ECEN 743: Reinforcement Learning

Markov Chains Review

Dileep Kalathil
Assistant Professor
Department of Electrical and Computer Engineering
Texas A&M University

References

- D. Bertsekas, and J. Tsitsiklis, *Introduction to Probability*, Athena Scientific, 2008.

Markov Chain

- Let $\{X_0, X_1, X_2, \ldots\}$ be a sequence of discrete random variables

Markov Chain

- Let $\{X_0, X_1, X_2, \ldots\}$ be a sequence of discrete random variables
- Assume that each X_t takes values from a finite set $\mathcal{X}$ (state space)

Markov Chain

- Let $\{X_0, X_1, X_2, \ldots\}$ be a sequence of discrete random variables
- Assume that each X_t takes values from a finite set $\mathcal{X}$ (state space)

Definition (Markov Chain)

The sequence $\{X_0, X_1, X_2, \ldots\}$ is a Markov chain if it satisfies the Markov condition

$$\mathbb{P}(X_t = y | X_{t-1} = x_{t-1}, X_{t-2} = x_{t-2}, \ldots, X_0 = x_0) = \mathbb{P}(X_t = y | X_{t-1} = x_{t-1})$$

for all $t \geq 1$ and for all $y, x_0, \ldots, x_{t-1} \in \mathcal{X}$

Markov Chain

- Let $\{X_0, X_1, X_2, \ldots\}$ be a sequence of discrete random variables
- Assume that each X_t takes values from a finite set $\mathcal{X}$ (state space)

Definition (Markov Chain)

The sequence $\{X_0, X_1, X_2, \ldots\}$ is a Markov chain if it satisfies the Markov condition

$$\mathbb{P}(X_t = y | X_{t-1} = x_{t-1}, X_{t-2} = x_{t-2}, \ldots, X_0 = x_0) = \mathbb{P}(X_t = y | X_{t-1} = x_{t-1})$$

for all $t \geq 1$ and for all $y, x_0, \ldots, x_{t-1} \in \mathcal{X}$

- Conditional distribution of any future state given the past states and current state is independent of the past states, and depends only on the current state

Transition Probability Matrix

- Homogeneous Markov chain: A Markov chain is homogeneous if

$$\mathbb{P}(X_t = y | X_{t-1} = x) = \mathbb{P}(X_1 = y | X_0 = x), \quad \forall t \geq 1, \forall x, y \in \mathcal{X}$$

Transition Probability Matrix

- Homogeneous Markov chain: A Markov chain is homogeneous if

$$\mathbb{P}(X_t = y | X_{t-1} = x) = \mathbb{P}(X_1 = y | X_0 = x), \quad \forall t \geq 1, \forall x, y \in \mathcal{X}$$

  - For a homogeneous Markov chain, transition probability depends only on the states, not on the time index t

Transition Probability Matrix

- Homogeneous Markov chain: A Markov chain is homogeneous if

$$\mathbb{P}(X_t = y | X_{t-1} = x) = \mathbb{P}(X_1 = y | X_0 = x), \quad \forall t \geq 1, \forall x, y \in \mathcal{X}$$

  - For a homogeneous Markov chain, transition probability depends only on the states, not on the time index t

- Transition probability matrix: For a homogeneous Markov chain, transition probability matrix $P = [p_{ij}]$ is defined as, $p_{ij} = \mathbb{P}(X_t = j | X_{t-1} = i)$

Transition Probability Matrix

- Homogeneous Markov chain: A Markov chain is homogeneous if

$$\mathbb{P}(X_t = y | X_{t-1} = x) = \mathbb{P}(X_1 = y | X_0 = x), \quad \forall t \geq 1, \forall x, y \in \mathcal{X}$$

  - For a homogeneous Markov chain, transition probability depends only on the states, not on the time index t
- Transition probability matrix: For a homogeneous Markov chain, transition probability matrix $P = [p_{ij}]$ is defined as, $p_{ij} = \mathbb{P}(X_t = j | X_{t-1} = i)$
- Properties of transition probability matrix P:

Transition Probability Matrix

- Homogeneous Markov chain: A Markov chain is homogeneous if

$$\mathbb{P}(X_t = y | X_{t-1} = x) = \mathbb{P}(X_1 = y | X_0 = x), \quad \forall t \geq 1, \forall x, y \in \mathcal{X}$$

