Vol. 36 (2005) *ACTA PHYSICA POLONICA B* No 9

POWER MAPPING AND NOISE REDUCTION FOR FINANCIAL CORRELATIONS*

Per-Johan Andersson, Andreas Öberg, Thomas Guhr

Matematisk Fysik, LTH, Lunds Universitet
Box 118, 22100 Lund, Sweden

(Received June 27, 2005)

The spectral properties of financial correlation matrices can show features known from completely random matrices. A major reason is noise originating from the finite lengths of the financial time series used to compute the correlation matrix elements. In recent years, various methods have been proposed to reduce this noise, *i.e.* to clean the correlation matrices. This is of direct practical relevance for risk management in portfolio optimization. In this contribution, we discuss in detail the *power mapping*, a new shrinkage method. We show that the relevant parameter is, to a certain extent, self-determined. Due to the "chirality" and the normalization of the correlation matrix, the optimal shrinkage parameter is fixed. We apply the power mapping and the well-known *filtering* method to market data and compare them by optimizing stock portfolios. We address the rôle of constraints by excluding short selling in the optimization.

PACS numbers: 89.65.Gh, 05.45.Tp

1. Introduction

Precise knowledge of the correlations between stocks or, more generally, risk elements in a portfolio is crucial for risk management [1]. However, the computation of the correlation coefficients $C_{kl},\ k,l = 1,\ldots,K$ for company k and l from the recorded time series does not always yield reliable results. An important reason is the finiteness of the time series which adds a specific type of *noise* to the correlation coefficients. This noise dressing was analyzed for empirical data in Refs. [2–4] by establishing a connection to random matrices [5]. It was shown that many spectral properties of financial correlation matrices C tend to be indistinguishable from those of purely random matrices. It should be emphasized that the random matrices in question belong

* Presented at the Conference on Applications of Random Matrices to Economy and Other Complex Systems, Kraków, Poland, May 25–28, 2005.

to the class of *chiral* random matrices which are important in the statistical analysis of lattice quantum chromo dynamics [6].

Numerous studies are devoted to noise dressing and a variety of methods were proposed for *noise reduction*. For a brief review, we refer the reader to Ref. [7]. Of the more recent work, we mention the interesting studies of Refs. [8] and [9]. Other contributions to these proceedings also address noise dressing, noise reduction and related issues. Here, we focus on the *power mapping* [7] and on a comparison between *filtering* [4, 10, 11] and power mapping.

For later purposes, we recall that the spectral density of a correlation matrix consists of, first, a generic *bulk part* which is well described by the analytically known spectral density for chiral random matrices and, second, a part reflecting the *industrial branches* [4, 10, 11]. This implies that only the information in the latter can directly be used, while the information in the bulk is buried under noise. In the filtering method, one determines a cut-off eigenvalue λ_{c}, separating the bulk and the large eigenvalues due to the industrial branches. Only the latter eigenvalues are re-expressed in the original basis, while the former are discarded. This gives the filtered, noise reduced correlation matrix $C^{(\mathrm{filtered})}$. The filtering has to be done in such a way that the normalization is preserved, *i.e.* one must have $C_{kk}^{(\mathrm{filtered})} = 1$.

2. Power mapping

The filtering is a powerful and robust method. However, some question arise. For example, if the dimension K of the correlation matrix is relatively small, the random matrix properties have not reached the ergodic limit, implying that the cut-off λ_c is ambiguous. Also, if some branches are small, the corresponding eigenvalues are small and might even slide into the bulk such that the filtering would remove relevant information. The power mapping proposed in Ref. [7] is an alternative method. Our main motivation in developing it was the observation that, to the best of our knowledge, all other noise reduction methods require some choice of parameters (such as the cut-off λ_c in the filtering) or other type of input. This introduces ambiguity. We will show that the power mapping is, to some extent, parameter-free. It is thus a *black box* method which does not even require to compute and inspect the eigenvalues or any other feature of the correlation matrix.

2.1. Idea and basic features

In the present context, "chirality" means that, by definition, every element C_{kl} of the correlation matrix C is a scalar product of the properly normalized time series for companies k and l. Importantly, this implies that each element contains the information about the length of the time series.

