ENGG7302: Advanced Computational Techniques in Engineering

Lecture 7: Markov Chains

Mandar Gujrathi

School of ITEE
UQ

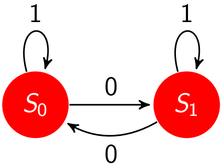
October 25, 2011

In this lecture..

- Markov chain
- Properties
- Definitions
- Examples

Simple State Machine

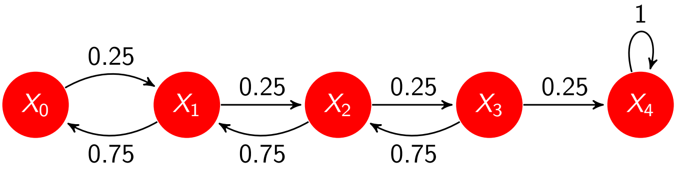
- A Simple state machine



- Conditions: Outcome of one state decides how the state machine proceeds.

Markov Process and Markov Property

- Application to conditional probability.



- A process can start in state X_i and moves from one state to another in steps with some probability.

$$P(X_3) = P(X_3|X_0 = 0.25, X_1 = 0.25, X_2 = 0.25)$$

- Markov property: Future states of a process can depend on the present state and not on how we arrived at this state.

Markov Chains

- A Markov chain is a discrete-time stochastic process, i.e. a sequence of random variables $X_1, X_2, X_3, ...$ with the Markov property.
- Conditional Probability distribution

$$P(X_n = s_n | X_1 = s_1, X_2 = s_2, \ldots, X_{n-1} = s_{n-1})$$

- So with the Markov condition

$$\begin{aligned} P(X_n = S | X_1 = s_1, X_2 = s_2, \ldots, X_{n-1} = s_{n-1}) \\ = P(X_n = s | X_{n-1} = s_{n-1}) \end{aligned}$$

- Take only the current state.

Transition Probabilities

- If **x** is a Markov chain with finite or infinite states
  - Probability at time t_m system occupies state s_i will be $p_i(m) = P(x_m = s_i)$
  - Probability that at time t_n the system goes from state s_i to s_j (at time t_m) will be

$$p_{ij}(m, n) = P(x_n = s_j | x_m = s_i)$$

  - One step probabilities are usually denoted simply by p_{ij}. Therefore,

$$p_{ij} = P(x_{n+1} = s_j | x_n = s_i)$$

  - We can develop a matrix of transition probabilities.

Example: Transition matrix

- Consider the following transition matrix.

$$P = \begin{array}{c c} & \begin{array}{ccc} \mathtt{R} & \mathtt{N} & \mathtt{S} \end{array} \\ \begin{array}{c} \mathtt{R} \\ \mathtt{N} \\ \mathtt{S} \end{array} & \begin{pmatrix} 0.5 & 0.25 & 0.25 \\ 0.5 & 0 & 0.5 \\ 0.25 & 0.25 & 0.5 \end{pmatrix} \end{array}$$

- Weather probabilities following rain, nice and snowy days.
- Given it is rain today lets find P(S in 2 days)

$$\begin{aligned} \mathtt{P(S\ in\ 2\ days)} &= \mathtt{P(RR)P(RS)+P(RN)P(NS)+P(RS)P(SS)}. \\ p_{13}^{(2)} &= p_{11}p_{13} + p_{12}p_{23} + p_{13}p_{33} \\ &= p(1,:)p(:,3) \end{aligned}$$

Example

$$
\mathbf{P}^{(1)} = \begin{array}{c} \\ \mathtt{R} \\ \mathtt{N} \\ \mathtt{S} \end{array} \begin{array}{c} \begin{array}{ccc} \mathtt{R} & \mathtt{N} & \mathtt{S} \end{array} \\ \begin{pmatrix} 0.5 & 0.25 & 0.25 \\ 0.5 & 0 & 0.5 \\ 0.25 & 0.25 & 0.5 \end{pmatrix} \end{array}
\quad
\mathbf{P}^{(2)} = \begin{array}{c} \\ \mathtt{R} \\ \mathtt{N} \\ \mathtt{S} \end{array} \begin{array}{c} \begin{array}{ccc} \mathtt{R} & \mathtt{N} & \mathtt{S} \end{array} \\ \begin{pmatrix} 0.438 & 0.188 & 0.375 \\ 0.375 & 0.25 & 0.375 \\ 0.375 & 0.188 & 0.438 \end{pmatrix} \end{array}
$$

