SPARSE MODEL STRUCTURES & STOCHASTIC COMPUTATION FOR HIGH-DIMENSIONAL MULTIVARIATE ANALYSIS

Mike West
Duke University, Durham, North Carolina, USA
(e-mail: mw@stat.duke.edu)

Abstract

I will discuss aspects of Bayesian modelling, prior specification and stochastic computation for a range of contexts involving "big" and "sparse" models. The contexts will be drawn from multivariate analysis using high-dimensional graphical models, regression variable model search/uncertainty with many candidate predictors, large-scale multivariate anova models, and large-scale, sparse factor models. Computation using varieties of MCMC methods and non-MCMC evolutionary stochastic search methods will be discussed and explored, as will emerging and practically important questions of Bayesian model/prior specification that are central to scalability to Bayesian technologies as models/parameter spaces increase in dimension. Much of our work across this range of high-dimensional, sparse model structures is at least partly responsive to programs of statistical analysis of in genomics with large and diverse sets of experimental and observational gene expression (and other) data, which will be noted as relevant. Open areas of core statistical research interest, including the need for novel computational methods and associated theory in problems where MCMC is simply too ponderous, and for improved implementations in distributed/cluster processing, will be mentioned.

References:
The work discussed represents collaborations with a number of students, postdocs and other colleagues, including Carlos Carvalho, Chris Hans, Quanli Wang, Adrian Dobra and Beatrix Jones.