Physica A 390 (2011) 3835–3841

Contents lists available at SciVerse ScienceDirect

Physica A

journal homepage: www.elsevier.com/locate/physa

Network analysis of a financial market based on genuine correlation and threshold method

A. Namaki [a], A.H. Shirazi [b], R. Raei [a], G.R. Jafari [c,*]

[a] Department of Financial Management, Faculty of Management, University of Tehran, Tehran, Iran

[b] Interdisciplinary Neuroscience Research Program (INRP), Tehran University of Medical Sciences, Tehran, Iran

[c] Department of Physics, Shahid Beheshti University, G.C., Evin, Tehran 19839, Iran

ARTICLE INFO

Article history:
Received 23 January 2011
Received in revised form 21 April 2011
Available online 24 June 2011

Keywords:
Stock correlation network
Topological structure
Random Matrix Theory

ABSTRACT

A financial market is an example of an adaptive complex network consisting of many interacting units. This network reflects market's behavior. In this paper, we use Random Matrix Theory (RMT) notion for specifying the largest eigenvector of correlation matrix as the market mode of stock network. For a better risk management, we clean the correlation matrix by removing the market mode from data and then construct this matrix based on the residuals. We show that this technique has an important effect on correlation coefficient distribution by applying it for Dow Jones Industrial Average (DJIA). To study the topological structure of a network we apply the removing market mode technique and the threshold method to Tehran Stock Exchange (TSE) as an example. We show that this network follows a power-law model in certain intervals. We also show the behavior of clustering coefficients and component numbers of this network for different thresholds. These outputs are useful for both theoretical and practical purposes such as asset allocation and risk management.

1. Introduction

In real world, a large number of systems can be described by complex networks [1,2]. After innovative works on complex networks [3,4], extensive empirical research has been carried out on topology characteristics of actual networks in different domains. All of these networks have both small-world and scale-free topological properties [2–9]. Based on market efficiency theory, there is no systemic way to exploit opportunities in the stock market superior gains. Operational efficiency is based on rationality of investors, incompleteness of information and heterogeneity of interpretation and decision rules. These differences have led to consider the stock exchange as a complex adaptive system to understand its behavior [10–12].

In essence, a financial market can be represented as a network where nodes represent financial entities such as stocks and the edges connecting them represent the correlations between their returns [13–18]. Mantegna was the first [13] who constructed networks of DJIA (Dow Jones Industrial Average) and S&P 500 (Standard and Poors 500) indexes based on stock price correlations. Onnela et al. [15] constructed the asset graph NYSE stocks (New York stock exchange) and studied its different topological properties. Tumminello et al. [16,17] investigated the planar maximally filtered graphs of the 300 most capitalized stocks of NYSE and its topological properties. Boginski et al. [18] constructed a stock market graph based on daily fluctuation of 6546 financial instruments in the US stock markets and found some of their properties. Wei-Qiang Huang et al. [9] investigated the topological properties of Chinese stock market. Several other studies on network topology of other emerging and mature markets have been carried out [19–21]. The study on such topological properties can help to understand correlation patterns among stocks. Thus it can be a guide for risk management.

* Corresponding author. Tel.: +98 21 29902773; fax: +98 21 22412889.
E-mail address: gjafari@gmail.com (G.R. Jafari).

0378-4371/$ – see front matter
doi:10.1016/j.physa.2011.06.033

In the field of Econophysics, there is an interest for studying the correlation matrices. many methods were proposed in order to investigate cross-correlations between either pairs of simultaneously recorded time series [22,23] or among a large number of them [24,25]. Random matrix theory is one of these methods [26–30]. RMT was developed by many researchers [29–37] in order to explain the statistics of the energy levels of complex quantum systems [27,38]. Recently, it has also been applied to studies of economic and financial data. Allowing us to compute the effect of uncertainty in the estimation of the correlation matrix makes RMT useful in financial applications. Therefore, it can be applied very effectively in portfolio management [31–36]. It is pointed that originally RMT was introduced in order to zero-lag cross-correlations in collective modes of empirical time series. however, recently, besides zero-lag cross-correlations, the time-lag RMT (TLRMT) was applied in finance and other disciplines [39,40].

