Jumping Into Markov Chains

A PRIMES Exposition

Ishita Goluguri, Toyesh Jayaswal
Mentor: Dr. Benjamin Landon
<table>
<thead>
<tr>
<th></th>
<th>Section Title</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Introduction</td>
</tr>
<tr>
<td>2</td>
<td>Hitting Times and Probabilities</td>
</tr>
<tr>
<td>3</td>
<td>Gambler’s Ruin</td>
</tr>
<tr>
<td>4</td>
<td>Recurrence and Transience</td>
</tr>
<tr>
<td>5</td>
<td>Random Walks on Lattices</td>
</tr>
<tr>
<td>6</td>
<td>Invariant Distributions and Consequences</td>
</tr>
</tbody>
</table>
A Markov process is characterized by the memoryless property that the future only depends on the current state and not on the previous states.

For example, a student may follow the chain below every 15 minutes.

Another example is radioactive decay where the time before the next particle decays does not depend on when the previous particles decayed.
**Definition (State Space)**

The state space $I$ is the set of all possible states of the Markov Chain.

**Definition (Measure and Distribution)**

A measure is a row vector $\lambda = (\lambda_i : i \in I)$ taking non-negative values in $\mathbb{R}$. A distribution is a measure with $\sum \lambda_i = 1$.

**Definition (Transition Matrix $P$)**

$P = (p_{ij} : i, j \in I)$, where $p_{ij}$ is the probability of jumping from state $i$ to state $j$. $p_{ij}^{(n)}$ is the probability of transitioning from $i$ to $j$ in $n$ steps and is the $ij$ entry of $P^n$.

**Definition (Markov Chain)**

A sequence of random variables $X_n$ taking values in $I$ is Markov($\lambda, P$) if $\mathbb{P}(X_0 = i) = \lambda_i$ and $\mathbb{P}(X_{n+1} = j | X_n = i) = p_{ij}$.
Hitting Times and Probabilities

Definition

For a certain subset $A \subset I$ and state $i \in I$, the hitting probability is defined as $h_i^A = \mathbb{P}_i$ (hit A) and the hitting time is defined as $k_i^A = \mathbb{E}_i$ (time to hit A).

The hitting probabilities satisfy

$$\begin{cases} h_i^A = 1 & i \in A \\ h_i^A = \sum_{j \in I} p_{ij} h_j^A & i \notin A \end{cases}$$

The hitting times satisfy

$$\begin{cases} k_i^A = 0 & i \in A \\ k_i^A = 1 + \sum_{j \in I} p_{ij} k_j^A & i \notin A \end{cases}$$

Moreover, $h_i^A$ and $k_i^A$ are the minimal non-negative solutions to these equations.
Proof for Hitting Probabilities

\[
\begin{cases}
h_i^A = 1 & i \in A \\ h_i^A = \sum_{j \in I} p_{ij} h_j^A & i \notin A
\end{cases}
\]

If \( i \in A \), \( h_i^A = 1 \) trivially.

for \( i \notin A \), let \( H_A(\omega) = \inf\{n | X_n(\omega) \in A\} \).

\[
h_i^A = \mathbb{P}_i(H_A < \infty)
\]

\[
\mathbb{P}_i(H_A < \infty | X_1 = j) = \mathbb{P}_j(H_A < \infty)
\]

\[
h_i^A = \sum_{j \in I} p_{ij} \mathbb{P}_j(H_A < \infty) = \sum_{j \in I} p_{ij} h_j^A
\]
In the chain above, what is the probability of getting to state 4 starting from state 2?

\[ h_4 = 1 \]
\[ h_1 = h_1 \]
\[ h_2 = \frac{h_1}{2} + \frac{h_3}{2} \]
\[ h_3 = \frac{h_2}{2} + \frac{h_4}{2} \]
\[ h_2 = \frac{1}{3} \]
Gamblers’ Ruin

Problem

Imagine that you enter a casino with a fortune of $i$ and gamble, $1$ at a time, with probability $p$ of doubling your stake and probability $q$ of losing it. What is the probability that you leave broke?

Let $h_i = P_i$(hitting 0).

We get the system $h_0 = 1, h_i = ph_{i+1} + qh_{i-1}$.

General solution: $h_i = A + B \left( \frac{q}{p} \right)^i$ If $p < q$, then $B = 0$, so $h_i = 1$. Similarly, if $p = q$, then $h_i = A + Bi$, and $B = 0$ once again. Thus, $h_i = 1$.