$$
\mathbf{P}^{(3)} = \begin{array}{c} \\ \mathtt{R} \\ \mathtt{N} \\ \mathtt{S} \end{array} \begin{array}{c} \begin{array}{ccc} \mathtt{R} & \mathtt{N} & \mathtt{S} \end{array} \\ \begin{pmatrix} 0.406 & 0.203 & 0.391 \\ 0.406 & 0.188 & 0.406 \\ 0.391 & 0.203 & 0.406 \end{pmatrix} \end{array}
\quad
\mathbf{P}^{(4)} = \begin{array}{c} \\ \mathtt{R} \\ \mathtt{N} \\ \mathtt{S} \end{array} \begin{array}{c} \begin{array}{ccc} \mathtt{R} & \mathtt{N} & \mathtt{S} \end{array} \\ \begin{pmatrix} 0.402 & 0.199 & 0.398 \\ 0.398 & 0.203 & 0.398 \\ 0.398 & 0.199 & 0.402 \end{pmatrix} \end{array}
$$

$$
\mathbf{P}^{(5)} = \begin{array}{c} \\ \mathtt{R} \\ \mathtt{N} \\ \mathtt{S} \end{array} \begin{array}{c} \begin{array}{ccc} \mathtt{R} & \mathtt{N} & \mathtt{S} \end{array} \\ \begin{pmatrix} 0.40 & 0.20 & 0.399 \\ 0.40 & 0.199 & 0.40 \\ 0.399 & 0.20 & 0.40 \end{pmatrix} \end{array}
\quad
\mathbf{P}^{(6)} = \begin{array}{c} \\ \mathtt{R} \\ \mathtt{N} \\ \mathtt{S} \end{array} \begin{array}{c} \begin{array}{ccc} \mathtt{R} & \mathtt{N} & \mathtt{S} \end{array} \\ \begin{pmatrix} 0.40 & 0.20 & 0.40 \\ 0.40 & 0.20 & 0.40 \\ 0.40 & 0.20 & 0.40 \end{pmatrix} \end{array}
$$

Example: Transition matrix

- May be given some transition matrix and might be asked to find $P^{(n)}$, for some n.
- For the given transition matrix with $a = 0, b = 0.5$ Find P^2, P^3.

$$\mathbf{P} = \begin{array}{c c} & \begin{array}{cc} \texttt{yes} & \texttt{no} \end{array} \\ \begin{array}{c} \texttt{yes} \\ \texttt{no} \end{array} & \begin{pmatrix} 1-a & a \\ b & 1-b \end{pmatrix} \end{array}$$

State Probability vector, $\mathbf{u}$

- Is a row vector represents the initial state of the Markov chain.
- i^{th} element of this vector represents the probability the chain starts in state s_i.
- Using $\mathbf{P}^{(n)}$ and $\mathbf{u}$, we can calculate the state of the chain after n steps
- Probability that the chain is in state s_i after n steps is just $\mathbf{u}^{(n)} = \mathbf{u}\mathbf{P}^{(n)}$.

Probability vector

- However as $n \to \infty$, $\mathbf{P}^n$ will become constant.
- No matter what $\mathbf{u}$ is as $n \to \infty$, $\mathbf{u}^n$ will also be constant. This is called as steady state probability distribution vector.
- $\lim_{n\to\infty} \mathbf{P}^n = \mathbf{W}$