RMT was introduced to sort out genuine correlations among stocks from spurious ones [36,41,42]. In each correlation matrix, the largest eigenvalue, develops an energy gap that separates it from the other eigenvalues. This specific eigenvalue is associated with a strongly delocalized eigenvector and is related to a collective evolution of a large group of stocks that usually determine the evolution of market global index and we call this eigenvector the market mode. From this perspective, the magnitude of the largest eigenvalue reflects how collective is the evolution of an analyzed market and it is used as an indicator for increase in cross-correlation of financial markets. Podobnik et al. [39] found for the members of S&P 500 index that the largest singular value of cross-correlation matrix versus time exhibits pronounced peaks in times of crisis. Also, some other researchers [43,44] by analyzing different markets found that there are peaks in times of market shocks and recessions for the largest eigenvalues of the asset correlation matrices.

In this paper, according to trading data of TSE as an emerging market [45] and DJIA as a mature one, we use Random Matrix Approach for removing the market mode from correlation matrices. Because of cleaning the correlation matrix from invaluable information, this technique is useful for a better risk management. In the next section, we construct a corresponding stock correlation network with a correlation threshold method. For showing the importance of our proposed technique, we tested this for DJIA index and found an important change in the mean of correlation coefficient distribution with respect to the situation where we have market mode. Then, we analyze the topological properties of this network. Different from other network construction arithmetic, the correlation threshold method is suitable for studying relationships between network characteristics and correlation threshold [9]. The organization of the paper is as follows. In Section 2, we describe how to construct the network. In Section 3, we present our data set in detail and the empirical results. Finally, in Section 4, we draw out our conclusions.

2. Cross-correlation matrix

In order to quantify correlations, we calculate the price change (return) of stocks $i = 1, \ldots, N$ over a time scale Δt

$$G_i(t) = \ln S_i(t) - \ln S_i(t - \Delta t), \tag{1}$$

where $S_i(t)$ denotes the price of stocks i at time t. we define a normalized return in order to standardize the different stock volatilities.

$$g_i(t) = \frac{G_i(t) - \langle G_i \rangle}{\sigma_i}, \tag{2}$$

where $\sigma_i = \sqrt{\langle G_i^2 \rangle - \langle G_i \rangle^2}$ is the standard deviation of G_i and $\langle \rangle$ denotes a time average over the period studied. Then, We compute the equal-time cross-correlation matrix C with elements [27,38]

$$C_{ij} = \langle g_i(t) g_j(t) \rangle. \tag{3}$$

The elements C_{ij} are restricted to the interval $[-1, 1]$, where $C_{ij} = 1$ defines perfect correlation, and $C_{ij} = -1$ corresponding to perfect anti-correlation. $C_{ij} = 0$ corresponds to uncorrelated pairs of stocks. Also, C is a symmetric matrix where $C_{ij} = C_{ji}$ [27,28].

We analyze Tehran Stock Exchange database that covers all transactions of securities of Tehran Stock Exchange. We extract from this database time series of 325 stock prices of Tehran Stock market, on the starting date April 1, 2005. We analyze daily change of this database over a period of 1291 consecutive trading days in 2005–2010.

3. Removing market effect from cross-correlation matrix

An important area of risk management is the estimation of correlations between the price movements of different assets in portfolios. So, studying the correlation matrices is an important topic of risk management. Applying RMT methods on correlation matrices shows that most of the eigenvalues of correlation matrices agree with RMT predictions, suggesting a considerable degree of randomness in the measured cross-correlations [27,28]. So, there is no information in this part. Also, as stated before, Application of RMT in the investigation of financial correlation matrices leads to the immediate observation that the largest eigenvalue is more greater than the other eigenvalues (in this research the largest eigenvalue of TSE correlation matrix is about 2 times greater than the previous largest eigenvalue and 65 times greater than the smallest eigenvalue) and based on Principle Component Analysis (PCA), it represents the maximum variance of the system [27–34].

