Testing the Markowitz Portfolio Optimization Method with Filtered Correlation Matrices*

Imre Gera
Institute of Informatics
P.O. Box 652
H-6701 Szeged, Hungary

Balázs Bánhelyi
Institute of Informatics
P.O. Box 652
H-6701 Szeged, Hungary

András London†
Institute of Informatics
P.O. Box 652
H-6701 Szeged, Hungary
london@inf.u-szeged.hu

ABSTRACT

In this work we analyze the performance of the Markowitz portfolio optimization method on the Budapest Stock Exchange data set using two different filtering techniques defined for correlation matrices. The results show that the estimated risk is much closer to the realized risk using filtering methods. Bootstrap analysis shows that ratio between the realized return and the estimated risk (Sharpe ratio) is also improved by filtering.

Categories and Subject Descriptors

I.6 [**Simulation and Modelling**]: Applications
; G.1.6 [**Optimization**]: *Constrained optimization, Nonlinear programming*

Keywords

Portfolio optimization, Markowitz model, Correlation matrices, Random matrix theory, Hierarchical clustering

1. INTRODUCTION

The portfolio optimization is one of the most important problem in asset management aims at reducing the risk of an investment by diversifying it into independently fluctuating assets [5]. In his seminal work [14], Markowitz formulated the problem through the criteria that given the expected return, the risk - measured by the variability of the return - has to be minimized. The classical model measures the risk as the variance of the asset returns resulting in a quadratic programming problem. Recently, the analysis of the correlation coefficient matrix, that appears through the covariance matrix in the objective function of the model, has become the focus of interest [2, 4, 9, 10, 17, 19]. Many attempts have been made in order to quantify the degree of statistical uncertainty present in the correlation matrix and filter the part of information which is robust against this uncertainty [2, 7, 9, 10, 11]. The filtered correlation matrices have been successfully used in portfolio optimization in terms of risk reduction [10, 17, 19]. In these studies, it was often assumed that the investor has perfect knowledge on the future returns.

In this work we investigate the portfolio selection problem using different filtering procedures applied to the correlation matrix. We measure the performance of the procedures in terms of both the predicted and realized risk and return, respectively. The future returns are not known at the time of the investment. In Section 2 we briefly describe the Markowitz portfolio optimization problem and two approaches for the correlation matrix filtering (Random Matrix Theory, Clustering). In Section 3 we present our results using standard performance measures on the return and risk, and finally, in Section 4 we draw some conclusions and indicate future work.

*This work was partially supported by the National Research, Development and Innovation Office - NKFIH, SNN-117879.
†Corresponding author

2. PORTFOLIO OPTIMIZATION

In Markowitz' formulation, the portfolio problem is a single period model of investment. At the beginning of the period (t_0), an investor allocates the capital among different assets. During the investment period ($[t_0, T]$), the portfolio produces a random rate of return and results a new value of the capital. In the original model of Markowitz, the risk of a single asset is measured by the variance of its returns, while the risk of the portfolio is measured via the covariance matrix of the returns of the assets in the portfolio. In this section we briefly introduce the Markowitz portfolio optimization problem and describe two filtering procedures of the covariance matrix in order to obtain less noisy matrix to decrease the statistical uncertainty it contains.

2.1 Markowitz's model

Given n risky assets, a portfolio composition is determined by the weights p_i ($i = 1, \dots, n$), such that $\sum_i^n p_i = 1$, indicating the fraction of wealth invested in asset i. The expected return and the variance of the portfolio $\mathbf{p} = (p_1, \dots, p_n)$ are

$$r_p = \sum_{i=1}^{n} p_i r_i = \mathbf{p}\mathbf{r}^T \tag{1}$$

and

$$\sigma_p^2 = \sum_{i=1}^{n} \sum_{j=1}^{n} p_i p_j \sigma_{ij} = \mathbf{p}\Sigma\mathbf{p}^T, \tag{2}$$

where r_i is the expected return of asset i, σ_{ij} is the covariance between asset i and j and $\mathbf{\Sigma}$ is the covariance matrix. Vectors are considered as row vectors int this paper.

