Annual Report 2

November 8th, 2012

Atul Mallik

Title: Adaptive threshold estimation using multi-scale tests

Abstract: We propose an upper confidence bound for the threshold at which a regression function takes off from its baseline value in a white noise model. The confidence bound is based on multi-scale test statistics which involve combining one-sided kernel estimates with varying bandwidths in a specific manner. The bound is adaptive and near-optimal in the sense that it does require knowing the smoothness of the underlying regression function in the vicinity of the threshold and attains the minimax convergence rate up to a factor of $\log(n)$, n being the scale parameter or the 'sample size'. Further, we extend our approach to Gaussian regression framework with a uniform fixed design.

Shawn Mankad

Title: Matrix Factorization Models for Dynamic Network and Three-Way Data

Abstract: Collections of data matrices (data arrays) are becoming increasingly available due to advances in acquisition technologies. Various matrix factorization techniques have been applied to such data for integrating them and extraction of common patterns. However, a direct application of such techniques to such structured data can be adversely affected by transient patterns, and thus fail to extract parsimonious ones. In this talk, we introduce a model for non-negative matrix factorization that captures how structure evolves through the data. The model decomposes the data into two factors: a basis common to all data matrices, and a coefficient matrix that varies for each data matrix. A regularization is utilized within the framework of non-negative matrix factorization to encourage local smoothness of the coefficient matrix. This improves interpretability and highlights the structural patterns underlying the data, while mitigating noise effects. The proposed methodology is illustrated on paper citation data.

Ashin Mukherjee

Title: Topics in Reduced Rank Regression

Abstract: In the first part we study the degrees of freedom of the reduced rank regression estimator in the framework of Stein's Unbiased Risk Estimation(SURE). We derive an unbiased estimator of the degrees of freedom of the reduced rank regression procedure and show that it is significantly different than the number of free parameters in the model which is often taken as a heuristic estimate of be the degrees of freedom of an estimation procedure.

The second part deals with a non-parametric extension to reduced rank regression under high dimensional setting where many of the predictors might be non-informative. We propose a two step penalized regression approach based on spline approximations that encourages both variable selection and rank reduction. We prove rank selection consistency and also provide error bounds for the proposed method.

Yeo Jung Park

Title: Discovering neural connectivity changes in EEG data

Abstract: A major challenge in neuroscience is uncovering the relationship between consciousness and brain activity. Electroencephalography (EEG) recordings made on human subjects who are given anesthesia for surgery provide an opportunity to directly study this relationship. The main focus in this area has been on changes in the connectivity between brain regions that occur as the consciousness state changes. Connectivity can be assessed in terms of the statistical dependence between EEG measurements from different recording sites on the scalp. Motivated by this scientific application, we consider two approaches for capturing changes in the dependence structure in pairs of time series. We first consider an approach that aims to identify frequencies at which the variation is correlated between the two series. We then consider an approach in which the two series are correlated after being transformed by local linear transformations. We optimize over a class of such transformations to maximize the correlation coefficient, leading to a new measure of dependence for serially observed data. Both of these correlation-based procedures are capable of identifying dependencies that are completely invisible to the simple correlation coefficient, and thus provide a more comprehensive assessment of connectivity and how it changes under different conditions (e.g. consciousness levels).