UNC
GILLINGS SCHOOL OF
GLOBAL PUBLIC HEALTH

Department of Biostatistics

Bernard G. Greenberg Distinguished Lecture Series

Presented by

Trevor Hastie, PhD
Professor of Statistics
Stanford University

Matrix Completion and Large-Scale SVD Computations

ABSTRACT

The Singular Value Decomposition (SVD) is a fundamental tool in all branches of data analysis - arguably one of the most widely used numerical tools. Over the last few years, partly inspired by the "Netflix problem", the SVD has again come into focus as a solution to the "matrix completion" problem. One partially observes a very large matrix, and would like to impute the values not observed. By assuming a low-rank structure, the SVD is one approach to the problem - a SVD with large amounts of missing data. In this talk we discuss an approach for building a path of solutions of increasing rank via nuclear-norm regularization. An integral part of this algorithm involves repeatedly computing low-rank SVDs of imputed matrices. We show how these tasks can be efficiently handled, and how to exploit distributed architectures, allowing the method to scale to very high-dimensional problems.

*joint work with Rahaul Mazumder

Place: Blue Cross Blue Shield Auditorium
Michael Hooker Research Center
Date: Wednesday, May 8, 2013
Time: 3:00 – 4:00 PM