Markov Chain Monte Carlo

Xiaobao Wu
Tsinghua University
xiaobao.wu1996@gmail.com

1 Markov Chain Monte Carlo

1.1 Stationary Distribution

The distribution at time t can be written as

$$\pi_t(x^*) = \sum_x \pi_{t-1}(x) k(x^*|x). \tag{1}$$

We want π_t and π_{t-1} are the same distributions, i.e., $\pi_t = \pi_{t-1} \triangleq \pi$. Thus, we have

$$\pi(x^*) = \sum_x \pi(x) k(x^*|x). \tag{2}$$

But it is hard to find a kernel function satisfying this. Fortunately, one sufficient condition for ensuring π invariant is called detailed balance, defined as

$$\pi(x)k(x^*|x) = \pi(x^*)k(x|x^*) \tag{3}$$

With detailed balance, we have

$$\sum_x \pi(x)k(x^*|x) = \sum_x \pi(x^*)k(x|x^*) = \pi(x^*) \sum_x k(x|x^*) = \pi(x^*) \tag{4}$$

1.2 Metropolis-Hastings Algorithm

Algorithm 1: Metropolis-Hastings Algorithm

```
Initialize x^0 ;
for iteration i = 1, 2 ... do
    u ~ Uniform(0, 1) ;
    x* ~ q(x*|x) ;
    α(x*) = min{1, π(x*)q(x|x*) / (π(x)q(x*|x))} ;
    if u < α(x*) then
        x^(i+1) ← x* ;
    end
    else
        x^(i+1) ← x ;
    end
end
```

q is the proposal distribution. α is called acceptance rate. $k(x^*|x) = q(x^*|x) \min\{1, \frac{\pi(x^*)q(x|x^*)}{\pi(x)q(x^*|x)}\}$ is the transition probability.

$$\begin{aligned}
\pi(x)k(x^*|x) &= \pi(x)q(x^*|x)\min\{1, \frac{\pi(x^*)q(x|x^*)}{\pi(x)q(x^*|x)}\} && (5)\\
&= \min\{\pi(x)q(x^*|x), \pi(x^*q(x|x^*))\} && (6)\\
&= \pi(x^*)q(x|x^*)\min\{\frac{\pi(x)q(x^*|x)}{\pi(x^*)q(x|x^*)}, 1\} && (7)\\
&= \pi(x^*)k(x|x^*) && (8)
\end{aligned}$$

So, during the transition from $x^{(i)}$ to $x^{(i+1)}$, x still follows the distribution $\pi(x)$.

A popular choice for the proposal distribution is $q(x^*|x) = g(x^* - x)$ where g is a symmetric distribution, thus

$$x^* = x + \epsilon, \quad \epsilon \sim g.$$

Since $g(\epsilon) = g(-\epsilon)$, then $q(x^*|x) = q(x|x^*)$.

1.3 Gibbs Sampling

Gibbs Sampling is a special case of Metropolis-Hastings algorithm, which replaces $q(x|x^*)$ with $\pi(x|x^*_{\neg n})$. Thus, Gibbs sampling follows

$$\begin{aligned}
\frac{\pi(x^*)q(x|x^*)}{\pi(x)q(x^*|x)} &= \frac{\pi(x^*_n|x^*_{\neg n})\pi(x^*_{\neg n})\pi(x_n|x^*_{\neg n})}{\pi(x_n|x_{\neg n})\pi(x_{\neg n})\pi(x^*_n|x_{\neg n})} && (9)\\
&= \frac{\pi(x^*_n|x_{\neg n})\pi(x_{\neg n})\pi(x_n|x_{\neg n})}{\pi(x_n|x_{\neg n})\pi(x_{\neg n})\pi(x^*_n|x_{\neg n})} && (10)\\
&= 1 && (11)
\end{aligned}$$

So the acceptance rate of Gibbs sampling is 1.