DYNAMICZNE MODELE EKONOMETRYCZNE
X Ogólnopolskie Seminarium Naukowe, 4–6 września 2007 w Toruniu
Katedra Ekonometrii i Statystyki, Uniwersytet Mikołaja Kopernika w Toruniu

Małgorzata Snarska
Akademia Ekonomiczna w Krakowie

Financial Applications of Random Matrix Theory – Dynamics of the Covariance Matrix

1. Portfolio Problem in Finance. Markowitz Solution vs. Noise

Modern Portfolio Theory (MPT) refers to an investment strategy that seeks to construct an optimal portfolio by considering the relationship between risk and return. The success of investment does not purely depend on return, but also on the risk, which has to be taken into account. Risk itself is influenced by the correlations between different assets, thus the portfolio selection process represents a complex optimization problem. As proposed by Markowitz (1952) the underlying stochastic process is multivariate normal with know returns and covariances between different assets. In practice however these parameters are determined from market quotations. Since the number of observations is limited, empirically determined parameters will always contain some uncertainty (i.e. noise).

In order to determine the optimal portfolio, one has to invert the covariance matrix (or equivalently correlation matrix). Any measurement error (noise) will get amplified and the resulting portfolio will be sensitive to noise.
In this paper we review some standard and more recent filtering techniques, based on Random Matrix Theory (RMT), that can reduce the "empirical" noise and slightly improve standard Markowitz model's predictions.

2. Empirical Correlation Matrices

Covariance or equivalently a Correlation Matrix plays an important role in the risk measurement and portfolio optimization. Empirical Correlation Matrices, built from historical data enclose such a high amount of noise, that at first look they can be treated as random. This means, that future risk and return of a portfolio are not well estimated and controlled. Only after the proper denoising procedure is made, one can construct an efficient portfolio using Markowitz's result.

In the RMT approach one computes the correlation matrix and finds the spectrum. One then computes the variance of the part not explained by the highest eigenvalues and uses the expected PDF of the low part of the spectrum to compute λ_{min} , λ_{max}. This information is used to "remove" all the eigenvalues that fits well to the "random" part of the spectrum. Deviations from the RMT might then suggest the presence of true information. Finally, one obtains a filtered correlation coefficient matrix by transforming back the filtered diagonal matrix. In order to obtain a meaningful correlation coefficient matrix one sets to one the diagonal elements of the filtered one.

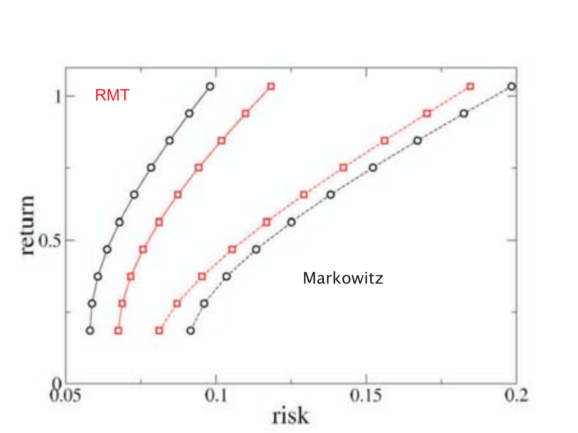

Fig.1. Estimated and effective Risk vs. Return according to the simple Markowitz model (black lines) and with RMT cleaning (red lines) (Bouchaud, Potters, 2003)

The difference (c.f.Fig.1) between the estimated (predicted) risk (left part of the diagram) and the effective one (right side of the diagram) is slightly smaller with help of the simplest RMT filtering procedure, than in the standard Markowitz case. The main goal of all the RMT techniques is to minimize the risk gap as far as it is possible.

2.1. Equal - Time Correlation Matrices

The simplest way to build empirical correlation matrix is to use N time series of quotations of length T, where in practice T is comparable to N. We have then:

$$E_{ij} = \frac{1}{T}\sum_t \frac{X_i^t X_j^t}{\sigma_i \sigma_j} \tag{1}$$

with empirical variance $\sigma^2{}_i = \frac{1}{T}\sum_t \left(X_i^2\right)^2$.