In the classical Markowitz model [14] the risk is measured by the variance providing a quadratic optimization problem which consists of finding a vector $\mathbf{p}$, assuming $\sum_{i=1}^{n} p_i = 1$, that minimizes σ_p^2 for a given "minimal expected return" value of r_p. Now, we assume that short selling is allowed and therefore p_i can be negative. The solution of this problem, found by Markowitz, is

$$\mathbf{p}^* = \lambda \mathbf{\Sigma}^{-1}\mathbf{1}^T + \gamma \mathbf{\Sigma}^{-1}\mathbf{r}^T, \tag{3}$$

where $\mathbf{1} = (1, \ldots, 1)$, while the other parameters are

$$\lambda = (C - r_p B)/D \text{ and } \gamma = (r_p A - B)/D,$$

where

$$A = \mathbf{1}\mathbf{\Sigma}^{-1}\mathbf{1}^T, B = \mathbf{1}\mathbf{\Sigma}^{-1}\mathbf{r}^T, C = \mathbf{r}\mathbf{\Sigma}^{-1}\mathbf{r}^T, D = AC - B^2.$$

Considering the daily price time series of n assets and denoting the closure price of asset i at time t ($t = 1, \ldots, T$) by $P_i(t)$, the daily logarithmic return of i is defined as

$$r_{it} = \log \frac{P_i(t)}{P_i(t-1)} = \log P_i(t) - \log P_i(t-1). \tag{4}$$

In case of stationary independent normal returns, which is usually assumed for asset prices, the maximum likelihood estimator is the sample mean of the past observations of r_i, is defined as

$$\hat{r}_i = \frac{1}{T}\sum_{t=1}^{T} r_{it}. \tag{5}$$

Hence, for the portfolio we define $\hat{\mathbf{r}} = (\hat{r}_1, \ldots, \hat{r}_n)$. The covariance σ_{ij} between assets i and j is estimated by

$$\hat{\sigma}_{ij} = \frac{1}{T-1}\sum_{t=1}^{T}(r_{it} - \hat{r}_i)(r_{jt} - \hat{r}_j) \tag{6}$$

and for the portfolio $\hat{\mathbf{\Sigma}} = (\hat{\sigma}_{ij})_{i,j}$. The correlation coefficient between asset i and j is defined as

$$\rho_{ij} = \sigma_{ij}/\sqrt{\sigma_{ii}\sigma_{jj}}, \tag{7}$$

where σ_{ii} is often called the volatility of asset i.

2.2 Random matrix theory and correlation matrices

A simple random matrix is a matrix whose elements are random numbers from a given distribution [15]. In context of asset portfolios random matrix theory (RMT) can be useful to investigate the effect of statistical uncertainty in the estimation of the correlation matrix [19]. Given the time series of length T of the returns of n assets and assuming that the returns are independent Gaussian random variables with zero mean and variance σ^2, then in the limit $n \to \infty$, $T \to \infty$ such that $Q = T/n$ is fixed, the distribution $\mathcal{P}_{rm}(\lambda)$ of the eigenvalues of the random correlation matrix (C_{rm}) is given by

$$\mathcal{P}_{rm}(\lambda) = \frac{Q}{2\pi\sigma^2}\frac{\sqrt{(\lambda_{\min} - \lambda)(\lambda_{\max} - \lambda)}}{\lambda}, \tag{8}$$

