Speaker: Art Owen, Stanford Statistics

Title: Estimation and Inference for Large-Scale Linear Mixed Effects Models with Crossed Random Effects

Abstract:
Linear mixed models with large imbalanced crossed random effects structures pose severe computational problems for maximum likelihood estimation and for Bayesian analysis. The costs can grow as fast as $N^{3/2}$ when there are $N$ observations. Such problems arise in any setting where the underlying factors satisfy a many to many relationship (instead of a nested one) and in electronic commerce applications, the $N$ can be quite large. Methods that do not account for the correlation structure can greatly underestimate uncertainty. We propose a method of moments approach that takes account of the correlation structure and that can be computed at $O(N)$ cost. The method of moments is very amenable to parallel computation and it does not require parametric distributional assumptions, tuning parameters or convergence diagnostics. For the regression coefficients, we give conditions for consistency and asymptotic normality as well as a consistent variance estimate. For the variance components, we give conditions for consistency and we use consistent estimates of a mildly conservative variance estimate. All of these computations can be done in $O(N)$ work. We illustrate the algorithm with some data from Stitch Fix where the crossed random effects correspond to clients and items.

This is joint work with Katelyn Gao.