Probability I
Lecture 11

Daniel Yu

October 29, 2024

Contents

1 Sojourn Time

Definition 1. For any state $i \in \Omega$, the sojurn time S_i is the first time the markov chain returns to state i if $x_0 = i$. Since this isn't deterministic, S_i is a random variable. We can try to describe the distribution of S_i i.e. $P[S_i = n] \forall n \geq 1$.

Note. This is hard

We will focus on finding the expectation $E[S_i \mid X_0 = i]$.

Remark. One approach is to consider the markov chain with S_i as an absorbing state then compute $E[T \mid X_0 = i]$= sum of row i in $N = (I - Q)^{-1}$ which is the expected number of steps until it is absorbed from state i. However, there is a far easier approach with stationary distributions.

Corollary 1. $E[S_i \mid X_0 = i] = \frac{1}{\pi_i}$. Additionally, $E[V(j,i) \mid X_0 = i] = \frac{\pi_j}{\pi_i}$ where π is the stationary distribution

Proof. Instead let's consider an indicator decomposition for $V(j,i)$.

1. $v(i,i) = 1$ with probability 1 ($x_0 = i$, and $x_n \neq i$ for any $n < S_i$, or in other words we stop after visiting i for the second time by problem construction) .
2. $v(j,i), i \neq j$. We can sum over time:

$$V(j,i) = \sum_{n}^{\infty} \nVdash |_{x_n = j, n < S_i} .$$

Using the linearity of expectation,

$$E[V(j,i) \mid X_0 = i] = \sum_{n=0}^{\infty} P[\{X_n = j\} \cap \{n < S_i\} \mid X_0 = i].$$

We can consider fixing $i \in \Omega$, and let V^i be a row vector such that $(V^i)_k = E[V(k,i) \mid X_0 = i] \forall k \in \Omega$.

Proposition 1. The claim is that for any state i, $v^i = v^i P$.

Proof. Let V^i be denoted as V for the notational purposes of this proof.

$$\begin{aligned}
(vP)_k &= [\begin{pmatrix} v_1 & v_2 & \dots & v_n \end{pmatrix} \cdot P]_k \\
&= \sum_{j \in \Omega} v_j \cdot P_{j,k} \\
&= \sum_{j \in \Omega} \left(\sum_{n=0}^{\infty} P[\{X_n = j\} \cap \{n < S_i\} \mid X_0 = i] \right) \cdot P_{j,k} \\
&= \sum_{j \in \Omega} \left(\sum_{n=0}^{\infty} P[\{X_n = j\}, \{X_{n+1}\}, n < S_i \mid X_0 = i] \right) \\
&= \sum_{n=0}^{\infty} P[\cup_j \{X_n = j\} \cap \{X_{n+1} = k\} \cap \{n < S_i\} \mid X_0 = i] \\
&= \sum_{n=0}^{\infty} P[1 \cap X_{n+1} = k, n < S_i \mid X_0 = i] \\
&= (vP)_k.
\end{aligned}$$

□

This breaks down into two cases:

1. Case 1: $k \neq i, \Rightarrow \{x_{n+1} = k \cap n < S_i\} = \{X_{n+1} = k \cap n+1 < S_i\}$. So, $(vP)_k = \sum_{n=0}^{\infty} P[X_{n+1} = k, n+1 < S_i \mid X_0 = i] = \sum_{m=1}^{\infty} P[X_m = k, m < S_i \mid X_0 = i] = \sum_{n=0}^{\infty} P[X_m = k, m < S_i \mid X_0 = i] = v_k$.
2. Case 2: $k = i \Rightarrow \{X_{n+1} = i, n < S_i\}$ with $S_i = n+1$, so $(vP)_i = \sum_{n=0}^{\infty} P[S_i = n+1 \mid X_0 = i] = P[S_i < \infty \mid X_0 = i] = 1 = v_i$

□

Recall that π is the unique row vector with $\pi = \pi P$ and $\sum_{i \in \Omega} \pi_i = 1$. Since $v = vP$ there must be a constant c such that $v = c \cdot \pi$. Since $\sum \pi_i = 1$,

$$\begin{aligned} c &= \sum_{j \in \Omega} v_j \\ &= \sum_{j \in \Omega} E[V(j,i) \mid X_0 = i] \\ &= E[\sum_{j \in \Omega} V(j,i) \mid X_0 = i] \\ &= E[S_i \mid X_0 = i]. \end{aligned}$$

