Discrete Stochastic Processes

Lecture 15

Countable-State Markov Chains – Chapter 5

Reversible Markov Chains

Burke’s Theorem

Semi-Markov Processes (very briefly)

M/G/1 Queues

Reversibility

Markov chains, looked at backwards in time, still have Markov property!

$$\boxed{P\left(X_{n+1} \middle| X_n, X_{n-1}, \ldots X_0\right) = P\left(X_{n+1} \middle| X_n\right)}$$

$$P\left(X_{n+j}, X_{n+j-1}, \ldots, X_{n+1} \middle| X_n, X_{n-1}, \ldots X_0\right) = P\left(X_{n+j}, X_{n+j-1}, \ldots, X_{n+1} \middle| X_n\right)$$

If E^+ is an event based on $X_{n+j}, X_{n+j-1}, \ldots, X_{n+1}$, and E^- is an event based on $X_{n-1}, \ldots, X_0$

then: $P\left(E^+ \middle| X_n E^-\right) = P\left(E^+ \middle| X_n\right)$, so $P\left(E^+ \middle| X_n E^-\right) P\left(E^- \mid X_n\right) = P\left(E^+ \middle| X_n\right) P\left(E^- \mid X_n\right) =$
(since for any three events A,B,C, P(A|B,C) P(C|B) = P(A,C|B))

$$P(E^+, E^- \mid X_n)$$

so

$$P\left(E^+, E^- \mid X_n\right) = P\left(E^+ \middle| X_n\right) P\left(E^- \mid X_n\right)$$

, i.e. past and future are conditionally independent, given the present,

so dividing by $P\left(E^+ \middle| X_n\right)$ we get $P\left(E^- \middle| X_n, E^+\right) = P\left(E^- \middle| X_n\right)$

$$P\left(X_{n-1} \middle| X_n, X_{n+1}, \ldots, X_{n+j}\right) = P\left(X_{n-1} \middle| X_n\right)$$

This is the Markov property, applied to ***backward time***.

By Bayes law,

$$P\left(X_{n-1}\middle|X_{n}\right)=P\left(X_{n}\middle|X_{n-1}\right)\left[P\left(X_{n-1}\right)/P\left(X_{n}\right)\right].$$

Consider the 2-state chain with $X_0 = 1$ and transition matrix $P = \begin{bmatrix} 0.9 & 0.1 \\ 0.1 & 0.9 \end{bmatrix}$

$\vec{p}(1) = (0.9,\ 0.1)$ $\vec{p}(2) = (0.82,\ 0.18)$ $\vec{p}(3) = (0.756,\ 0.244)$

$$\left[P(X_0 \mid X_1)\right]_{\mathrm{ij}} = \mathrm{P}(X_0 = j \mid X_1 = i)$$

$$\left[P(X_0 \mid X_1)\right] = \begin{bmatrix} 1 & 0 \\ 1 & 0 \end{bmatrix}$$

$$\left[P(X_1 \mid X_2)\right] = \begin{bmatrix} 0.9 \bullet (.9/.82) & 0.1 \bullet (.1/.82) \\ 0.1 \bullet (.9/.18) & 0.9 \bullet (.1/.18) \end{bmatrix} = \begin{bmatrix} 0.988 & 0.012 \\ 0.5 & 0.5 \end{bmatrix}$$

By Bayes law, $P(X_{n-1}|X_n) = P(X_n|X_{n-1})P(X_{n-1})/\ P(X_n)$. Or, being more explicit about the initial distribution of states, we could write this as

$$P(X_{n-1}|X_n, X_0) = P(X_n|X_{n-1}, X_0)P(X_{n-1} \mid X_0)/P(X_n \mid X_0) =$$
$$P(X_n|X_{n-1})\left[P(X_{n-1} \mid X_0)/P(X_n \mid X_0)\right]$$

Thus backward running chain is not necessarily a homogeneous Markov chain just because the forward chain is homogeneous.

If we visualize the (forward running) Markov chain as starting at time $-\infty$ and being in steady state (or if we start the forward chain with steady state probabilities, then $P(X_{n-1} = i) = P(X_n = i) = \pi_i$. Then the transition probabilities for the backward Markov chain are (let $i = X_n, j = X_{n-1}$)

$$P_{ij}^* = P(X_{n-1} = j|X_n = i) = \pi_j P_{ji} / \pi_i$$

Definition: A Markov chain is **reversible** in steady state if $P_{ij}^* = P_{ij}$ for all i, j.

