Physica A 433 (2015) 92–99

Contents lists available at ScienceDirect

Physica A

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

Random matrix theory and portfolio optimization in Moroccan stock exchange

Marwane El Alaoui *

Department of Management — Faculty of Legal, Economic and Social Sciences — Agdal, Mohammed V University of Rabat, Avenue des Nations-Unies, B.P. 721 Agdal — Rabat, Morocco

HIGHLIGHTS

- We studied the cross-correlation among stocks of Casablanca Stock Exchange portfolio.
- We used Marčenko–Pastur distribution to analyze eigenvalues.
- We analyzed distribution of eigenvectors components.
- We used the inverse participation ratio to measure the deviation degree of eigenvectors.
- We found that more than 11% of eigenvalues might contain the pertinent information.

ARTICLE INFO

Article history:
Received 11 November 2014
Received in revised form 3 February 2015
Available online 3 April 2015

Keywords:
Random matrix theory
Correlation matrix
Inverse participation ratio

ABSTRACT

In this work, we use random matrix theory to analyze eigenvalues and see if there is a presence of pertinent information by using Marčenko–Pastur distribution. Thus, we study cross-correlation among stocks of Casablanca Stock Exchange. Moreover, we clean correlation matrix from noisy elements to see if the gap between predicted risk and realized risk would be reduced. We also analyze eigenvectors components distributions and their degree of deviations by computing the inverse participation ratio. This analysis is a way to understand the correlation structure among stocks of Casablanca Stock Exchange portfolio.

1. Introduction

Due to financial globalization, markets become more and more connected and dynamic. So, investors should use methods that allow them to maximize their expected returns in these markets. There exist numerous methods for this end, the most known is Markowitz's model [1]. It estimates risk and expected returns based on the standard deviation and the expected value of returns. Other methods have aroused to manage and optimize portfolios. Correlation seems to be an important element to study portfolio management; we should have a way to understand interactions among matrices of returns. Numerous methods were proposed to study cross-correlation among series [2–6]. Besides, cross-correlation was also studied among several financial series [6–13].

In this paper, we are interested on another interesting method called Random Matrix Theory (RMT) to study cross-correlations among stocks of one portfolio. This method was used in nuclear physics by Wigner [14]. It was also used by Dyson and Mehta [15] to explain the energy levels of complex nuclei [16].

RMT has been used to analyze correlation in the finance area and specially to improve portfolio management. By using RMT, Pafka and Kondor [17] found that the effect of noise in correlation matrices determined from financial series can indeed

* Tel.: +212 610269922.
E-mail address: marwane.elalaoui@gmail.com.

http://dx.doi.org/10.1016/j.physa.2015.03.081
0378-4371/

be large that the filtering based on random matrix theory is particularly powerful in this respect. Laloux et al. [18,19] found that the empirical correlation matrix leads to a dramatic underestimation of the real risk, by overinvesting in artificially low-risk eigenvectors. They found that less than 6% of the eigenvectors, which are responsible of 26% of the total volatility, appear to carry some information. Pafka and Kondor [20] found that "realized" risk is a good proxy for "true" risk in all cases of practical importance and that "predicted" risk is always below, whereas "realized" risk is above the "true" risk. Using a simulation-based approach they show that for parameter values typically encountered in practice the effect of noise on the risk of the optimal portfolio may not necessarily be as large as one might expect. Wang et al. [21] investigated the statistical properties of cross-correlations in the US stock market. They found that the DCCA coefficient method has similar results and properties with Pearson's Correlation Coefficient, such as the properties of the largest eigenvalue and the corresponding eigenvector. Daly et al. [22] found that RMT-based filtering can, in the most cases, improve the realized risk of minimum risk portfolios. Plerou et al. [16] analyzed cross-correlations between price fluctuations of different stocks using methods of random matrix theory (RMT). They concluded that the deviating eigenvectors are useful for the construction of optimal portfolios that have a stable ratio of risk to return.

In this paper, we use Random Matrix Theory to study cross-correlation among stocks of Casablanca Stock Exchange. We clean the correlation matrix to observe if the difference between predicted risk and realized risk will be reduced. We also analyze eigenvectors through their distributions and by computing the inverse participation ratio. This paper is organized as follows, in Section 2 we present a brief description of data. Then, we expose in Section 3 the theoretical background of RMT. In Section 4, we show the main empirical results and finally we conclude.

2. Data

The data used include 62 securities listed in the Casablanca Stock Exchange.[1] We chose the period from the 1st January 2008 to the 3rd January 2014, we have then 1492 daily closing prices and 1491 logarithmic returns.

We designate by p_t the closing price of the index on day t. In the present paper, the method applied to the natural logarithmic returns of the index is defined by:

$$r_t = \ln\left(\frac{p_{t+1}}{p_t}\right). \tag{1}$$

Then, we compute mean return and standard deviation of each of these securities before establishing the portfolio selection process.

3. Theoretical background

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

$$G_i(t) \equiv \ln S_i(t + \Delta t) - \ln S_i(t),$$

where $S_i(t)$ denotes the price of stock i. Since different stocks have varying levels of volatility (standard deviation), we define a normalized return

$$g_i(t) \equiv \frac{G_i(t) - \langle G_i \rangle}{\sigma_i},$$

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

$$C_{ij} \equiv \langle g_i(t) g_j(t) \rangle.$$

By construction, the elements C_{ij} are restricted to the domain $-1 \le C_{ij} \le 1$, where $C_{ij} = 1$ corresponds to perfect relations, $C_{ij} = -1$ corresponds to perfect anti-correlations, and $C_{ij} = 0$ corresponds to uncorrelated pairs of stocks.

The difficulties in analyzing the significance and meaning of the empirical cross-correlation coefficients C_{ij} are due to several reasons, which include the following:

(i) Market conditions change with time and the cross-correlations that exist between any pair of stocks may not be stationary.
(ii) The finite length of time series available to estimate cross-correlations introduces "measurement noise".

If we have N returns with the same length equal to L, then, the empirical cross-correlation matrix $\mathbf{C}$ could be computed by C_{ij}. In our case, we have $N = 62$ and $L = 1491$. By diagonalizing matrix $\mathbf{C}$, we obtain

$$\mathbf{C}\mathbf{u}_k = \lambda_k \mathbf{u}_k.$$

In matrix notation, the correlation matrix can be expressed as

$$C = \frac{1}{L} G G^T$$

[1] http://www.casablanca-bourse.com/.

where G is an $N \times L$ matrix with elements $\{g_{im} \equiv g_i(m\Delta t); i = 1, \ldots, N; m = 0, \ldots, L-1\}$, and G^T denotes transpose of G. Therefore, we consider a random correlation matrix

$$R = \frac{1}{L} A A^T,$$

where A is an $N \times L$ matrix containing N time series of L random elements a_{im} with zero mean and unit variance that are mutually uncorrelated.

Statistical properties of random matrices such as R are known [23,24]. Particularly, in the limit $N \to \infty$, $L \to \infty$ such that $Q \equiv L/N (> 1)$ is fixed, the probability density function $P_{rm}(\lambda)$ of eigenvalues λ of the random correlation matrix R is given by

$$P_{rm}(\lambda) = \frac{Q}{2\pi\sigma^2} \frac{\sqrt{(\lambda_+ - \lambda)(\lambda - \lambda_-)}}{\lambda}.$$

For λ within the bounds $\lambda_- \le \lambda_i \le \lambda_+$, where λ_- and λ_+ are the minimum and maximum eigenvalues of R, respectively, given by

$$\lambda_\pm = \sigma^2 \left(1 + \frac{1}{Q} \pm 2\sqrt{\frac{1}{Q}}\right)$$

where σ^2 is equal to the mean of eigenvalues of the correlation matrix [25].

– **Distribution of eigenvector components**

The distribution of the components $\{u_k(l) | l = 1, 2, \ldots, N\}$ of an eigenvector $\mathbf{u}_k$ of a random correlation matrix $\mathbf{R}$ should obey the standard normal distribution with zero mean and unit variance [16],

$$P_R(u) = \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{u^2}{2}\right).$$

– **Inverse participation ratio**

In order to quantify the number of components that participates significantly in each eigenvector, we use the inverse participation ratio [26,27,16]. It also reflects the deviation degree of the distribution of eigenvectors from RMT results. It distinguishes between one eigenvector with approximately equal components and another with a small number of large components.

The IPR of the eigenvector u^k is defined as

$$I^k \equiv \sum_{l=1}^{N} [u_l^k]^4,$$

where u_l^k, $l = 1, \ldots, 1000$ are the components of eigenvector $\mathbf{u}^k$. The physical meaning of I^k can be illustrated by two limiting cases:

(i) a vector with identical components $u_l^k \equiv 1/\sqrt{N}$ has $I^k = 1/N$, whereas
(ii) a vector with one component $u_l^k = 1$ and the remainder zero has $I^k = 1$.

