Rcpp: An Introduction

Adam Peterson

January 15, 2020

Motivation

Problem

- Need to fit a big fancy model with lots of parameters and data.
- How to do it?

Possible Coding Languages

- Fortran, C
- R, Python , ... Go?
- C++

Language Considerations

- "Easy" user interface
  - R or Python
- Library Support / Development Community
  - R , C++ (Statistics)
  - Python (Machine Learning)
- Fast Computation
  - C++ / C / Fortran

Rcpp: A Good Mix of Everything

- "Wrap" functions written in C++ with R code.
- Get speed of a C++ function in an R session
- Downside? You have to learn how to write C++ code

Why is it faster?

	C++	R
Typing	Static	Dynamic
Compilation	Pre-RunTime	At Runtime (interpreted)
Call by reference	Can do	Can not (without oo package)

Demo

Suppose we have

$$\boldsymbol{X}_i = \begin{bmatrix} X_{i1} \\ X_{i2} \end{bmatrix} \overset{iid}{\sim} \mathcal{MVN}_2(\boldsymbol{\mu}, \Sigma) \quad i = 1, .., n$$

and want:

$$p(\boldsymbol{\mu}|\boldsymbol{X}, \Sigma) \propto p(\boldsymbol{X}|\boldsymbol{\mu}, \Sigma)p(\boldsymbol{\mu})$$

A simple use case:

Let's write a simple Rcpp function to sample from this posterior and walk through its syntax

Demo (cont'd)

How to sample from the posterior?

- Conjugate priors - closed form distribution
- Metropolis Hastings (MH) Sampler
  - Classical
  - Gibbs Sampler
  - Hamiltonian Monte Carlo

Demo (cont'd)

How to sample from the posterior?

- Conjugate priors - closed form distribution
- Metropolis Hastings (MH) Sampler
  - Classical
  - Gibbs Sampler
  - Hamiltonian Monte Carlo

Gibbs Sampler

A MH sampler with acceptance probability 1, the Gibbs sampler proceeds by drawing samples from the conditional distribution of each parameter.

Gibbs Sampler Math

$$p(\boldsymbol{\mu}) \propto 1 \qquad \text{(use an improper prior)}$$

$$\Rightarrow \boldsymbol{\mu}|\boldsymbol{X} \sim \text{MVN}(\bar{\boldsymbol{x}}, \Sigma/n)$$

$$\Rightarrow \mu_1|\mu_2, \boldsymbol{x} \sim N(\bar{x}_1 + \Sigma_{12}\Sigma_{22}^{-1}(\mu_2 - \bar{x}_2), \frac{1}{n}(\Sigma_{11} - \Sigma_{12}\Sigma_{22}^{-1}\Sigma_{21}))$$

$$\mu_2|\mu_1, \boldsymbol{X} \sim N(\bar{x}_2 + \Sigma_{12}\Sigma_{11}^{-1}(\mu_1 - \bar{x}_1), \frac{1}{n}(\Sigma_{22} - \Sigma_{12}\Sigma_{11}^{-1}\Sigma_{12}))$$

Go to github.com/Biostatistics4SocialImpact/dragonsfirstrcpp

Resources

- Advanced R: Chapter 25
- Rcpp Homepage
- Rcpp for Everyone
- Debugging with Rcpp: Dirk's Notes
- Debuggin Rcpp blog post
- C++ syntax and annotations
- RcppEigen Documentation/Tutorial
- RcppArmadillo Documentation/Tutorial
- bendr - an Rcpp R package using functional programming
- rstap2 - an Rcpp R Package using OOP and Templates