Example

$$\begin{pmatrix} 1/2 & 0 & 1/2 \end{pmatrix} \begin{array}{c c} & \begin{array}{ccc} \mathtt{R} & \mathtt{N} & \mathtt{S} \end{array} \\ \begin{array}{c} \mathtt{R} \\ \mathtt{N} \\ \mathtt{S} \end{array} & \begin{pmatrix} 0.40 & 0.20 & 0.40 \\ 0.40 & 0.20 & 0.40 \\ 0.40 & 0.20 & 0.40 \end{pmatrix} \end{array} = \begin{pmatrix} 0.4 & 0.2 & 0.4 \end{pmatrix}$$

$$\begin{pmatrix} 1/3 & 1/3 & 1/3 \end{pmatrix} \begin{array}{c c} & \begin{array}{ccc} \mathtt{R} & \mathtt{N} & \mathtt{S} \end{array} \\ \begin{array}{c} \mathtt{R} \\ \mathtt{N} \\ \mathtt{S} \end{array} & \begin{pmatrix} 0.40 & 0.20 & 0.40 \\ 0.40 & 0.20 & 0.40 \\ 0.40 & 0.20 & 0.40 \end{pmatrix} \end{array} = \begin{pmatrix} 0.4 & 0.2 & 0.4 \end{pmatrix}$$

$$\begin{pmatrix} 1 & 0 & 0 \end{pmatrix} \begin{array}{c c} & \begin{array}{ccc} \mathtt{R} & \mathtt{N} & \mathtt{S} \end{array} \\ \begin{array}{c} \mathtt{R} \\ \mathtt{N} \\ \mathtt{S} \end{array} & \begin{pmatrix} 0.40 & 0.20 & 0.40 \\ 0.40 & 0.20 & 0.40 \\ 0.40 & 0.20 & 0.40 \end{pmatrix} \end{array} = \begin{pmatrix} 0.4 & 0.2 & 0.4 \end{pmatrix}$$

Stationary Probability Vector

- As $n \to \infty$ each row of $\mathbf{P}^{(n)}$ remains the same
- i.e. as $n \to \infty$ the transition matrix will converge to a constant matrix, $\mathbf{W}$.
- Also as $n \to \infty$ the state probability vector approaches steady state probability vector, $\mathbf{w}$.
- Stationary Probability Vector

Ergodic Markov Chain

- A Markov chain is called an ergodic chain if it is possible to go from every state to every other state not necessarily in one move.

$$\mathbf{P} = \begin{array}{c c} & \begin{array}{ccccc} 0 & 1 & 2 & 3 & 4 \end{array} \\ \begin{array}{c} 0 \\ 1 \\ 2 \\ 3 \\ 4 \end{array} & \begin{pmatrix} 0 & 1 & 0 & 0 & 0 \\ \frac{1}{4} & 0 & \frac{3}{4} & 0 & 0 \\ 0 & \frac{1}{2} & 0 & \frac{1}{2} & 0 \\ 0 & 0 & \frac{3}{4} & 0 & \frac{1}{4} \\ 0 & 0 & 0 & 1 & 0 \end{pmatrix} \end{array}$$

Regular Markov Chains

- A Markov chain is called a regular chain if some power of the transition matrix has only positive elements.
- It can be possible for a regular Markov chain to have a transition matrix that has zeros.
- example P^2 becomes a regular Markov chain.

$$\mathbf{P}^{(2)} = \begin{array}{c c} & \begin{array}{ccc} \mathtt{R} & \mathtt{N} & \mathtt{S} \end{array} \\ \begin{array}{c} \mathtt{R} \\ \mathtt{N} \\ \mathtt{S} \end{array} & \begin{pmatrix} 0.44 & 0.19 & 0.37 \\ 0.38 & 0.25 & 0.37 \\ 0.37 & 0.19 & 0.4 \end{pmatrix} \end{array}$$

Classification of states

- Given 2 states s_i and s_j and if $p_{i,j}^{(n)} > 0$ then there is positive probability of reaching state s_j from s_i in n steps. s_j is **accessible** from s_i.
- If s_j is accessible from s_i with a positive probability, then s_i **communicates** with s_j

$$\mathbf{P} = \begin{array}{c} \\ 0 \\ 1 \\ 2 \\ 3 \\ 4 \end{array} \begin{array}{c} \begin{array}{ccccc} 0 & 1 & 2 & 3 & 4 \end{array} \\ \begin{pmatrix} 0 & 1 & 0 & 0 & 0 \\ \frac{1}{4} & 0 & \frac{3}{4} & 0 & 0 \\ 0 & \frac{1}{2} & 0 & \frac{1}{2} & 0 \\ 0 & 0 & \frac{3}{4} & 0 & \frac{1}{4} \\ 0 & 0 & 0 & 1 & 0 \end{pmatrix} \end{array}$$

