Statistica Sinica **35** (2025), 1-23
doi:https://doi.org/10.5705/ss.202022.0377

INVARIANCE PRINCIPLE AND CLT FOR THE SPIKED EIGENVALUES OF LARGE-DIMENSIONAL FISHER MATRICES AND APPLICATIONS

Dandan Jiang, Zhiqiang Hou, Zhidong Bai and Runze Li*

Xi'an Jiaotong University, Shandong University of Finance and Economics, Northeast Normal University and Pennsylvania State University at University Park

Abstract: This paper aims to derive the asymptotic distributions of the spiked eigenvalues of large-dimensional spiked Fisher matrices, without imposing Gaussian assumptions or restrictive assumptions on covariance matrices. We first establish an invariance principle for the spiked eigenvalues of the Fisher matrix. That is, we show that the limiting distributions of the spiked eigenvalues are invariant over a broad range of population distributions satisfying certain conditions. Utilizing this invariance principle, we establish a central limit theorem (CLT) for the spiked eigenvalues, and further explore some interesting applications by using the CLT to derive the power functions of the Roy Maximum Root test for linear hypotheses in linear models, as well as the test in signal detection. To evaluate the effectiveness of the newly proposed test, we conduct Monte Carlo simulation studies and compare its performance with existing tests.

Key words and phrases: Random matrix theory, Roy Maximum Root test, spiked model, two-sample covariance problem.

1. Introduction

Motivated by several applications of hypothesis on two-sample covariance matrices and linear hypothesis on regression coefficient matrix in linear models, we consider the following two-sample spiked model. Let $\mathbf{\Sigma}_1$ and $\mathbf{\Sigma}_2$ be the covariance matrices from two p-dimensional populations, and let $\mathbf{S}_1$ and $\mathbf{S}_2$ be the corresponding sample covariance matrices with sample sizes n_1 and n_2. The two-sample spiked model assumes that $\mathbf{\Sigma}_2 = \mathbf{\Sigma}_1 + \mathbf{\Delta}$, where $\mathbf{\Delta}$ is a $p \times p$ matrix of finite rank M. It is of great interest to study statistical inference on the spikes, including, but not limited to, testing the presence of the spikes, testing the number of the spikes, and calculating the power under the alternative hypothesis in two-sample testing problems. Thus, it is critical to establish the asymptotic properties of the spiked eigenvalues of a Fisher matrix $\mathbf{F} = \mathbf{S}_1\mathbf{S}_2^{-1}$. It is of particular interest to derive the asymptotic distribution of $\lambda_{\max}(\mathbf{F})$, the largest eigenvalue of $\mathbf{F}$. However, the existing related works are limited due to imposed

*Corresponding author. E-mail: rzli@psu.edu