CS70 - Lecture 25 Notes

Name: Felix Su SID: 25794773

Spring 2016 GSI: Gerald Zhang

Markov Chain Review

1. Markov Chain:
   - Finite MC set $\mathscr{X}$
   - Initial Distribution π_0
   - Transition Probabilities $P = \{P(i,j), i,j \in \mathscr{X}\}$
   - $\Pr[X_0 = i] = \pi_0(i), i \in \mathscr{X}$
   - $\Pr[X_{n+1} = j | X_0, \ldots, X_n = i] = P(i,j), i,j \in \mathscr{X}, n \geq 0$
   - Note: $\Pr[X_0 = i_0, X_1 = i_1, \ldots, X_n = i_n] = \pi_0(i_0)P(i_0,i_1)P(i_{n-1},i_n)$.
2. First Passage Time:
   - $A \cap B = \emptyset; \beta(i) = E[T_A | X_0 = i]; \alpha(i) = \Pr[T_A < T_B | X_0 = i]$
   - $\beta(i) = 1 + \sum_j P(i,j)\beta(j); \alpha(i) = \sum_j P(i,j)\alpha(j)$

First Passage Time:
Given disjoint sets of states $A \cap B = \emptyset$
Expected Timesteps to get to state in A

$$\beta(i) = E[T_A | X_0 = i] = 1 + \sum_j P(i,j)\beta(j) \tag{1}$$

Probability of reaching A before B

$$\Pr[T_A < T_B | X_0 = i] \tag{2}$$

Distribution of X_n

1. Use $\pi_n = \pi_0 P^n$ function to check if it converges to a vector that does depend on π_0 or not

- Let $\pi_m(i) = \Pr[X_m = i], i \in X$. Note that
- $\Pr[X_{m+1} = j] = \sum_i \Pr[X_{m+1} = j, X_m = i]$
  - $= \sum_i \Pr[X_m = i]\Pr[X_{m+1} = j | X_m = i]$
  - $= \sum_i \pi_m(i)P(i,j)$
  - Hence, $\pi_{m+1}(j) = \sum_i \pi_m(i)P(i,j), \forall j \in X$.
- 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, $\pi_n = \pi_0 P^n, n \geq 0$

Distribution of X_n
Given that $\pi_m(i) = \Pr[X_m = i], i \in \mathscr{X}$

$$\pi_{m+1}(j) = \sum_i \pi_m(i) P(i,j), \forall j \in X \tag{3}$$

With π_m, π_{m+1} as row vectors

$$\pi_{m+1} = \pi_m P \tag{4}$$

General case of π_n

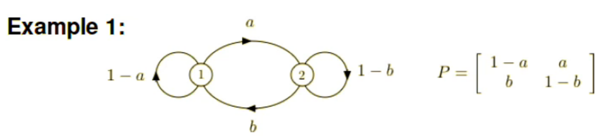
$$\pi_n = \pi_0 P^n, n \geq 0 \tag{5}$$

Balance Equations

1. A distribution π_0 such that $\pi_m = \pi_0, \forall m$ is said to be an invariant distribution ($\pi_0 P = \pi_0$).

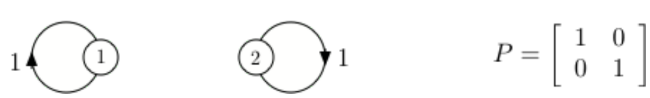
- **Theorem** A distribution π_0 is **invariant** iff $\pi_0 P = \pi_0$. These equations are called the balance equations.
- **Proof:** $\pi_n = \pi_0 P^n$, so that $\pi_n = \pi_0, \forall n$ iff $\pi_0 P = \pi_0$
  - π_0 is invariant $\implies$ the distribution of X_n is always equal to X_0.
  - This does not mean that X_n does not move. It means that the probability that it leaves a state i is equal to the probability that it enters state i.
  - The balance equations say that $\sum_j \pi_(j) P(j,i) = \pi_(i)$
  - That is, $\sum_{j \neq i} \pi_(j) P(j,i) = \pi_(i)(1 - P(i,i)) = \pi_(i) \sum_{j \neq i} P(i,j)$.
  - Thus, $\Pr[\text{enter } i] = \Pr[\text{leave } i]$.

Example 1:



$\pi P = \pi \iff \begin{bmatrix} \pi(1) & \pi(2) \end{bmatrix} \begin{bmatrix} 1-a & a \\ b & 1-b \end{bmatrix} = \begin{bmatrix} \pi(1) & \pi(2) \end{bmatrix}$
$\iff \pi(1)(1-a) + \pi(2)b = \pi(1)$ and $\pi(1)a + \pi(2)(1-b) = \pi(2)$
$\iff \pi(1)a = \pi(2)b$
Equations are redundant, so add an equation: $\pi(1) + \pi(2) = 1$. Then we find
$\pi = \begin{bmatrix} \frac{b}{a+b} & \frac{a}{a+b} \end{bmatrix}$

Example 2:



$\pi P = \pi \iff \begin{bmatrix} \pi(1), \pi(2) \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$
$= \begin{bmatrix} \pi(1), \pi(2) \end{bmatrix} \iff \pi(1) = \pi(1)$ and $\pi(2) = \pi(2)$
Every distribution is invariant for this Markov chain. This is obvious, since $X_n = X_0$ for all n. Hence, $\Pr[X_n = i] = \Pr[X_0 = i], \forall (i, n)$

Irreducibility

> 1. MC is **irreducible** if it can go from every state i to every state j in any amount of steps

Existence and Uniqueness of Invariant Distribution

- **Theorem:** A **finite irreducible** Markov chain has **one and only one invariant distribution**.
  - There is a unique positive vector $\pi = \begin{bmatrix}\pi(1) & \dots & \pi(K)\end{bmatrix}$ such that $\pi P = \pi$ and $\sum_k \pi(k) = 1$
- **Fact:** If a Markov chain has **two different invariant distributions** π and ν, then it has **infinitely many invariant distributions**.
  * $p\pi + (1-p)\nu$ is then invariant since $[p\pi + (1-p)\nu]P = p\pi P + (1-p)\nu P = p\pi + (1-p)\nu$

Finite irreducible Markov chain has one and only one invariant distribution
There is a unique positive vector $\pi = \begin{bmatrix}\pi(1) & \dots & \pi(K)\end{bmatrix}$ such that

$$\pi P = \pi \text{ and } \sum_k \pi(k) = 1 \tag{6}$$

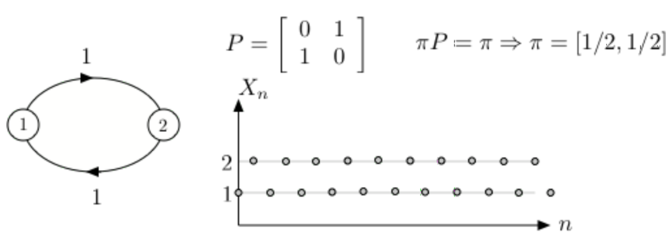
Long Term Fraction of Time in States

- **Theorem** Let X_n be an irreducible Markov chain with invariant distribution π.
- Then, for all i, $\frac{1}{n}\sum_{m-0}^{n-1} 1\{X_m = i\} \to \pi(i)$, as $n \to \infty$.
- The left-hand side is the fraction of time that $X_m = i$ during steps $0, 1, \dots, n-1$. Thus, this fraction of time approaches $\pi(i)$.

Long Term Fraction of Time in States

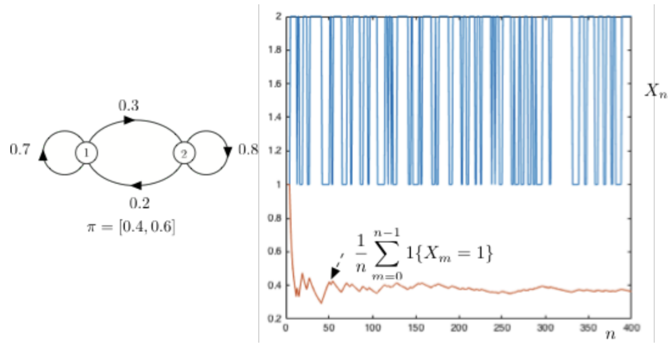
$$\text{for all } i, \frac{1}{n}\sum_{m-0}^{n-1} 1\{X_m = i\} \to \pi(i) \text{ as } n \to \infty \tag{7}$$

Example 1



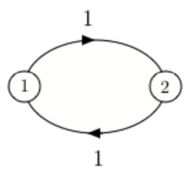
- The fraction of time in state 1 converges to 1/2, which is $\pi(1)$.