Thus, the IPR quantifies the reciprocal of the number of eigenvector components that contribute significantly [26,28–33].

If deviations at edges of the eigenvalue spectrum are considerably larger than the mean of IPR $\langle I \rangle$, it suggests that the vectors are localized (localization theory) [28–33].

In the context of localization theory, one frequently finds "random band matrices" [28–32].

4. Empirical results

In this section, we start by analyzing the empirical distribution and the theoretical distribution of Marčenko–Pastur of eigenvalues (see Fig. 1). We observe that there are deviations from the interval of eigenvalues $[\lambda_-, \lambda_+]$ predicted by Random Matrix Theory. Then, these deviating values might contain pertinent informations in the market and they are not noisy elements.

It is found that theoretical eigenvalues bound (maximum and minimum) are $\lambda_{\max} = 1.4497$ and $\lambda_{\min} = 0.6337$. We have 62 (N) equities and 1491 (L) daily returns for each equity. Then, the value of Q is equal to $\frac{L}{N} = 24.0484$.

By analyzing results, we observed that seven eigenvalues deviate from the RMT interval of predictions. These deviations represent 11.29% of the total of eigenvalues. Laloux et al. [18] found that there is less than 6% of eigenvalues that might contain the pertinent information. In our case, the percentage is very important and then, only 88.71% of eigenvalues deals with random matrix theory distribution. Moreover, the maximum of empirical value of eigenvalues ($\lambda_1 = 4.2456$) exceeds what is predicted by random matrix theory $\lambda_{\max} = 1.4497$.

Consequently, it is important to analyze these pertinent informations to exploit them because they could be very useful for portfolio management.

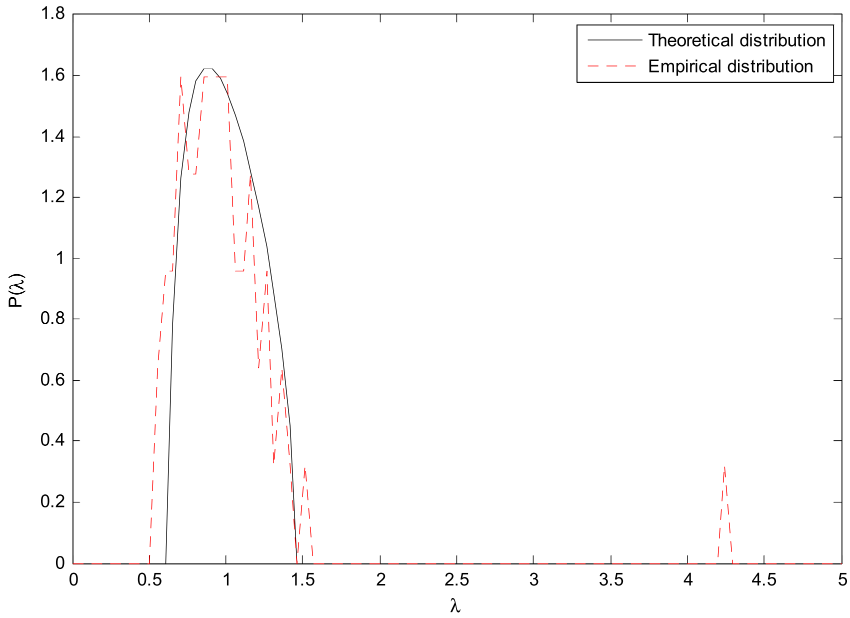

Fig. 1. Theoretical (Marčenko–Pastur) and empirical distributions of eigenvalues.

