Unbiased MCMC

Rmk: Parallel Bayesian Computation is quite difficult due to the nature of the M.C.

Context: π a distribution of interest on $\mathbb{R}^d$

Suppose we have a Markov Kernel P such that $\pi P = \pi$

Use $P(\cdot,\cdot)$ to propose new states

$$X_{t+1} \mid X_t \sim P(X_t, \cdot)$$

and for large n and some function of interest $h: \mathbb{R}^d \mapsto \mathbb{R}$

$$\frac{1}{n} \sum_{i=1}^{n} h(x_i) \simeq \pi(h) \equiv \int h(x)\, \pi(x)\, dx$$

and we have $\pi(h) = \lim_{n \mapsto \infty} \mathbb{E}[h(X_n)]$

We can reformulate this problem as follows

Define $Y_0 = h(X_0), Y_1 = h(X_1) - h(X_0), \dots, Y_k = h(X_k) - h(X_{k-1})$

So that $\sum_{k=0}^{n} \mathbb{E}[Y_k] = \mathbb{E}[h(X_n)]$ and more importantly

$$\boxed{\pi(h) = \sum_{k=0}^{\infty} \mathbb{E}(Y_k)}$$

Assumption: Interchange $\mathbb{E}\left(\sum_{k=0}^{\infty} Y_k\right) = \sum_{k=0}^{\infty} \mathbb{E}(Y_k)$

then $\pi(h) = \mathbb{E}\left[\sum_{k=0}^{\infty} Y_k\right]$

We know how to build unbiased estimators for this.

Draw $N \sim q$. and estimate $\pi(h)$ by Y_N / q_N

Alg: (i) Draw $N \sim q$

(ii) Run MC $X_0, \dots, X_N$

(iii) Return $\dfrac{h(X_N) - h(X_{N-1})}{q_N}$.

$$\mathbb{E}\left[Y_N / q_N \right] = \mathbb{E}\left[\mathbb{E}\left[\frac{Y_N}{q_N} \,\middle|\, (X_k)_{k=1}^{\infty} \right] \right] = \mathbb{E}\left[\sum_{i=0}^{\infty} \frac{Y_i}{q_i} q_i \right] = \mathbb{E}\left[\sum_{i=0}^{\infty} Y_i \right] = \pi(h)$$

A better estimate of $\sum_{k=0}^{\infty} X_k$ is $N \sim q$ and take

$$\sum_{k=0}^{N} \frac{x_k}{\mathbb{P}(N \geq k)}.$$

In terms of variance and stability.

Alg: (i) Draw $N \sim q$

(ii) Build $(X_n)_{n=0}^{N}$ a MC.

(iii) $\hat{\pi}(h) = \sum_{k=0}^{N} \dfrac{h(x_k) - h(x_{k-1})}{\mathbb{P}(N \geq k)}$

$\Rightarrow$ Really useful in SDE

Doesn't really work in general due to the assumption we made.

We can interchange if $\sum_{k=0}^{\infty} \mathbb{E}\left(|h(X_k)-h(X_{k-1})|\right) < +\infty$

But in MCMC for RWM $X_k = X_{k+1} + \sigma Z_k$ oftentimes

$$|h(X_k)-h(X_{k-1})| = O(\sigma|Z_k|) = O(1)$$

for Lipshitz h $\Rightarrow \sum \mathbb{E}\,|h(X_k)-h(X_{k-1})| = +\infty$

But if $\{\bar{Y}_k\}_{k=0}^{\infty}$ is s.t. $\mathbb{E}(\bar{Y}_k) = \mathbb{E}(Y_k)$

then $\pi(h) = \sum_{k=0}^{\infty} \mathbb{E}(\bar{Y}_k)$

We build a bivariate MC $\{X_0, (X_k, Z_{k-1}), k \geq 1\}$ s.t.

$\{X_k, k \geq 0\}$ is a M.C. with trans. kernel $P(\cdot,\cdot)$. and $X_0 \sim \mu$.

and $\{Z_k, k \geq 0\}$ with tran. Kernel $P(\cdot,\cdot)$ and $Z_0 \sim \mu$.

Hence $\mathbb{E}[h(X_k) - h(Z_{k-1})] = \mathbb{E}[h(X_k) - h(X_{k-1})]$

We build the chain so that $X_k = Z_{k-1}$ after a while.

We can interchange sum and expectation so that

$$\pi(h) = \mathbb{E}\left[h(X_0) + \sum_{k=1}^{\infty} h(X_k) - h(Z_{k-1})\right]$$

If $\tau = \inf\{k \geq 1 : X_k = Z_{k-1}\}$

$$\pi(h) = \mathbb{E}\left[h(X_0) + \sum_{k=1}^{\tau-1} h(X_k) - h(Z_{k-1})\right]$$

$$\pi(h) = \lim_{n\to\infty} \mathbb{E}[h(X_n)] = \mathbb{E}\left[h(X_0) + \sum_{k=1}^{\infty} \mathbb{E}(Y_k)\right]$$

<u>Alg:</u> (1) Build a BMC $\{X_0, (X_k, Z_{k-1})_{k=1}^{\infty}\}$ s.t.

τ = first time $(X_n)_{k=0}^{\infty}$ $(Z_n)_{k=0}^{\infty}$ are two identical MC with initial dist μ t.k. P

$= \inf\{k \geq 1 : X_k = Z_{k-1}\}$

return $\hat{\pi}(h) = h(X_0) + \sum_{k=1}^{\tau-1} [h(X_k) - h(Z_{k-1})]$

<u>Construction of BMC.</u>

Basic idea: Coupling. Let p, q be 2 distributions

We can sample $X \sim p$, $Y \sim q$ s.t. $\mathbb{P}(X=Y) = \int p \wedge q$

<u>Alg:</u> Draw $X \sim p$, $u \sim \text{Unif}(0, q(x))$

If $u \leq p(x)$: (X, X)

Else $Y^* \sim q$, $V \sim U(0,1)$ until $V \geq \min\left(1, \frac{p(Y^*)}{q(Y^*)}\right)$

return (X, Y^*)