UNIVERSITÄT BIELEFELD

Faculty of Physics
Faculty of Mathematics

THE UNIVERSITY OF MELBOURNE

Seminar

Bielefeld - Melbourne Random Matrices

Johannes Heiny

Ruhr Universität Bochum

Recent advances in large sample correlation matrices and their applications

Many fields of modern sciences are faced with high-dimensional data sets. In this talk, we investigate the spectral properties of large sample correlation matrices. First, we consider a p-dimensional population with iid coordinates in the domain of attraction of a stable distribution with index $\alpha\in (0,2)$. Since the variance is infinite, the sample covariance matrix based on a sample of size n from the population is not well behaved and it is of interest to use instead the sample correlation matrix R. We find the limiting distributions of the eigenvalues of R when both the dimension p and the sample size n grow to infinity such that $p/n\to \gamma$. The moments of the limiting distributions $H_{\alpha,\gamma}$ are fully identified as the sum of two contributions: the first from the classical Marchenko-Pastur law and a second due to heavy tails. Moreover, the family $\{H_{\alpha,\gamma}\}$ has continuous extensions at the boundaries $\alpha=2$ and $\alpha=0$ leading to the Marchenko-Pastur law and a modified Poisson distribution, respectively. A simulation study on these limiting distributions is also provided for comparison with the Marchenko-Pastur law.

In the second part of this talk, we assume that the coordinates of the p-dimensional population are dependent and $p/n \le 1$. Under a finite fourth moment condition on the entries we find that the log determinant of the sample correlation matrix R satisfies a central limit theorem. In the iid case, it turns out the central limit theorem holds as long as the coordinates are in the domain of attraction of a stable distribution with index $\alpha>3$, from which we conjecture a promising and robust test statistic for heavy-tailed high-dimensional data. The findings are applied to independence testing and to the volume of random simplices.

Reference:
Limiting distributions for eigenvalues of sample correlation matrices from heavy-tailed populations

Wednesday, 05 May 2021, 0900 hrs CEST

Zoom Conference Call— Please contact Anas Rahman (anas.rahman@live.com.au) for details regarding access

SFB 1283

→ www.physik.uni-bielefeld.de