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

Statistics and Data Science Seminar

Leveraging Algorithms for Logistic Regression with Massive Data

Haiying Wang (University of New Hampshire)

Abstract: For massive data with super-large sample size n, it is computationally infeasible to obtain maximum likelihood estimates for unknown parameters, especially when the estimator does not have a close-form solution. This paper proposes fast leveraging algorithms to efficiently approximate the maximum likelihood estimates of unknown parameters in logistic regression models with binary responses, one of the most commonly used models in practice for classification. We theoretically prove the consistency of the leveraging algorithms, develop nearly optimal two-step leveraging strategies, and evaluate the performance of the proposed methods using synthetic and real data sets.

Wednesday, February 18 at 4:00 PM in SEO 636