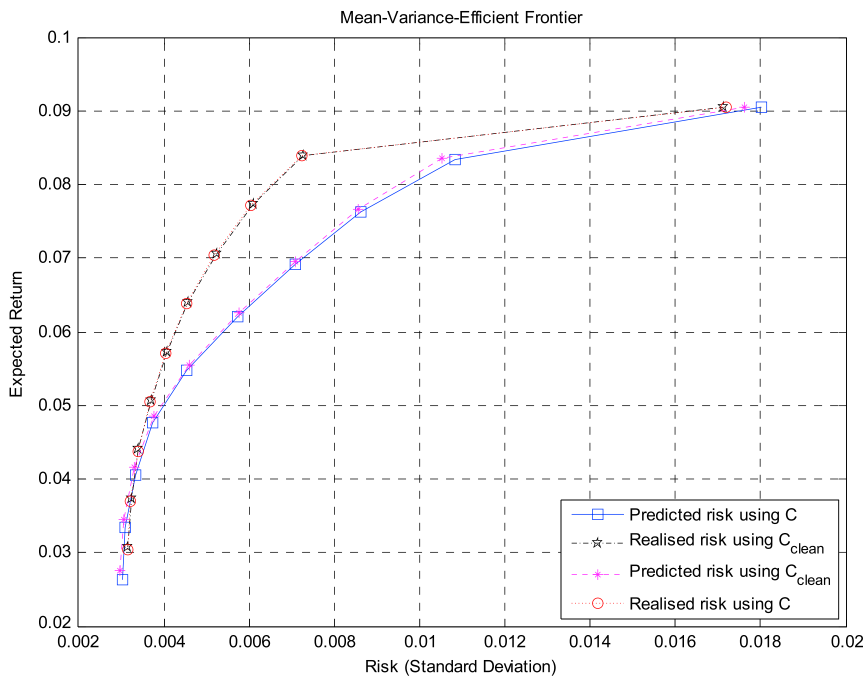

Fig. 2. Predicted risk and realized risk using the correlation matrix and the cleaned correlation matrix.

4.1. Cleaning correlation matrix

In order to see the effect of noisy elements belonging to random matrix theory, we clean matrix correlation from these elements. For this reason, we divide the series into two equal sub-periods for the 62 equities. The first sub-period presents the "predicted risk" and the second one presents the "realized risk".

First, we separate eigenvalues that are noisy and those that are non-noisy elements. These last are situated out of the bound of RMT predictions. Then, we hold deviating elements from $[\lambda_-, \lambda_+]$ and we replace the others by their average to maintain the same matrix trace. We will have identity matrix of these elements. At this moment, we could construct the new correlation matrix cleaned from noisy elements [18].

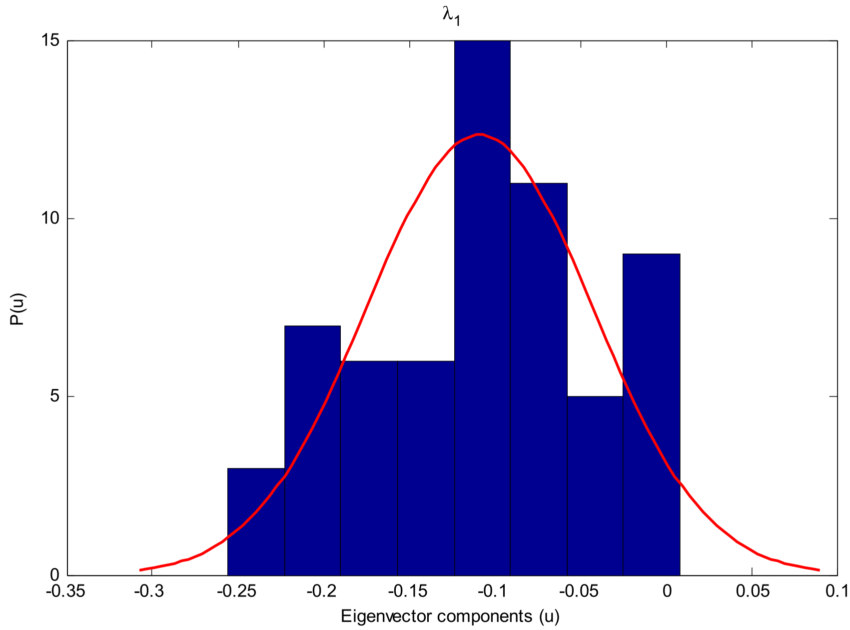

Fig. 3. Distribution of eigenvector components for the largest eigenvalue (λ_1).