By definition, for general chain: $\pi_i P_{ij}^* = \pi_j P_{ji}$

For a reversible chain: $\pi_i P_{ij} = \pi_j P_{ji}$

Theorem: For an irreducible Markov chain, if set of numbers $\{\pi_i;\, i \geq 0\}$ exists satisfying $\pi_i \geq 0$ for all *i*, $\sum_i \pi_i = 1$, and $\pi_i P_{ij} = \pi_j P_{ji}$ for all *i*, *j*, the $\{\pi_i;\, i \geq 0\}$ is set of steady state probabilities and the chain is reversible.

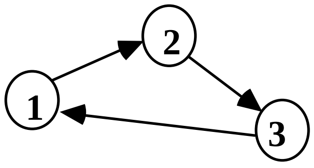
Proof: If $\pi_i P_{ij} = \pi_j P_{ji}$ for all *i*,*j*, then $\sum_j \pi_i P_{ij} = \sum_j \pi_j P_{ji} = \pi_i$, so steady state equations are satisfied. As in Birth Death, $P_{i,j} = P^*_{i,j}$

General Guessing theorem: If $\{\pi_i;\, i \geq 0\}$ is a probability vector, if $\{P^*_{ij}\}$ is a set of transition probabilities, and if $\pi_i P_{ij} = \pi_j P^*_{ji}$ for all *i*,*j*, then $\{\pi_i;\, i \geq 0\}$ is steady state probability vector for $\{P_{ij}\}$ and $\{P^*_{ij}\}$ is set of backward transition probabilities.

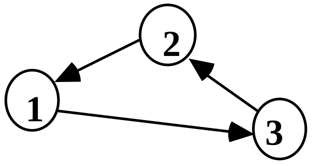
These results often allow one to "guess" the solution to steady state probabilities.

For any MC in steady state, the backward running chain is also a homogeneous MC. Usually, it is a MC with different transition probabilities. A reversible chain is a special case with same transition probabilities.

Example of MC that is not reversible:



Forward



Backward

Every Birth Death chain in steady state is reversible!

From Birth Death property $P_{i,i+1} = \pi_{i+1} P_{i+1,i} / \pi_i$

From calculation of reverse chain probabilities $P_{ij}^* = \pi_j P_{ji} / \pi_i$

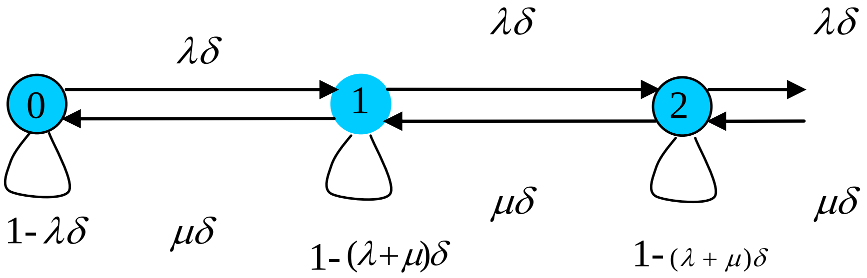
So, $P_{i,i+1} = P_{i,i+1}^*$ Likewise, $P_{i,i-1} = P_{i,i-1}^*$ Also $P_{ii} = P_{ii}^*$ and $P_{ij} = P_{ij}^* = 0$ for $|i-j| > 1$

More generally, every Markov chain whose graph is a tree or forest (no loops) is reversible!

Simple tests for non-reversibility: $P_{ij} > 0,\ P_{ji} = 0$ for some ij; or $P_{ij} P_{jk} P_{ki} \neq P_{ik} P_{kj} P_{ji}$

Sampled M/M/1 Queue

The state of the chain at iteration n is the total number of customers in an M/M/1 queue (i.e., in queue and in service) at time t = n. We choose $\delta \ll \min$ {mean interarrival time, mean service time}



