Uncertainty of identification of some network structures in stock market network

Kalyagin V.A., Koldanov A.P., Pardalos P.M.
National Research University Higher School of Economics, Nizhny Novgorod,
Laboratory of Algorithms and Technologies for Networks Analysis

Problem statement.

Network analysis became, in our days, a useful tool for stock market investigation (Boginski at al, 2005, 2006), (Tumminello at al, 2005, 2010). The following network structures of stock market network attract a growing attention: maximum spanning tree (MST), planar maximally filtered graph (PMFG), market or threshold graph (MG), cliques and independent sets in the market graph. These network structures are investigated in the literature for different stock markets (Martia at al, 2017). However, one important question remains still less investigated. This question is: how reliable are the results of market network analysis? The present paper is a contribution to the study of this question.

Proposed approach and methodology.

We relate reliability of the results of market network analysis with the uncertainty of network structure identification algorithms. To study uncertainty we use a general framework of random variables network (Kalyagin at al, 2017). Random variables network is a pair (X, γ), where $X=(X_1, X_2, \ldots, X_p)$ is a random vector, and γ is a pairwise measure of similarity (dependence, association,...) between random variables. Random variables network generates a network model. Network model for the random variable network (X, γ) is the complete weighted graph (V, Γ) with p nodes where $V=\{1,2,\ldots,p\}$ is the set of nodes, and $\Gamma= (\gamma_{i,j})$ is the matrix of weights, $\gamma_{i,j} = \gamma(X_i, X_j)$. Network structure in the network model (V, Γ) is an unweighted graph (U,E), where $U\subset V$, and E is a set of edges between nodes in U. In this paper we consider two popular network structures, maximum spanning tree (MST), and planar maximally filtered graph (PMFG). MST is a spanning tree of a complete weighted graph with the maximal total weight. PMFG is a planar subgraph of a complete weighted graph with the maximal total weight. It was noted in (Kalyagin at al, 2014), that uncertainty of MST and PMFG identification by the Kruskal (MST) and Kruskal type (PMFG) algorithms is rather high, when we use Pearson correlation as the measure of similarity between stock's returns. In the present paper we study uncertainty of MST and PMFG identification for the Kruskal and Kruskal type algorithms for alternative measures of similarity: sign similarity, Kendall, and Spearman correlations. To measure uncertainty we use appropriate loss functions, which is in our case a linear combination of the number of Type I (False Positive) and of the number of Type II (False Negative) errors of MST and PMFG identification respectively. Expected value of the loss function is called risk function and uncertainty is defined as the number of observations needed to achieve a given level of risk.

Results

First we prove theoretically, that MST and PMFG, as network structures, are the same in a class of elliptical distributions for three measures of similarity: sign similarity, Pearson and Kendall correlations. For Spearman correlation these structures differ from the common structures for sign similarity, Pearson and Kendall correlations. Second, we study and compare numerically uncertainty of MST and PMFG identification for all these measures of similarity. We show, that for the sign similarity, uncertainty of MST and PMFG identification does not depend on distribution from the described class. For Kendall and Spearman correlations we observe a weak dependence of uncertainty on distributions, but in these cases uncertainty is much lower. Some practical applications of the obtained results will be discussed.

This work is supported by RFFI grant 18-07-00524.

References:

1. Boginski V., Butenko S., Pardalos P.M. (2005) Statistical analysis of financial networks, Computational Statistics and Data Analysis. 48 (2), 431--443.

2. Boginski V., Butenko S., Pardalos P.M. (2006) Mining market data: a network approach J. Computers and Operations Research. 33 (11) 3171--3184.

3. Kalyagin, V.A., Koldanov, A.P., Koldanov, P.A. (2017) Robust identification in random variable networks, Journal of Statistical Planning and Inference, 181, pp. 30-40.

4. Kalyagin V.A., Koldanov A.P., Koldanov P.A., Pardalos P.M., Zamaraev V.A. (2014) Measures of uncertainty in market network analysis, Physica A: Statistical Mechanics and its Applications, v. 413, No. 1, pp. 59-70.

5. Martia G., Nielsen F., Bikowski M., and Donnat F. (2017) A review of two decades of correlations, hierarchies, networks and clustering in financial markets. arXiv: 1703.00485v1, 2017.

6. Tumminello M., Aste T., Matteo T.D., Mantegna R.N. (2005) A tool for filtering information in complex systems. Proceedings of the National Academy of Sciences. 102 (30), 10421--10426 (2005).

6. Tumminello M., Lillo F., Mantegna R.N. (2010). Correlation, Hierarchies and Networks in Financial Markets, J. of Econ. Behavior Organization. Vol. 75. P. 40-58.