RUTGERS UNIVERSITY
DEPARTMENT OF STATISTICS AND BIOSTATISTICS
HILL CENTER #501, BUSCH CAMPUS, PISCATAWAY

[www.stat.rutgers.edu](http://www.stat.rutgers.edu)

### Seminar

**Speaker:**   Hyonho Chun, Yale University

**Title:**   Spare Partial Least Squares Regression for Simultaneous Dimension Reduction and Variable Selection with Applications to High Dimensional Genomic Data

**Date:**   Friday, January 29, 2010

**Time:**   12 Noon

**Place:**   552 Hill Center

### Abstract

Partial least squares (PLS) regression has been an alternative to ordinary least squares for handling multicollinearity in several areas of scientific research since 1960s.  It has recently gained much attention in the analysis of high dimensional genomic data.  In this talk, I will present that the asymptotic consistency of the PLS estimator does not hold with the very large p and small n paradigm, and then I will propose a spares partial least squares (SPLS) formulation which aims to simultaneously achieve good predictive performance and variable selection by producing sparse linear combinations of the original predictors.  An efficient implementation of the SPLS regression will be provided and the SPLS regression will be compared to well known variable selection and dimension reduction approaches via simulation experiments.  In the second part of the talk, I will illustrate the advantages of SPLS regression with applications to genomic data including expression quantitative trait loci (eQTL) mapping and multi-marker association study.