Using the Markowitz optimization one needs to find a portfolio with maximum expected return for a given risk or equivalently minimum risk for a given return P. The result in matrix notation:

$$w_C = P\frac{C^{-1}p}{p^T C^{-1}p}. \tag{2}$$

There are many different techniques to reduce the noise, all however follow the steps:

1. Diagonalize the N x N empirical correlation matrix with help of spectral decomposition $E = U^{-1}DU$, where D is a diagonal matrix and U is the matrix of eigenvectors.
2. Remove the noisy eigenvalues D = diag ($\lambda_1, \dots, \lambda_c, \lambda_{c+1}, \dots, \lambda_\kappa$).
3. Keep branch eigenvalues to obtain $D^{(filtered)}$.
4. Obtain filtered correlation matrix $E^{(filtered)}=U^{-1}D^{(filtered)}U$.
5. Restore normalization $E_{ii}^{(filtered)} = 1$.

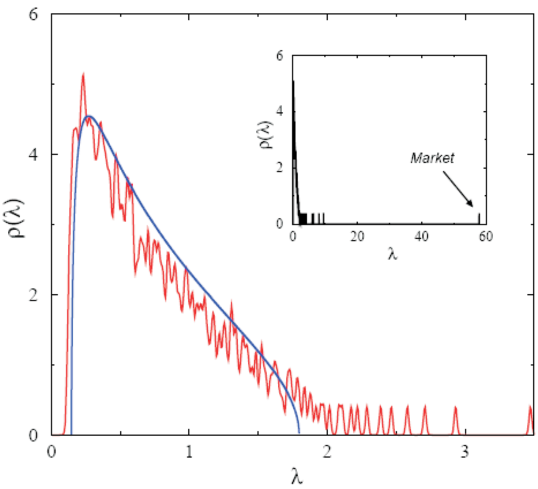

Fig.2. Empirical Correlation Matrix (Bouchaud, Potters, 2003) with Marčenko Pastur fit

It is common, that histogram of eigenvalues (Fig.2.) for any empirical correlation/covariance matrix built from financial data has one eigenvalue much bigger than the other ones (we call this eigenvalue the market one, cause the corresponding eigenvector represents all the stocks that are present in the considered market), there are also some big eigenvalues which correspond to the market sectors, and the huge amount of eigenvalues is concentrated not very far from zero.

In the RMT approach one believes, that in the lower part of the spectrum there exists a nontrivial information (i.e. one or more important eigenvalues) but blurred by noise.
Suppose that, there is only one eigenvalue behind the noise (C = I) and use the Marčenko, Pastur (1967) result:

$$\rho(\lambda)=\frac{\sqrt{(\lambda-\lambda_-)(\lambda_+-\lambda)}}{2\pi\mu_2 r\lambda}, \qquad \lambda_\pm=\mu_2(1+\varepsilon r)+\mu_1 r\pm\mu_2\sqrt{2r\varepsilon}. \tag{3}$$

Filtering technique squeezes the random part of the spectrum to the single eigenvalue (see Snarska, Krzych, 2006).
If one assumes that there exists more than one significant eigenvalue in the lower part of the spectrum, the solution is then obtained by the replica trick or by summing over planar diagrams (Burda et al., 2003):

$$ZG_E(z)=ZG_C(Z), \qquad Z=\frac{z}{1+r(zG_E(z)-1)}. \tag{4}$$

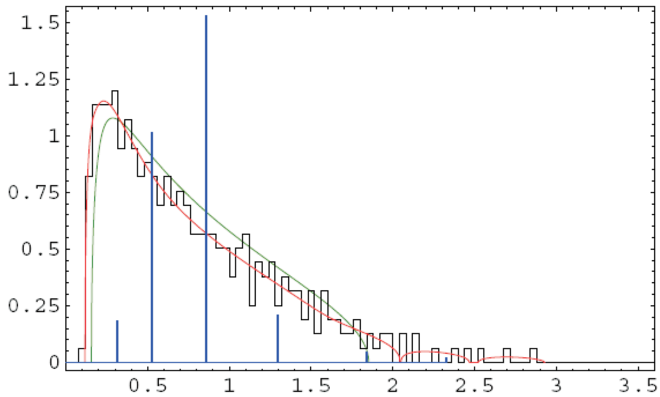

Fig.3. Empirical Correlation Matrix (Burda et al., 2003) with 6 eigenvalues in the random part of the spectrum