We now show that this can be exploited. We map the correlation matrix C via the power mapping

$$C_{kl} \qquad \longrightarrow \qquad \mathrm{sign}\,(C_{kl})\,|C_{kl}|^q \;=\; C_{kl}^{(q)} \tag{2.1}$$

onto another correlation matrix $C^{(q)}$. In this matrix, the noise-dressed information is partly recovered [7]. The optimal power $q \approx 1.5$ is automatically determined by the very definition and normalization of the correlation matrix. The effect is illustrated in Fig. 1 for the bulk part of the spectral density. These correlation matrices were generated from a one-factor-model.

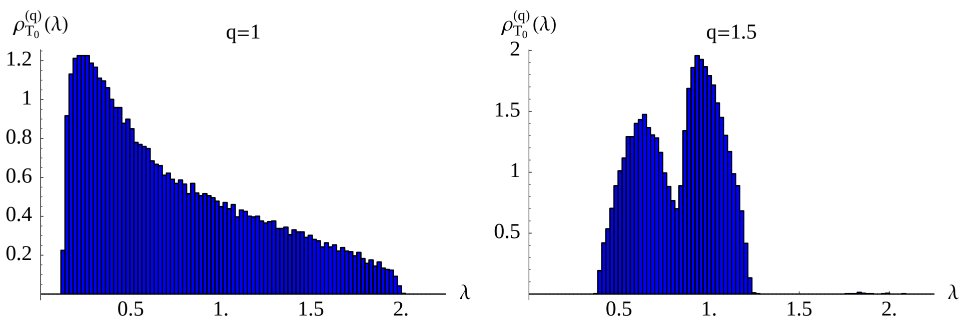

Fig. 1. Bulk part of the spectral density for correlation matrices, generated with a one-factor-model, before (left) and after (right) the power mapping. Taken from Ref. [7].

After the power mapping, two peaks are seen: the left one stems from the true correlations, while the right one is produced by remainders of the noise around the unit normalization of all diagonal elements — before and after the power mapping. This clearly shows that even the information in the bulk can partly be reconstructed. The power mapping can be viewed as an "effective prolongation" of the time series.

We stress that the power mapping affects the entire correlation matrix and thus changes all eigenvalues. In the sequel, we will mainly discuss how the bulk changes. However, this does not mean that the larger eigenvalues due to the industrial branches are removed by the power mapping or otherwise omitted. Their numerical value is a bit lowered, but they are always there and contribute significantly to the power mapped correlation matrix $C^{(q)}$. We focus on the bulk because it turns out that some important properties of the power mapping, in particular the dependence on the length of the time series, can conveniently be deduced from there.

Heuristically, one can explain the effect as follows. Some elements C_{kl} comprise a *true* part u, say, and a *noisy* part v which scales with $1/\sqrt{T}$

where T is the length of the time series. Purely noisy elements have no true part. The power mapping gives

$$\begin{aligned} \left(u + \frac{v}{\sqrt{T}}\right)^q &= u^q + q\frac{u^{q-1}v}{\sqrt{T}} + \mathcal{O}\left(\frac{1}{T}\right), \\ \left(\frac{v}{\sqrt{T}}\right)^q &= \frac{v^q}{T^{q/2}}\,. \end{aligned} \tag{2.2}$$

Thus, the elements which only contain noise are stronger suppressed than those with a true part, if $q > 1$. The power mapped correlation matrices $C^{(q)}$ are used as they stand instead of the original ones for risk management. We mention that the $1/\sqrt{T}$ scaling of the noise results from the Gaussian statistics which we assume in our one-factor-model. If one uses other statistics, it might change to $1/T^\alpha$ with some parameter α. This does not spoil our line of arguing, because one simply has to replace $1/\sqrt{T}$ with $1/T^\alpha$ everywhere in Eq. (2.2). Thus, the noise goes with $1/T^{q\alpha}$ after the power mapping, while the correction to the true correlation still goes with $1/T^\alpha$. We conclude that the power mapping can handle any statistics. This is important for the applications to be presented later on. Here, we only aim at clarifying the mechanisms of how the power mapping works. Hence, it is sufficient to keep the simple Gaussian statistics for the discussion.

2.2. Parameter-free "effective prolongation" of the time series

At first sight, Eq. (2.2) seems to imply that a value of q higher than $q \approx 1.5$ would give even better results. This is not so, as can be seen in Fig. 2 which shows spectral densities for increasing values of q. Beyond $q \approx 1.5$, the two peaks that emerged because of the power mapping, grow together again. The reason is simply that the true part u in Eq. (2.2) is bound, *i.e.* $|u| < 1$. A q value too high suppresses the true part too much. Consequently, there is an optimal value for q which, by visual inspection, is found to be $q \approx 1.5$. In Ref. [7], more quantitative arguments are given. However, very recently, it has been shown that there is a dependence of the parameter q on controllable overall properties of the correlation matrix, but not on its internal structure [12]. This independence is the important one. In any case, more work on these issues is called for.

