

Extracting Principal Components from Pseudo-Random Data by Using Random Matrix Theory

Mieko Tanaka-Yamawaki

Department of Information and Knowledge Engineering

Graduate School of Engineering

Tottori University, Tottori, 680-8552 Japan

Mieko@ike.tottori-u.ac.jp

Abstract

In a stock market, numerous stock prices move under a high level of randomness and some regularity. Some stocks exhibit strong correlation to other stocks. A strong correlation among eminent stocks should result in a visible global pattern. However, the networks of such correlation are unstable and the patterns are only temporal. In such a condition, a detailed description of the network may not be very useful, since the situation quickly changes and the past knowledge is no longer valid under the new environment. If, however, we have a methodology to extract, in a very short time, major components that characterize the motion of the market, it should give us a powerful tool to describe temporal characteristics of the market and help us to set up a time varying model to predict the future move of such market.

Recently, there have been wide interest on a possible candidate for such a methodology using the eigenvalue spectrum of the equal-time correlation matrix between pairs of price time series of different stocks, in comparison to the corresponding matrix computed by means of random time series [1-4]. Plerau, et. al. [1,2] applied this technique on the daily close prices of stocks in NYSE and S&P500. We carry on the same line of study used in Ref. [1] for the intra-day price correlations on American stocks to extract principal components.

We clarify the process in an explicit manner to set up our algorithm of RMT_PCM to be applied on intra-day price correlations. Based on this approach, we show how we track the trend change based on the results from year by year analysis. Here we extract significant principal components by picking a few distinctly large eigenvalues of cross correlation matrix of stock pairs in comparison to the known spectrum of corresponding random matrix derived in the random matrix theory (RMT). The criterion to separate signal from noise is the maximum value of the theoretical spectrum of We test the method using 1 hour data extracted from NYSE-TAQ database of tickwise stock prices, as well as daily close price and show that the result correctly reflect the actual trend of the market.

References

1. Plerou, V., Gopikrishnan, P., Rosenow, B., Amaral, L.A.N., and Stanley, H.E., “Random matrix approach to cross correlation in financial data”, *Physical Review E*, vol. 65, 066126, American Institute of Physics (2002)
2. Plerou, V., Gopikrishnan, P., Rosenow, B., Amaral, L.A.N., and Stanley, H.E., *Physical Review Letters*, vol. 83, pp. 1471-1474, American Institute of Physics (1999)
3. Laloux, L., Cizeaux, P., Bouchaud, J.-P., and Potters, M., vol.83, pp.1467-1470, American Institute of Physics (1999)
4. Bouchaud, J.-P. and Potters, M., “Theory of Financial Risks”, Cambridge University Press, (2000)
5. Mantegna, R.N. and Stanley, H.E., “An Introduction to Econophysics: Correlations and Complexity in Finance”, Cambridge University Press (2000)
6. Mehta, M.L., “Random Matrices”, 3rd edition, Academic Press (2004)
7. Sengupta, A.M, and Mitra, P.P., “Distribution of singular values for some random matrices” , *Physical Review E* vol.60, pp.3389- (1999)
8. Mieko Tanaka-Yamawaki, “Extracting Principal Components from Pseudo-Random Data by Using Random Matrix Theory”, to appear in Proceedings of KES2010, Cardiff, UK (2010)