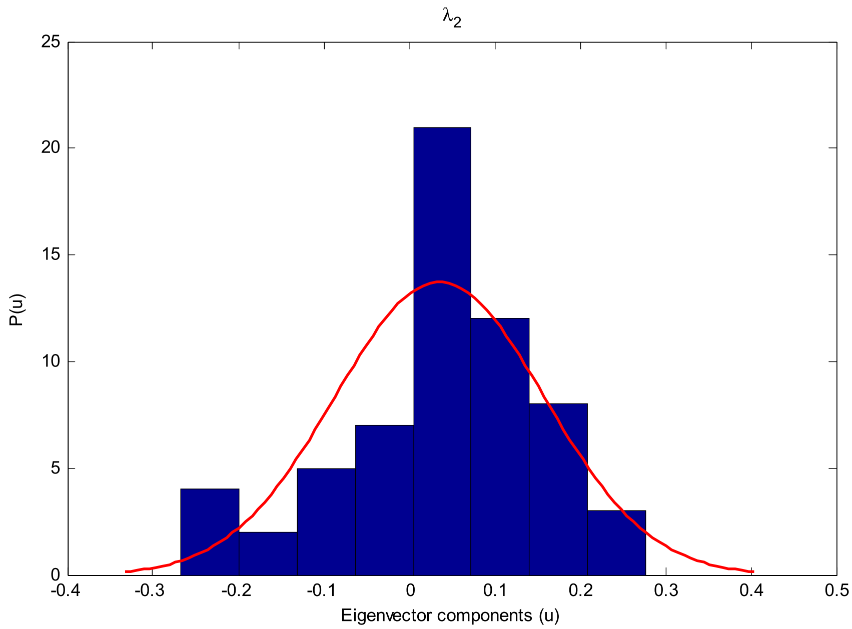

Fig. 4. Distribution of eigenvector components of an eigenvalue (λ_2) deviating from the interval predictions of RMT.

In Fig. 2, we find that the gap, existing between the predicted risk and the realized risk, is reduced slightly when the correlation matrix is cleaned.

Therefore, the cleaning procedure of matrix correlation from noisy elements improves marginally the quality of prediction by reducing a little the difference between predicted efficient frontier and realized one.

4.2. Distribution of eigenvector components

To analyze the distribution of eigenvectors, we compare distribution of eigenvector components that are included in the bounds of RMT predictions $[\lambda_-, \lambda_+]$ with those that are outside.

In Fig. 3, the normal distribution of the largest eigenvector components (λ_1) presents a negative mean that is different to zero. Then, there is an asymmetric distribution in the left.

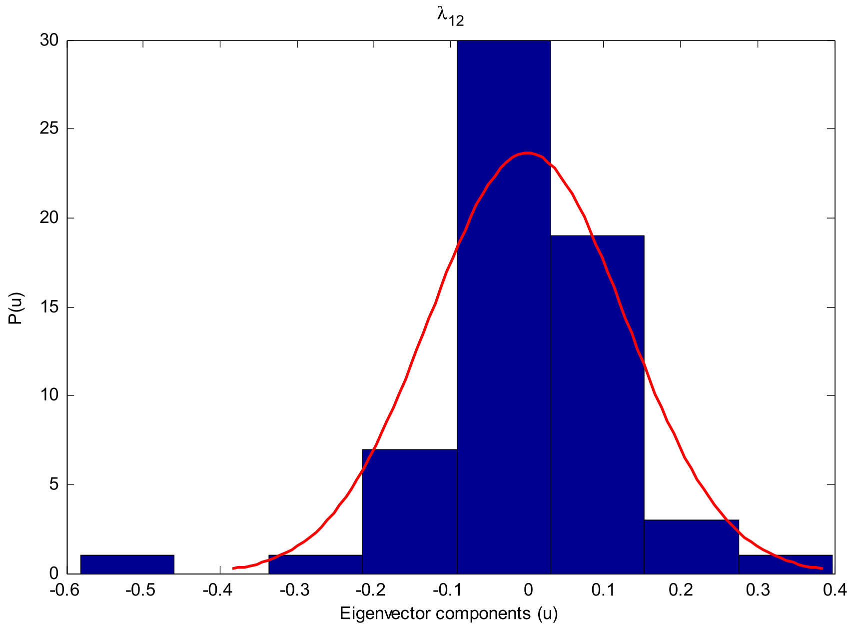

Fig. 5. Distribution of eigenvector components of an eigenvalue (λ_{12}) deviating from the interval predictions of RMT.

We show in Fig. 4 eigenvector components distribution of an eigenvalue (λ_2) deviating from the RMT interval of predictions. We see that the distribution is slightly asymmetric in the right and data are not well fitted.

For the case of eigenvector components distribution of eigenvalue λ_{12} (Fig. 5) that deviate from predictions interval of RMT, we see that the distribution is clustered in the center. Some observations are not well fitted by the normal law distribution as it is observed in the left. Moreover, the values of $P(u)$ are seemed to be high.

However, the distribution of eigenvector components of an eigenvalue (λ_{50}) that is included in the interval predictions of RMT seems to be well adjusted to the normal distribution (see Fig. 6). Thus, the distribution is centered in zero and it has a relatively constant standard deviation.

