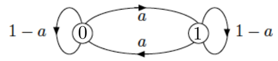
CS70: Lecture 24.

Markov Chains

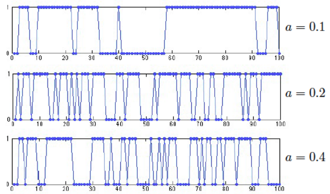
1. Examples
2. Definition
3. First Passage Time
4. Distribution

Two-State Markov Chain

Here is a symmetric two-state Markov chain. It describes a random motion in $\{0,1\}$. Here, a is the probability that the state changes in the next step.

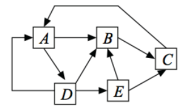


Let's simulate the Markov chain:

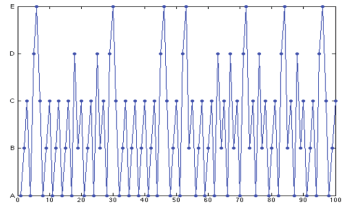


Five-State Markov Chain

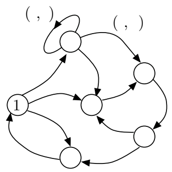
At each step, the MC follows one of the outgoing arrows of the current state, with equal probabilities.



Let's simulate the Markov chain:



Finite Markov Chain: Definition



- A finite set of states: $\mathcal{X} = \{1, 2, \ldots, K\}$
- A probability distribution π_0 on $\mathcal{X} : \pi_0(i) \geq 0, \sum_i \pi_0(i) = 1$
- Transition probabilities: $P(i,j)$ for $i, j \in \mathcal{X}$

  $P(i,j) \geq 0, \forall i,j; \sum_j P(i,j) = 1, \forall i$
- $\{X_n, n \geq 0\}$ is defined so that

  $Pr[X_0 = i] = \pi_0(i), i \in \mathcal{X}$ (initial distribution)

  $Pr[X_{n+1} = j \mid X_0, \ldots, X_n = i] = P(i,j), i, j \in \mathcal{X}.$

Poll

Xn is a MC with transition matrix P. Select all true statements.

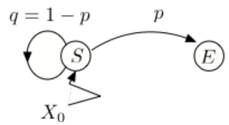
- Row sum for each row of P is 1.
- Column sum of each column is 1.
- Sum of all elements of P is equal to the number of states of the MC.
- X3 is independent of X0

First Passage Time - Example 1

Let's flip a coin with $Pr[H] = p$ until we get H. How many flips, on average?

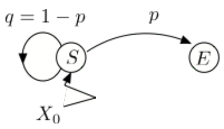
Let's define a Markov chain:

- $X_0 = S$ (start)
- $X_n = S$ for $n \geq 1$, if last flip was T and no H yet
- $X_n = E$ for $n \geq 1$, if we already got H (end)



First Passage Time - Example 1

Let's flip a coin with $Pr[H] = p$ until we get H. How many flips, on average?



Let $\beta(S)$ be the average time until E, starting from S.

Then,

$$\beta(S) = 1 + q\beta(S) + p0.$$

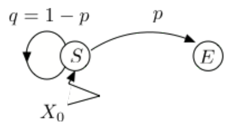
(See next slide.) Hence,

$$p\beta(S) = 1, \text{ so that } \beta(S) = 1/p.$$

Note: Time until E is $G(p)$. We have rediscovered that the mean of $G(p)$ is $1/p$.

First Passage Time - Example 1

Let's flip a coin with $Pr[H] = p$ until we get H. How many flips, on average?



Let $\beta(S)$ be the average time until E.
Then,

$$\beta(S) = 1 + q\beta(S) + p0.$$

Justification: Let N be the random number of steps until E, starting from S. Let also N' be the number of steps until E, after the second visit to S. Finally, let $Z = 1\{\text{first flip } = H\}$. Then,

$$N = 1 + (1 - Z) \times N' + Z \times 0.$$

Now, Z and N' are independent. Also, $E[N'] = E[N] = \beta(S)$. Hence, taking expectation,

$$\beta(S) = E[N] = 1 + (1 - p)E[N'] + p0 = 1 + q\beta(S) + p0.$$

First Passage Time - Example 2

Let's flip a coin with $Pr[H] = p$ until we get two consecutive Hs. How many flips, on average?

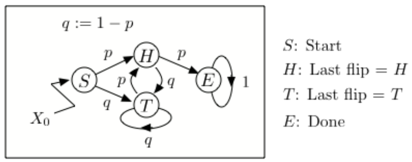
$$H\ T\ H\ T\ T\ T\ H\ T\ H\ T\ H\ T\ T\ T\ H\ T\ H\ H$$

Let's define a Markov chain:

- $X_0 = S$ (start)
- $X_n = E$, if we already got two consecutive Hs (end)
- $X_n = T$, if last flip was T and we are not done
- $X_n = H$, if last flip was H and we are not done

First Passage Time - Example 2

Let's flip a coin with $Pr[H] = p$ until we get two consecutive Hs. How many flips, on average? Here is a picture:



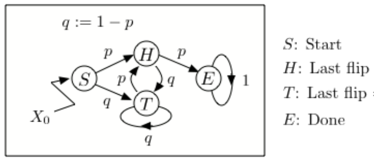
Let $\beta(i)$ be the average time from state i until the MC hits state E.

