Testing homogeneity of high-dimensional covariance matrices

Abstract

Testing homogeneity of multiple high-dimensional covariance matrices is becoming more critical in multivariate statistical analysis owing to the emergence of big data. Many existing homogeneity tests for high-dimensional covariance matrices mainly focus on two populations, and they often target at some specific situations, for example, either sparse alternatives or dense alternatives, thus the available methods are not suitable for general cases with multiple groups. To accommodate various situations, we propose a power-enhancement high-dimensional test for multi-sample comparisons of covariance matrices, which includes the homogeneity testing of two matrices as a special case. Not only do the proposed tests require no distributional assumption, but they can also handle both sparsity and non-sparsity structures. Based on random matrix theories, the asymptotic normality properties of our tests are established under both the null and alternative hypotheses. Numerical studies demonstrate substantial gain in power for our proposal, and the new method is illustrated with a gene expression dataset from the breast cancer study.