Example of MC for Continuous State Space

Notes 2, Sta624

The MC we discussed so far are "discrete state space", i.e. the random variable take values in a discrete space. We shall discuss here briefly MC with continuous state space, in particular, the random variables take value in the real line R^1.

Recall: A Markov Chain is a sequence of random variables $X_0, X_1, X_2, \ldots$ taking values in a space $\mathcal{X}$. The main property of the chain is that the past is conditionally independent of the future given the present. Here $\mathcal{X} = R^1$

• Transition probability matrix needs to be replaced by a transition probability function/kernel: p_{ij} becomes $p(x|y)$;

which is a probability density function for any given y value. Or a conditional density.

Interpretation: $p(x|y)dx$ is the probability of going to x given it is at y now.

Example 1 MC for continuous state space (take values as continuous random variables). But still discrete time (i.e. a chain).

In the following example $p(x|y)$ is taken to be Uniform $(1-y, 1)$. State space is the interval $(0, 1)$.

$X_0 \sim$ any number between 0 and 1 (or could be from any distribution on (0,1) interval)

$X_1 \sim$ Unif $(1 - X_0, 1)$

$X_2 \sim$ Unif $(1 - X_1, 1)$

......

$X_n \sim$ Unif $(1 - X_{n-1}, 1)$; ...

This is an MC.

If the distribution of X_n is convergent at all (here it does), the stationary/limiting distribution (density $f(x)$) must satisfy the following (integral) equation:

$$f(x) = \int f(y)p(x|y)dy \quad \text{for any } x \tag{1}$$

We may check that the probability density function $f(x) = 2xI[0 < x < 1]$ solves the above equation. [with $p(x|y) = U(1 - y, 1)$].

Therefore X_n, when n large, will have a density approximately equal to $f(x) = 2x$ for $0 < x < 1$.

Remark: Compare (1) to the equation for a discrete limiting distribution $\pi_j = \lim P^n_{ij}$:

$$\pi_j = \sum_i \pi_i P_{ij} \quad \text{for any } j$$

we see that (1) is just the continuous version of the above equation.

Detail verifications: Since

$$p(x|y) = \frac{I_{[1-y<x<1]}}{y}$$

We compute, using $f(y) = 2y$,

$$\int_0^1 p(x|y)f(y)dy = \int I_{[1-y<x<1]} 2dy = 2\int_{1-x}^1 dy = 2x \ .$$

Notice here we only used the random variables from uniform distributions and end up with a random variable with density $f(x) = 2x$.

The idea of MCMC is to use an easy $p(x|y)$ in the iteration but end up with a random variable with (complicated) $f(x)$ that we desire.

Example 2 Suppose
$X_0 \sim$ any number between 0 and 1
$X_1 \sim$ Unif $(0, 1 - X_0)$
$X_2 \sim$ Unif $(0, 1 - X_1)$
......
$X_n \sim$ Unif $(0, 1 - X_{n-1})$, $\cdots$
Find the stationary distribution $f(x)$ of this MC. This will also be the approximate distribution of X_n for large n.

Usually the checking or solving of the equation (1) is not easy. But the following Theorem offers a special easier case.

Theorem 1 *If a probability density function $f(x)$ and a transition probability kernel $p(x|y)$ satisfy the so called 'detailed balance equation':*

$$f(x)p(y|x) = f(y)p(x|y) \tag{2}$$

then the equation (1) above is satisfied.

We usually shall check (2) instead of (1). In the example 1, $f(x) = 2xI[0 < x < 1]$ and $p(x|y) = \frac{I_{[1-y<x<1]}}{y}$, therefore

$$f(x)p(y|x) = 2xI[0 < x < 1]\frac{I[1-x < y < 1]}{x} = 2I[0 < x < 1; 1-x < y < 1],$$

and

$$f(y)p(x|y) = 2yI[0 < y < 1]\frac{I_{[1-y<x<1]}}{y} = 2I[0 < y < 1, 1-y < x < 1].$$

Draw two pictures, and you will convince yourself the two right hand side are the same region.

Metropollis Type Chains:

At stage n, suppose a MC takes value y_n. We generate the 'candidate' value Y using the uniform random walk: $Y \sim Unif[y_n - a, y_n + a]$. Finally we accept the generated Y value as y_{n+1} if

$$\frac{f(Y)}{f(y_n)} > U(0,1)$$

(where $U(0,1)$ denote another independently generated random variable) otherwise the MC do not move, i.e. $y_{n+1} = y_n$.

We can show that this transition kernel function, and density function $f(\cdot)$ pair satisfy the 'detailed balance equation'.

The transition probability density function is

$$\frac{1}{2a} I[y_n - a < Y < y_n + a] Pr\left(\frac{f(Y)}{f(y_n)} > U\right)$$

since U is from uniform (0,1), therefore we have

$$= \frac{1}{2a} I[|Y - y_n| < a] \min\left\{1, \frac{f(Y)}{f(y_n)}\right\}$$

From here you can proof the "detailed balance equation" hold.

Remark: If we want to construct an MC that has limiting distribution $f(\cdot)$ by Metropollis, we need to be able to compute the ratio

$$\frac{f(Y)}{f(y_n)}.$$

This imply we do not need to worry about the constant if $f(t) = Cg(t)$.

Application in Bayesian Analysis

In statistical problem of Bayesian inference, we are interested in the posterior distribution/density, or the mean of the posterior etc.

posterior density = constant $\times$ prior density function $\times$ likelihood function.

Example 0: $X_i \sim N(\theta, \sigma^2)$ and the prior on θ is $N(\mu, \tau^2)$. Here σ^2, τ^2, μ are all known.

Example 1: Using the R package `mcmc`. Two dim parameter. In the package, we input the posterior as "log of un-normalized posterior".

Here prior density is

$$g(\alpha, \lambda) = C \times \frac{\sqrt{\alpha\ trigamma(\alpha) - 1}}{\lambda}$$

and the likelihood function is a two parameter gamma distribution

$$\prod_i \frac{\lambda^\alpha}{\Gamma(\alpha)} x_i^{\alpha-1} \exp(-\lambda x_i).$$

```
> lupost <- function(theta) {
+ stopifnot(is.numeric(theta))
+ stopifnot(is.finite(theta))
+ stopifnot(length(theta) == 2)
+ alpha <- theta[1]
+ lambda <- theta[2]
+ if (alpha <= 0) return(- Inf)
+ if (lambda <= 0) return(- Inf)
+ logl <- sum(dgamma(x, shape = alpha, rate = lambda, log = TRUE))
+ lpri <- (1 / 2) * log(alpha * trigamma(alpha) - 1) - log(lambda)
+ return(logl + lpri)
+ }
```

The only tricky bit is that we define this function to be $-\infty$ when parameter is off the allowed space.

```
> out <- metrop(out, blen = 200, nbatch = 500)
> alpha <- out$batch[ , 1]
> lambda <- out$batch[ , 2]
> t.test(alpha)
```