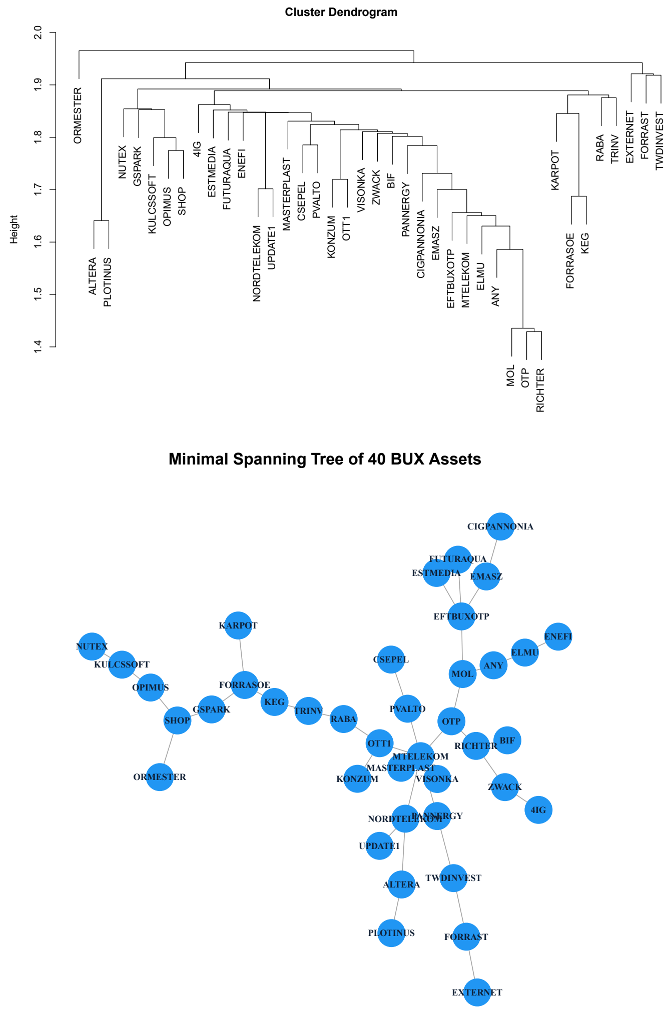

Figure 1: Indexed hierarchical tree - obtained by the single linkage procedure - and the associated MST of the correlation matrix of 40 assets of the Budapest Stock Exchange

where $\lambda_{\min}$ and $\lambda_{\max}$ are the minimum and maximum eigenvalues, respectively [18], given in the form

$$\lambda_{\max,\min} = \sigma^2\left(1 + \frac{1}{Q} \pm 2\sqrt{\frac{1}{Q}}\right). \tag{9}$$

Previous studies have pointed out that the largest eigenvalue of correlation matrices from returns of financial assets is completely inconsistent with Eq. 8 and refers to the common behavior of the stocks in the portfolio [9, 16]. Since Eq. 8 is strictly valid only for $n \to \infty$, $T \to \infty$, we constructed random matrices for certain n and T values of the data sets that are used and compare the largest eigenvalues and the spectrum with $\mathbf{C}$. We found high consistency with Eq. 8. Since Trace($\mathbf{C}$) $= n$ the variance of the part not explained by the largest eigenvalue can be quantified as $\sigma^2 = 1 - \lambda_{\text{largest}}/n$. Using this, we can recalculate $\lambda_{\min}$ and $\lambda_{\max}$ in Eq. 9 and construct a filtered diagonal matrix $\mathbf{C}_{RMT}$, that we get by setting all eigenvalues of $\mathbf{C}$ smaller than $\lambda_{\max}$ to zero and transform it to the basis of $\mathbf{C}$ with setting the diagonal elements to one (and using singular value decomposition). A possible RMT approach for portfolio optimization, following [17], is to use $\mathbf{\Sigma}_{RMT}$ (that can be easily calculated form $\mathbf{C}_{RMT}$) instead of $\mathbf{\Sigma}$ in the Markowitz model.