  - For a homogeneous Markov chain, transition probability depends only on the states, not on the time index t
- Transition probability matrix: For a homogeneous Markov chain, transition probability matrix $P = [p_{ij}]$ is defined as, $p_{ij} = \mathbb{P}(X_t = j | X_{t-1} = i)$
- Properties of transition probability matrix P:
  - P has non-negative entries

Transition Probability Matrix

- Homogeneous Markov chain: A Markov chain is homogeneous if

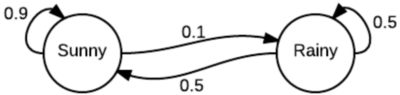
$$\mathbb{P}(X_t = y | X_{t-1} = x) = \mathbb{P}(X_1 = y | X_0 = x), \quad \forall t \geq 1, \forall x, y \in \mathcal{X}$$

  - For a homogeneous Markov chain, transition probability depends only on the states, not on the time index t
- Transition probability matrix: For a homogeneous Markov chain, transition probability matrix $P = [p_{ij}]$ is defined as, $p_{ij} = \mathbb{P}(X_t = j | X_{t-1} = i)$
- Properties of transition probability matrix P:
  - P has non-negative entries
  - P has row sums equal to one, $\sum_j p_{ij} = 1$

Markov Chain Representation

- Markov chains with small state space are often represented by directed graphs

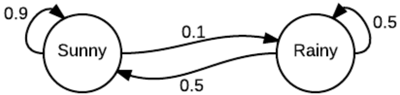
$$P = \begin{bmatrix} 0.9 & 0.1 \\ 0.5 & 0.5 \end{bmatrix}$$



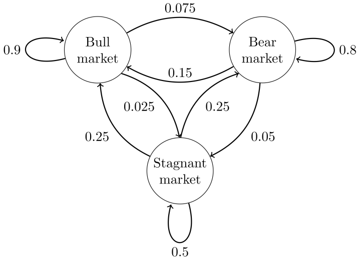
Markov Chain Representation

- Markov chains with small state space are often represented by directed graphs

$$P = \begin{bmatrix} 0.9 & 0.1 \\ 0.5 & 0.5 \end{bmatrix}$$



$$P = \begin{bmatrix} 0.9 & 0.075 & 0.025 \\ 0.15 & 0.8 & 0.05 \\ 0.25 & 0.25 & 0.5 \end{bmatrix}$$



[Figures are taken from Wikipedia]

One Step to Many Steps

- Let $p(t)$ be the row vector such that $p_i(t) = \mathbb{P}(X_t = i)$
- Given $p(t)$, how do we compute $p(t+1), p(t+2), \ldots, p(t+m)$?

One Step to Many Steps

- Let $p(t)$ be the row vector such that $p_i(t) = \mathbb{P}(X_t = i)$
- Given $p(t)$, how do we compute $p(t+1), p(t+2), \ldots, p(t+m)$?

Proposition

Let P be the transition probability matrix of a Markov chain with state space $\mathcal{X} = \{1, 2, \ldots, n\}$. Let $p(t)$ be the row vector such that $p_i(t) = \mathbb{P}(X_t = i)$. Then,

$$p(t+1) = p(t)\ P, \quad p(t+m) = p(t) P^m$$

One Step to Many Steps

- Let $p(t)$ be the row vector such that $p_i(t) = \mathbb{P}(X_t = i)$
- Given $p(t)$, how do we compute $p(t+1), p(t+2), \ldots, p(t+m)$?

Proposition

Let P be the transition probability matrix of a Markov chain with state space $\mathcal{X} = \{1, 2, \ldots, n\}$. Let $p(t)$ be the row vector such that $p_i(t) = \mathbb{P}(X_t = i)$. Then,

$$p(t+1) = p(t)\ P, \quad p(t+m) = p(t)P^m$$

Proof:

$$p_j(t+1) = \mathbb{P}(X_{t+1} = j) = \sum_{i \in \mathcal{X}} \mathbb{P}(X_{t+1} = j | X_t = i)\mathbb{P}(X_t = i)$$

$$= \sum_{i \in \mathcal{X}} p_{ij}\ p_i(t) = p(t)P$$

$\square$

Classification of States

Classification of States

- Reachable states: State j is reachable from state i if $\mathbb{P}(X_t = j | X_0 = i) > 0$ for some $t \geq 1$. Equivalently, if $P_{ij}^t > 0$ for some $t \geq 1$.

