Joint Statistics/ORIE Seminar

Wednesday, February 3, 2016
4:15 pm
G01 Biotechnology

Weichen Wang
Princeton University

Semiparametric Factor Models and Projected Principal Component Analysis

Factor analysis is one of the most useful tools for modeling common dependence among multivariate outputs. We propose a flexible semiparametric factor model, which decomposes the factor loading matrix into the component that can be explained by subject-specific covariates and the orthogonal residual component. The method Projected Principal Component Analysis was introduced to recover the latent structure. We show that the unobserved factors and the smooth factor loading matrices can be more accurately estimated than the conventional PCA if the projection is genuine. The convergence is achieved even when the sample size is finite. This leads us to develop nonparametric tests on whether observed covariates have explaining powers on the loadings. The proposed method is illustrated by both simulated data and the returns of the S&P 500 constituents.

Furthermore, as one application of the semiparametric model, we consider the problem of heterogeneity adjustment when aggregating datasets from multiple sources. We propose a generic framework named ALPHA (Adaptive Low-rank Principal Heterogeneity Adjustment) to model, estimate, and adjust heterogeneity from the original data. As an illustrative application of this generic framework, we conduct graphical model inference for a brain imaging network based on multiple datasets.

Refreshments will be served after the seminar in 1181 Comstock Hall.