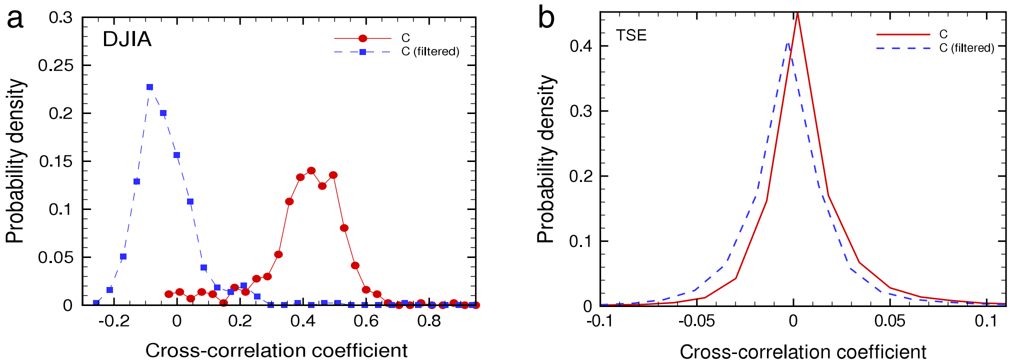

Fig. 1. $P(C_{ij})$ for (a) DJIA and (b) TSE before and after removing the market mode.

Since a majority of the components participate in the eigenvector corresponding to the largest eigenvalue, it can represent an influence that is common to the whole market. This eigenvector describes the common behavior of all stocks and quantifies the qualitative concept that certain market events affect all stocks alike [27,38]. So, this eigenvector is the market mode of the correlation matrix and shows the average behavior of the stock exchange. Hence, it does not have any new information for studying the risk management.

We filter out this market-wide effect from the correlation matrix, because its effect overwhelms the other synchronizations due to the activities of investors based on business sectors [46]. For this purpose, we adopt the regression method forming the basis of widely used economic models, such as capital asset pricing model (CAPM) and Market model [47]. The decomposition is given as

$$G_i(t) = \alpha_i + \beta_i M(t) + \epsilon_i(t), \tag{4}$$

where α_i and β_i are stock-specific constants and $\langle \epsilon_i(t) \rangle = 0$. $M(t)$ is a common factor giving rise to correlations between stocks, and $\langle M(t)\epsilon(t) \rangle = 0$ [27]. So, by considering one factor, we approximate $M(t)$ (market effect) with projection $G^{325}(t)$ defined as

$$M(t) = G^{325}(t) = \sum_{i=1}^{325} u_j^{325} G_j(t). \tag{5}$$

By definition, $G^{325}(t)$ shows the return of the portfolio defined by u^{325} (the largest eigenvector). As stated above, components of the largest eigenvector quantify market-wide influences on all of its stocks. As we know Market model has three parts: (1) β_i: the part of return of stock i that depends on market effect. (2) α_i: the part of return of stock i that shows the return's trend during time. (3) $\epsilon_i(t)$: the part of return of stock i that demonstrates the fluctuations return term after removing the return's trend. The parameters α_i and β_i can therefore be estimated by an ordinary least squares regression. Then, in order to construct a filtered correlation matrix $C^{filtered}$, we estimate $\epsilon_i(t)$ by applying linear square regression and use these residuals for constructing the correlation matrix. For showing the importance of our proposed technique, we check this for DJIA data from 1st January 2000 until 1st January 2010. As we can see in Fig. 1(a), it is obvious that the mean value of the correlation coefficient distribution of DJIA is much sensitive to removing the market mode from data. Hence, it seems that market mode is a characteristic of stock markets. In the following sections for constructing the TSE correlation network, we use this filtered correlation matrix.

4. Threshold method

The main idea of the threshold method is as follows. The stocks represent the vertices of the network. Also we specify a certain threshold value θ, $-1 \le \theta \le +1$. Hence, if the correlation coefficient C_{ij} is greater than or equal to θ, we add an undirected edge connecting the vertices i and j. So, different values of θ define the networks with the same set of vertices, but different sets of edges [9].

5. Network construction

Based on the data of TSE market, we calculate the cross-correlation coefficients C_{ij} between each pair of stocks using Eq. (3). Then, by removing the market mode based on RMT, we construct cross-correlation matrix from the residuals of Eq. (5). Fig. 1(b) shows the distribution $P(C_{ij}^{filtered})$ after removing the market effect. Its mean is significantly smaller than average value C_{ij} (before removing the market effect). This shows that a large number of high correlations contained in C are the effect of the market mode [27,28].