2.3 Clustering

The correlation matrix $\mathbf{C}$ has $n(n-1)/2 \sim n^2$ distinct elements therefore it contains a huge amount of information even for a small number of assets considered in the portfolio selection problem. As shown by Mantegna and later many others [3, 8, 12, 19, 20], the single linkage clustering approach [6] (closely related to minimal spanning trees (MST), Fig. 1) provides economically meaningful information using only $n-1$ distinct elements of the correlation matrix. To construct the filtered matrix, the correlation matrix $\mathbf{C}$ is converted into a distance matrix $\mathbf{D}$, for instance following [12, 13], using $d_{ij} = \sqrt{2(1-\rho_{ij})}$ ultrametric distance[1]. The distance matrix $\mathbf{D}$ can be seen as a fully connected graph of the assets with edge weights d_{ij} representing similarity between time series of them. Then the filtered correlation matrix $\mathbf{C}_{MST}$ is constructed with just $n-1$ distinct correlation coefficients by converting the filtered ultrametric distance matrix back. It was proven that the ultrametric correlation matrix obtained by the single linkage clustering method is always positive definite if all the elements of the obtained ultrametric correlation matrix are positive [1]. This condition has been observed for all correlation matrices we used. Then, for portfolio optimization, we can use the obtained $\mathbf{\Sigma}_{MST}$ instead of $\mathbf{\Sigma}$ in the Markowitz model.

3. RESULTS

3.1 Data set

To compare the performance of the methods we analyze the data set of $n = 40$ stocks traded in the Budapest Stock Exchange (BSE) in the period 1995-2016, using 5145 records of daily returns per stock.

We consider $t = t_0$ as the time when the optimization is performed. Since the covariance matrix has $\sim n^2$ elements while the number of records used in the estimation is nT, the length of the time series need to be $T >> n$ in order to get small errors on the covariance. On the other hand, for large T the non-stationarity of the time series likely appears. This problem is known as the curse of dimensionality. Because of this, we compute the covariance matrix and expected returns using the $[-T, 0]$ interval, i.e. using $T = 50 \approx n$, $T = 100 > n$ and $T = 500 >> n$ days preceding $t = 0$. Furthermore, filtering techniques are able the filter the part of the covariance matrix which is less affected by statistical uncertainty. To quantify and compare the different methods are considered, we use the measures described below.

3.2 Performance evaluation

To measure the performance of the portfolios determined by the different models, we use the following quantities for the estimated return and risk at the time of investment and the realized risk and returns after the investment period. For

[1]Ultrametric distances are such distances that satisfy the inequality $d_{ij} \leq \max\{d_{ik}, d_{kj}\}$, which is a stronger assumption that the standard triangular inequality.

