Assignment 6

Bayesian models & data analysis

Exercise 1: Hamiltonian Monte Carlo

You are given 500 data points that are assumed to come from a normal distribution with mean μ and variance σ^2,

```
true_mu <- 800
true_var <- 100 #sigma^2
y <- rnorm(500,mean=true_mu,sd=sqrt(true_var))
hist(y)
```

Copy and paste the above chunk of code to generate data y.

Model:

Let y_i be i^{th} data point,

$y_i \sim Normal(\mu, \sigma^2)$

$\mu \sim Normal(m = 1000, s = 100)$

$\sigma \sim Normal(a = 10, b = 2)$

You have to estimate the posterior distributions of μ and σ using Hamiltonian Monte Carlo sampler.

You can directly use the following HMC sampler code to estimate the parameters of the above model.

```
#Gradient functions
gradient <- function(mu,sigma,y,n,m,s,a,b){
  grad_mu <- (((n*mu)-sum(y))/(sigma^2))+((mu-m)/(s^2))
  grad_sigma <- (n/sigma)-(sum((y-mu)^2)/(sigma^3))+((sigma-a)/(b^2))
  return(c(grad_mu,grad_sigma))
}
#Potential energy function
V <- function(mu,sigma,y,n,m,s,a,b){
  nlpd <- -(sum(dnorm(y,mu,sigma,log=T))+dnorm(mu,m,s,log=T)+dnorm(sigma,a,b,log=T))
  nlpd
}

#HMC sampler
HMC <- function(y,n,m,s,a,b,step,L,initial_q,nsamp,nburn){
  mu_chain <- rep(NA,nsamp)
```

```
  sigma_chain <- rep(NA,nsamp)
  reject <- 0
  #Initialization of Markov chain
  mu_chain[1] <- initial_q[1]
  sigma_chain[1] <- initial_q[2]
  #Evolution of Markov chain
  i <- 1
  while(i < nsamp){
    q <- c(mu_chain[i],sigma_chain[i])   # Current position of the particle
    p <- rnorm(length(q),0,1)            # Generate random momentum at the current posi
    current_q <- q
    current_p <- p
    current_V = V(current_q[1],current_q[2],y,n,m,s,a,b) # Current potential energy
    current_T = sum(current_p^2)/2                       # Current kinetic energy
    # Take L leapfrog steps
    for(l in 1:L){
      # Change in momentum in 'step/2' time
      p <- p-((step/2)*gradient(q[1],q[2],y,n,m,s,a,b))
      # Change in position in 'step' time
      q <- q + step*p
      # Change in momentum in 'step/2' time
      p <- p-((step/2)*gradient(q[1],q[2],y,n,m,s,a,b))
    }
    proposed_q <- q
    proposed_p <- p
    proposed_V = V(proposed_q[1],proposed_q[2],y,n,m,s,a,b)  # Proposed potential energ
    proposed_T = sum(proposed_p^2)/2                         # Proposed kinetic energy
    accept.prob <- min(1,exp(current_V+current_T-proposed_V-proposed_T))
    # Accept/reject the proposed position q
    if(accept.prob>runif(1,0,1)){
      mu_chain[i+1] <- proposed_q[1]
      sigma_chain[i+1] <- proposed_q[2]
      i <- i+1
    }else{
      reject <- reject+1
    }
  }
  posteriors <- data.frame(mu_chain,sigma_chain)[-(1:nburn),]
  posteriors$sample_id <- 1:nrow(posteriors)
  posteriors
}
```

Exercise 1.1 Use the following values for the internal parameters of the HMC sampler,

- Total number of samples, $nsamp = 6000$
- Total number of burn-in samples, $nburn = 2000$ (a certain number of initial samples

that you want to throw away)

- Step-size parameter, $step = 0.02$
- Number of leapfrog steps, $L = 12$
- Initializing value of the *mu* and *sigma* chain, $initial_q = c(1000, 11)$

Estimate and plot the posteriors for μ and σ.

(Hint: You can directly run the HMC function as follows:)

```
df.posterior <- HMC(y=y,n=length(y),          # data
                    m=1000,s=20,a=10,b=2,     # priors
                    step=0.02,                # step-size
                    L=12,                     # no. of leapfrog steps
                    initial_q=c(1000,11),      # Chain initialization
                    nsamp=6000,               # total number of samples
                    nburn=2000)               # number of burn-in samples
```

Exercise 1.2 Check posterior sensitivity to the total number of samples. How do the posteriors change with change in total number of samples?

Estimate and compare the posteriors obtained for the following values of total number of samples, *nsamp* -

- $nsamp = 100$
- $nsamp = 1000$
- $nsamp = 6000$

Change burn-in samples *nburn* accordingly. For example, you can set $nburn = nsamp/3$. Keep all other parameters same as in Expercise 1.1.

Exercise 1.3 How do the posteriors change with change in step-size parameter?

Estimate and compare the posteriors obtained when step-size had the ollowing values -

- $step = 0.001$
- $step = 0.005$
- $step = 0.02$

Keep all other parameter same as in Exercise 1.1.

Exercise 1.4 Visually inspect the *mu* and *sigma* chains obtained in expercise 1.3. Do you find anything problematic?

Exercise 1.5 Check the prior sensitivity for the μ parameter. Estimate and compare the posterior distribution of μ when the prior on μ are -

- $\mu \sim Normal(m = 400, s = 5)$
- $\mu \sim Normal(m = 400, s = 20)$
- $\mu \sim Normal(m = 1000, s = 5)$
- $\mu \sim Normal(m = 1000, s = 20)$

- $\mu \sim Normal(m = 1000, s = 100)$

(Keep all other parameters same as in Exercise 1.1).