Lab 7 STA 360

Rebecca C. Steorts

Agenda

In this lab, we will deriving conditional distributions, code a Gibbs sampler, and analyze the output of the Gibbs sampler.

Consider the following Exponential model for observation(s) $\mathbf{x} = (\mathbf{x_1}, \ldots, \mathbf{x_n})$.[1]:

$$p(x|a, b) = ab\exp(-abx)I(x > 0),$$

where the x_i are assumed to be iid for $i = 1, \ldots n$. and suppose the prior is

$$p(a, b) = \exp(-a - b)I(a, b > 0).$$

You want to sample from the posterior $p(a, b|x_{1:n})$. You may assume that

$$a = 0.25, b = 0.25$$

when coding up your Gibbs sampler.

1. Find the conditional distributions needed for implementing a Gibbs sampler.
2. Code up your own Gibbs sampler using part (1).
3. Run the Gibbs sampler, providing convergence diagnostics.
4. Plot a histogram or a density estimate of the estimated posterior using (2) and (3).
5. How do you know that your estimated posterior in (3) is reliable? **If your traceplot for example or running average plot is not "flattening out," then try running more Gibbs iterations.**

Task 1

Consider the following Exponential model for observation(s) $x = (x_1, \ldots, x_n)$.[2]:

$$p(x|a, b) = ab\exp(-abx)I(x > 0)$$

and suppose the prior is

$$p(a, b) = \exp(-a - b)I(a, b > 0).$$

You want to sample from the posterior $p(a, b|x)$.

It is easy to show that the posterior distribution is intractable, hence, we derive the conditional distributions:

$$\begin{aligned}
p(\boldsymbol{x}|a, b) &= \prod_{i=1}^{n} p(x_i|a, b) \\
&= \prod_{i=1}^{n} ab\exp(-abx_i) \\
&= (ab)^n \exp\left(-ab\sum_{i=1}^{n} x_i\right).
\end{aligned}$$

[1] Please note that the observed data can be found in data-exponential.csv.
[2] The observed data can be found in data-exponential.csv.

The function is symmetric for a and b, so we only need to derive $p(a|\boldsymbol{x}, b)$.

This conditional distribution satisfies

$$
\begin{aligned}
p(a|\boldsymbol{x}, b) &\propto_a p(a, b, \boldsymbol{x}) \\
&= p(\boldsymbol{x}|a, b)p(a, b) \\
&= (ab)^n \exp\left(-ab\sum_{i=1}^{n} x_i\right) \times \exp(-a-b) I(a, b > 0) \\
&\underset{a}{\propto} p(x, a, b) \underset{a}{\propto} a^n \exp(-abn\bar{x} - a)\mathbb{1}(a > 0) = a^{n+1-1}\exp(-(bn\bar{x}+1)a)\mathbb{1}(a > 0) \underset{a}{\propto} \mathrm{Gamma}(a \mid n+1, bn\bar{x}+1).
\end{aligned}
$$

Therefore, $p(a|b, x) = \mathrm{Gamma}(a \mid n+1, bn\bar{x}+1)$ and by symmetry, $p(b|a, x) = \mathrm{Gamma}(b \mid n+1, an\bar{x}+1)$.

Task 2

We now provide code to implement the Gibbs sampler.

```
knitr::opts_chunk$set(cache=TRUE)
library(MASS)
data <- read.csv("data-exponential.csv", header = FALSE)
#######################################
# This function is a Gibbs sampler
#
# Args
#   start.a: initial value for a
#   start.b: initial value for b
#   n.sims: number of iterations to run
#   data:  observed data, should be in a
            # data frame with one column
#
# Returns:
#   A two column matrix with samples
     #   for a in first column and
# samples for b in second column
#######################################
```

```
knitr::opts_chunk$set(cache=TRUE)
sampleGibbs <- function(start.a, start.b, n.sims, data){
  # get sum, which is sufficient statistic. note: sum(x) = n*x_bar.
  x_sum <- sum(data)
  # get n
  n <- nrow(data)
  # create empty matrix, allocate memory for efficiency
  res <- matrix(NA, nrow = n.sims, ncol = 2)
  res[1,] <- c(start.a,start.b)
  for (i in 2:n.sims){
    # sample the values
    res[i,1] <- rgamma(1, shape = n+1,
                       rate = res[i-1,2]*x_sum+1)
    res[i,2] <- rgamma(1, shape = n+1,
                       rate = res[i,1]*x_sum+1)
  }
  return(res)
}
```

Task 3

We now run the Gibbs sampler and produce some results. Finish the rest of this for homework.

```r
# run Gibbs sampler
n.sims <- 10000
res <- sampleGibbs(.25,.25,n.sims,data)
head(res)
```

```
##          [,1]      [,2]
## [1,] 0.250000 0.2500000
## [2,] 2.465476 0.2261853
## [3,] 1.461120 0.2578814
## [4,] 1.684417 0.2964077
## [5,] 1.960193 0.3557249
## [6,] 1.471124 0.2957595
```

```r
dim(res)
```

```
## [1] 10000     2
```

```r
res[1,1]
```

```
## [1] 0.25
```

Task 4 (Finish for homework)

Task 5 (Finish for homework)