Mathematics, Statistics, and Computer Science   **@ UIC**

Statistics and Data Science Seminar

Envelope-based Sparse Partial Least Squares

Zhihua Su (University of Florida)

Abstract: Sparse partial least squares (SPLS) is widely used in applied sciences as a method that performs dimension reduction and variable selection simultaneously in linear regression. Several implementations of SPLS have been derived, among which the SPLS proposed in Chun and Keleş (2010) is very popular and highly cited. However, for all of these implementations, the theoretical properties of SPLS are largely unknown. In this paper, we propose a new version of SPLS, called the envelope-based SPLS, using a connection between envelope models and partial least squares (PLS). We establish the consistency, oracle property and asymptotic normality of the envelope-based SPLS estimator. The large-sample scenario and high-dimensional scenario are both considered. We also develop the envelope-based SPLS estimators under the context of generalized linear models, and discuss its theoretical properties including consistency, oracle property and asymptotic distribution. Numerical experiments and examples show that the envelope-based SPLS estimator has better variable selection and prediction performance over the existing SPLS estimators.

Wednesday, September 4 at 4:00 PM in 636 SEO