Steady-State Probabilities

$$\lambda\delta\pi_0 = \mu\delta\pi_1$$

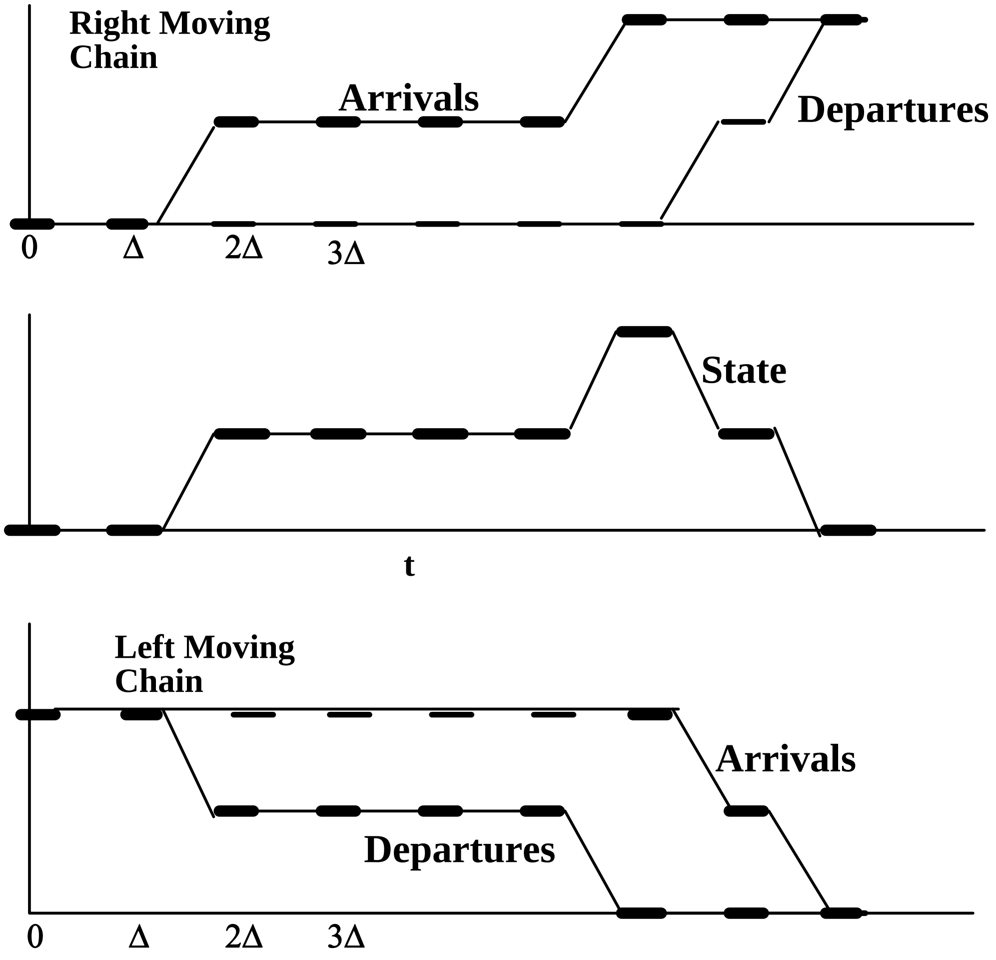
$$\pi_1 = ({}^{\lambda}\!/\!{}_{\mu})\pi_0 = \rho\pi_0$$

$$\pi_2 = ({}^{\lambda}\!/\!{}_{\mu})\pi_1 = \rho\pi_1 = \rho^2\pi_0$$

$$\pi_k = \rho^k\pi_0$$

$$\pi_k = (1-\rho)\rho^k$$

M/ M/ 1 Markov chain & Burke's theorem:



Correspondences Between Left-Moving and Right-Moving Chains

A) Both chains have sampled time intervals labelled in increasing order from left to right. But they are traversed from left to right for the right-moving chain and from right to left for the left-moving chain.

B) ***Departures*** for right-moving chain (transitions from i to i-1) are ***arrivals*** for left-moving chain. Arrivals for right-moving chain are departures for left-moving chain.

C) Any sequence of states for right-moving chain has ***same probability*** as that sequence of states for the left-moving chain, since chain is reversible and is in steady state.

Since left moving chain obeys same probabilistic description as right moving chain (because of reversibility), the arrivals in the left moving chain are Bernoulli, with rate λ.

Since each sample function of right moving chain corresponds to sample of left moving chain, probabilities associated with the departure process of right moving chain are the same as the probabilities associated with the arrivals of the left moving chain so **the departures process is Bernoulli with rate λ.**

That is, given state $i > 0$, the conditional probability of departure in next δ interval is $\mu\delta$, but unconditionally, the probability of departure is $\lambda\delta$.

