<u>Jie Cheng</u>

Title: High-dimensional Mixed Graphical Model

Abstract:

In this project we propose a mixed graphical model that allows us to model data sets with both continuous and discrete variables. The proposed model reduces the complexity of a complete conditional Gaussian (CG) density yet maintains its flexibility. It also allows us to develop a computationally efficient regression based algorithm by focusing on the conditional log-likelihood of each variable given the rest. Sparsity is obtained by incorporating a weighted l1 penalty. We demonstrate the effectiveness of the proposed method through an extensive simulation study as well as compare it with some alternative approaches. Our method shows promising results in an analysis of music annotation data set (CAL500) leading to a sparse interpretable graphical structure. We also illustrate that the proposed methodology can readily be extended to the general case of multi-level discrete variables.

<u>Zhen (Adelle) Wang</u>

Title: p-MIF in estimating DSGE model

Abstract:

Iterated ltering is based on a sequence of particle ltering, which could facilitates likelihood-based inference in Dynamic Stochastic General Equilibrium (DSGE) models. Economics produces examples of quantities with low measurement error (e.g., GDP, unemployment rate), which will lead to technical difficulty. Numerous researchers have studied some examples on ltering dynamic economic models. Recent economic turmoil makes reassessment of structural models an urgent problem. Testing new macroeconomic models against recent data is relevant in the context of the recent macroeconomic crisis. We will compare Particle Filter within Markov Chain Monte Carlo (PMCMC) and Iterated Filtering (MIF) in estimating DSGE model using simulated data.

There is a trade-o between numbers of parameter values sampled each ltering and the number of ltering operation needed. PMCMC is at one extreme of this (only 1 new parameter value per ltering operation; thousands of ltering operations needed). Traditional MIF is at the other extreme (1 new parameter per particle per time point, 50 ltering operations needed). We will propose p-MIF (p $(0 < p < 1)$ new parameter per particle per time point on average, 50 ltering operations needed) schemes as an intermediate algorithm between these two very different extremes. This is shown to perform better for this sort of problem than either existing methods. We also will apply p-MIF to re-evaluate DSGE model using US data.

<u>Tom Brown</u>

Title: Problems in Spatial Statistics

Abstract:

Spatial statistics is a field in which data is analyzed through its spatial properties. Common examples include geographic and astronomical data. When attempting to perform spatial smoothing or prediction, a common, but often unsatisfied assumption is that the data is stationary. Methods have been proposed to analyze nonstationarity, including transforming the spatial aspects of the data to obtain a stationary process. When data is abundant, like with satellite imagery, infill asymptotics is a method that can be applied to analyze nonstationary spatial data.

Ming-Chi Hsu

Title: Estimating effect size distribution in high dimensional data

Abstract:

An effect size is an index that measures the magnitude of a statistical relationship. When a large number of related effect sizes are observed, it may be of interest to focus on aspects of their distribution rather than focusing on their individual values. This contrasts with traditional statistical analysis focusing on a small number of effects, where it is more common to consider whether individual effects are significantly different from zero. We plan to develop approaches for studying effect size distributions.

Juan Zhang

Title: Statistical assessment of relationships between marginal features of variables and their external correlations.

Abstract:

Many genomic studies aim to identify characteristics of individuals in a population that are associated with an external trait. Then the external correlations between characteristics and the external trait are calculated and often used to identify a subset of characteristics that meet some level of statistical confidence such as false discovery rate in a traditional "screening analysis". But in this chapter, we focus on approaches for identifying certain global trends in the data that help us to understand what types of associations may be present and what marginal features of characteristics may be relevant to these associations. We illustrate that this can be accomplished even when the power is too low to attribute these trends to specific characteristics.