So, $v^i = E[S_i \mid X_0 = i] \cdot \pi$. Then, $1 = (V^i)_i = E[S_i \mid X_0 = i] = \pi_i$. At $k \neq i, E[V(k,i) \mid X_0 = i] = (V^i)_k = E[S_i|X_0 = i] \cdot \pi_j = \frac{\pi_j}{\pi_i}$

2 Reversible Markov Chains

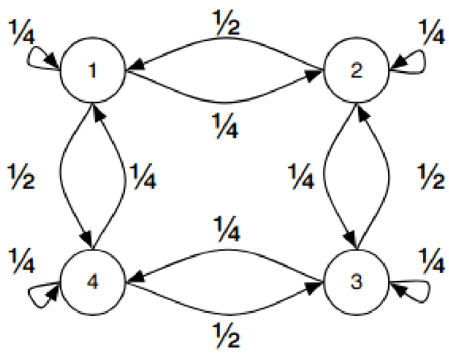

Figure 2: A non-reversible Markov chain

Note. Despite being irreducible and aperiodic the above is not a reversible markov chain.

Definition 2. A reversible markov chain is defined specifically as: let $P[X_0 = i] = \pi_i$ where π is the stationary distribution. Then,

$$P[X_{n-1} = j \mid X_n = i] = \frac{P[X_{n-1} = j, X_n = i]}{P[X_n = i]} = \frac{P[X_n = i \mid X_{n-1} = i] \cdot P[X_{n-1} = j]}{P[X_n = i]} = \frac{P_{i,j} \cdot \pi_j}{\pi_i}.$$

If $\frac{P_{j,i}\pi_j}{\pi_i} = P_{i,j}$. We call the chain reversible.

Corollary 2. An irreducible, aperiodic markov chain is reversible if $\pi_i P_{i,j} = \pi_j P_{j,i} \forall i, j \in \Omega$. flow in = flow out at equilibrium.

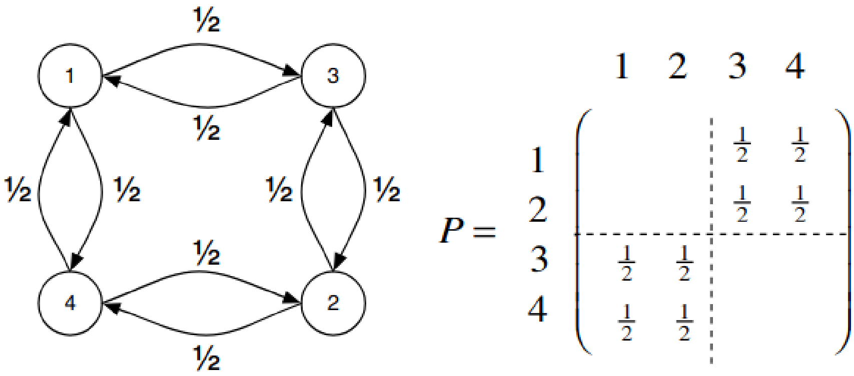

Figure 1: A reversible Markov chain and its transition probability matrix $\boldsymbol{P}$

Proposition 2. Let w be a row vector such that:

$$w_i P_{i,j} = w_j P_{j,i} \forall i, j \in \Omega.$$

Then $w = c\pi$ where $c = \sum_{i=1}^{n} w_i$. and the markov chain is reversible.

Proof. From a previous theorem it suffices to show that $w = wP$. Then,

$$\begin{aligned}
(wP)_k &= \sum_{j \in \Omega} w_j P_{j,k} \\
&= \sum_{j \in \Omega} w_k P_{k,j} \\
&= w_k \cdot \left(\sum_j P_{k,j} \right) \\
&= w_k.
\end{aligned}$$

w is the row eigenvector and a multiple of the stationary distribution! □

Note. Thus, we can find if a markov chain is reversible and find its stationary distribution by constructing w through some guess for w_i. The process will fail if the markov chain is not reversible