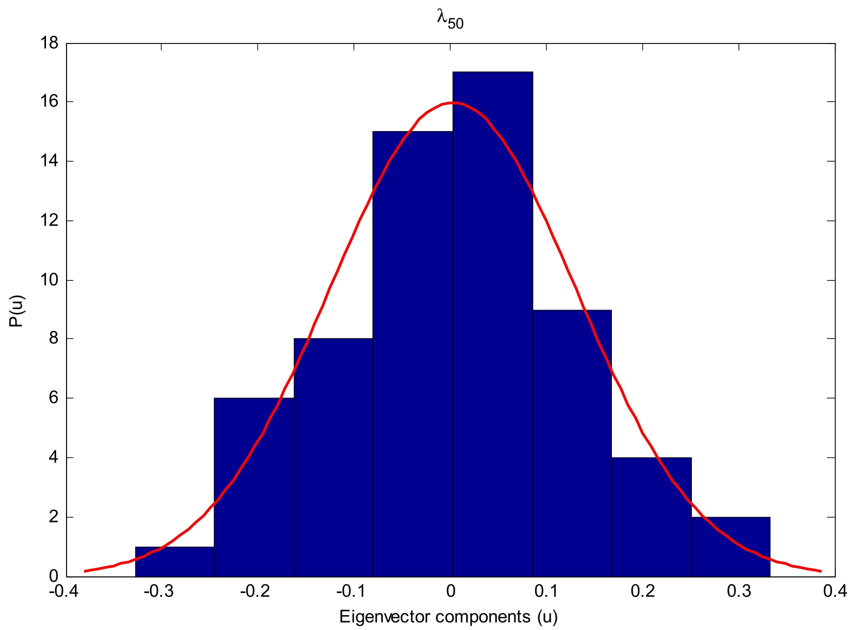

Fig. 6. Distribution of eigenvector components of an eigenvalue (λ_{50}) that is included in the interval predictions of RMT.

Comparatively to the other results corresponding to deviated eigenvalues from RMT interval of predictions, the values of $P(u)$ are not so high and the normal distribution fitting is best for eigenvector components of λ_{50} than λ_1, λ_2 and λ_{12}.

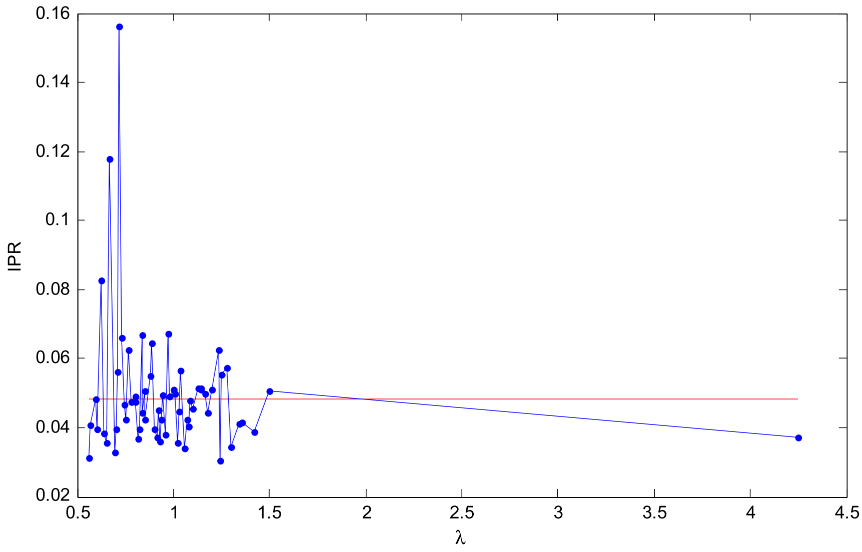

Fig. 7. Inverse participation ratio and eigenvalues.

