Available online at www.sciencedirect.com
SCIENCE DIRECT®
Physica A 319 (2003) 487–494
PHYSICA A
www.elsevier.com/locate/physa

Noisy covariance matrices and portfolio optimization II

Szilárd Pafka[a,*], Imre Kondor[a,b]

[a] *Department of Physics of Complex Systems, Eötvös University, Pázmány P. sétány 1/a, H-1117 Budapest, Hungary*
[b] *Market Risk Research Department, Raiffeisen Bank, Akadémia u. 6, H-1054 Budapest, Hungary*

Received 29 July 2002

Abstract

Recent studies inspired by results from random matrix theory (Galluccio et al.: Physica A 259 (1998) 449; Laloux et al.: Phys. Rev. Lett. 83 (1999) 1467; Risk 12 (3) (1999) 69; Plerou et al.: Phys. Rev. Lett. 83 (1999) 1471) found that covariance matrices determined from empirical financial time series appear to contain such a high amount of noise that their structure can essentially be regarded as random. This seems, however, to be in contradiction with the fundamental role played by covariance matrices in finance, which constitute the pillars of modern investment theory and have also gained industry-wide applications in risk management. Our paper is an attempt to resolve this embarrassing paradox. The key observation is that the effect of noise strongly depends on the ratio $r = n/T$, where n is the size of the portfolio and T the length of the available time series. On the basis of numerical experiments and analytic results for some toy portfolio models we show that for relatively large values of r (e.g. 0.6) noise does, indeed, have the pronounced effect suggested by Galluccio et al. (1998), Laloux et al. (1999) and Plerou et al. (1999) and illustrated later by Laloux et al. (Int. J. Theor. Appl. Finance 3 (2000) 391), Plerou et al. (Phys. Rev. E, e-print cond-mat/0108023) and Rosenow et al. (Europhys. Lett., e-print cond-mat/0111537) in a portfolio optimization context, while for smaller r (around 0.2 or below), the error due to noise drops to acceptable levels. Since the length of available time series is for obvious reasons limited in any practical application, any bound imposed on the noise-induced error translates into a bound on the size of the portfolio. In a related set of experiments we find that the effect of noise depends also on whether the problem arises in asset allocation or in a risk measurement context: if covariance matrices are

* Corresponding author.
E-mail addresses: syl@complex.elte.hu (S. Pafka), ikondor@raiffeisen.hu (I. Kondor).

PII: S0378-4371(02)01499-1

used simply for measuring the risk of portfolios with a fixed composition rather than as inputs to optimization, the effect of noise on the measured risk may become very small.

PACS: 87.23.Ge; 05.45.Tp; 05.40.−a

Keywords: Noisy covariance matrices; Random matrix theory; Portfolio optimization; Risk management

1. Introduction

Covariance matrices of financial returns play a crucial role in several branches of finance such as investment theory, capital allocation or risk management. For example, these matrices are the key input parameters to Markowitz's classical portfolio optimization problem [6], which aims at providing a recipe for the composition of a portfolio of assets such that risk (quantified by the standard deviation of the portfolio's return) is minimized for a given level of expected return. For any practical use of the theory it would therefore be necessary to have reliable estimates for the volatilities and correlations of the returns on the assets making up the portfolio (i.e., for the elements of the covariance matrix), which are usually obtained from historical return series. However, the finite length T of the empirical time series inevitably leads to the appearance of noise (measurement error) in the covariance matrix estimates. It is clear that this noise becomes stronger and stronger with increasing portfolio size n, until at a certain n one overexploits the available information to such a degree that the positive definiteness of the covariance matrix (and with that the meaning of the whole exercise) is lost.

This long known difficulty has been put into a new light by Galluccio et al. [1], Laloux et al. [2] and Plerou et al. [3] where the problem has been approached from the point of view of random matrix theory. These studies have shown that empirical correlation matrices deduced from financial return series contain such a high amount of noise that, apart from a few large eigenvalues and the corresponding eigenvectors, their structure can essentially be regarded as random. In Ref. [2], e.g., it is reported that about 94% of the spectrum of correlation matrices determined from return series on the S&P 500 stocks can be fitted by that of a random matrix. One wonders how, under such circumstances, covariance matrices can be of any use in finance. Indeed, in Ref. [2] the authors conclude that "Markowitz's portfolio optimization scheme based on a purely historical determination of the correlation matrix is inadequate".

Two subsequent studies [4,5] found that the risk level of optimized portfolios could be improved if prior to optimization one filtered out the lower part of the eigenvalue spectrum of the covariance matrix, thereby removing the noise (at least partially). In both of these studies, portfolios have been optimized by using the covariance matrix extracted from the first half of the available empirical sample, while risk was measured as the standard deviation of the return on these portfolios in the second half of the sample. Laloux et al. [4] and Plerou et al. and Rosenow et al. [5] found a significant discrepancy between "predicted" risk (as given by the standard deviation of the optimal portfolio in the first half of the sample) and "realized" risk (given by its