From Fig. 2 which demonstrates the edge density of the network with respect to θ, it can be seen that the number of edges in the stockcorrelation network decreases as the threshold value θ increases. We can see that edge density drops sharply

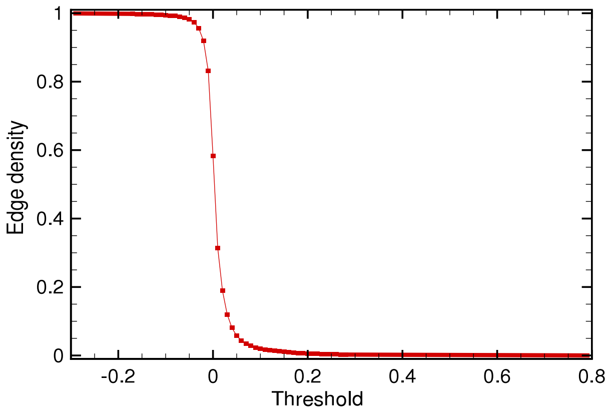

Fig. 2. Edge density of the stock correlation network for different values of the correlation thresholds.

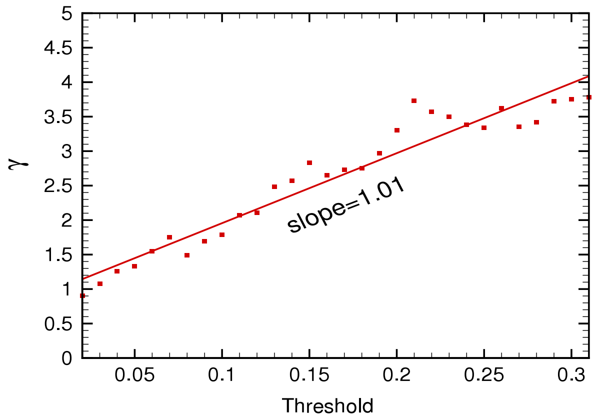

Fig. 3. The network degree distribution parameter for different thresholds. In this interval, the network is scale-free.

from 0.9 to 0.15 as θ increases from -0.02 to 0.02. This is because most of the correlation coefficients of stock returns are at this threshold interval.

6. Network topology structure

6.1. Degree distribution

The degree of vertex i is $k_i = \sum_{j \neq i} e_{ij}$ which equals the number of vertices connecting to i. The vertex degree distribution function P_k is the probability of a selected vertex being connected with k edges [9]. Many empirical studies on actual networks such as the WWW, movie actor collaboration network, etc., show that a power law [2] $P_k \propto k^{-\gamma}$ governs the vertex degree distribution where γ is the parameter. These networks that have a power-law degree distribution are called scale-free networks where most of the vertices have a small degree and only a few of the vertices have a large degree. Nodes that have a large degree are called "Hub" vertices. These results show that they have correlations with many other stocks in the sense of price fluctuations.

In this section, we show that TSE correlation network also obeys the power-law model. In this network, it should be noted that since each network corresponds to a certain value of θ, the degree distributions will be different for each θ. The results of our empirical study show that for $\theta \in [0.02, 0.31]$ the degree distribution is approximately a straight line in the log–log scale, which is exactly the power-law distribution. Fig. 3 demonstrates the degree distribution parameter γ with respect to different θ in mentioned range. Obviously, this parameter increases with θ. For other correlation thresholds, the distribution of the degrees does not have any well-defined structure. These values of correlation thresholds lead to high dense or sparse networks.

From the point of price fluctuation influence, the stocks that have large degrees, generally have a higher status to influence the market [9]. So the scale-free property of TSE correlation network demonstrates that most of the stocks are at the same level and a few stocks have higher status in the market, that play a very important role in the price fluctuation. Figs. 4 and 5 demonstrate the degree distributions of the networks for some values of the correlation threshold (0.05, 0.27).

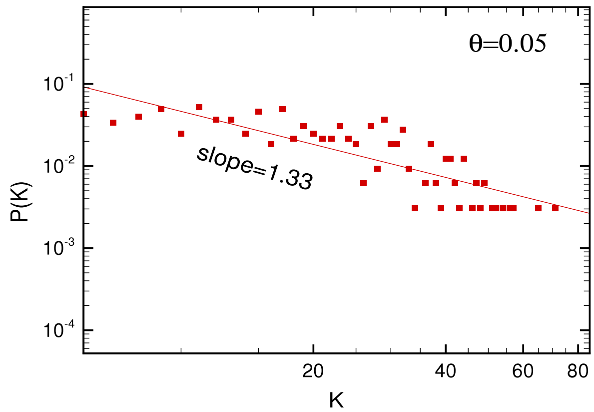

Fig. 4. Log–log linear fitting of network degree distribution for $\theta = 0.05$.

