Comparison of MCMC Algorithms Using Convergence Diagnostics

Aaron N. Baker
Department of Statistics
Iowa State University

Overview

- Basic Motivation/Concept
- Goals
- Example Use
- Output
- Future Work

Basic Motivation/Concept

- Compare specific MCMC methods using useful diagnostic methods for chain convergence
  - ACF plots
    - Lower ACF means less dependence, more valuable variance estimates for chains.
  - Gelman-Rubin (GR) Diagnostics
    - A value less than one suggests chain convergence
  - Geweke Statistics
    - Checks to see if the end and beginning sections of a chain are behaving similarly

Goals

- Improve former work by
  - Reproducability
  - Parallelization
  - Improving former code via good practices (removing inefficiencies, naming)
  - Provide diagnostics for more than any method desired

Reproducibility

- Take functions and turn them into easily executable functions
- Post these to github for reproducibility purposes

Parallelization

- Use DoParallel, Foreach, to run multiple chains at the same time.
- learned that
  - Parallelization can be annoying as all get out
  - Windows sucks
  - To ensure that functions and libraries are explicitly loaded onto each cluster
  - And honestly I just have no idea why but I couldn't get multidplyr to work

Improving Code

- Got rid of ugly for loops
- Annotated the code
- Used better saving spots (github) and naming

Improving Code-Example

- Former Code:
  - for(i in 1 : $length(y)$){
    $if\,(ys[i] * X[i,j] > 0)\{$
    $mincount = mincount + 1$
    $minim[mincount] = -(\eta[i] + X[i,-j]\beta[-j,])/X[i,j]$
    }
    $if\,(ys[i] * X[i,j] < 0)\{$
    $maxcount = maxcount + 1$
    $maxim[maxcount] = -(\eta[i] + X[i,-j]\beta[-j,])/X[i,j]$
    }
    }
- One example of a for loop inside a for loop. Eegads!

Improving Code-Example

- New Code:
  - bound = bounds2(y, β, x, n)
    $phi = rtruncnorm(1, mean = c((x\beta)), sd = 1, a = bound[, 1], b = bound[, 2])$
- Parallelized main loop, and destroyed the inner. For the win!

Example Use and Output

- To the R (studio, of course, who uses base R?)!!

Future Work

- Make sure there aren't any bugs for multiple parameter groups
- Include more diagnostics
- Other bells and whistles and stuff.

Questions?