Laboratory 5: MCMC and SLR

Introduction

We will make use of some simple coin – die simulations to motivate the MCMC algorithm. The simulations will begin with tactile examples, move to R functions and finally to JAGS using the package rjags in order to form the posterior estimates of the parameters of a SLR problem.

We will also compare the results with classical least squares regression.

This is a seminal lab and will need to be completely mastered.

Objectives:

1. Learn how to carry out 2 state MCMC simulations with a coin and die.
2. Do the same with R functions and learn how to predict deterministic parts of the algorithm.
3. Make transition diagrams and fill out probabilities
4. Create transition matrix and find stationary distribution.
5. Discover Markov chain properties of MCMC chains.
6. Learn about the GIBBS sampler – make a function that will carry out GIBBS sampling for a two parameter density.

Each lab has at least one file to download from CANVAS. Sometimes I will include a second R file (not this time).

Create an R file in RStudio that is well hash commented. Call it Lab5.R

Complete the lab by creating an RMarkdown document. All code needed to answer the questions should be put in r chunks and all mathematical equations should be put into Latex using $$ inline or mainline $$ $$.

The document should read so that all parts connect with the questions and objectives of the lab.

Please note that some questions are open ended "improve the plots" etc – this means that you can be creative and employ more sophisticated packages to make new and better plots and output – all plots will need to be interpreted in the mark down document. Do not "make" and NOT understand!!

You will need OpenBUGS installed in your computer – http://www.openbugs.net/w/Downloads

If you have a MAC then you will need an emulator http://www.davideagle.org/r-2/bayesian-modeling-using-winbugs-and-openbugs/running-openbugs-on-mac-using-wine

Task 1: Make coin-die output using an R function

1. Use the function coin die Bayes' box `cdbbox()` to make some useful output for coin die simulation.
    a. Suppose we wish to make a prior for a two state Bayes' box that corresponds to an acceptance set that has 2 values in it, x=4, n=10 in a Binomial experiment. The parameter values are 0.4 and 0.8.
        i. Place the plot here:
        ii. Place the output matrix here:
        iii. What would be a suitable acceptance set for going from high to low h values?
    b. Take the function `cdbbox()` and improve the graphics in some way. Call the same function as above and place the new graphic here:
2. Derive the result shown in the code snippet of `cdbbox()` put the derivation in your R markdown document using Latex.

```
ifelse(h1=="s",pi1<-k/6*lik2/(lik1+k/6*lik2), pi1<-lik2/(lik2+k/6*lik1))
```

Task 2: Make coin-die simulations in R and interpret them

1. Use the function `coindie()` to make a number of iterations.
    a. Use n=10,h=c(0.6,0.4),E2=c(2,3,4,5) to make some MCMC output.
    b. Paste the above simulation output here:
    c. Improve the graphics in some way and say what you did!
2. Use the output of `cdbbox()` as inputs to the `coindie()` function that you altered – use any examples you wish – explain the input and output.

Task 3: Make a simulation with any number of discrete theta values.

1. In the context of the function `simR()` explain the code snippet

```
u[i]=runif(1)
if(u[i]<=alpha[i]){post[i]<-prop[i]}
else{post[i]<-post[i-1]}
}
```

2. Using a uniform prior and 40 values of theta, x=4, n=10 binomial experiment create a simulated posterior histogram – place here using Rmd:
3. Improve the graphical output by editing the function – place your new graphic here using Rmd:

Task 4: Use different proposals

1. Use simRQ() to trial different proposals
2. Make a proposal that is peaked in the middle with say 11 values.
3. x=4, n=10 as before, prior uniform.
4. Show the first 20 iterations.
5. Improve the plot in the function.
6. Make sure the plot will appear in the knitted documents

Task 5: Make simulations from a continuous parameter with any proposal.

1. We will use the function `simRC()`
2. Improve the function so that it will make informative plots containing the proposal, prior, likelihood and posterior (exact and simulated).
3. Use your function to produce plots for the case where a uniform prior is used and a alpha=3, beta =4 proposal with x=4,n=10 Binomial experiment and theta continuous.
4. Make sure the plot will appear in the knitted documents

Task 6: Use JAGS to carry out a Gibbs sampler for SLR.

1. Describe what Gibbs sampling is and give the algorithm – see https://towardsdatascience.com/gibbs-sampling-explained-b271f332ed8d

2. Now use OpenBUGS make a doodle for a SLR. You can use the model $y_i = \beta_0 + \beta_1 x_i + \epsilon_i$ where $\epsilon_i \sim N(0, \sigma^2)$.
3. Place into Rmd
4. Once the model is made you could use pretty print and insert the code into the exemplar code file "Jags-ExampleScript.R" found in JK's folder of scripts.
5. Use SPRUCE.csv **`Height Vs BHDiameter.`**
6. What are your point and interval estimates?
    a. Diagnose the chains (should use 3 chains) – choose shrinkage.
    b. Is there evidence that they have converged to stationarity?
    c. Give trace and history plots.
7. Compare with classical tests by using the linear model function `lm()`
8. Now fit model y ~ x + I(x^2) use a Bayesian and classical analysis.
9. Compare results!!

MATH 4803/5803 SPECIAL TOPIC: Bayesian Statistics: