STA447/2006 Lecture 5

(Midterm 1 week, 1-hr lecture).

Recall from last time: "SLLN" for Markov chains.

$$\lim_{n\to+\infty} \frac{N_i(n)}{n} = \frac{1}{\mathbb{E}_i[T_i]} = \pi_i \quad (a.s.)$$

Extend to general functionals (instead of just counts).

Thm: Suppose $f: S \to \mathbb{R}$, MC positive recurrent, irreducible

π is a stationary distribution,

Assume $\mathbb{E}_\pi\left[|f(X)|\right] < +\infty$ $\left(\sum_{i\in S} \pi_i |f(i)| < +\infty \right)$.

For any initial state,

$$\frac{1}{n}\sum_{t=1}^{n} f(X_t) \xrightarrow{a.s.} \mathbb{E}_\pi[f(X)].$$

e.g. use case: MCMC in statistics

Goal: to compute $\mathbb{E}_\pi[f(X)] = \sum_{i\in S} \pi_i f(i)$

for some complicated distribution π.

(e.g. posterior distribution)

Direct integration / summation is computationally impossible
(i.e. when $|S|$ infinite / exponentially large)

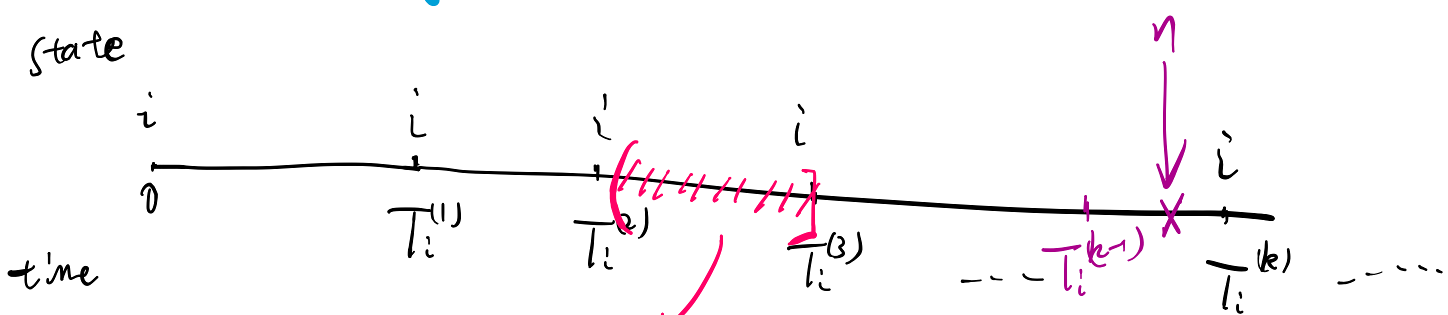
"MC" "Monte Carlo"

Markov chain "MC"

- Construct a MC based on π
- Run it from any starting state.
- Take trajectory average $\longrightarrow$ $\mathbb{E}_\pi[f(X)]$.

Proof of theorem.

Fix $i \in S$ — starting state. Let $T_i^{(k)} :=$ k-th visit time to i starting from time $t=1$.



Let $Y_k := \sum_{t=T_i^{(k-1)}+1}^{T_i^{(k)}} f(X_t)$. iid random varibles.

$\boxed{\text{Relate } \mathbb{E}(Y_k) \text{ to } \pi \text{ ?}}$

Recall construction of stationary measure,

$$\mu_i(j) = \mathbb{E}[\#\text{ visits to } j \text{ before returning to } i]$$

$(\text{for } j \neq i)$ $\quad \mu_i(i) = 1.$

Positive recurrence $\Rightarrow A = \sum_{j \in S} \mu_i(j) < +\infty$

μ_i / A is stationary distribution, $A = \frac{1}{\pi_i} = \mathbb{E}_i[T_i]$.

