

①

Refs: Ch 7 of Lawler for ds time, but also applies to discrete time
: Robert + Casella; Monte Carlo Statistical Methods, Springer, 1999.

Lecture 7

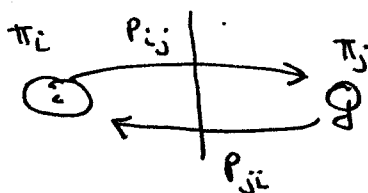
April 20, 2005

REVERSIBLE MARKOV CHAIN

S = discrete state space for discrete time MC X

Suppose π is a probability distribution on S st.

$$\pi_i P_{ij} = \pi_j P_{ji} \quad \forall i, j \in S$$



Then MC is said to be reversible with respect to π .

1. Note: π is a stationary distn:

$$\sum_i \pi_i P_{ij} = \sum_i \pi_j P_{ji} = \pi_j \quad \forall j.$$

2. Consider $\bar{X}_n = X_{N-n}$, $n = 0, 1, \dots, N$ for some fixed $N < \infty$
where X is initialized with distn π .

FACT

$\{\bar{X}_n, n = 0, 1, \dots, N\}$ is Markov (check) with transition probs:

$$\tilde{P}_{ij} = P(\bar{X}_{n+1} = j \mid \bar{X}_n = i) \quad \text{assuming } P(\bar{X}_n = i) \neq 0$$

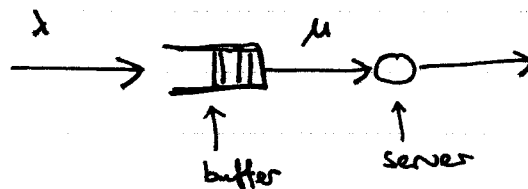
$$= P(\bar{X}_{n+1} = j, \bar{X}_n = i) / P(\bar{X}_n = i)$$

$$= P(X_{N-n-1} = j, X_{N-n} = i) / P(X_{N-n} = i)$$

$$= \pi_j P_{ji} / \pi_i$$

$$= P_{ij} = P(X_{n+1} = j \mid X_n = i)$$

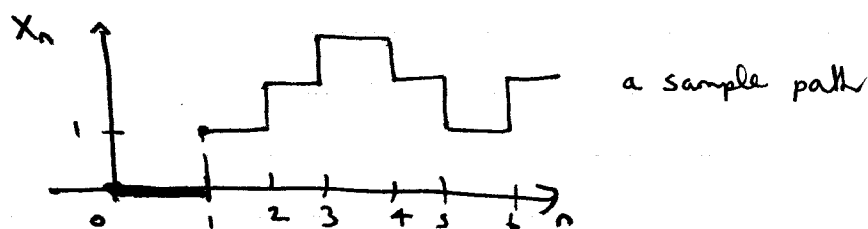
2/

EXAMPLE - QUEUEING MODEL

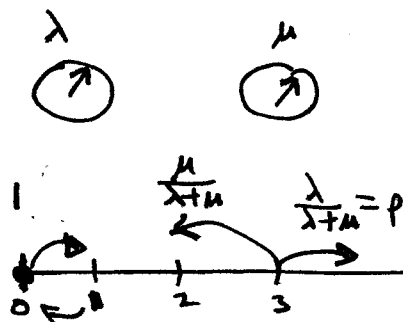
$M/M/1$
 ↑ arrivals Poisson: (service times iid exponential, mean $\frac{1}{\mu}$)
 iid exponential interarrival times with mean $\frac{1}{\lambda}$

$X_n = \# \text{ jobs in system at } n^{\text{th}} \text{ transition (new arrival or departure)}$

(discrete event system) - uses memoryless property of exponentials to show that $\{X_n, n=0,1,2,\dots\}$ is a Markov chain (discrete time, discrete state)



Alarm clocks - indept, exponential



Prob. get new arrival before new departure = $\frac{\lambda}{\lambda+\mu}$

transition probs for $\{X_n\}$

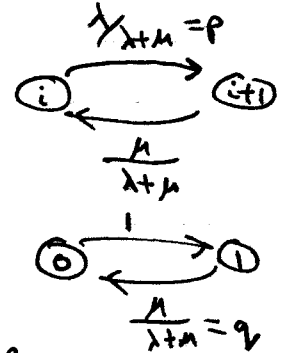
③

Does this satisfy detailed balance?

Seek prob. distn $\pi = (\pi_i)_{i=0}^{\infty}$ s.t. for $i=1,2,\dots$

$$\pi_i P_{i,i+1} = \pi_{i+1} P_{i+1,i}$$

$$+ \pi_0 \cdot 1 = \pi_1 \frac{\mu}{\lambda + \mu}$$



$$\text{i.e. } \pi_i \rho = \pi_{i+1} q \Leftrightarrow \pi_{i+1} = \frac{\rho}{q} \pi_i$$

$$\pi_1 = \frac{1}{q} \pi_0$$

$$\text{So } \pi_i = \left(\frac{\rho}{q}\right)^{i-1} \frac{1}{q} \pi_0, \quad i=1,2,\dots$$