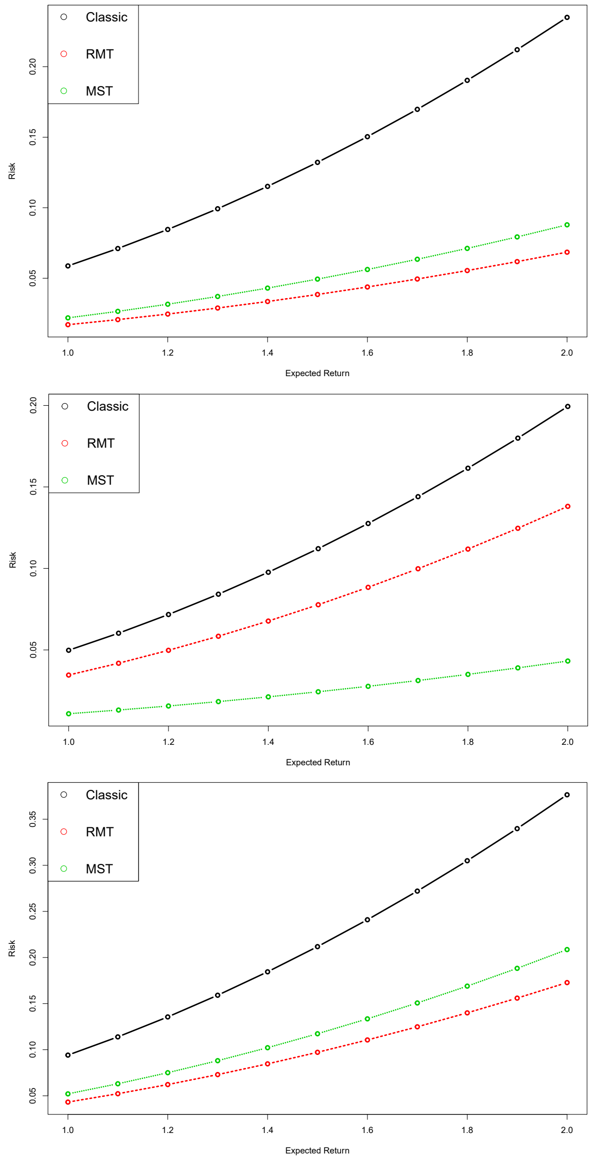

Figure 2: The ratio of the realized risk σ_r^2 and the predicted risk $\hat{\sigma}_p^2$ as the function of expected portfolio return r_p for the different procedures as $T = 50, 100, 500$ (top-down). The data set contains 40 BSE stocks in the period

portfolio p, the ex-ante Sharpe ratio measures the excess return per unit of risk:

$$S_p = \frac{\hat{r}_p - r_f}{\sigma_p}, \tag{10}$$

while the ex-post Sharpe ratio uses the same equation but with the realized return r_p. Here, r_f is the risk-free rate of return. The portfolio risk, due to the estimation of the correlation matrix is calculated as

$$R_p = \frac{|\sigma_r^2 - \hat{\sigma}_p^2|}{\hat{\sigma}_p^2} \tag{11}$$

where $\hat{\sigma}_p^2$ is the predicted risk, while σ_r^2 is the realized risk of the portfolio.

Table 1: Bootstrap experiments using 50 random samples for each value of T in case of 120% expected return

$r_p = 1.2$		Original	RMT	MST
T=50	Return	0.145 (0.330)	0.180 (0.425)	0.186 (0.348)
	S_p	0.009	0.180	0.186
	R_p	16.66	0.99	0.99
T=100	Return	0.319 (0.332)	0.315 (0.541)	0.362 (0.418)
	S_p	0.036	0.315	0.364
	R_p	8.954	0.99	0.99
T=500	Return	-0.185 (0.928)	-0.313 (1.234)	0.264 (0.724)
	S_p	-0.077	-0.313	0.264
	R_p	2.415	0.99	0.99

3.3 Experiments

Fig. 2 shows the ratio of the ratio of the realized risk σ_r^2 and the predicted risk $\hat{\sigma}_p^2$ as the function of the estimated return r_p obtained by the different procedures. For each T, the investment time t_0 and the set of stocks were the same. The ratio is significantly smaller in case of the portfolios that obtained by using filtering. Interestingly, for $T = 100$ the MST method gave better results than the RMT.

To check the robustness of the methods, we performed a bootstrap experiment as follows. We considered 50 random initial times to solve the optimization problem using the time series on the intervals $[-T, t_0]$ ($T = 50, 100, 500$). For each portfolio, we computed the predicted risk using Eq. 2 for expected returns $r_p = 1, 1.1, \ldots, 2$ ($0 - 100\%$ gain). We further constrained p_i to the interval $[-1, 1]$ and used the Lagrange multiplier method for the optimization. In all cases, the portfolios with realized returns in the top and bottom 10% were neglected. We computed the realized risk using the calculated stock weights at t_0 and the realized covariance matrix on $[t_0, T]$. We also computed the realized returns by comparing the value of the portfolio at t_0 and T. The average S_p, R_p values and returns with standard deviations for $r_p = 1.2$ are shown in Tab. 1. It can be seen, the R_p values are significantly smaller in case of the RMT and MST than in case of the original method for each T confirming the reliability of the filtering methods. The post-ante Sharpe ratio, however it is much smaller than 1 in every case, also shows the that the RMT and MST methods outperforms the original method. We note, interestingly, that the highest average return was obtained for $T = 100$ (and not for $T = 500$) using the BSE data set.