If $p > q$, then $h_i = \left( \frac{q}{p} \right)^i + A \left( 1 - \left( \frac{q}{p} \right)^i \right)$, with the minimal nonnegative solution being $h_i = \left( \frac{q}{p} \right)^i$.
Recurrence and Transience

Definition (Recurrence)
A state $i$ is recurrent if $P_i(\{t \geq 0 : X_t = i\} \text{ is unbounded}) = 1$

Definition (Transience)
A state $i$ is transient if $P_i(\{t \geq 0 : X_t = i\} \text{ is unbounded}) = 0$

Definition (Communicating States)
State $i$ communicates with state $j$ if $P_i(X_n = j \text{ for } n \geq 0) > 0$ and $P_j(X_n = i \text{ for } n \geq 0) > 0$

Communicating is an equivalence relation, and partitions the state space into communicating classes. If $I$ is a single class, $P$ is said to be irreducible.

Definition (Closed Communicating Class)
A communicating class is closed if $i \in C$ and $i$ communicates with $j$ implies that $j \in C$
Properties of Recurrence and Transience

Figure: Communicating classes: {1,2,3}, {4}, {5, 6}

Definition (First Passage Time to State $i$)

$$T_i(\omega) = \inf\{n \geq 1 : X_n(\omega) = i\}$$

Definition (Return probability)

$$f_i = \mathbb{P}_i(T_i < \infty)$$
Definition (Number of visits $V_i$)

\[ V_i = \sum_{n=0}^{\infty} 1_{X_n = i}, \mathbb{E}_i(V_i) = \sum_{n=0}^{\infty} p_{ii}^{(n)} \]

Theorem

- If $f_i = 1$, then $i$ is recurrent and $\sum_{n=1}^{\infty} p_{ii}^{(n)} = \infty$
- If $f_i < 1$, then $i$ is transient and $\sum_{n=1}^{\infty} p_{ii}^{(n)} < \infty$

Proof.

If $\mathbb{P}_i(T_i < \infty) = 1$, then $\mathbb{P}_i(V_i = \infty) = 1$, so $i$ is recurrent, and $\sum_{n=0}^{\infty} p_{ii}^{(n)} = \mathbb{E}_i(V_i) = \infty$.

If $f_i = \mathbb{P}_i(T_i < \infty) < 1$, then

\[ \sum_{n=0}^{\infty} p_{ii}^{(n)} = \mathbb{E}_i(V_i) = \sum_{r=0}^{\infty} \mathbb{P}_i(V_i > r) = \sum_{r=0}^{\infty} f_i^r = \frac{1}{1 - f_i} < \infty \]
Theorem

All states in a communicating class are either recurrent or transient

Proof.

Take $i, j \in C$ and assume $i$ is transient. Thus, there exist $n, m \geq 0$ such that $p_{ij}^{(n)} > 0$ and $p_{ji}^{(m)} > 0$. For all $r \geq 0$,

$$p_{ii}^{(m+n+r)} \geq p_{ij}^{(n)} p_{jj}^{(r)} p_{ji}^{(m)} ,$$

so

$$\sum_{r=0}^{\infty} p_{jj}^{(r)} \leq \frac{1}{p_{ij}^{(n)} p_{ji}^{(m)}} \sum_{r=0}^{\infty} p_{ii}^{(n+r+m)} < \infty$$
Given an odd sequence, $p_{00}^{(2n+1)} = 0$ for all $n$.
Given an even sequence of length $2n$, the probability of having $n$ steps up and $n$ steps down is \( \binom{2n}{n} p^n q^n \).
By Stirling’s formula,
\[
n! \approx \sqrt{2\pi n} \left(\frac{n}{e}\right)^n \quad \text{as } n \to \infty
\]
Thus,
\[
p_{00}^{(2n)} = \frac{(2n)!}{(n!)^2} (pq)^n \approx \frac{(4pq)^n}{A \sqrt{n/2}} \quad \text{as } n \to \infty
\]
For $p = q = \frac{1}{2}$,
\[
p_{00}^{(2n)} \geq \frac{1}{2\sqrt{2\pi n}} \quad \text{so } \sum_{n=N}^{\infty} p_{00}^{(2n)} \geq \frac{1}{2\sqrt{2\pi}} \sum_{n=N}^{\infty} \frac{1}{\sqrt{n}} = \infty
\]
so the random walk on $\mathbb{Z}$ is recurrent.
If $p \neq q$, $4pq < 1$, so

$$\sum_{n=N}^{\infty} p_{00}^{(2n)} \leq \frac{1}{\sqrt{2\pi}} \sum_{n=N}^{\infty} (4pq)^n < \infty$$

so this walk is transient.
Random Walk on $\mathbb{Z}^3$

$$p_{ij} = \begin{cases} 
\frac{1}{6} & \text{if } |i - j| = 1 \\
0 & \text{otherwise}
\end{cases}$$

Again, with an odd step sequence, $p_{00}^{(2n+1)} = 0$.