$$= \left(\frac{\lambda}{\mu}\right)^{i-1} \left(\frac{\lambda+\mu}{\mu}\right) \pi_0 = \left(\frac{\lambda}{\mu}\right)^i \left(\frac{\lambda+\mu}{\lambda}\right) \pi_0$$

$$\text{Need } \sum_{i=0}^{\infty} \pi_i = 1 \Leftrightarrow \pi_0 \left[1 + \sum_{i=1}^{\infty} \left(\frac{\lambda}{\mu}\right)^i \left(\frac{\lambda+\mu}{\lambda}\right) \right] = 1$$

$$\Leftrightarrow \pi_0 \left[1 + \frac{\frac{\lambda}{\mu}}{1 - \frac{\lambda}{\mu}} \left(\frac{\lambda+\mu}{\lambda}\right) \right] = 1$$

iff $\frac{\lambda}{\mu} < 1$

$$\Leftrightarrow \pi_0 \left[1 + \frac{\lambda+\mu}{\mu-\lambda} \right] = 1$$

$$\Leftrightarrow \pi_0 \left[\frac{2\mu}{\mu-\lambda} \right] = 1$$

$$\text{So } \pi_i = \left(\frac{\lambda}{\mu}\right)^i \left(\frac{\lambda+\mu}{\lambda}\right) \left(\frac{\mu-\lambda}{2\mu}\right) \pi_0 = \frac{\mu-\lambda}{2\mu}$$

($\mu > \lambda$ here).

④

MARROW CHAIN MONTE CARLO

- Consider finite state space S , otherwise have to be careful about finiteness of expectat.
- Suppose π is probability distribution on S and want to approximate the average value of some function ϕ on S under π , i.e., $\phi: S \rightarrow \mathbb{R}$, want $\sum_j \phi_j \pi_j$
 $\phi_j = \phi(j)$

Monte Carlo simulation:

Let $X_1, X_2, \dots, X_n$ be iid with distribution π

Then

$$\sum_j \phi_j \pi_j = E_{\pi}[\phi(X_1)] \underset{\substack{\uparrow \\ \text{SLLN}}}{\approx} \frac{1}{n} \sum_{k=1}^n \phi(X_k)$$

It may not be so easy to sample from distn π .

Markov chain Monte Carlo - use simulations of a

Markov chain with stationary distn π to approximately simulate π .

Suppose $\{X_n\}_{n=0}^{\infty}$ is an irreducible, positive recurrent Markov chain with stationary distn π .

From last time, a.s., as $n \rightarrow \infty$,

$$\frac{1}{n+1} \sum_{i=0}^n 1_{\{X_i=j\}} \rightarrow \pi_j$$

So

$$\sum_{j \in S} \left(\frac{1}{n+1} \sum_{i=0}^n 1_{\{X_i=j\}} \right) \phi_j \xrightarrow{n \rightarrow \infty} \sum_{j \in S} \phi_j \pi_j \quad (\text{finitely many states})$$

$$\begin{aligned} \text{(Time average)} \rightarrow \frac{1}{n+1} \sum_{i=0}^n \phi(X_i) &\xrightarrow{n \rightarrow \infty} \sum_{j \in S} \phi_j \pi_j \leftarrow \text{(ensemble average)} = E^{\pi}[\phi(X_1)] \end{aligned}$$

⑤

Rates of convergence are important for practical use - lots of work in 1990s.

Metropolis-Hastings algorithm (version of MCMC that is especially good if don't know normalizing constant, i.e., only know π up to a scalar constant used to make it a prob. distr.).

HISTORY

- Metropolis-Hastings algorithm has its origins in statistical physics, but it is now used widely in other contexts, eg., in pedigree analysis in genetics.
- Gibbs sampler - particular case of Metropolis-Hastings in some sense.
 - more specialized, needs more knowledge of distrs.

¹⁹⁵³ ¹⁹⁷⁰
METROPOLIS (Hastings discussed efficiency + generalized to non-symmetric Q).

- Suppose π is prob. distr. on S with $\pi_i > 0 \forall i$.
- Let Q be a symmetric transition matrix on Q ($Q_{ij} = Q_{ji}$) where Q is irreducible (Q gives proposal distr.) (Want Q to be simple so can simulate it, eg., uniform)
- Define

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

(make move with prob. < 1 if $\pi_j < \pi_i$)

$$P_{ii} = 1 - \sum_{j \neq i} P_{ij}$$

[If Q not symmetric,

$$P_{ij} = Q_{ij} \cdot \min\left(1, \frac{\pi_j Q_{ji}}{\pi_i Q_{ij}}\right)$$

$$\pi_i P_{ij} = \pi_i Q_{ij} \min\left(1, \frac{\pi_j Q_{ji}}{\pi_i Q_{ij}}\right) = Q_{ij} \min(\pi_i, \pi_j)$$

$$\pi_j P_{ji} = \pi_j Q_{ji} \min\left(1, \frac{\pi_i Q_{ij}}{\pi_j Q_{ji}}\right) = Q_{ji} \min(\pi_j, \pi_i) \underset{\substack{\uparrow \\ \text{symmetric} \\ Q}}{=} Q_{ij} \min(\pi_i, \pi_j)$$

Check
detailed
balance

for
symmetric Q