2.2. Filtering Techniques – a Comparison

One can very easily check, how well the method works. In order to do so one has to simulate the original known and clean covariance or correlation matrix. Adding Gaussian or any other type of noise by generating finite price – change time series will produce an "empirical" correlation (covariance) matrix. The next step is to reconstruct the best estimate of the clean covariance matrix using the investigated cleaning technique and finally one compares the risk of the cleaned, empirical portfolio to the risk of the clean portfolio using the Markowitz optimization procedure.

2.3. EWMA Correlation Matrix

Now we are ready to introduce dynamics to the model. The most trivial approach is to change the definition of the empirical covariance matrix. Going back to the case with single eigenvalue hidden in the lower part of the spectrum, the empirical matrix is now computed using an exponentially weighted moving average with:

$$E_{ij} = (1-\lambda)\sum_{k=0}^{\infty}\lambda^k X_i^k X_j^k, \qquad \left\langle X_i^k X_j^l \right\rangle = \delta_{ij}\delta_{kl}, \tag{5}$$

where $\lambda = 1 - r/N$.

As the result we get (Bouchaud, Potters, 2005, Pafka et al., 2004):

$$\rho(\lambda) = \frac{1}{\pi}\operatorname{Im} G(\lambda), \tag{6}$$

where G(λ) solves $\lambda r G = r - \log(1 - rG)$.

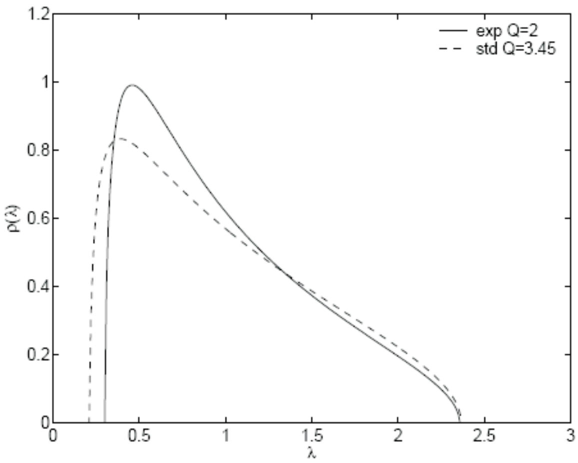

Fig.4. Spectrum of exponentially weighted random matrix compared to the spectrum of a standard Wishart matrix (Pafka et al., 2004)

2.4. Non-Equal Time Correlation Matrices and Cross Correlations

This simple model can be then further complicated by taking into the consideration the existence of nontrivial cross correlations in time, which are not observed in the standard Markowitz model.

One stock's price behavior does not only affect the other stock's price behavior in the same trading day. One stock can also mimic the behavior of the another one within next few trading days. The effect is easily seen, when taking under consideration more dense data i.e. not only daily quotations, but also e.g. 15 – minutes data.

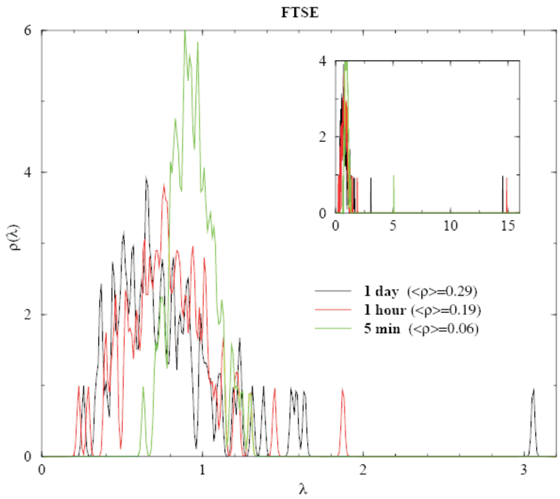

Fig.5. Eigenvalue density for different time scales of quotations

That is why now we build the empirical correlation matrix with:

$$E_{ij}^{\tau} = \frac{1}{T}\sum_{t}\frac{X_i^t X_j^{t+\tau}}{\sigma_i \sigma_j} \tag{7}$$

with E still N x N but not symmetric.
The final result (Bouchaud et. al., 2007) is similar to the Marčenko – Pastur case:

$$\rho(s) = (1-n, 1-m)^{+}\delta(s) + (m+n-1)^{+}\delta(s-1) + \frac{\sqrt{(s^2-\gamma_{-})(\gamma_{+}-s^2)}}{\pi s(1-s^2)}, \tag{8}$$

where

$$\gamma_{\pm} = n + m - 2mn \pm 2\sqrt{mn(1-n)(1-m)}, \qquad 0 \le \gamma_{\pm} \le 1. \tag{9}$$

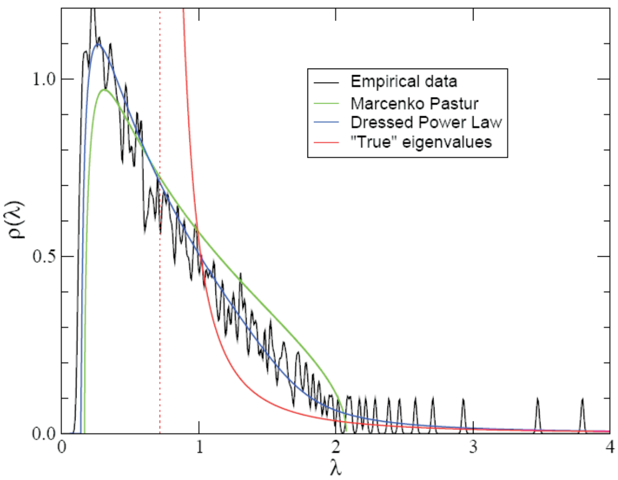

Fig.6. Wishart – type distributions for a non – equal time correlation matrix (Bouchaud et. al., 2007)

3. Conclusions

The aim of this paper was to present and review some old and more recent methods of the Random Matrix Theory approach to the Portfolio Problem in Finance.

There are still however problems that need to be considered, like the behavior of the highest eigenvalue when the data are more dense or fat tails, usually present in the financial time series.

Furthermore there is a strong need in making the methods more efficient not only in the artificial environment and it is then crucial to evaluate the methods that can not only eliminate the noise from the eigenvalue spectrum, but can also clean the corresponding eigenvectors.

References

Bouchaud J. P., et. al. (2007), Talk presented At the Random Matrix Theory: From Fundamental Physics to Applications ESF Exploratory Workshop, 2–6 May 2007, Kraków (to be published).

Bouchaud J. P., Potters M., Laloux, L. (2005), Financial Applications of Random Matrix Theory: Old Laces and New Pieces, APPB, Vol. 36 (9), 2767–2784.

Bouchaud, J. P., Potters, M. (2003), *Theory of Financial Risk. From Statistical Physics to Risk Management 2nd ed.*, Cambridge University Press, Cambridge.

Burda, Z., Görlich, A., Jarosz, A., Jurkiewicz, J. (2003), Signal and Noise in Correlation Matrix, http://arxiv.org/pdf/cond-mat/0305627.

Burda, Z., Jurkiewicz, J. (2004), Signal and Noise in Financial Correlation Matrices, http://arxiv.org/pdf/cond-mat/0312496.

Marčenko, V.A., Pastur, L.A. (1967), Distribution of Eigenvalues for Some Sets of Random Matrices, *Math USSR Sbornik*, 1, 457–483.

Markowitz, H. (1952), Portfolio Selection, *The Journal of Finance*, Volume 7 No. 1, 77–91.

Pafka Sz., Potters M., Kondor I., (2004), Exponential Weighting and Random-Matrix-Theory-Based Filtering of Financial Covariance Matrices for Portfolio Optimization, http://arxiv.org/pdf/cond-mat/0402573.

Snarska M., Krzych J., (2006), Automatic Trading Agent. RMT Based Portfolio Theory and Portfolio Selection, APPB Vol. 37 (11), 3145–3160.