Markov Chain Monte Carlo Ⅲ

Woojoo Lee

So far, we studied when Markov chain has the stationary distribution and the limiting distribution. But, the real problem is the reverse of this logic !

MCMC approach is to construct a Markov chain having the following properties:

- the chain has a unique stationary distribution
- the transition probabilities of the chain are simple to implement.

When the target distribution π is given, our task is to construct an "easily-simulated" ergodic Markov chain with transition matrix $\mathbf{P}$ where the stationary distribution is π.

We want to compute:

$$\theta = \int h(x)f(x)dx.$$

We will construct a Markov chain whose stationary distribution is $f(x)$. Under regularity conditions,

$$\frac{1}{n}\sum_{i=1}^{n} h(X_i) \rightarrow E(h(X))$$

where X is a Markov Chain whose stationary distribution is f.

We say that π satisfies "detailed balance" condition if

$$\pi_i p_{ij} = \pi_j p_{ji}.$$

Theorem: If the detailed balance holds for π, then π is a stationary distribution.

Q) Prove this.

This concept is very important in implementing MCMC !

The Metropolis-Hastings Algorithm

Let $q(y|x)$ be a "easy-to-use" conditional density in the sense that we know how to sample from it. This is called "proposal distribution".

Then, MH algorithm is as follows: Choose X_0 first.

- Generate a random variable from $Y \sim q(y|X_i)$.
- Compute

  $$r(x, y) = \min\left(\frac{f(y)q(x|y)}{f(x)q(y|x)}, 1\right)$$

  at $x = X_i$ and $y = Y$.
- $X_{i+1} = Y$ with probability $r(x, y)$, otherwise $X_{i+1} = X_i$.
- Repeat the above procedure.

Remark) A common choice of $q(y|x)$ is $N(x, \alpha^2)$ for some $\alpha > 0$. Here, α is related to the efficiency of the chain.

From now on, to deal with a continuous state Markov chain, we change notations a little. Let $p(x, y)$ denote the probability of transition from x to y. And we use $f(x)$ instead of π.

- f is a stationary distribution if $f(x) = \int f(y)p(y, x)dy$.
- detailed balance holds if $f(x)p(x, y) = f(y)p(y, x)$.

Q) Show that if detailed balance holds, f becomes a stationary distribution.

But, we do not have $p(x, y)$, instead, we have $q(y|x)$ where

$$f(x)q(y|x) < f(y)q(x|y)$$

or

$$f(x)q(y|x) > f(y)q(x|y).$$

Suppose that

$$f(x)q(y|x) > f(y)q(x|y),$$

and $r(x, y) = \min(\frac{f(y)q(x|y)}{f(x)q(y|x)}, 1)$.

Q) Construct $p(x, y)$ from the description of MH algorithm, and show that detailed balance holds.

Example) Consider the Cauchy distribution:

$$f(x) = \frac{1}{\pi} \frac{1}{1+x^2}.$$

Q) Construct a Markov chain whose stationary distribution is f.

R-code for the previous example:

```
n<-10000
xvec<-rep(0,n)
b<-5

for (i in 1:(n-1)){

y<-rnorm(1,xvec[i],b)

r<-min((1+xvec[i]^2)/(1+y^2),1)

u<-runif(1)

if (u<=r) {xvec[i+1]<-y}
if (u>r)  {xvec[i+1]<-xvec[i]}

}

plot(xvec)
```

In practice, there are difficult issues to be addressed:

- burn-in time
- multiple short Markov chains vs a single long Markov chain
- Criteria to guarantee that MCMC converges