PHYSA: 15218 Model 3Gsc pp. 1–16 (col. fig: NIL)

ARTICLE IN PRESS

Physica A xx (xxxx) xxx–xxx

Contents lists available at ScienceDirect

Physica A

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

Not all that glitters is RMT in the forecasting of risk of portfolios in the Brazilian stock market

Q1 Leonidas Sandoval Jr. *, Adriana Bruscato Bortoluzzo, Maria Kelly Venezuela

Insper Instituto de Ensino e Pesquisa, Rua Quatá, 300, São Paulo, SP, 04546-2400, Brazil

HIGHLIGHTS

- Use of Random Matrix Theory and the Single Index Model in the building of portfolios.
- We compare combinations of techniques in the cleaning of the correlation matrix.
- Cleaning the correlation matrix is not always advisable for times of high volatility.

ARTICLE INFO

Article history:
Received 6 October 2013
Received in revised form 20 March 2014
Available online xxxx

Keywords:
Portfolio building
Covariance matrix
Random Matrix Theory
Single Index Model
BM&F-Bovespa

ABSTRACT

Using stocks of the Brazilian stock exchange (BM&F-Bovespa), we build portfolios of stocks based on Markowitz's theory and test the predicted and realized risks. This is done using the correlation matrices between stocks, and also using Random Matrix Theory in order to clean such correlation matrices from noise. We also calculate correlation matrices using a regression model in order to remove the effect of common market movements and their cleaned versions using Random Matrix Theory. This is done for years of both low and high volatility of the Brazilian stock market, from 2004 to 2012. The results show that the use of regression to subtract the market effect on returns greatly increases the accuracy of the prediction of risk, and that, although the cleaning of the correlation matrix often leads to portfolios that better predict risks, in periods of high volatility of the market this procedure may fail to do so. The results may be used in the assessment of the true risks when one builds a portfolio of stocks during periods of crisis.

1. Introduction

Modern portfolio theory is largely based on Markowitz's ideas, where portfolios of various equities are built on the principle of minimizing risk given some expected returns, allowing one to obtain an efficient frontier of risk and returns of portfolios. Risk is assessed as the volatility of each stock that made up the portfolio, as well as their covariances. The covariance matrix is used to predict the risk of a portfolio, and it is usually different from the realized risk of the same portfolio, since the matrix is built using the stock returns of past data.

Three problems arise from this approach. The first one is that past data reflect the market as it was, and not as it will be. So, the theory assumes the hypothesis that future events shall mimic past events, which is usually not true, since it does not incorporate news releases, or the current mood of the market. There is not much that can be done about this, but to minimize effects of events that might change the behavior of a market, one cannot use past data that is too far in the past.

* Corresponding author. Tel.: +55 11 46362034; fax: +55 11 986994965.
E-mail addresses: Leonidassj@insper.org.br, leonidassj@insper.edu.br (L. Sandoval Jr.), AdrianaB@insper.edu.br (A.B. Bortoluzzo), MariaKV@insper.edu.br (M.K. Venezuela).

http://dx.doi.org/10.1016/j.physa.2014.05.006
0378-4371/

This leads us to the second problem, which are the deviations associated with the finite sample effect, that arises purely from the fact that the available data are finite. Since one cannot go back in time indefinitely, and even if one could, it would not be advisable given the discussion in the preceding paragraph, there is only a limited amount of data (in our case, price quotations) from which to build a covariance matrix. The problem gets even more severe if we think that an efficient portfolio should be built from many and diverse equities, while maintaining a fairly recent scope of historical data, since that leads to more finite sample effects due to a smaller ratio between the number of days in the historical data and the number of stocks in the portfolio.

A third problem is the statistical noise that emerges from the complex interactions between the many elements of a stock market: news, foreign markets, crisis, and the very prices of stocks interact in order to guide the price of a stock. Those interactions are usually too complex to be accommodated by any econometric model.