We now show that the power mapping can indeed be viewed as an "effective prolongation" of the time series. To this end, we fit the two peaks of the bulk with two Gaussians. We do that for $q = 1, 1.25, 1.5$ and for different length T of the time series. In Fig. 3, the standard deviations of these two Gaussians are displayed as functions of T. The predicted scaling $1/\sqrt{T}$ for the correlation peak and $1/T^{q/2}$ for the noise peak is realized. One should keep in mind that the predictions of the scaling were only rough estimates.

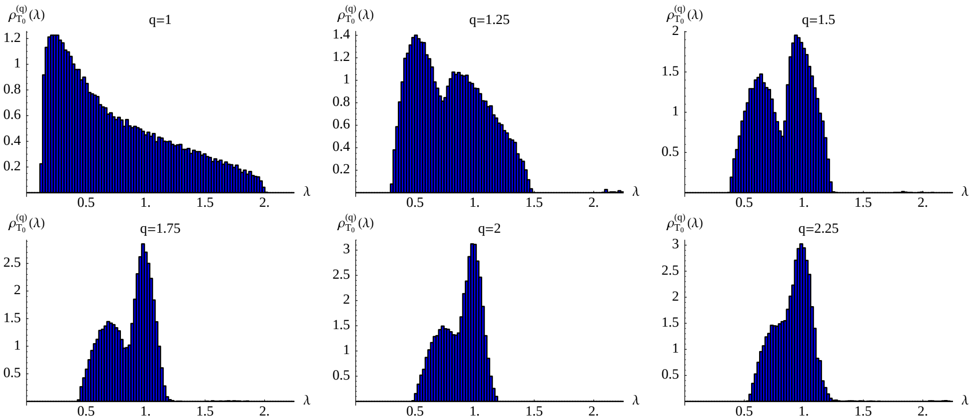

Fig. 2. Bulk part of the spectral density for correlation matrices, generated with a one-factor-model, for different values of the parameter $q = 1, 1.25, 1.5$ (top) and $q = 1.75, 2, 2.25$ (bottom). Taken from Ref. [7].

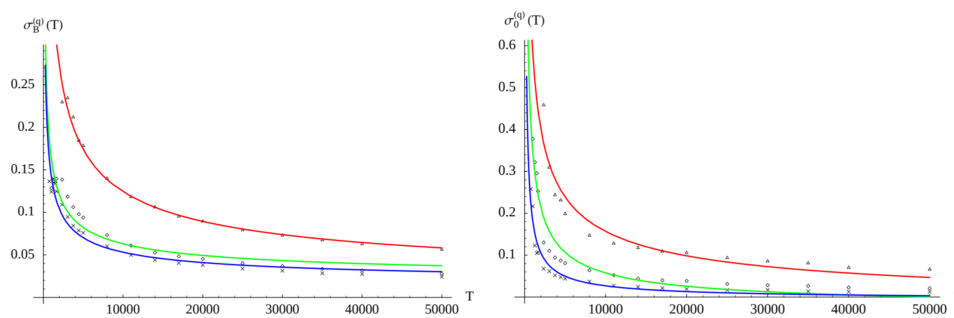

Fig. 3. Standard deviations of the fitted Gaussians as function of the length T of the time series for $q = 1, 1.25, 1.5$ from top to bottom. Results form the fits (crosses, diamonds, triangles) compared with the scaling behavior (solid lines). The standard deviations of the left peaks stemming from the true correlations scale with $1/\sqrt{T}$ (left), while those of the right peaks stemming from the noise scale with $1/T^{q/2}$ (right). Taken from Ref. [7].

Nevertheless, the trend clearly confirms our claim: the standard deviation that one would expect for a correlation matrix computed from much longer time series is effectively reached for much shorter time series.