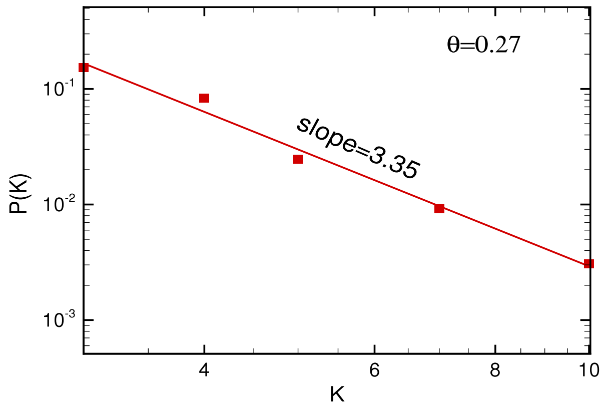

Fig. 5. Log–log linear fitting of network degree distribution for $\theta = 0.27$.

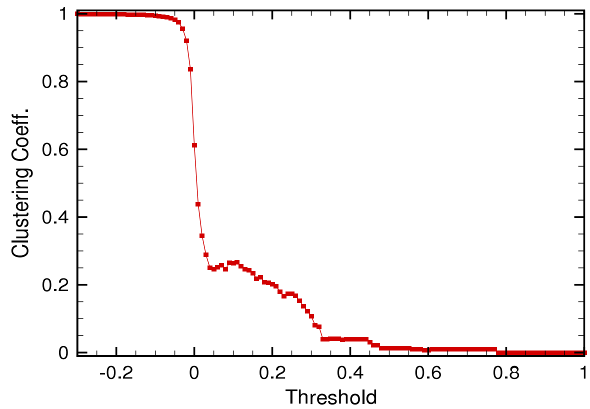

Fig. 6. The clustering coefficients of the stock correlation network.

6.2. Clustering coefficient

As we know, every vertex *i* has *k* nearest neighbors. For calculating the clustering coefficient of a vertex *i*, we divide the number of edges among these neighbors to the maximum edges possible among them. The network clustering coefficient is the average of the clustering coefficient of all vertices. The network clustering coefficient of a stock correlation network demonstrates the clustering property of stocks in the sense of price fluctuation correlation [9]. Fig. 6 shows the clustering coefficients of TSE correlation network and these coefficients become smaller with an increase of thresholds. Integrating with the above results, we understand that the Tehran stock correlation network, as many other complex systems, is scale-free and highly clustered when $\theta \in [0.02, 0.31]$.

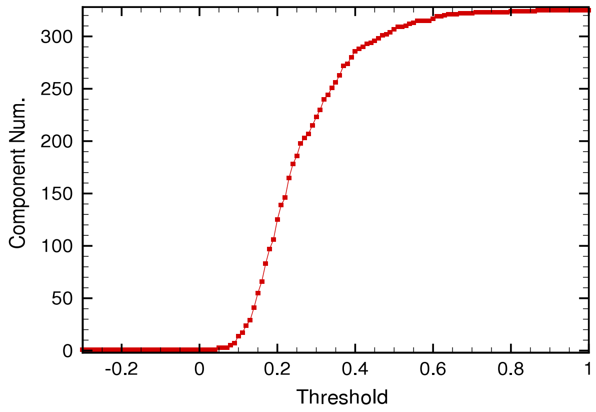

Fig. 7. The component number in the stock correlation network under different correlation thresholds.

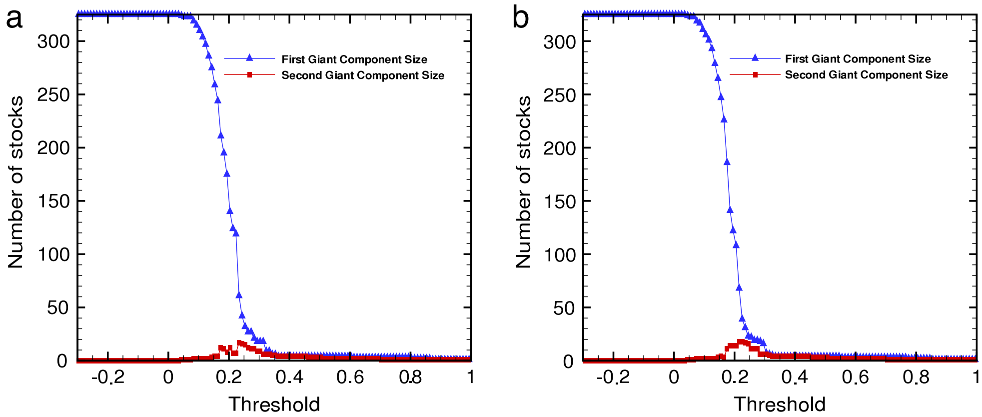

Fig. 8. The first and second giant component sizes of stock correlation network under different correlation thresholds (a) before and (b) after removing the market mode.