So, all these effects are incorporated into the covariance matrix that is used in the attempt to forecast the risk of a particular portfolio, and if one can remove some of those from the matrix, one is then able to make better risk predictions. Some authors made studies on the influence of noise and other factors on the covariance matrix in the building of portfolios [1–6]. Most of the approaches for solving them involve the reduction of the dimensionality of the covariance matrix by introducing some structure into it, obtained by principal component analysis, and the separation of stocks into economic sectors, among other means [7,8].

A technique that has been applied to a number of complex systems, and, particularly, to financial markets, is Random Matrix Theory [9]. Of the many results that were obtained, the building of portfolios that most closely resemble the realized risk of the future market, based on past data, is one of them [10–12], and it has been successfully applied to stocks [13,14], and to hedge funds [15].

Random Matrix Theory had its origins in 1953, in the work of the Hungarian physicist Eugene Wigner [16,17]. He was studying the energy levels of complex atomic nuclei, such as uranium, and had no means of calculating the distances between those levels. He then assumed that those distances between energy levels should be similar to the ones obtained from a random matrix which expressed the connections between the many energy levels. Surprisingly, he could then be able to make sensible predictions about how the energy levels related to one another by removing the results due to a random matrix. The theory was later developed, with many and surprising results arising. Of particular importance for our study are the results obtained by Marčenku and Pastur [18] on Random Matrix Theory applied to correlation matrices, better described in the section on methodology.

Today, Random Matrix Theory is applied to quantum physics, nanotechnology, quantum gravity, the study of the structure of crystals, and may have applications in ecology, linguistics, and many other fields where a huge amount of apparently unrelated information may be understood as being somehow connected. The theory has also been applied to finance in a series of works dealing with the correlation matrices of stock prices, as well as with risk management in portfolios [19–23] (for a recent review on the subject, see Ref. [24]).

Another technique that can be used to better estimate the real relations among the components of the matrix correlation is to use a regression model to remove the market effect on the asset returns, i.e., to estimate the relationship between returns and an asset that represents the market (like the BM&F-Bovespa index, in Brazil's case) and use only the residue of this model, thus eliminating the common variations of all stocks due to market movements. This procedure allows the estimation of the correlation matrix with greater precision, since there is just a part of the dependence which is due to the assets, which generates more reliable forecasts for the risk of a portfolio, being a large part of it due to the collective responses of the market to news or to other factors. This procedure is standard in many models in finance, most importantly in the CAPM (Capital Asset Pricing Model), and it is called Single Index Model (SIM), based on the idea that the majority of the systemic risk is captured by a single market index. The use of SIM is similar to the use of one component in the RMT filter, since the highest eigenvalue of the correlation matrix corresponds with the market.

Other models can be used to remove internal or external effects, the so called factor models, that defend the hypothesis that the systemic risk is due to a number of factors, which may include statistical, macroeconomic, or fundamentalist influences, that also can be used to remove noise. Ross [25], as an example, presents a model, called Arbitrage Pricing Model (APT), which uses more than one factor to explain systemic risk. According to Campbell, Lo and McKinley [26], the APT model provides an approximate relation for expected asset returns with an unknown number of unidentified factors. In the same way as Rosenow [27] selected the number of factors to be used in a MV-GARCH model based on the number of eigenvectors of the correlation matrix outside the Wishart (noise) region, RMT may be used in order to decide how many factors should be used in a multifactor model of the APT type.

Previous works on the stock exchanges of emerging markets using Random Matrix Theory have been conducted for South Africa [28,29], India [30], Sri Lanka [31], and Mexico [32]. Their results show some differences between the stock exchanges of emerging markets and the stock exchanges of more developed ones, such as less liquidity for the stocks, and less integration of different sectors.

Recent results on the application of Random Matrix Theory to financial data are basically concerned with the actual calculation of the optimal portfolios, which involve the inversion of the correlation matrix of log-returns of the time series of the stocks that may take part in the portfolio [33,34] and on a better formalization of the theory [35], with recent results [36] that claim to outperform the usual cleaning procedures used in this article.

Following a similar methodology as ours, the authors in Ref. [37] studied the stock market from Chile using Random Matrix Theory and performing an analysis of the eigenvalues and of the eigenvectors of the correlation matrix, and also