3. Application to market data

3.1. Procedure

We compare the two noise reduction methods by optimizing stock portfolios. To this end, we use the standard Markowitz theory. The correlation matrices for our empirical data are computed by sampling over a certain (longer) period. Then, two noise reduced correlation matrices are obtained by applying the two methods. The portfolio is evaluated over a certain (shorter) period with the noise reduced correlation matrices, and for additional comparison, with the original, not noise reduced correlation matrix (referred to as *sample*). Historical data are employed, implying that we can compare, at the end of the evaluation period, risk and return involving the correlation matrices without noise reduction with risk and return after noise reduction has been applied. To reach a basic statistical significance, the lengths of sample and evaluation period are chosen considerably shorter than the total length of the available time series. Hence, the procedure can be applied several times by moving through the available data. The data are daily Swedish stock returns [13] for 197 companies from July 12$^{\text{th}}$, 1999, to July 18$^{\text{th}}$, 2003, sampling period one year, evaluation period one week.

3.2. Results

To always achieve a solution of the Markowitz optimization, we set the desired return to 0.3% per week. Higher values were not possible in this study, because the performance of the market was in general weak during that time. We estimate the expected returns on the individual stocks by fitting their drift. We compute the daily risk and the monthly return as functions of time. In a first step, we impose no constraints in the Markowitz optimization, the results are shown in Fig. 4. The yearly actual risk which amounts to 20.7% without noise reduction is considerably lowered by the noise reduction, we find 11.3% for power mapping and 11.4% for filtering. Due to the uncertainties stemming from the expected returns for the individual stocks, statements about the return on the portfolio are, in general, less meaningful in such a study. Nevertheless, we believe that they are still of interest as relative information, when comparing the results for the two noise reduction methods — and the latter with the results without noise reduction. The yearly actual return is 11.1% without noise reduction and is given by 5.0% and 10.5% for power mapping and filtering, respectively. Thus both methods reduce the risk very efficiently, but the filtering seems to be at an advantage regarding the return. In a second step, we impose a constraint by forbidding short selling. The results are quite different, they are shown in Fig. 5. The yearly actual risk of 10.1% without noise reduction is only very slightly lowered to 9.9% for both noise reduction methods. The

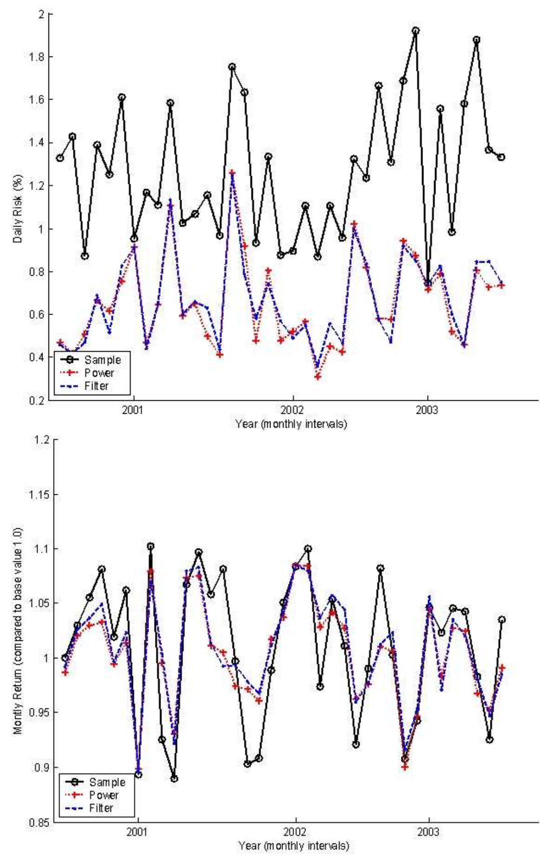

Fig. 4. Daily risk (top) and monthly return (bottom) for the Swedish market data, no constraints imposed: no noise reduction, *i.e.* plain sample (circles), power mapping (crosses) and filtering (dashes).

yearly actual return is 0.5% without noise reduction and is improved to 1.1% for power mapping and 0.7% for filtering. Being aware of the ambiguities in evaluating returns in this context, we find it interesting that the power mapping seems to perform better than the filtering in this case

The remarkable difference between the results in these two steps without and with constraints becomes clearer by looking at the fractions of wealth invested in the individual companies as obtained form the Markowitz optimization. Without constraints, the fractions scatter almost symmetrically around zero. Hence, there are many negative fractions and short selling is very important to reach the optimal portfolio. When short selling is forbidden and all weights are confined to non-negative values, the situation is dramatically changed. The power mapping handles this situation well by diversifying with less fluctuations than filtering.

