Accepted Manuscript

Cleaning large correlation matrices: Tools from random matrix theory

Joël Bun, Jean-Philippe Bouchaud, Marc Potters

PII: S0370-1573(16)30333-7
DOI: http://dx.doi.org/10.1016/j.physrep.2016.10.005
Reference: PLREP 1931

To appear in: *Physics Reports*

Accepted date: 24 October 2016

Please cite this article as: J. Bun, J.-P. Bouchaud, M. Potters, Cleaning large correlation matrices: Tools from random matrix theory, *Physics Reports* (2016), http://dx.doi.org/10.1016/j.physrep.2016.10.005

Cleaning large Correlation Matrices: tools from Random Matrix Theory

Joël Bun

Capital Fund Management, 23–25, rue de l'Université, 75 007 Paris

LPTMS, CNRS, Univ. Paris-Sud, Université Paris-Saclay, 91405 Orsay, France

Jean-Philippe Bouchaud, Marc Potters

Capital Fund Management, 23–25, rue de l'Université, 75 007 Paris

Abstract

This review covers recent results concerning the estimation of large covariance matrices using tools from Random Matrix Theory (RMT). We introduce several RMT methods and analytical techniques, such as the Replica formalism and Free Probability, with an emphasis on the Marčenko-Pastur equation that provides information on the resolvent of multiplicatively corrupted noisy matrices. Special care is devoted to the statistics of the eigenvectors of the empirical correlation matrix, which turn out to be crucial for many applications. We show in particular how these results can be used to build consistent "Rotationally Invariant" estimators (RIE) for large correlation matrices when there is no prior on the structure of the underlying process. The last part of this review is dedicated to some real-world applications within financial markets as a case in point. We establish empirically the efficacy of the RIE framework, which is found to be superior in this case to all previously proposed methods. The case of additively (rather than multiplicatively) corrupted noisy matrices is also dealt with in a special Appendix. Several open problems and interesting technical developments are discussed throughout the paper.

Keywords: Random matrix theory, High dimensional statistics, Correlation matrix, Spectral decomposition, Rotational invariant estimator

Contents

Email addresses: joel.bun@gmail.com (Joël Bun), jean-philippe.bouchaud@cfm.fr (Jean-Philippe Bouchaud), marc.potters@cfm.fr (Marc Potters)