based on RDC preserves the important characteristics of the trees based on Pearson’s correlation. The main consideration is the much larger concentration of links between companies from the same sector (intrasector links) than in a random or full graph. These results are important as they cannot be reproduced by simulating a market. We have
many trees created for both measures, so we compare them using kernel density, as shown in Figure 1. The dashed line represents the percentage of intrasector links in all links within the MST based on correlation, while the solid line represents the same for RDC-based MST. We can see that (save for a brief period in 2007) the two methods give comparable results, which are around 15%. This is satisfactory given 125 companies divided into 26 sectors (~6% in fully connected graph).

![Figure 1. Percentage of intrasector links in all links within constructed MST over time](source: author’s calculations)

We turn to presenting the changes in topology of the MST based on RDC over time. In Figure 2 we show how the diameter of the tree changes over time, and compare it with how average log return (averaged over the window and over all studied companies) changes over the same period. We can see that the diameter of the network tends to be lower in times of crisis (seen as the times when average log returns are negative around 2009–2009 and 2011–2012). This means that the Warsaw Stock Exchange is characterised by a much more concentrated topology in bad times, which makes it more vulnerable due to the relatively larger influence of a few important stocks on the whole market at that time. In times of better economic prosperity, the networks are more spread around, and the risks are lower.

To complete this picture we also present the correlation between MST diameter and the average log returns (Figure 3), and between MST diameter and the standard deviation of the log returns (Figure 4). The Pearson’s correlation
between the diameter and the average log returns is equal to 0.42, while the correlation between the diameter and the standard deviation of the log returns is equal to –0.16. The former confirms our analysis above, the latter hints that it’s further true that the more diverse times are associated with more concentrated networks as well.

Figure 2. Changes in MST diameter and average log returns over time
Source: author’s calculations.

Figure 3. Scatterplot of average log returns vs MST diameter
Source: author’s calculations.
Figure 4. Scatterplot of MST diameter vs st. dev. of log returns
Source: author’s calculations.

Degree distribution is an important characteristic of a network, and particularly important to analysis and controllability of social and financial networks. In Figure 5 we present how the maximum degree within the MST has changed in the studied period, and also how the difference between the highest and the second highest degree changed within the MST over time. The latter is a quantile measure of how concentrated on one stock is the network in a given time. We see that the latter is mostly driven by the highest degree. There are times when these spike, which are of obvious interest to the analysts and the market practitioners. We will investigate the largest spike, which happened in 2011. The largest degree at the time belongs to a company called IGROUP.

To analyse the highest spike in Figure 5 we present the price and degree of IGROUP around the time of this spike, in Figure 6. It is interesting to see that the degree of the stock in the network has been rising before the large price fluctuations, hinting that this can constitute an early warning system for the market participants.
Figure 5. Changes in highest degree and difference between highest and second highest degree within the MST over time

Source: author’s calculations.

Figure 6. Changes in price and degree of IGROUP in the MST

Source: author’s calculations.
To finish the analysis, in Figure 7 we show the degree distributions for two windows, together with power law and lognormal distributions fitted. The top distribution belongs to a tree with the largest difference between first and second degree (25, for 15/07/2011), while the bottom one belongs to the one where this difference is minimal (0, for 25/08/2006). With these we confirm that the presented networks show the desired and expected characteristic of preferential attachment, though they are not strictly scale free networks.

![Degree distributions for two specific MST with power law and log-normal distributions fitted](image)

Figure 7. Degree distributions for two specific MST with power law and log-normal distributions fitted

Source: author’s calculations.

4. CONCLUSIONS

In this study we have analysed the time evolution of minimally spanning trees for securities traded on the Warsaw Stock Exchange between 2006 and 2013, based on a methodology allowing us to use narrow running window and further allowing us to include non-linear behaviour in the analysis. We have shown that this methodology retains the useful characteristics of the standard methodology based on Pearson’s correlation, but at the same time, due to the general nature of the RDC, it can capture more interdependencies. In particular,
we have shown that the topology of the market in Warsaw is much more concentrated in the times of crisis, further increasing the risks associated with it. We have also shown that the presented methodology may be a useful tool for early warnings of large fluctuations on the market. Further studies should look into the tuning of the RDC with respect to other non-linear maps, and also to the application of this methodology to other markets, both geographically and objectively.

REFERENCES


Analysis of the Time Evolution of Non-Linear Financial Networks


Pawel Fiedor

ANALIZA EWOLUCJI NIELINIOWYCH SIECI FINANSOWYCH


**Słowa kluczowe:** sieci finansowe, zależności nieliniowe, współczynnik największej korelacji, korelacja kanoniczna.