Hamiltonian MC Exercise

DUE: in class on Friday October 11

Developed by Michael Lerch

Intro

You are already familiar with Metropolis-Hastings and Gibbs sampling routines in the context of MCMC sampling. The `stan` software package utilizes a different routine called NUTS which is more generally based on Hamiltonian Monte Carlo (HMC). This assignment will get you familiar with the basics of HMC.

Hamiltonian MC Idea

Assume we want draws from the posterior distribution $p(\phi|y)$. The principle of HMC is equivalent to rolling a large number of balls on a potential field of $-\log(p(\phi|y))$ and sampling the locations (ϕ values) of those balls. Those sampled locations are samples from $p(\phi|y)$. In the case that ϕ is one or two dimensional, it is pretty easy to think about: $-\log(p(\phi|y))$ is a deformed surface or hill on which a ball can roll. The analogue is the same in higher dimensions, but not easy to envision.

1. Consider an observation y from the binomial likelihood, $y \sim \text{binomial}(n, \pi)$ with n known. If $\phi = \log\left(\frac{\pi}{1-\pi}\right)$ (ie the exponential family canonical parameter of the binomial likelihood), and we choose a prior on ϕ of a normal distribution with a mean of μ and standard deviation of σ, show that the posterior of ϕ based on an observation of y is proportional to

   $$f(\phi) = \exp(\phi)^y \left(1 + \exp(\phi)\right)^{-n} \exp\left(-\frac{(\phi - \mu)^2}{2\sigma^2}\right)$$

   Write an R function that calculates the above expression given inputs ϕ, y, n, μ, and σ. Plot it as a function of ϕ for $y = 5,\ n = 10,\ \mu = 0,\ \sigma = 2$. (Does this problem sound familiar?)

2. Plot $(-\log(f(\phi)))$.

3. A ball sitting on the potential field $-\log(p(\phi|y))$ at position ϕ will accelerate at a rate

proportional to the gradient of the field. In one-dimension we simply find this by:

$$\frac{d}{d\phi}\left(-\log(p(\phi|y))\right).$$

Typically, we just write down $p(\phi|y)$ up to a multiplicative constant like our $f(\phi)$. Show that we only need to know the posterior up to a multiplicative factor to calculate the derivative above. Calculate the derivative for our scenario in terms of y, n, μ, and σ. Write an R function that evaluates the derivative and plot it.

4. If you have written R functions to calculate f (problem 1) and the derivative of the negative log of f (problem 3) and both take as inputs ϕ, y, n, μ, and σ in that order, you can run the following two lines of code to watch a ball roll around:

```
source("roll.R")  #check directory
roll(0, grad, dens)
```

where 0 is the initial ϕ value (location to start the ball), `grad` is the name of your function from problem 3 and `dens` is the name of your function from problem 1.