With an even step sequence, we must have $i$ steps up and down, $j$ steps north and south, $k$ steps east and west with $i + j + k = n$.

$$p_{00}^{(2n)} = \sum_{i+j+k=n} \frac{(2n)!}{(i!j!k!)^2} \left( \frac{1}{6} \right)^{2n}$$

$$= \binom{2n}{n} \left( \frac{1}{2} \right)^{2n} \sum_{i+j+k=n} \binom{n}{ijk} \left( \frac{1}{3} \right)^{2n}$$

For $n = 3m$, $\binom{n}{ijk} \leq \binom{n}{mmm}$, so

$$p_{00}^{(2n)} \leq \binom{2n}{n} \left( \frac{1}{2} \right)^{2n} \binom{n}{mmm} \left( \frac{1}{3} \right)^n \approx \frac{1}{2\sqrt{2\pi^3}} \left( \frac{6}{n} \right)^{3/2} \text{ as } n \to \infty$$
Random Walk on $\mathbb{Z}^3$ Continued

\[ \sum_{n=0}^{\infty} p_{00}^{(6m)} < \infty \text{ because } \sum n^{-3/2} \text{ converges. Since } p_{00}^{(6m)} \geq \left(\frac{1}{6}\right)^4 p_{00}^{(6m-4)} \text{ and} \]

\[ \sum_{n=0}^{\infty} p_{00}^{(n)} < \infty \]

and this walk is transient.
A measure $\pi$ is called invariant if $\pi P = \pi$.

For a process $X$ which is Markov($\pi$, $P$) for an invariant distribution $\pi$, $X_n$ is also Markov($\pi$, $P$) for all $n$.

For a fixed state $k$, let $\gamma_i^k = \mathbb{E}_k \sum_{n=0}^{T_k-1} 1\{X_n=i\}$ be the expected number of visits to $i$ between visits to $k$. $\gamma^k$ turns out to be an invariant measure with $\gamma_k^k = 1$.

If $\sum_i \gamma_i^k = m_k$, which is the expected return time to $k$, is finite (positive recurrence), $\gamma^k / m_k$ is an invariant distribution.

If a chain is irreducible and positive recurrent, the invariant measure turns out to be unique up to scaling and in this case, $\pi_k = \frac{1}{m_k}$.

**Theorem (Convergence to Equilibrium)**

If $P$ is irreducible and aperiodic, $\mathbb{P}(X_n = j) \to \pi_j$ as $n \to \infty$ for all $j$ regardless of the initial distribution.

Periodic case: $\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$
Theorem (Ergodic Theorem)

Let $P$ be irreducible and positive recurrent and let $X$ be Markov($\lambda$, $P$). Then, for any bounded function $f : I \rightarrow \mathbb{R}$,

$$P \left( \frac{1}{n} \sum_{k=1}^{n-1} f(X_k) \rightarrow \bar{f} \text{ as } n \rightarrow \infty \right) = 1$$

where $\bar{f} = \sum_{i \in I} \pi_i f(i)$ regardless of the initial distribution.
An opera singer is due to perform a long series of concerts. She is liable to pull out each night with probability $1/2$. The promoter sends her flowers every day until she returns. Flowers costing $x$ thousand dollars, $0 \leq x \leq 1$, bring about a reconciliation with probability $\sqrt{x}$. The promoter stands to make $750$ from each successful concert. How much should he spend on flowers?

$$P = \begin{pmatrix} 1/2 & 1/2 \\ \sqrt{x} & 1 - \sqrt{x} \end{pmatrix}$$

$$\lambda_1 = \lambda_1/2 + \sqrt{x}\lambda_2$$

$$\lambda_2 = \lambda_1/2 + (1 - \sqrt{x})\lambda_2$$

$$\lambda_1 + \lambda_2 = 1$$

$$\lambda_1 = \frac{2\sqrt{x}}{2\sqrt{x} + 1}, \lambda_2 = \frac{1}{2\sqrt{x} + 1}$$

$$\frac{1}{n} \sum_{k=0}^{n-1} f(X_k) \to 750\lambda_1 - 1000x\lambda_2 = \frac{1500\sqrt{x} - 1000x}{2\sqrt{x} + 1}$$

$$x = 1/4, \mathbb{E}(f) \to \$250$$
• We would like to thank our parents for their endless support
• We would like to thank Dr Slava Gerovitch, Dr Tanya Khovanova, Dr Pavel Etingof, and the entire PRIMES program
• We would like to thank our mentor Dr. Benjamin Landon for teaching and mentoring us