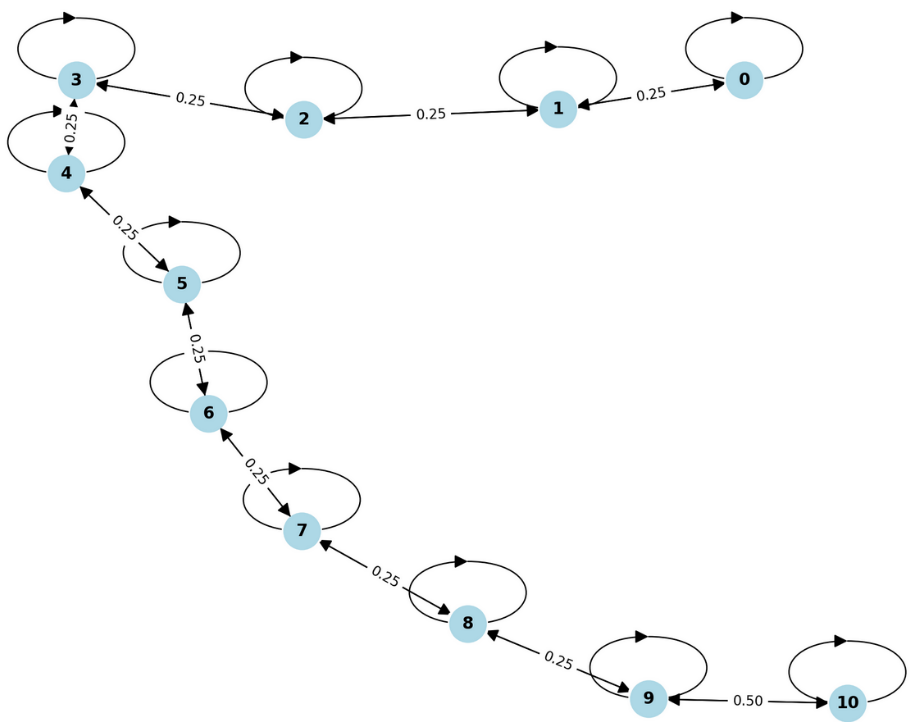

Figure 1: Reversible Markov Chain Problem P_{10}

We can construct the stationary distribution using w to satisfy $w_j P_{i,j} = w_j P_{j,i}$. We have a degree of freedom so let's let $w_0 = 1$. Then,

$$\begin{aligned} &w_0 \cdot P_{0,1} = w_1 \cdot P_{1,0} \\ &1 \cdot \frac{1}{2} = w_1 \cdot \frac{1}{4} \Rightarrow w_1 = 2 \\ &\ldots P_{0,3} = P_{3,0} = P_{n,0} = P_{0,n} = 0. \end{aligned}$$

Then generalizing,

$$w_1 \frac{1}{4} = w_2 \frac{1}{4} \Rightarrow w_2 = 2.$$

and so $w_1 = w_2 = \ldots = w_{n-1} = 2$. Then,

$$\begin{aligned} w_{n-1} \cdot P_{n,n-1} &= w_n P_{n,n-1} \\ 2 \cdot \frac{1}{4} &= w_n \frac{1}{2} \Rightarrow w_n = 1. \end{aligned}$$

Then, normalizing w to π, we get $\pi = \left(\frac{1}{2n}, \frac{1}{n}, \ldots, \frac{1}{n}, \frac{1}{2n}\right)$ is the stationary distribution P_n.

3 Classes of Reversible Markov chains

Definition 3. Let G a graph on n vertices $V = \{1, \ldots, n\}$ and $E =$ undirected pairs of vertices. Given G, we define markov chain where if $x_n = v$, $x_n + 1$ is uniformly chosen amongst the neighbors of v i.e.

$$P_{v,w} = \frac{1}{\text{edges adjacent to v}} = \frac{1}{\text{deg(v)}} \text{ if } (v, w) \in E.$$

Proposition 3. Every random walk on a graph is reversible.

Proof. Let $w_v =$ number of edges incident to v is . Then by construction,

$$w_v \cdot P_{v,w} = \begin{cases} 0, & \text{if w is not a neighbor of v} \\ 1, & \text{if } (v,w) \in E = w_w P_{w,v} \end{cases} .$$

And in particular,

$$\pi_v = \frac{w_v}{\sum_{w_v}} = \frac{\text{number of edges incident to v}}{2 \cdot \text{ simple edges } + \text{ self loops}}.$$

□