Classification of States

- Reachable states: State j is reachable from state i if $\mathbb{P}(X_t = j | X_0 = i) > 0$ for some $t \geq 1$. Equivalently, if $P_{ij}^t > 0$ for some $t \geq 1$.
- Communicating states: States i and j are communicating if i is reachable from j and j is reachable from i.

Classification of States

- Reachable states: State j is reachable from state i if $\mathbb{P}(X_t = j | X_0 = i) > 0$ for some $t \geq 1$. Equivalently, if $P_{ij}^t > 0$ for some $t \geq 1$.
- Communicating states: States i and j are communicating if i is reachable from j and j is reachable from i.
- Communicating class: A communicating (recurrent) class is a *maximal* set of states C such that every pair of states in C communicates with each other.

Classification of States

- Reachable states: State j is reachable from state i if $\mathbb{P}(X_t = j | X_0 = i) > 0$ for some $t \geq 1$. Equivalently, if $P_{ij}^t > 0$ for some $t \geq 1$.
- Communicating states: States i and j are communicating if i is reachable from j and j is reachable from i.
- Communicating class: A communicating (recurrent) class is a *maximal* set of states C such that every pair of states in C communicates with each other.
- Irreducible Markov chain: A Markov chain is irreducible if its state space is a single communicating class

Classification of States

- Reachable states: State j is reachable from state i if $\mathbb{P}(X_t = j | X_0 = i) > 0$ for some $t \geq 1$. Equivalently, if $P_{ij}^t > 0$ for some $t \geq 1$.
- Communicating states: States i and j are communicating if i is reachable from j and j is reachable from i.
- Communicating class: A communicating (recurrent) class is a *maximal* set of states C such that every pair of states in C communicates with each other.
- Irreducible Markov chain: A Markov chain is irreducible if its state space is a single communicating class
- Absorbing state: State i is absorbing if $P_{ij} = 0$ for all $j \in \mathcal{X}$

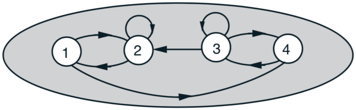
Classification of States

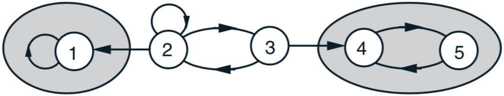
- Reachable states: State j is reachable from state i if $\mathbb{P}(X_t = j | X_0 = i) > 0$ for some $t \geq 1$. Equivalently, if $P_{ij}^t > 0$ for some $t \geq 1$.
- Communicating states: States i and j are communicating if i is reachable from j and j is reachable from i.
- Communicating class: A communicating (recurrent) class is a *maximal* set of states C such that every pair of states in C communicates with each other.
- Irreducible Markov chain: A Markov chain is irreducible if its state space is a single communicating class
- Absorbing state: State i is absorbing if $P_{ij} = 0$ for all $j \in \mathcal{X}$
- Transient and recurrent state: State i transient if starting at state i there is a nonzero probability of never returning to i. State i is called recurrent if it is not transient.

Classification of States

Proposition

(i) A Markov chain can be decomposed into one or more communicating (recurrent) classes, and possibly some transient states.
(ii) A recurrent state is accessible from all states in its class, but is not accessible from recurrent states in other classes.
(iii) A transient state is not accessible from any recurrent state.
(iv) At least one, possibly more, recurrent states are accessible from a given transient state.





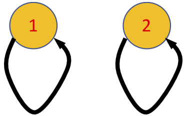
Stationary Distribution

- Stationary distribution: The vector π is called a stationary distribution of a Markov chain with transition probability matrix P if $\pi = \pi P$ and $\sum_j \pi_j = 1$

Stationary Distribution

- Stationary distribution: The vector π is called a stationary distribution of a Markov chain with transition probability matrix P if $\pi = \pi P$ and $\sum_j \pi_j = 1$

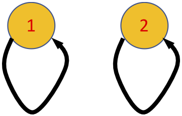
- $P = \left[\begin{array}{cc} 1 & 0 \\ 0 & 1 \end{array} \right]$. Does this Markov chain have a stationary distribution? Is it unique?



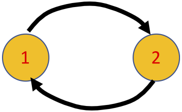
Stationary Distribution

- Stationary distribution: The vector π is called a stationary distribution of a Markov chain with transition probability matrix P if $\pi = \pi P$ and $\sum_j \pi_j = 1$

- $P = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$. Does this Markov chain have a stationary distribution? Is it unique?



- $P = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}$. Does this Markov chain have a stationary distribution? Is it unique?



