Markov Chains

A stochastic process is a collection of random variables over a probability space. In this class, we look at discrete time: $X_0, X_1, X_2, \ldots$

X_0: state of process at time 0

$\vdots$

X_n: state of process at time n

The random variables $X_0, X_1, \ldots$ represent the states of the process. We call the set of possible states the state space and denote it as S.

Def (Markov Property) A process $X_0, X_1, \ldots$ obeys the Markov Property if for every possible sequence of values $i_0, i_1, \ldots, i_n, i_{n+1}$:

$$\mathbb{P}(X_{n+1} = i_{n+1} \mid X_0 = i_0, X_1 = i_1, \ldots, X_n = i_n) = \mathbb{P}(X_{n+1} = i_{n+1} \mid X_n = i_n)$$

That is, for each $n \geq 0$, the distribution of X_{n+1} given $X_0, X_1, \ldots, X_n$ depends only on X_n.

Ex Consider the process of flipping a coin that flips heads with probability p until we see two consecutive heads.

This process obeys the Markov Property, so it is a Markov chain.

$S = \{H, T, HH\}$

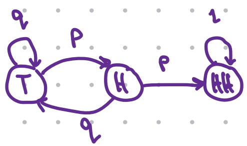
$\mathbb{P}(T \mid H) = 1 - p = q$

$\mathbb{P}(T \mid T) = 1 - p = q$

$\mathbb{P}(H \mid T) = p$

$\mathbb{P}(HH \mid H) = p$

$\mathbb{P}(HH \mid HH) = 1$



Claim For $X_0, X_1, \ldots$ a Markov chain on S and $i_0 i_1 \ldots i_n$ a sequence of states visited,

$$\mathbb{P}(X_0 = i_0, X_1 = i_1, \ldots, X_n = i_n) = \mathbb{P}(X_0 = i_0) \cdot \mathbb{P}(X_1 = i_1 \mid X_0 = i_0) \cdot \mathbb{P}(X_2 = i_2 \mid X_0 = i_0, X_1 = i_1) \cdot \ldots \cdot \mathbb{P}(X_n = i_n \mid X_0 = i_0, \ldots, X_{n-1} = i_{n-1})$$

$$= \underbrace{\mathbb{P}(X_0 = i_0)}_{\text{initial distribution}} \cdot \mathbb{P}(X_1 = i_1 \mid X_0 = i_0) \cdot \mathbb{P}(X_2 = i_2 \mid X_1 = i_1) \cdot \ldots \cdot \mathbb{P}(X_n = i_n \mid X_{n-1} = i_{n-1})$$

Ex Consider the process of flipping a coin that flips heads with probability p until we see two consecutive heads.

$\mathbb{P}(X_0 = H) = p \qquad \mathbb{P}(X_0 = T) = q$

$\mathbb{P}(HTTHH) = p \cdot q \cdot q \cdot p \cdot p$

Transition Matrix

Def (Transition Matrix) The one-step transition matrix of a chain is the matrix $\mathbb{P}$ such that
$$\mathbb{P}(i,j) = \mathbb{P}(X_1 = j \mid X_0 = i) \quad \text{(probability of } i \to j \text{ in one step)}$$

Note:

① $\mathbb{P}$ is a square matrix

② Each row of $\mathbb{P}$ is a distribution:
$$\sum_j \mathbb{P}(i,j) = 1 \quad \text{for any } i$$

Ex Consider the process of flipping a coin that flips heads with probability p until we see two consecutive heads.

$$\mathbb{P} = \begin{array}{c} \\ T \\ H \\ HH \end{array} \begin{array}{c} \begin{array}{ccc} T & H & HH \end{array} \\ \begin{bmatrix} q & p & 0 \\ q & 0 & p \\ 0 & 0 & 1 \end{bmatrix} \end{array}$$

Claim The n-step transition matrix $\mathbb{P}_n$ is given by $\mathbb{P}_n(i,j) = \mathbb{P}(X_n = i \mid X_0 = j)$ (probability of $i \to j$ in n steps). Then
$$\begin{aligned}
\mathbb{P}_2(i,j) &= \mathbb{P}(X_2 = j \mid X_0 = i) \\
&= \sum_{k \in S} \mathbb{P}(X_1 = k, X_2 = j \mid X_0 = i) \\
&= \sum_{k \in S} \mathbb{P}(X_1 = k \mid X_0 = i) \cdot \mathbb{P}(X_2 = j \mid X_0 = i, X_1 = k) = \sum_k \mathbb{P}(i,k) \cdot \mathbb{P}(k,j) \\
&= \mathbb{P}^2(i,j)
\end{aligned}$$

