Online High Dimensional Covariance Change Point Detection

Clifford Lam

London School of Economics and Political Science
E-mail: C.Lam2@lse.ac.uk

Abstract: Detecting changes in the covariance structure in time series data is an important problem which finds applications in, e.g., clustering, identification of disease related genes in bioinformatics, or portfolio and risk management in finance, to name but a few areas. In high dimensional time series data where the dimension can be larger than the sample size, this problem can be extremely difficult. For one, the typical statistic for change detection involves the calculation of sample covariance matrix of two different parts of data, which can be affected by the poor performance of sample covariance matrix under high dimensional setting. More importantly, as far as we know there are no covariance detection methods so far that can have very imbalanced sample sizes for the two parts of data involved, which unfortunately is exactly the setting for online covariance detection, where one part can have good number of data points but the "newer" part of the data may have only finite number of data points.

We propose a series of statistics which can be powerful in covariance change detection under the online setting, with the dimension of each observation vector grows together with or even faster than the sample size, while the "newer" part of the data under suspicion of change can have only finite sample size. Asymptotic normality of these statistics under both no change and change scenarios are proved and demonstrated, while modifications for variance reduction and power boosting are also illustrated with numerical examples. Incorporation of thee into full change points detection algorithm is also discussed.