Stationary Distribution

- The vector π is called a stationary distribution of a Markov chain with transition probability matrix P if $\pi = \pi P$ and $\sum_j \pi_j = 1$

Proposition

A finite state irreducible Markov chain has a unique stationary distribution, i.e., there exists a unique vector π such that $\pi = \pi P$.

Limiting Distribution

- Limiting distribution: $\lim_{t\to\infty} p(t)$

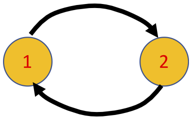
Limiting Distribution

- Limiting distribution: $\lim_{t \to \infty} p(t)$
  - Does it exist? Is it unique? Does it depend on the initial distribution? How is it related to the stationary distribution?

Limiting Distribution

- Limiting distribution: $\lim_{t\to\infty} p(t)$
  - Does it exist? Is it unique? Does it depend on the initial distribution? How is it related to the stationary distribution?

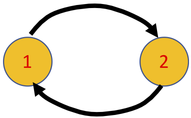
- $P = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}$. $\pi = \begin{bmatrix} 0.5 & 0.5 \end{bmatrix}$
- Assume $P(X_0 = 1) = 1.0$
- What is $\lim_{t\to\infty} p(t)$?



Limiting Distribution

- Limiting distribution: $\lim_{t \to \infty} p(t)$
  - Does it exist? Is it unique? Does it depend on the initial distribution? How is it related to the stationary distribution?

- $P = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}$. $\pi = \begin{bmatrix} 0.5 & 0.5 \end{bmatrix}$
- Assume $P(X_0 = 1) = 1.0$
- What is $\lim_{t \to \infty} p(t)$?



$$p(t) = \begin{cases} (0,1) & \text{if } t \text{ is odd} \\ (1,0) & \text{if } t \text{ is even} \end{cases}$$

- Limiting distribution may not exist

Limiting Distribution

- Period of a state: State i is said to have period $d(i)$ if it is only possible to return to state i in $t \geq 1$ steps, where t is a multiple of $d(i)$. Formally,

$$d(i) = \text{gcd}\{t : \mathbb{P}(X_t = i | X_0 = i) > 0\}$$

Limiting Distribution

- Period of a state: State i is said to have period $d(i)$ if it is only possible to return to state i in $t \geq 1$ steps, where t is a multiple of $d(i)$. Formally,

  $$d(i) = \text{gcd}\{t : \mathbb{P}(X_t = i | X_0 = i) > 0\}$$

  State i is aperiodic if $d(i) = 1$.

Limiting Distribution

- Period of a state: State i is said to have period $d(i)$ if it is only possible to return to state i in $t \geq 1$ steps, where t is a multiple of $d(i)$. Formally,

  $$d(i) = \text{gcd}\{t : \mathbb{P}(X_t = i | X_0 = i) > 0\}$$

  State i is aperiodic if $d(i) = 1$.
  A Markov chain is aperiodic if all states are aperiodic.

Limiting Distribution

- Period of a state: State i is said to have period $d(i)$ if it is only possible to return to state i in $t \geq 1$ steps, where t is a multiple of $d(i)$. Formally,

$$d(i) = \text{gcd}\{t : \mathbb{P}(X_t = i | X_0 = i) > 0\}$$

  State i is aperiodic if $d(i) = 1$.
  A Markov chain is aperiodic if all states are aperiodic.

Proposition

A finite state irreducible aperiodic Markov chain has a unique stationary distribution, i.e., there exists a unique π such that $\pi = \pi P$. Moreover, $\lim_{t \to \infty} p(t) = \pi$.

Limiting Distribution

- Period of a state: State i is said to have period $d(i)$ if it is only possible to return to state i in $t \geq 1$ steps, where t is a multiple of $d(i)$. Formally,

$$d(i) = \text{gcd}\{t : \mathbb{P}(X_t = i | X_0 = i) > 0\}$$

State i is aperiodic if $d(i) = 1$.
A Markov chain is aperiodic if all states are aperiodic.

Proposition

A finite state irreducible aperiodic Markov chain has a unique stationary distribution, i.e., there exists a unique π such that $\pi = \pi P$. Moreover, $\lim_{t \to \infty} p(t) = \pi$.

Proposition

For an irreducible and aperiodic MC, the stationary probability π satisfy $\pi_j = \lim_{t \to \infty} \frac{n_{ij}(t)}{t}$, where $n_{ij}(t)$ is the expected value of the number of visits to state j within the first t transitions, starting from state i.