By induction, $\mathbb{P}_n(i,j) = \mathbb{P}^n(i,j)$, so $\mathbb{P}_n = \mathbb{P}^n$

Ex Consider the process of flipping a coin that flips heads with probability p until we see two consecutive heads.

$$\mathbb{P} = \begin{array}{c} \\ T \\ H \\ HH \end{array} \begin{array}{c} \begin{array}{ccc} T & H & HH \end{array} \\ \begin{bmatrix} q & p & 0 \\ q & 0 & p \\ 0 & 0 & 1 \end{bmatrix} \end{array} \qquad \mathbb{P}_2 = \mathbb{P}^2 = \begin{bmatrix} q & p & 0 \\ q & 0 & p \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} q & p & 0 \\ q & 0 & p \\ 0 & 0 & 1 \end{bmatrix} = \begin{array}{c} \\ T \\ H \\ HH \end{array} \begin{array}{c} \begin{array}{ccc} T & H & HH \end{array} \\ \begin{bmatrix} q^2 + pq & qp & p^2 \\ q^2 & qp & p \\ 0 & 0 & 1 \end{bmatrix} \end{array}$$

Distribution over Time

Thm Let π_0 be the initial distribution over the state space written as a row vector:

$$\pi_0(i) = \mathbb{P}(X_0 = i)$$

For example, with $S = \{1, 2, \dots, N\}$,

$$\pi_0 = [\mathbb{P}(X_0=1) \quad \mathbb{P}(X_0 = 2) \quad \dots \quad \mathbb{P}(X_0 = N)]$$

And let π_n be the distribution over the state space after $n \geq 0$ steps.

For example, with $S = \{1, 2, \dots, N\}$,

$$\pi_n = [\mathbb{P}(X_n=1) \quad \mathbb{P}(X_n = 2) \quad \dots \quad \mathbb{P}(X_0 = N)]$$

Then

$$\pi_n = \pi_0 \mathbb{P}_n = \pi_0 \mathbb{P}^n$$

Pf For $i \in S$,

$$\pi_n(i) = \sum_{k \in S} \mathbb{P}(X_0 = k) \cdot \mathbb{P}(X_n = i \mid X_0 = k)$$

$$= \sum_{k \in S} \pi_0(k) \cdot \mathbb{P}_n(k, i)$$

$$= (\pi_0 \mathbb{P}_n)(i)$$

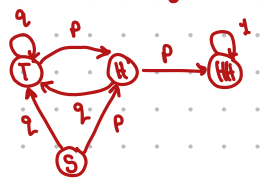
Note To specify a Markov chain, you need

- S, the state space
- $\mathbb{P}$, the transition probabilities
- π_0, the initial distribution.

Hitting Time

Q Suppose you repeatedly flip a coin with probability p of heads until you see two consecutive heads. What is the expected number of flips it will take?

The process is given by the Markov chain



Let R be the number of flips it takes to see two consecutive heads.

We can find the expected time until we see HH by conditioning. Let $\beta(i)$ be the expected time until HH starting from state i, i.e. $\beta(i) = \mathbb{E}[R \mid X_0 = i]$

$$\left.\begin{aligned}
\beta(HH) &= 0 \\
\beta(H) &= 1 + p\cdot\beta(HH) + q\cdot\beta(T) \\
\beta(T) &= 1 + p\cdot\beta(H) + q\cdot\beta(T) \\
\beta(S) &= 1 + p\cdot\beta(H) + q\cdot\beta(T)
\end{aligned}\right\} \text{First Step Equations}$$

① $\beta(H) = 1 + p\cdot 0 + q\beta(T) = 1 + q\beta(T)$

$\beta(T) - q\beta(T) = p\beta(T) = 1 + p\cdot\beta(H)$, so $\beta(T) = \frac{1}{p} + \beta(H)$

$\beta(S) = \beta(T)$

② $\beta(H) = 1 + q\left(\frac{1}{p} + \beta(H)\right) = \frac{p}{p} + \frac{q}{p} + q\beta(H) = \frac{1}{p} + q\beta(H)$