6.3. Component structure

In the stock correlation network, components represent stocks that are correlated with each other in the sense of price fluctuations [9]. We find that the component structure of TSE correlation network depends on threshold values. In Fig. 7, we see that the component number of the network under different thresholds increases by increasing the threshold. Fig. 8(a) is the size of the first and second giant components of the correlation network before removing the market effect under different thresholds. This reflects the connectivity and community structures of a stock correlation network. With comparing these two (Figs. 8(a) and 7), we can see that the greater a component number is, the smaller the maximum component size will be. We can see that there is just one major community which almost all stocks (some core and some peripheral) belong to it. To check the origin of these results, we remove the market dominant mode (described in Section 3) to eliminate the effect of macroeconomic and governmental supportive policies. Fig. 8(b) shows the same plot after market mode removal, as we can see that the results are almost the same. Hence, we can conclude that this effect (the presence of one major component) is the intrinsic property of this market. In the context of risk management, there is an interest for measuring the systemic risk, a concept which applies to measure any broad-based breakdown in the financial system. Systemic risk can be described as a series of correlated defaults among financial system, occurring over a short-time span and causing the withdrawal of liquidity and widespread loss of confidence in the whole financial system [48]. The number of stocks in giant components is a measure of connectivity of the stock market and also is a measure of systemic risk and so in TSE market, by increasing the threshold, the risk decreases.

7. Conclusion

In this paper, we used Random Matrix Approach notion for removing the common factor among all stocks and then constructing the stock correlation network of Tehran Stock Market based on it. We checked our proposed technique (removing the market mode from data) for DJIA and found that the mean value of cross-correlation coefficients had a great change. Because of cleaning the correlation matrix from invaluable information, This technique is appropriate for a better risk management. Next, we presented a detailed study of TSE correlation network. This approach is a good tool for analyzing the market structure by classifying the stocks into different groups. It would be helpful for investors to make decisions, about their portfolios. Therefore, this technique is useful for both theoretical and practical purposes. The statistical analysis of this network has shown that this network is scale-free in a specific interval.

For more research, we propose to construct the network based on time-lag cross-correlations of stocks or we can study stock correlation network based on time series of other parameters such as liquidity instead of time series of returns. It would be very interesting to study the properties of these networks and compare them with the network considered in this paper.

Acknowledgment

The authors would like to thank Mrs. Sara Zohoor for helping to edit the manuscript.