$\lambda\delta < \mu\delta$ (since steady state exists) and unconditional result averages over empty state (with no departures) and busy states (with more rapid departures). P(departure next time) = P(departure next time | non-empty queue) P(non-empty queue) + P(departure next time | empty queue) P(empty queue) $= \mu\delta \cdot \rho + 0 \cdot (1-\rho) = \mu\delta \cdot \frac{\lambda}{\mu} = \lambda\delta$

From reversibility, we also have seen that these departures (unconditional on state) are independent of each other.

Since left moving chain is M/M/1, the arrivals in the left moving chain to the left of time n are independent of X_n.

Thus, for right moving chain, departures to left of time n are independent of X_n. Thus, state at time n is independent of the departure process previous to time n.

This is very **counter-intuitive**. It says that if I observe an unusually large number of departures in the interval before n, it tells me **nothing** about the state of the queue at n.

It does, of course, tell me that the queue was probably rather full during the interval with the large number of departures.

Burke's Theorem: Given a discrete-time M/M/1 chain in steady state, the departure process is Bernoulli and the state at time $n\delta$ is independent of the departures previous to $n\delta$.

Semi-Markov Process

This is a generalization of both a Markov chain and a renewal process. It has a countable or finite set of "states," $\{0, 1, \ldots\}$ or $\{1, \ldots, J\}$, and the times between state transitions are random variables.

Let $\{S_1, S_2, \ldots\}$ be the epochs at which transitions occur and let $U_n = S_n - S_{n-1}$ be the nth inter-transition interval.

Let X_n be the state entered at epoch S_n, and let $X(t) = X_n$ for $S_n \le t < S_{n+1}$.

The process is described stochastically by a set of transition probabilities, $P_{ij} = P(X_n = j \mid X_{n-1} = i)$. These describe a Markov chain known as the **embedded chain**. Also U_n is a non-negative rv with a conditional distribution

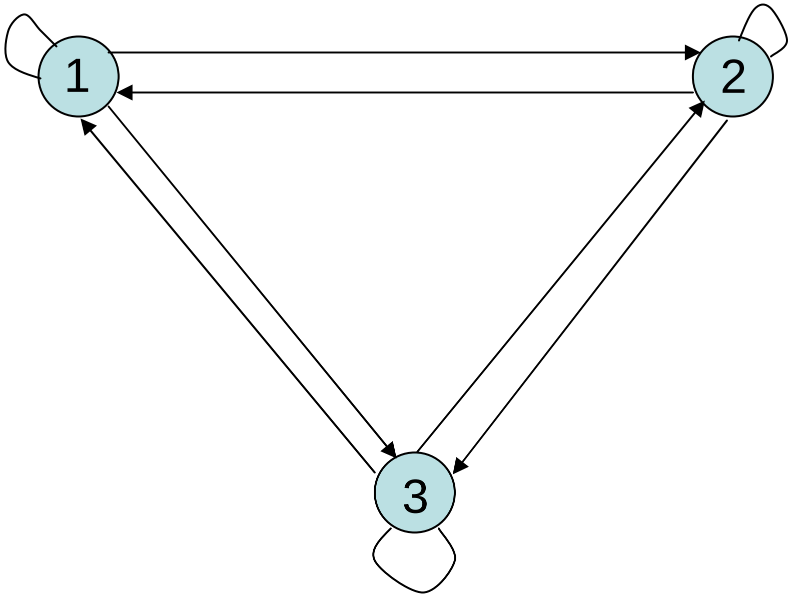
$$G_{ij}(u) = P(U_n \le u \mid X_{n-1} = i, X_n = j)$$

(Perhaps U_{ij} would be a better notation, but Gallager uses U_n.)

Visualize starting in state X_0 at time 0. The next state X_1 is chosen according to embedded chain, and then the time U_1 chosen, conditional on X_0 and X_1. The process remains in state X_0 for $0 \le t < U_1$ and then transitions to X_1 at $t = S_1 = U_1$. Next X_2 is chosen according to embedded chain, and then U_2 is chosen, conditional on X_1 and $X_2, \cdots$.

If there is only one state, then we have a simple renewal process. If the transition times are deterministic and equal, then we have a Markov chain. Alternatively, if we ignore transition times and only ask for the *sequence* of states, we also have a Markov chain.

