

## Control of the False Discovery Rate by Covariance Dependence Thresholding

Jianqing Fan, Princeton University

Multiple hypothesis testing is fundamental in high dimensional inference, with wide applications in scientific fields. In genome-wide association study, tens of thousands of tests are performed simultaneously to find if any genes are associated with some traits and those tests are correlated.

When test statistics are correlated, false discovery control becomes very challenging under arbitrarily dependent. In the current paper, we propose a new method based on principal factor approximations, which successfully substracts the common dependence and weakens significantly the correlation structure, to deal with an arbitrary dependence structure. We derive the theoretical distribution for false discovery proportion (FDP) in large scale multiple testing when a common threshold is used and provide a consistent FDP.

This result has interesting applications in controlling FDR and FDP. Our estimate of FDP compares favorably with Efron (2007)'s approach, as demonstrated by in the simulated examples. Our approach is further illustrated by some real data applications. (Joint work with Xu Han and Weijie Gu)