$$Y_k = \sum_{j \in S} f(j) \cdot \underbrace{\#\text{ visits to } j \text{ in } [T_i^{(k-1)}+1, T_i^{(k)}]}$$

$$\sum_{j \in S} |f(j)| \mu_i(j) < +\infty$$

Fubini thm

$$\mathbb{E}[Y_k] = \sum_{j \in S} f(j) \cdot \mathbb{E}[\#\text{ visits to } j \text{ in } [T_i^{(k-1)}+1, T_i^{(k)}]]$$

$$= \sum_{j \in S} f(j) \mu_i(j) = \frac{\sum_{j \in S} f(j) \pi_j}{\pi_i}$$

$$\text{SLLN} \Rightarrow \frac{1}{n} \sum_{k=1}^{n} Y_k \xrightarrow{a.s.} \frac{\mathbb{E}_\pi[f(X)]}{\pi_i}.$$

$$\text{SLLN} \Rightarrow \frac{1}{k} T_i^{(k)} \xrightarrow{a.s.} \frac{1}{\pi_i}$$

(from last time)

Use sandwich trick from last time

$$\lim_{n\to+\infty} \frac{1}{n}\sum_{1}^{n} f(X_t) = \frac{\lim_{n\to+\infty} \frac{1}{n}\sum_{1}^{n} Y_k}{\lim_{n\to+\infty} \frac{1}{n} T_i^{(n)}} \overset{(a.s.)}{=} \mathbb{E}_\pi[f(X)].$$

Question: how to construct the Markov chain?

A popular choice: Metropolis - Hastings algorithm.

- A general recipe of construction. "proposal distribution".
- Idea: start w/ $\boxed{\text{some transition kernel}}$ (may not have the correct stationary).
  use a correction step to make sure it converges to the correct one.

---

eg simple case. If $q(i,j) = q(j,i)$.

(satisfied by SRW).

Define transition prob.

$$P_{ij} = \begin{cases} q(i,j)\cdot \min\left(1, \frac{\pi_j}{\pi_i}\right) & (j\neq i) \\ 1 - \sum_{\ell\neq i} P_{i\ell} & (j=i) \end{cases}$$

Operationally, given X_t at t-th step

— Generate proposal $Y_t \sim q(X_t, \cdot)$.

— If $\boxed{\pi(Y_t) > \pi(X_t)}$, accept w.p. 1

Otherwise — accept w.p. $\boxed{\frac{\pi(Y_t)}{\pi(X_t)}}$.

— If $\begin{cases} \text{accept}, & X_{t+1} = Y_t. \\ \text{reject}, & X_{t+1} = X_t. \end{cases}$

require knowledge about π up to normalization const

Claim: π is a stationary distribution of MH chain.

$$\pi_i p_{ij} = q(i,j) \cdot \min\left(1, \frac{\pi_j}{\pi_i}\right) \cdot \pi_i$$

$$= q(i,j) \min(\pi_i, \pi_j) = q(j,i) \min(\pi_i, \pi_j) = \pi_j p_{ji}.$$

So P is reversible w.r.t. π.

- Irreducibility: need $\begin{cases} q \text{ be irreducible}, \\ \pi_i > 0 \quad \forall i \in S \end{cases}$

  If possible to get j from i under q also possible under p.

- Aperiodicity: as long as rejection is possible.

In general, asymmetric proposal q.

$$P_{ij} = q_{ij} \min\left\{1, \frac{\pi_j q(j,i)}{\pi_i q(i,j)}\right\}$$

(i.e. starting from i, accept proposal j w.p. $\min\left\{1, \frac{\pi_j q(j,i)}{\pi_i q(i,j)}\right\}$).

$$\pi_i P_{ij} = \pi_i q_{ij} \cdot \min\left\{1, \frac{\pi_j q(j,i)}{\pi_i q(i,j)}\right\}$$

$$= \min\{\pi_i q(i,j), \pi_j q(j,i)\} = \pi_j P_{j,i}.$$

Irreducible & Aperiodic: similar to symmetric case.

In practice, design principles for q:

- Rejection does not happen too often
- q needs to move fast