We claim that (these are called the first step equations)

$$\beta(S) = 1 + p\beta(H) + q\beta(T)$$
$$\beta(H) = 1 + p0 + q\beta(T)$$
$$\beta(T) = 1 + p\beta(H) + q\beta(T).$$

Solving, we find $\beta(S) = 2 + 3qp^{-1} + q^2p^{-2}$. (E.g., $\beta(S) = 6$ if $p = 1/2$.)

First Passage Time - Example 2



Let us justify the first step equation for $\beta(T)$. The others are similar.

Let $N(T)$ be the random number of steps, starting from T until the MC hits E. Let also $N(H)$ be defined similarly. Finally, let $N'(T)$ be the number of steps after the second visit to T until the MC hits E. Then,

$$N(T) = 1 + Z \times N(H) + (1 - Z) \times N'(T)$$

where $Z = 1\{\text{first flip in } T \text{ is } H\}$. Since Z and $N(H)$ are independent, and Z and $N'(T)$ are independent, taking expectations, we get

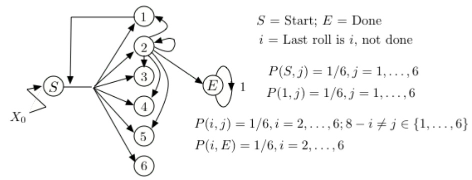
$$E[N(T)] = 1 + pE[N(H)] + qE[N'(T)],$$

i.e.,

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

First Passage Time - Example 3

You roll a balanced six-sided die until the sum of the last two rolls is 8. How many times do you have to roll the die, on average?



The arrows out of $3, \ldots, 6$ (not shown) are similar to those out of 2.

$$\beta(S) = 1 + \frac{1}{6}\sum_{j=1}^{6}\beta(j); \beta(1) = 1 + \frac{1}{6}\sum_{j=1}^{6}\beta(j); \beta(i) = 1 + \frac{1}{6}\sum_{j=1,\ldots,6; j\neq 8-i}\beta(j), i = 2, \ldots, 6.$$

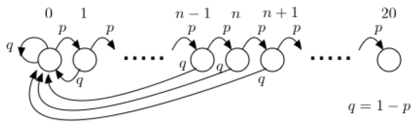
Symmetry: $\beta(2) = \cdots = \beta(6) =: \gamma$. Also, $\beta(1) = \beta(S)$. Thus,

$$\beta(S) = 1 + (5/6)\gamma + \beta(S)/6; \quad \gamma = 1 + (4/6)\gamma + (1/6)\beta(S).$$

$$\Rightarrow \cdots \beta(S) = 8.4.$$

First Passage Time - Example 4

You try to go up a ladder that has 20 rungs. At each time step, you succeed in going up by one rung with probability $p = 0.9$. Otherwise, you fall back to the ground. How many time steps does it take you to reach the top of the ladder, on average?



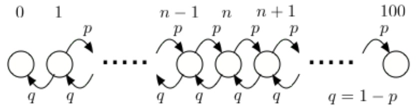
$$\beta(n) = 1 + p\beta(n+1) + q\beta(0), 0 \le n < 19$$
$$\beta(19) = 1 + p0 + q\beta(0)$$

$$\Rightarrow \beta(0) = \frac{p^{-20} - 1}{1 - p} \approx 72.$$

See MC Note for algebra.

First Passage Time - Example 5

You play a game of "heads or tails" using a biased coin that yields 'heads' with probability $p < 0.5$. You start with \$10. At each step, if the flip yields 'heads', you earn \$1. Otherwise, you lose \$1. What is the probability that you reach \$100 before \$0?



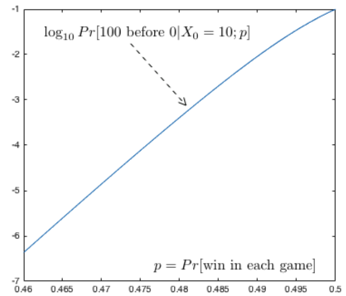
Let $\alpha(n)$ be the probability of reaching 100 before 0, starting from n, for $n = 0, 1, \ldots, 100$.

$$\alpha(0) = 0; \alpha(100) = 1.$$
$$\alpha(n) = p\alpha(n+1) + q\alpha(n-1), 0 < n < 100.$$

$$\Rightarrow \alpha(n) = \frac{1 - \rho^n}{1 - \rho^{100}} \text{ with } \rho = qp^{-1}. \text{ (See MC Note)}$$

First Passage Time - Example 5

You play a game of "heads or tails" using a biased coin that yields 'heads' with probability 0.48. You start with \$10. At each step, if the flip yields 'heads', you earn \$1. Otherwise, you lose \$1. What is the probability that you reach \$100 before \$0?



Morale of example: Be careful!