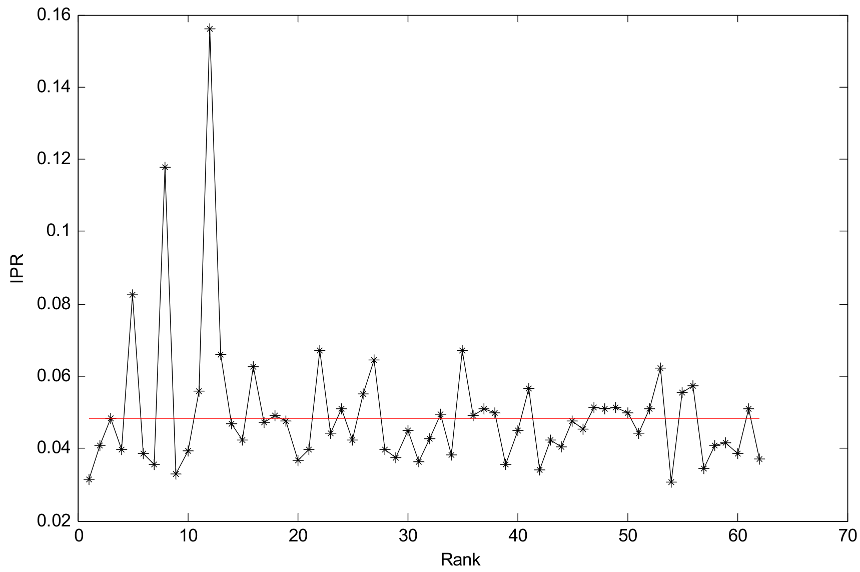

Fig. 8. Inverse participation ratio and their rank.

Besides, we remark that asymmetry is more sharp for λ_1 than that of λ_{50}. By considering the fact that λ_1 may represent the market, we could say that the market reacts more to the negative variations than to the positive ones.

4.3. *Inverse participation ratio*

We use the inverse participation ratio to examine elements that participate significantly in each eigenvector. This ratio measures their degree of deviations.

By analyzing Fig. 7, we observe a significant deviation of inverse participation ratio (IPR) for the first eigenvalues. Thus, the deviation of the largest value relatively to the smallest one is more than 5 times. This means that eigenvectors are localized. The red line shows the level of noise that is equal to the average of IPR. In our case, we put the average of IPR as $\langle I \rangle \approx 3/N$.

When IPR I^k values are estimated to be so high relatively to their mean $\langle I \rangle$, it implies that only some companies contribute in the eigenvectors.

We see in Fig. 8 that IPR values still close to the level of noise except the important deviations observed for the first elements.

5. Conclusion

To conclude, we could say that RMT allows us to analyze in detail the correlation structure of a portfolio of equities. In this effect, Marčenko–Pastur distribution presented the theoretical interval of RMT predictions to observe which eigenvalues are deviating by plotting their empirical distribution. These deviating eigenvalues might contain important information about market and they represents about 11% of the studied eigenvalues of Casablanca Stock Exchange stocks.

By observing the largest eigenvector components distribution, we see that there is a sharp asymmetry in the left, which means that the market reacts more to bad events than good events. Portfolio managers should consider this when they construct their portfolios.

In addition, the cleaning procedure of correlation matrix reduced slightly the gap between predicted and realized risks. This procedure could be helpful for practitioners by reducing their errors of predictions.

Moreover, the analysis of eigenvectors components distributions of eigenvalues showed that normal distribution fitting is not very suitable for elements that are outside of the range of RMT predictions, which confirms that they are not noisy elements.

Finally, the inverse participation ratio gives more precision about deviation degree of eigenvalues elements in order to understand better the correlation structure of the portfolio.

Acknowledgments

The authors gratefully acknowledge helpful comments and suggestions from the editor and referees.