References

[1] R. Albert, A.L. Barabasi, Reviews of Modern Physics 74 (2002) 47.
[2] M.E.J. Newman, SIAM Review 45 (2003) 167.
[3] D.J. Watts, S.H. Strogatez, Nature 393 (1998) 440.
[4] A.L. Barabasi, R. Albert, Science 286 (1999) 440.
[5] M.E.J. Newman, D.J. Watts, Physics Letters A 263 (1999) 341.
[6] A. Barrat, M. Weigt, The European Physical Journal B 13 (2000) 547.
[7] G. Bianconi, A.L. Barabasi, Physical Review Letters 65 (2001) 5632.
[8] X. Li, G. Chen, Physica A 328 (2003) 274.
[9] W.Q. Huang, X.T. Zhuang, S. Yao, Physica A 388 (2009) 2956.
[10] D. Kahneman, A. Tverskey, Econometrica 47 (1979) 263.
[11] T. Lux, M. Marchesi, Nature (London) 397 (1999) 498.
[12] A.H. Shirazi, G. Reza Jafari, J. Davoudi, J. Peinke, M. Reza Rahimi Tabar, Muhammad Sahimi, Journal of Statistical Mechanics (2009) P07046.
[13] R.N. Mantegna, The European Physical Journal B 11 (1999) 193.
[14] H.J. Kim, Y. Lee, B. Kahng, I.M. Kim, Journal of Physical Society of Japan 71 (2002) 2133.
[15] J.P. Onnela, K. Kaski, J. Kertesz, The European Physical Journal B 38 (2004) 353.
[16] M. Tumminello, T. Aste, T.D. Matteo, R.N. Mantegna, Proceeding of National Academy of Sciences 102 (2005) 10421.
[17] M. Tumminello, T.D. Matteo, R.N. Mantegna, The European Physical Journal B 55 (2007) 209.
[18] V. Boginski, S. Butenko, P.M. Pardalos, Computational Statistics & Data Analysis 48 (2005) 431.
[19] B.M. Tabak, M.Y. Takami, D.O. Cajueiro, A. Petiting, Physica A 388 (2009) 59.
[20] R.K. Pan, S. Sinha, Physical Review E 76 (2007) 046116.
[21] W.S. Jung, S. Chae, J.S. Yang, H.T. Moon, Physica A 361 (2006) 263.
[22] B. Podobnik, H.E. Stanley, Physical Review Letters 100 (2008) 084102;
B. Podobnik, D. Horvatic, A.M. Peterson, H.E. Stanley, Proceedings of the National Academy of the USA 106 (2009) 22079.
[23] S. Arianos, A. Carbone, Journal of Statistical Mechanics (2009) P03037.
[24] I.T. Jollife, Principal Component Analysis, Springer-Verlag, New York, 1986.
[25] T. Guhr, B. Kalber, Journal of Physics A 36 (2003) 3009.
[26] L. Laloux, P. Cizeaue, J.P. Bouchaud, M. Potters, Physical Review Letters 83 (1999) 1467.
[27] V. Plerou, P. Gopikrishnan, B. Rosenow, L.A.N. Amaral, H.E Stanley, Physical Review Letters 83 (1999) 1471.
[28] V. Plerou, P. Gopikrishnan, B. Rosenow, L.A.N. Amaral, T. Guhr, H.E. Stanley, Physical Review E 65 (2002) 066126.
[29] M.L. Mehta, Random Matrices, Academic Press, Boston, 1999.
[30] A.M. Sengupta, P.P. Mitra, Physical Review E 60 (1999) 3389.
[31] L. Laloux, P. Cizeau, P. Potters, J. Bouchaud, International Journal of Theoretical and Applied Finance 3 (2000) 391.
[32] P. Gopikrishnan, B. Rosenow, V. Plerou, H.E. Stanley, Physical Review E 64 (2001) 035106R.
[33] B. Rosenow, V. Plerou, P. Gopikrishnan, H.E. Stanley, Europhysics Letters 59 (2002) 500.
[34] S. Sharifi, M. Crane, A. Shamaie, H.J. Ruskin, Physica A 335 (2004) 629.
[35] D. Wilcox, T. Gebbie, Physica A 344 (2004) 294.
[36] Z. Burda, A. Gorlich, A. Jarosz, J. Jurkiewicz, Physica A 344 (2004) 295.
[37] F.J. Dyson, M.L. Mehta, Journal of Mathematical Physics 4 (1963) 701.
[38] A. Namaki, R. Raei, G.R. Jafari, International Journal of Modern Physics C 22 (4) (2011) 371–383.
[39] B. Podobnik, D. Wang, D. Horvatic, I. Grosse, H.E. Stanley, Europhysics Letters 90 (2010) 68001.
[40] K.B.K. Mayya, R.E. Amritkar, AAPPS Bulletin 17 (2007) 96.
[41] A. Utsugi, K. Ino, M. Oshikawa, Physical Review E 70 (2004) 026110.
[42] R. Muirhead, Aspects of Multivariate Statistical Theory, Wiley, New York, 1982.
[43] A. Namaki, G.R. Jafari, R. Raei, Physica A 390 (2011) 3020.
[44] A. Sharkasi, M. Crane, H.J. Ruskin, J.A. Matos, Physica A 368 (2006) 511.
[45] P. Norouzzadeh, G.R. Jafari, Physica A 356 (2005) 609–627;
M. Vahabi, G.R. Jafari, Physica A 385 (2007) 583590;
M. Vahabi, G.R. Jafari, Physica A 388 (2009) 3859–3865.
[46] G. Lim, et al., Physica A 388 (2009) 3851.
[47] J. Campbell, A.W. Lo, A.C. MacKinlay, The Econometrics of Financial Markets, Princeton University Press, Princeton, 1997.
[48] M. Billio, M. Getmansky, A.W. Lo, L. Pelizzon, MIT Sloan School Working Paper, 4774–10, 2010.