



Testing for Independence of Large Dimensional Vectors

August 2017

TARAS BODNAR^a, HOLGER DETTE^b AND NESTOR PAROLYA^c

^a *Department of Mathematics, Stockholm University, SE-10691 Stockholm, Sweden*

e-mail: taras.bodnar@math.su.se

^b *Department of Mathematics, Ruhr University Bochum, D-44870 Bochum, Germany*

e-mail: holger.dette@ruhr-uni-bochum.de

^c *Institute of Statistics, Leibniz University Hannover, D-30167 Hannover, Germany*

e-mail: parolya@statistik.uni-hannover.de

Abstract

In this paper new tests for the independence of two high-dimensional vectors are investigated. We consider the case where the dimension of the vectors increases with the sample size and propose multivariate analysis of variance-type statistics for the hypothesis of a block diagonal covariance matrix. The asymptotic properties of the new test statistics are investigated under the null hypothesis and the alternative hypothesis using random matrix theory. For this purpose we study the weak convergence of linear spectral statistics of central and (conditionally) non-central Fisher matrices. In particular, a central limit theorem for linear spectral statistics of large dimensional (conditionally) non-central Fisher matrices is derived which is then used to analyse the power of the tests under the alternative.

The theoretical results are illustrated by means of a simulation study where we also compare the new tests with several alternative, in particular with the commonly used corrected likelihood ratio test. It is demonstrated that the latter test does not keep its nominal level, if the dimension of one sub-vector is relatively small compared to the dimension of the other sub-vector. On the other hand the tests proposed in this paper provide a reasonable approximation of the nominal level in such situations. Moreover, we observe that one of the proposed tests is most powerful under a variety of correlation scenarios.

AMS 2010 subject classifications: 60B20, 60F05, 62H15, 62H20, 62F05

Keywords: testing for independence, large dimensional covariance matrix, non-central Fisher random matrix, linear spectral statistics, asymptotic normality