Simulating random samples using Hamiltonian Monte Carlo and the abcHMC package

October 22, 2015

1 The HMC function

The function HMC generates a random sample from a specified distribution. It takes as arguments: total.samples, the number of samples to be simulated after burnin; q.density, the density the samples should be drawn from; M, the mass matrix; q, the starting values for the simulation; epsilon, the stepsize to be used in the leapfrog; L, the number of leapfrog steps to be made; diff.density, the derivative of the density to be drawn from; burnin the number of samples at the beginning of the simulation to be discarded. The following code will give a random sample of size 1,000 from a univariate standard normal distribution using a step-size (ϵ) of 0.05 and 20 leapfrog iterations, L.

```
HMC(total.samples = 10000,
     q.density = function(x) dnorm(x,0,1),
     M=1,
     q = 0,
     epsilon = 0.05,
     L = 20,
     diff.density = function(x) x,
     burnin = 100)
```

This can be extended to the bivariate Gaussian by specifying a bivariate density function. The following will simulate a bivariate Gaussian distribution with highly-correlated covariates.

```
bivariate.density <- function(l) {
  dmvnorm(l, c(0, 0),matrix(c(1, 0.95, 0.95, 1), 2, 2))
}
bivariate.diff <- function(x) {
  solve(matrix(c(1, 0.95, 0.95, 1), 2, 2))%*%as.matrix(x)
}
HMC(total.samples = 10000,
     q.density = bivariate.density,
     q = c(-2,-2),
     M = diag(2),
     epsilon = 0.18,
     L = 20,
     diff.density = bivariate.diff,
     burnin = 0)
```

Sampling from higher higher dimensions can be done in a similar manner. The following will simulate 500 samples from a 150-dimensional distribution with independent covariates.

```
multi.density <- function(l) {
  dmvnorm(l,rep(0,150),diag(seq(from=0.02,to=1,length=150)^2))
}
```

```
multi.diff <- function(x) {
  solve(diag(seq(from=0.02,to=1,length=150)^2))%*%as.matrix(x)
}
out.multidimension <- HMC(total.samples = 500,
                          q.density = multi.density,
                          q = rep(0,150),
                          M=diag(150),
                          epsilon = 0.014,
                          L = 100,
                          diff.density = multi.diff,
                          burnin = 0)
```

2 The WordPrint function

The function *WordPrint* is used to plot a random sample from a distribution whose probability density resembles a chosen word. It takes two arguments, word: the word you would like plotted, given as a character string; and samples: the number of simulated points to be used in the plot. Before running this function the dataset specifying the underlying models must be loaded.

```
letter.models <- data(”letter.models”)
WordPrint(word = ”abcHMC”, samples = 2500)
```