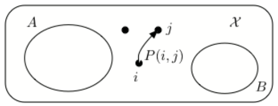
Summary of First Step Equations

Let X_n be a MC on $\mathscr{X}$ and $A, B \subset \mathscr{X}$ with $A \cap B = \emptyset$. Define

$$T_A = \min\{n \ge 0 \mid X_n \in A\} \text{ and } T_B = \min\{n \ge 0 \mid X_n \in B\}.$$

Let

$$\beta(i) = E[T_A \mid X_0 = i] \text{ and } \alpha(i) = Pr[T_A < T_B \mid X_0 = i], i \in \mathscr{X}.$$



The FSE are

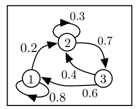
$$\beta(i) = 0, i \in A$$
$$\beta(i) = 1 + \sum_j P(i,j)\beta(j), i \notin A$$
$$\alpha(i) = 1, i \in A$$
$$\alpha(i) = 0, i \in B$$
$$\alpha(i) = \sum_j P(i,j)\alpha(j), i \notin A \cup B.$$

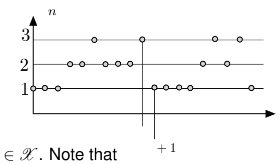
Poll

Xn is a MC. Select all true statements.

- FSEs can be used for calculating hitting time and absorption probabilities.
- FSEs for hitting time are of the form of 1 + weighted sum of hitting times from the neighbors.
- FSEs for absorption probabilities are of the form of 1 + weighted sum of absorption probabilities from the neighbors.

Distribution of X_n





Let $\pi_m(i) = Pr[X_m = i], i \in \mathscr{X}$. Note that

$$\begin{aligned} Pr[X_{m+1} = j] &= \sum_i Pr[X_{m+1} = j, X_m = i] \\ &= \sum_i Pr[X_m = i] Pr[X_{m+1} = j \mid X_m = i] \\ &= \sum_i \pi_m(i) P(i,j). \end{aligned}$$

Hence,

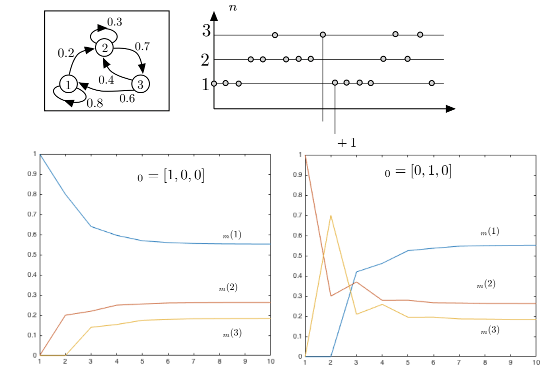
$$\pi_{m+1}(j) = \sum_i \pi_m(i) P(i,j), \forall j \in \mathscr{X}.$$

With π_m, π_{m+1} as a row vectors, these identities are written as $\pi_{m+1} = \pi_m P$.

Thus, $\pi_1 = \pi_0 P$, $\pi_2 = \pi_1 P = \pi_0 PP = \pi_0 P^2, \ldots$. Hence,

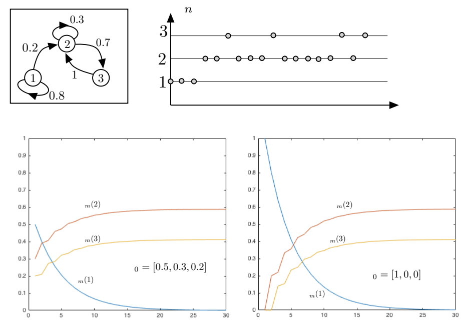
$$\pi_n = \pi_0 P^n, n \ge 0.$$

Distribution of X_n



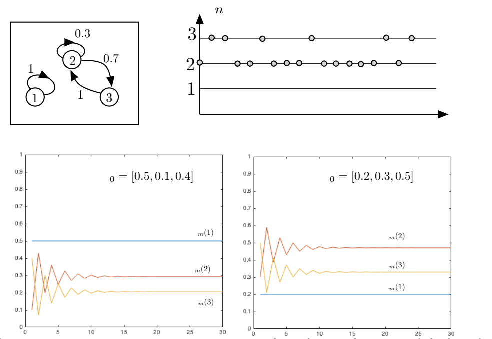
As m increases, π_m converges to a vector that does not depend on π_0.

Distribution of X_n



As m increases, π_m converges to a vector that does not depend on π_0.

Distribution of X_n



As m increases, π_m converges to a vector that depends on π_0 (obviously, since $\pi_m(1) = \pi_0(1), \forall m$).

Poll

Xn is a MC. Select all true statements.

- ”Steady State” probabilities are always independent of the initial probability distribution.
- Under certain conditions, ”Steady State” probabilities are indpependent of the initial probability distribution.

Summary

Markov Chains

- Definition: Markov Chains: State Space, Initial Distribution, Transition Probabilities
- FSEs: Hitting Times, Probability of A before B
- Distribution at time n: May or may not depend on the initial distribution.