Example 2

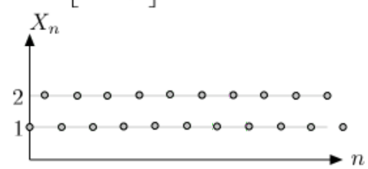


Convergence to Invariant Distribution

- Assuming the MC is irreducible π_n does not necessarily approach a unique invariant distribution π
- **Example:**



$$P = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} \qquad \pi P = \pi \Rightarrow \pi = [1/2, 1/2]$$



> Assume $X_0 = 1$. Then $X_1 = 2, X_2 = 1, X3_{=}2, \ldots$
> Thus, if $\pi 0 = [1, 0], \pi 1 = [0, 1], \pi 2 = [1, 0], \pi 3 = [0, 1]$, etc.
> Hence, π_n does not converge to $\pi = [1/2, 1/2]$.

Periodicity

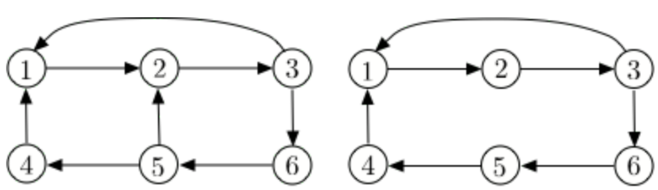
> 1. If the Markov chain is irreducible, $d(i)$ is the same for all i. Check for 1 state.

- **Definition** If $d(i) = 1$, the Markov chain is said to be **aperiodic**.
  - Otherwise, it is periodic with period $d(i)$.
- **Theorem** (see below)
  - Gcd of the set of all numbers of steps it takes to go from from state i, back to state i where the probability of that path is greater that 0
  - Proof: See Lecture notes 24.

Theorem: Periodicity
Assume that the MC is irreducible

$$d(i) := g.c.d.\{n > 0 | \Pr[X_n = i | X_0 = i] > 0\} \text{ has the same value for all states } i \tag{8}$$

Example



> $[A] : \{n > 0 | \Pr[X_n = 1 | X_0 = 1] > 0\} = \{3, 6, 7, 9, 11, ...\} \implies d(1) = 1.$
> $\{n > 0 | \Pr[X_n = 2 | X_0 = 2] > 0\} = \{3, 4, ...\} \implies d(2) = 1.$
> $[B] : \{n > 0 | \Pr[X_n = 1 | X_0 = 1] > 0\} = \{3, 6, 9, ...\} \implies d(i) = 3.$
> $\{n > 0 | \Pr[X_n = 5 | X_0 = 5] > 0\} = \{6, 9, ...\} \implies d(5) = 3.$

Convergence of π_n

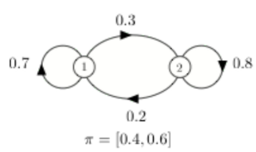
- Irreducible MC $\implies$ fraction of time spent in state i is equal to the invariant probability of that state
- Irreducible + Aperiodic MC $\implies$ fraction of time spent in state i is equal to and converges to the invariant probability of that state
- **Proof:** See EE126, or Lecture notes 24.

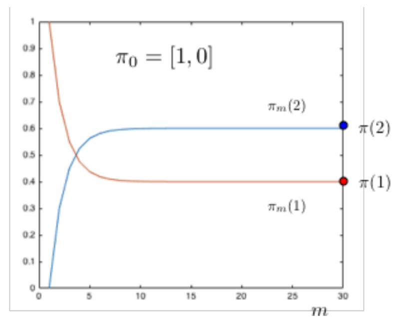
Theorem: Convergence of π_n
Let X_n ba an irreducible and aperiodic MC with invariant distribution π

$$\text{For all } i \in X, \pi_n(i) \to \pi(i), \text{ as } n \to \infty \tag{9}$$

Example 1

- Irreducible + Aperiodic MC $\implies$ fraction of time spent in state i is equal to and converges to the invariant probability of that state