4. CONCLUSIONS

In this study, we performed portfolio optimization using filtered correlation matrices obtained by two different procedures, namely a random matrix theory approach and the single linkage clustering. A large set of experiments have shown that using filtered covariance matrices the original Markowitz solution is outperformed in terms of standard portfolio performance measures.

In the future, it would be interesting to analyze portfolio optimization using various estimators of expected returns together with different filtering procedures and check the methods using various stock exchange data sets and also varying the number of stocks considered.

5. REFERENCES

[1] M. R. Anderberg. Cluster analysis for applications. monographs and textbooks on probability and mathematical statistics, 1973.

[2] T. Conlon, H. J. Ruskin, and M. Crane. Random matrix theory and fund of funds portfolio optimisation. *Physica A: Statistical Mechanics and its applications*, 382(2):565–576, 2007.

[3] T. Di Matteo, T. Aste, and R. N. Mantegna. An interest rates cluster analysis. *Physica A: Statistical Mechanics and its Applications*, 339(1):181–188, 2004.

[4] M. El Alaoui. Random matrix theory and portfolio optimization in moroccan stock exchange. *Physica A: Statistical Mechanics and its Applications*, 433:92–99, 2015.

[5] E. J. Elton, M. J. Gruber, S. J. Brown, and W. N. Goetzmann. *Modern portfolio theory and investment analysis*. John Wiley & Sons, 2009.

[6] J. C. Gower and G. Ross. Minimum spanning trees and single linkage cluster analysis. *Applied statistics*, pages 54–64, 1969.

[7] T. Guhr and B. Kälber. A new method to estimate the noise in financial correlation matrices. *Journal of Physics A: Mathematical and General*, 36(12):3009, 2003.

[8] L. Kullmann, J. Kertész, and K. Kaski. Time-dependent cross-correlations between different stock returns: A directed network of influence. *Physical Review E*, 66(2):026125, 2002.

[9] L. Laloux, P. Cizeau, J.-P. Bouchaud, and M. Potters. Noise dressing of financial correlation matrices. *Physical review letters*, 83(7):1467, 1999.

[10] L. Laloux, P. Cizeau, M. Potters, and J.-P. Bouchaud. Random matrix theory and financial correlations. *International Journal of Theoretical and Applied Finance*, 3(03):391–397, 2000.

[11] Y. Malevergne and D. Sornette. Collective origin of the coexistence of apparent random matrix theory noise and of factors in large sample correlation matrices. *Physica A: Statistical Mechanics and its Applications*, 331(3):660–668, 2004.

[12] R. N. Mantegna. Hierarchical structure in financial markets. *The European Physical Journal B-Condensed Matter and Complex Systems*, 11(1):193–197, 1999.

[13] R. N. Mantegna and H. E. Stanley. *Introduction to econophysics: correlations and complexity in finance*. Cambridge university press, 1999.

[14] H. Markowitz. Portfolio selection: Efficient diversification of investments. cowles foundation monograph no. 16, 1959.

[15] M. L. Mehta. *Random matrices*, volume 142. Academic press, 2004.

[16] V. Plerou, P. Gopikrishnan, B. Rosenow, L. A. N. Amaral, and H. E. Stanley. Universal and nonuniversal properties of cross correlations in financial time series. *Physical Review Letters*, 83(7):1471, 1999.

[17] B. Rosenow, V. Plerou, P. Gopikrishnan, and H. E. Stanley. Portfolio optimization and the random magnet problem. *EPL (Europhysics Letters)*, 59(4):500, 2002.

[18] A. M. Sengupta and P. P. Mitra. Distributions of singular values for some random matrices. *Physical Review E*, 60(3):3389, 1999.

[19] V. Tola, F. Lillo, M. Gallegati, and R. N. Mantegna. Cluster analysis for portfolio optimization. *Journal of Economic Dynamics and Control*, 32(1):235–258, 2008.

[20] M. Tumminello, T. Di Matteo, T. Aste, and R. Mantegna. Correlation based networks of equity returns sampled at different time horizons. *The European Physical Journal B*, 55(2):209–217, 2007.