$\Rightarrow \beta(H) - q\beta(H) = p\beta(H) = \frac{1}{p}$

$\Rightarrow \beta(H) = \frac{1}{p^2}$

$\Rightarrow \beta(T) = \frac{1}{p} + \frac{1}{p^2} = \beta(S)$

So $\mathbb{E}[R] = \frac{1}{p} + \frac{1}{p^2}$

Note Let $X_0, X_1, \ldots$ be a finite Markov chain on state space S with transition matrix $\mathbb{P}$.
Let $A \subset S$ and $\beta(i)$ be the expected time to reach a state in A from state i.
Then

$$\beta(i) = 0 \quad \text{if} \quad i \in A$$

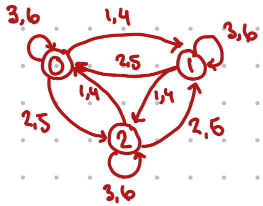
$$\beta(i) = 1 + \sum_{j \in S} \mathbb{P}(i,j)\cdot\beta(j)$$

A Before B

Q We repeatedly roll a six-sided die and sum the rolls modulo 3 as we go.
What is the chance our sum hits 1 before it hits 2?

$S = \{0, 1, 2\}$ the value of the sum

all transitions have probability $\frac{2}{6} = \frac{1}{3}$



Leverage Markov Property: let $\alpha(i) = \mathbb{P}(1 \text{ before } 2 \mid \text{in state } i)$.

$\alpha(0) = \frac{1}{3}\cdot\alpha(0) + \frac{1}{3}\cdot\alpha(1) + \frac{1}{3}\cdot\alpha(2) \quad \Rightarrow \quad \frac{2}{3}\alpha(0) = \frac{1}{3}$

$\alpha(1) = 1 \quad \Rightarrow \quad \alpha(0) = \frac{1}{2}$

$\alpha(2) = 0$

Q Consider a sequence of iid trials, each of which results in n mutually exclusive categories outcomes. On each trial, let the chance of category i be $p_i > 0$.
What is the chance category i appears before category j?

$S = \{i, j, k\}$ the category, where k is any category other than i or j

For $\alpha(m) = \mathbb{P}(i \text{ before } j \mid \text{in state } m)$,

$\alpha(i) = 1$

$\alpha(j) = 0$

$\alpha(k) = p_i\alpha(i) + p_j\alpha(j) + (1 - p_i - p_j)\alpha(k)$

$\Rightarrow \alpha(k) = \dfrac{p_i}{p_i + p_j}$

Note Let $X_0, X_1, \dots$ be a finite Markov chain on state space S with transition matrix $\mathbb{P}$.
Let $A, B \subset S$ be mutually exclusive and $\alpha(i)$ be the probability of hitting A before B from state i.
Then

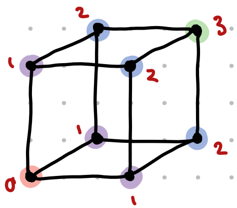
$\alpha(i) = 1 \quad \text{if } i \in A$

$\alpha(i) = 0 \quad \text{if } i \in B$

$\alpha(i) = \sum_{k \in S} \mathbb{P}(i,k)\,\alpha(k) \quad \text{otherwise}$

Examples

Q An ant is sitting at the corner of a cube. At each timestep, she traverses an edge uniformly at random. What is the expected time until she reaches the other end of the cube?



Rather than defining a state for each corner, define a state as her distance from the start.
$S = \{0, 1, 2, 3\}$;

Then

$\beta(0) = 1 + \beta(1)$

$\beta(1) = 1 + \frac{1}{3}\beta(0) + \frac{2}{3}\beta(2)$

$\beta(2) = 1 + \frac{2}{3}\beta(1) + \frac{1}{3}\beta(3)$

$\beta(3) = 0$

$\Rightarrow \beta(1) = 1 + \frac{1}{3} + \frac{1}{3}\beta(1) + \frac{2}{3}\beta(2)$

$\Rightarrow \beta(1) = 2 + \beta(2)$

$\Rightarrow \beta(2) = 1 + \frac{4}{3} + \frac{2}{3}\beta(2) + 0$

$\Rightarrow \beta(2) = 3 + 4 = 7$

$\Rightarrow \beta(1) = 2 + 7 = 9$

$\Rightarrow \beta(0) = 10$