useR! 2016 Tutorial Proposal

Tutorial Title

Programming with models: An introduction to NIMBLE, a BUGS-compatible system for fitting and programming with hierarchical statistical models using MCMC and more

Instructor Details

Name: Chris Paciorek

Institution: University of California, Berkeley

Address: Department of Statistics, 367 Evans Hall, Berkeley CA 94720

Email: paciorek@stat.berkeley.edu

Short Instructor Biography

Christopher Paciorek is an associate research statistician, lecturer, and statistical computing consultant in the Department of Statistics at UC Berkeley. He has 15 years of experience in applications of Bayesian statistics with a wide variety of published analyses in environmental and public health applications (http://www.stat.berkeley.edu/~paciorek/cv/cv.html).He has extensive experience developing and presenting workshops on statistical computing topics, including intensive two-day bootcamps on R (https://github.com/berkeley-scf/r-bootcamp-2015) and shorter workshops on parallel programming (https://github.com/berkeley-scf/parallel-scf-2015), using Spark and MapReduce (https://github.com/berkeley-scf/spark-workshop-2014), using GPUs (https://github.com/berkeley-scf/gpu-workshop-2014), and using C/C++ from R http://statistics.berkeley.edu/computing/cpp. He has taught the graduate-level statistical computing class at UC Berkeley for five years (https://github.com/berkeley-stat243/stat243-fall-2015). He is a co-developer of the NIMBLE package (http://R-nimble.org and http://arxiv.org/abs/1505.05093) and the developer of the bigGP (http://www.jstatsoft.org/v63/i10/) and spectralGP packages (http://www.jstatsoft.org/v19/a2).

Brief Description of Tutorial

This tutorial will introduce attendees to the NIMBLE system for programming with hierarchical models in R. The tutorial will first show how to specify a hierarchical statistical model using BUGS syntax and fit that model using MCMC. Attendees will learn how to customize the MCMC for better performance and how to specify one's own statistical distributions and functions for use within the BUGS syntax. Next the tutorial will demonstrate how to use other non-MCMC algorithms to fit and evaluate models. Finally, the tutorial will show how to write algorithms for use with hierarchical models.

Goals

1. Be able to specify hierarchical statistical models using BUGS
2. Be able to fit and assess models using NIMBLE's default MCMC system
3. Be able to customize the MCMC for a model
4. Be able to specify user-provided distributions and functions to extend BUGS

5. Be able to use other algorithms such as MCEM and particle filtering
6. Be able to code one's own algorithms using NIMBLE's flexible system for programming algorithms

Detailed Outline

Detailed outline text.

1. A complete worked example to illustrate the system
2. Specifying models with BUGS and manipulating models
   (a) Writing a model in BUGS
   (b) Creating a NIMBLE model
   (c) Working with the model in R
3. Fitting models with MCMC
   (a) Building a default MCMC
   (b) Using the MCMC output
   (c) Customizing your MCMC: blocking parameters and choosing samplers
4. Using other algorithms provided by NIMBLE
5. Writing your own algorithms using nimbleFunctions
   (a) A basic nimbleFunction for fast execution of R code
   (b) Using your own distributions and functions to extend BUGS
   (c) Writing an algorithm as a nimbleFunction

Justification

Fitting hierarchical models is at the heart of many statistical and machine learning workflows in a wide variety of fields, including statistics, machine learning/robotics/AI, biology (ecology, phylogenetics), social science (economics, education research, psychology, quantitative political science), public health (epidemiology, health services research, global health), and more. BUGS-based software tools (WinBUGS, JAGS) have been very popular but have limitations, specifically in only providing a user a black box MCMC. NIMBLE is much more flexible and opens up a variety of options for those fitting models and those hoping to provide algorithms that others can then use on their own models.

Background Knowledge

Familiarity with R, hierarchical statistical models/Bayesian statistics, and basic understanding of Markov chain Monte Carlo (MCMC)

Expected Number of Attendees

15-30