Spectral statistics of large dimensional Spearman's rank correlation matrix and its application

Liang-Ching Lin*
National Sun Yat-sen University

Zhigang Bao, Guangming Pan
Nanyang Technological University

Wang Zhou
National University of Singapore

Abstract

Let $\mathbf{Q} = (Q_1, \ldots, Q_n)$ be a random vector drawn from the uniform distribution on the set of all $n!$ permutations of $\{1, 2, \ldots, n\}$. Let $\mathbf{Z} = (Z_1, \ldots, Z_n)$, where Z_j is the mean zero variance one random variable obtained by centralizing and normalizing Q_j, $j = 1, \ldots, n$. Assume that $\mathbf{X}_i, i = 1, \ldots, p$ are i.i.d. copies of $\frac{1}{\sqrt{p}}\mathbf{Z}$ and $X = X_{p,n}$ is the $p \times n$ random matrix with $\mathbf{X}_i$ as its i-th row. Then $S_n = XX^*$ is called the $p \times n$ Spearman's rank correlation matrix which can be regarded as a high dimensional extension of the classical non-parametric statistic Spearman's rank correlation coefficient between two independent random variables. In this paper we will establish a CLT for the linear spectral statistics of this non-parametric random matrix model in the scenario of high dimension supposing that $p = p(n)$ and $p/n \to c \in (0, \infty)$ as $n \to \infty$. Relying on this CLT we then construct a distribution-free statistic to test complete independence for components of random vectors. Owing to the non-parametric property, we can use this test on generally distributed random variables including the heavy-tailed ones.

Keywords: Spearman's rank correlation matrix, linear eigenvalue statistics, central limit theorem.