

Colloquium

Mahlet Tadesse

Georgetown University

Friday, October 7, 2011 at 3 pm in Bell Hall 143

A stochastic partitioning method to associate high-dimensional datasets

Several procedures have been developed to associate high-dimensional covariate data to univariate outcomes. In recent years, there has been a growing interest in relating data sets in which both the number of regressors and response variables are substantially larger than the sample size. For example, in an attempt to gain new insights into molecular processes, many efforts are being carried out to integrate data from various high-throughput genomic experiments. We propose a Bayesian stochastic partitioning method to identify sets of covariates associated with correlated outcomes. The procedure provides a unified framework for uncovering response variables with similar expression profiles and determining subsets of predictors that modulate these patterns. We illustrate the method with applications to genomic data sets.