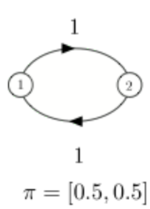


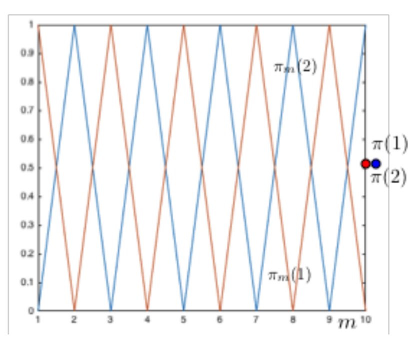


Example 2

- Irreducible MC $\implies$ fraction of time spent in state i is equal to the invariant probability of that state

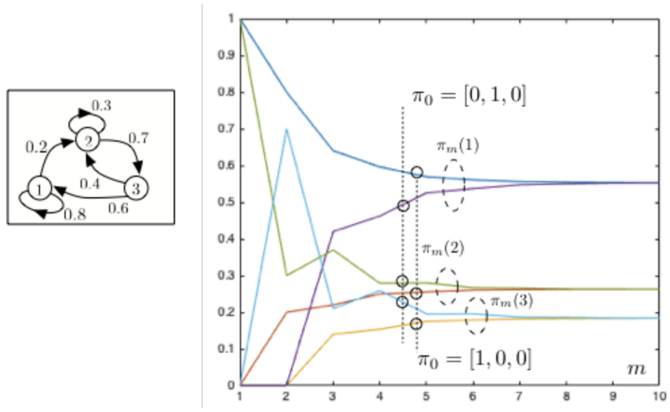
m





Example 3

- Loop implies aperiodicity



Calculating π

Method

> 1. Let P be irreducible
> 2. $\pi P = \pi \implies \pi[P - I] = 0$
> 3. Replace the last equation with ones $\pi 1 = 1$ to get $\pi P_1 = [0,0,1]$
>   - Observe the sum of the columns of $P - I = 0$, which shows the equations are redundant, which means the equations are redundant
> 4. Solve $\pi = [0,0,1]P_1^{-1}$

Example:

Let P be irreducible. Find π where $P = \begin{bmatrix} 0.8 & 0.2 & 0 \\ 0 & 0.3 & 0.7 \\ 0.6 & 0.4 & 0 \end{bmatrix}$

> One has $\pi P = \pi$, i.e., $\pi[P - I] = 0$ where I is the identity matrix:
> $\pi \begin{bmatrix} 0.8-1 & 0.2 & 0 \\ 0 & 0.3-1 & 0.7 \\ 0.6 & 0.4 & 0-1 \end{bmatrix} = [0,0,0].$
> However, the sum of the columns of $P - I$ is 0. This shows that these equations are redundant: If all but the last one hold, so does the last one. Let us replace the last equation by $\pi 1 = 1$, i.e., $\sum_j \pi(j) = 1$:
> $\pi \begin{bmatrix} 0.8-1 & 0.2 & 1 \\ 0 & 0.3-1 & 1 \\ 0.6 & 0.4 & 1 \end{bmatrix} = [0,0,1].$
> Hence, $\pi = [0,0,1] \begin{bmatrix} 0.8-1 & 0.2 & 1 \\ 0 & 0.3-1 & 1 \\ 0.6 & 0.4 & 1 \end{bmatrix}^{-1} \approx [0.55, 0.26, 0.19]$

Summary: Markov Chains

> 1. Markov Chain: $\Pr[X_n + 1 = j | X_0, \dots, X_n = i] = P(i,j)$
> 2. FSE: $\beta(i) = 1 + \sum_j P(i,j)\beta(j); \alpha(i) = \sum_j P(i,j)\alpha(j)$.
> 3. $\pi_n = \pi_0 P^n$
> 4. π is invariant iff $\pi P = \pi$

5. Irreducible $\implies$ one and only one invariant distribution π
6. Irreducible $\implies$ fraction of time in state i approaches $\pi(i)$
7. Irreducible + Aperiodic $\implies$ $\pi_n \to \pi$.
8. Calculating π: One finds $\pi = [0, 0. \ldots, 1]Q - 1$ where $Q = \cdots$.