Example

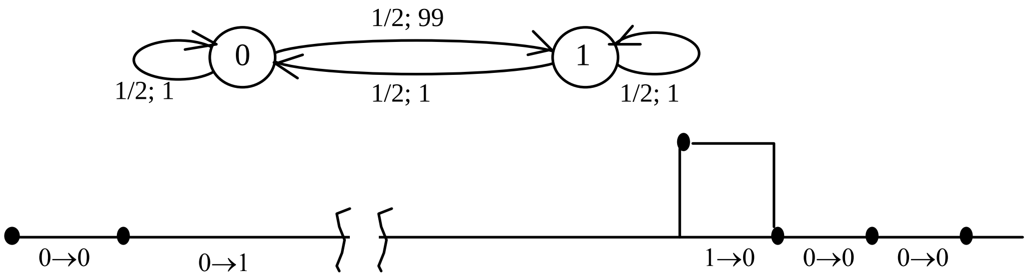


A taxi serves 3 cities. When it arrives in city 1 a new customer immediately asks to be taken to a location in city 1 with probability P_{11}, to a location in city 2 with probability P_{12} and to a location in city 3 with probability P_{13}. The durations of the taxi rides are the random variables U_{11}, U_{12} and U_{13}, respectively. We assume a new customer enters the moment the trip is up. The taxi location during the trip is counted as its city of origin, i.e., $X(t)=1$ for all t from arrival in city 1 up to the time it first arrives in another city.

Question – For the 2-state semi-Markov process below, is the present state the only thing worth knowing about the past for purposes of predicting the future? If you know the present state is 0, do you learn anything more by knowing it has been in state 0 for 95 seconds? The correct answer to this explains why these are called *semi*-Markov processes.

Fraction of Time in the State (p_j) versus Fraction of Transitions to the State (π_j)

Let π_i be steady state probability of state i in embedded chain. Let p_i be the steady state fraction of time spent in state i. This is <u>not</u> typically the same as π_i.



The M/G/1 Queue as a Semi-Markov Process

Consider the state of an M/G/1 queue as the number of customers left behind by a departing customer. Could we choose the number of customers in the system (queue + service) at each customer arrival time as the state of the queue (at arrival times)?

No, because you would also need to know the length of time the present customer has been in service to predict the time to the next departure, which you would need to know to have a complete statistical distribution.

Could we choose the number of customers in the system (queue + service) immediately after each departure as the state?

Consider the state of an M/G/1 queue as the number of customers left behind by a departing customer. We turn our thinking of the Semi-Markov Process around as we think of first picking an intertransition time and then deciding what state to go to. When $i \geq 1$ the distribution of the time spent between transitions is the service time distribution, call it $G(u)$. In this case, for $j \geq i-1$

$$P\left(X_{n+1}=j \mid X_n=i, U_n=u\right)=\begin{cases} \dfrac{(\lambda u)^{j-(i-1)} e^{-\lambda u}}{\left[j-(i-1)\right]!}, & \text{j} \geq \text{i-1} \\ 0 \qquad , & \text{j<i-1} \end{cases}$$

The unconditional probability P_{ij} is obtained by multiplying by the density $g(u)$ of U and integrating:

$$P_{ij}=\begin{cases} \displaystyle\int_0^{\infty} \frac{g(u)(\lambda u)^{j-(i-1)} e^{-\lambda u}}{\left[j-(i-1)\right]!} du, & j \geq i-1 \\ 0 \qquad , & \text{j<i-1} \end{cases}$$

$$P_{ij} = \begin{cases} \int_0^\infty \frac{g(u)(\lambda u)^{j-(i-1)} e^{-\lambda u}}{\left[j-(i-1)\right]!} du, & j \geq i-1,\ \mathrm{i} \geq 1 \\ 0 & , \qquad \mathrm{j<i-1} \end{cases}$$

When $i = 0$ the distribution of the time spent between transitions is the distribution of the sum of two random variables: a memoryless interarrival time and a service time. In this case, for $j \geq 0$ (i.e., for *j* additional arrivals during service time of first customer),

$$P(X_{n+1} = j \mid X_n = 0, U_n = u) = \frac{(\lambda u)^j e^{-\lambda u}}{j!}$$

$$\bar{U}(0) = 1/\lambda + \int_0^\infty \left[1 - G(u)\right] du$$

However, since the first customer departs at the next state-measurement time, the number of customers left is the number who arrive ***during*** the first customer's service time, and therefore

$$P_{oj} = \int_0^\infty \frac{g(u)(\lambda u)^j e^{-\lambda u}}{j!} du, \quad j \geq 0$$