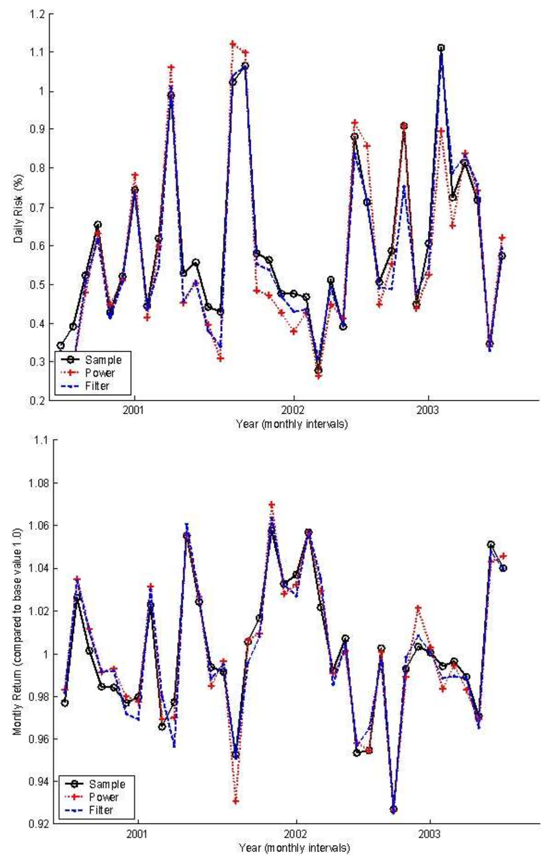

Fig. 5. Daily risk (top) and monthly return (bottom) for the Swedish market data, no short selling allowed: no noise reduction, *i.e.* plain sample (circles), power mapping (crosses) and filtering (dashes).

4. Conclusions

We applied the power mapping to our market data as a *black box*, with the parameter $q = 1.5$ and no further adjustments whatsoever. Given this simplicity, we are very pleased with the results. The risk is reduced in exactly the same way as for the filtering. The return is worse in one, but better in another case. We are now experimenting with modifications of the power mapping and with hybrid methods, including variations of q. Some very recent findings [12] seem to make this even more interesting.

The two noise reduction methods are based on completely different ideas. Thus, it is not surprising that also the results produced are different. Nevertheless, comparing the curves in Figs. 4 and 5, one sees that power mapping and filtering do rather similar things, more similar than what one would naively have expected. At present, the theoretical understanding of the power mapping is not deep enough to explain this in detail. Work is in progress.

TG thanks Marc Potters for fruitful discussion and helpful comments on the results presented in this contribution. We acknowledge financial support from Det Svenska Vetenskapsrådet.

REFERENCES

[1] F. Black, R. Litterman, *Financial Analysts J.* **48**, 28 (1992); E.J. Elton, M.J. Gruber, *Modern Portfolio Theory and Investment Analysis*, Wiley, New York 1995.

[2] L. Laloux, P. Cizeau, J.P. Bouchaud, M. Potters *Phys. Rev. Lett.* **83**, 1467 (1999).

[3] V. Plerou, P. Gopikrishnan, B. Rosenow, L.A.N. Amaral, H.E. Stanley, *Phys. Rev. Lett.* **83**, 1471 (1999).

[4] J.P. Bouchaud, M. Potters, *Theory of Financial Risks*, Cambridge University Press, Cambridge 2000.

[5] T. Guhr, A. Müller-Groeling, H.A. Weidenmüller, *Phys. Rep.* **299**, 189 (1998).

[6] J.J.M. Verbaarschot, T. Wettig, *Ann. Rev. Nucl. Part. Sci.* **50**, 343 (2000).

[7] T. Guhr, B. Kälber, *J. Phys. A* **36**, 3009 (2003).

[8] S. Pafka, I. Kondor, *Physica A* **343**, 623 (2004); S. Pafka, I. Kondor, *Physica A* **319**, 487 (2003).

[9] M. Tumminello, T. Aste, T. Di Matteo, R.N. Mantegna, `cond-mat/0501335`.

[10] P. Gopikrishnan, B. Rosenow, V. Plerou, H.E. Stanley *Phys. Rev.* **E64**, 035106(R) (2001).

[11] V. Plerou, P. Gopikrishnan, B. Rosenow, L.A.N. Amaral, T. Guhr, H.E. Stanley, *Phys. Rev.* **E65**, 066126 (2002).

[12] M. Potters, private communication, 2005.

[13] `www.stockholmsborsen.se` (2003), web download.