References

[1] H. Markowitz, Portfolio selection, J. Finance 7 (1952) 77–91.
[2] B. Podobnik, H. Stanley, Detrended cross-correlation analysis: A new method for analyzing two nonstationary time series, Phys. Rev. Lett. 100 (2008) 084102.
[3] W.-X. Zhou, Multifractal detrended cross-correlation analysis for two nonstationary signals, Phys. Rev. E 77 (6) (2008) 066211.
[4] G.F. Zebende, DCCA cross-correlation coefficient: Quantifying level of cross-correlation, Physica A 390 (2011) 614.
[5] L.-Y. He, S.-P. Chen, A new approach to quantify power-law cross-correlation and its application to commodity markets, Physica A 390 (2011) 3806–3814.
[6] G.-J. Wang, C. Xie, L.-Y. He, S. Chen, Detrended minimum-variance hedge ratio: A new method for hedge ratio at different time scales, Physica A 405 (2014) 70–79.
[7] B. Podobnik, D. Horvatic, A.M. Petersen, H.E. Stanley, Cross-correlations between volume change and price change, Proc. Natl. Acad. Sci. USA 106 (2009) 22079–22084.
[8] D.-H. Wang, Y.-Y. Suo, X.-W. Yu, M. Lei, Price-volume cross-correlation analysis of CSI300 index futures, Physica A 392 (2013) 1172–1179.
[9] G. Cao, X. Longbing, J. Cao, Multifractal detrended cross-correlations between the Chinese exchange and stock market, Physica A 391 (2012) 4855–4866.
[10] G.-H. Mu, W.-X. Zhou, W. Chen, J. Kertész, Long-term correlations and multifractality in trading volumes for Chinese stocks, Physics Procedia 3 (2010) 1631–1640.
[11] M. El Alaoui, S. Benbachir, Multifractal detrended cross-correlation analysis in the MENA area, Physica A 392 (2013) 5985–5993.
[12] G. Cao, J. Cao, L. Xu, L.-Y. He, Detrended cross-correlation analysis approach for assessing asymmetric multifractal detrended cross-correlations and their application to the Chinese financial market, Physica A 393 (2014) 460–469.
[13] L.-Y. He, S.-P. Chen, Nonlinear bivariate dependency of price–volume relationships in agricultural commodity futures markets: A perspective from multifractal detrended cross-correlation analysis, Physica A 390 (2011) 297–308.
[14] E. Wigner, On a class of analytic functions from the quantum theory of collisions, Ann. of Math. 53 (1951) 36–67.
[15] F. Dyson, Statistical theory of the energy levels of complex systems I, II and III, J. Math. Phys. 3 (1) (1962) 140–175.
[16] V. Plerou, P. Gopikrishnan, B. Rosenow, L.A. Amaral, T. Guhr, H.E. Stanley, Random matrix approach to cross correlations in financial data, Phys. Rev. E 65 (2002) 066126.
[17] S. Pafka, I. Kondor, Estimated correlation matrices and portfolio optimization, Physica A 343 (2004) 623–634.
[18] L. Laloux, P. Cizeau, M. Potters, J.-P. Bouchaud, Random matrix theory and financial correlations, Int. J. Theor. Appl. Finance 3 (2000) 1–6.
[19] L. Laloux, P. Cizeau, J.-P. Bouchaud, M. Potters, Noise dressing of financial correlation matrices, Phys. Rev. Lett. 83 (1999) 1467–1470.
[20] S. Pafka, I. Kondor, Noisy covariance matrices and portfolio optimization II, Physica A 319 (2003) 487–494.
[21] G.-J. Wang, C. Xie, S. Chen, J.-J. Yang, M.-Y. Yang, Random matrix theory analysis of cross-correlations in the US stock market: Evidence from Pearson's correlation coefficient and detrended cross-correlation coefficient, Physica A 392 (2013) 3715–3730.
[22] J. Daly, M. Crane, H. Ruskin, Random matrix theory filters in portfolio optimisation: A stability and risk assessment, Physica A 387 (2008) 4248–4260.
[23] F. Dyson, Distribution of eigenvalues for a class of real symmetric matrices, Rev. Mexicana Fís 20 (1971) 231–237.
[24] A.M. Sengupta, P.P. Mitra, Distributions of singular values for some random matrices, Phys. Rev. E 60 (1999) 3389–3392.
[25] J.-P. Bouchaud, M. Potters, Theory of Financial Risk and Derivative Pricing: From Statistical Physics to Risk Management, Cambridge University Press, Cambridge, 2003.
[26] T. Guhr, A. Müller-Groeling, H. Weidenmüller, Random matrix theory in quantum physics, Phys. Rep. 299 (1998) 198–425.
[27] V. Plerou, P. Gopikrishnan, B. Rosenow, L.A. Amaral, H.E. Stanley, Universal and nonuniversal properties of cross correlations in financial time series, Phys. Rev. Lett. 83 (1999) 1471–1474.
[28] Y.V. Fyodorov, A.D. Mirlin, Analytical derivation of the scaling law for the inverse participation ratio in quasi-one-dimensional disordered systems, Phys. Rev. Lett. 69 (1992) 1093.
[29] Y. Fyodorov, A. Mirlin, Level-to-level fluctuations of the inverse participation ratio in finite 1D disordered systems, Phys. Rev. Lett. 71 (1993) 412.
[30] Y. Fyodorov, A. Mirlin, Statistical properties of eigenfunctions of random quasi 1D one-particle Hamiltonians, Internat. J. Modern Phys. B 8 (1994) 3795.
[31] A. Mirlin, Y. Fyodorov, The statistics of eigenvector components of random band matrices: analytical results, J. Phys. A: Math. Gen. 26 (1993) 551.
[32] E.P. Wigner, Characteristic vectors of bordered matrices with infinite dimensions, Ann. of Math. 62 (1955) 548–564.
[33] P. Lee, T. Ramakrishnan, Disordered electronic systems, Rev. Modern Phys. 57 (1985) 287.