Classification of states

- If s_i is accessible from s_j and s_j from s_i, the states intercommunicate with each other.
- If every state in Markov chain is accessible from every other state then the chain and the transition matrix is **irreducible**

Transient, Persistent states

- Closed Set
    - If C is a set of states such that no state outside C can be reached from any state in C, then set C is a closed set.
    - Example: If $s_i \in C$ and $s_j \notin C$ then $p_{ij} = 0$
- If starting from s_j return to the same state is not certain then that state called transient.
- If the return is certain then it is called persistent/ recurrent.

Periodic and aperiodic states

- If state s_j is said to be periodic with a period T if return to that state is possible only in multiples of T time steps. Consider the following example of 4×4 transition matrix.

$$\mathbf{P} = \begin{array}{c} \\ 0 \\ 1 \\ 2 \\ 3 \end{array} \begin{array}{c} \begin{array}{cccc} 0 & 1 & 2 & 3 \end{array} \\ \begin{pmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 \end{pmatrix} \end{array}$$

- This one has a period 4.
- A state is aperiodic is no such T exists.

Example

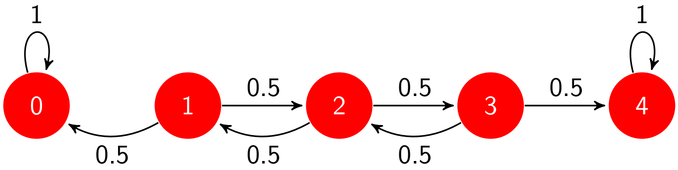
- Is this Periodic

$$\mathbf{P} = \begin{array}{c c} & \begin{array}{cccc} 0 & 1 & 2 & 3 \end{array} \\ \begin{array}{c} 0 \\ 1 \\ 2 \\ 3 \end{array} & \begin{pmatrix} 0 & 0.5 & 0 & 0.5 \\ 0.5 & 0 & 0.5 & 0 \\ 0 & 0.5 & 0 & 0.5 \\ 0.5 & 0 & 0.5 & 0 \end{pmatrix} \end{array}$$

- Period is ?

Absorption State

- A state is called as absorbing state if it is impossible to leave this state, e.g.



Example

- A man walks along a four-block stretch of Park Avenue. If he is at corner 1, 2, or 3, then he walks to the left or right with equal probability. He continues until he reaches corner 4, which is a bar, or corner 0, which is his home. If he reaches either home or the bar, he stays there.

  Identify the transient states.

  Identify the absorption states.

  Find the time to absorption

  Find the absorption probabilities

Example: Develop the transition matrix

- A man walks along a four-block stretch of Park Avenue. If he is at corner 1, 2, or 3, then he walks to the left or right with equal probability. He continues until he reaches corner 4, which is a bar, or corner 0, which is his home. If he reaches either home or the bar, he stays there. Identify the transient and absorption states.

$$\mathbf{P} = \begin{array}{c} \\ 0 \\ 1 \\ 2 \\ 3 \\ 4 \end{array} \begin{array}{c} \begin{array}{ccccc} 0 & 1 & 2 & 3 & 4 \end{array} \\ \begin{pmatrix} 1 & 0 & 0 & 0 & 0 \\ \frac{1}{2} & 0 & \frac{1}{2} & 0 & 0 \\ 0 & \frac{1}{2} & 0 & \frac{1}{2} & 0 \\ 0 & 0 & \frac{1}{2} & 0 & \frac{1}{2} \\ 0 & 0 & 0 & 0 & 1 \end{pmatrix} \end{array}$$

- Absorption states are 0 and 4...2 states
- Transient states are 1, 2 and 3....3 states

Canonical Form

- Renumber the states with transient states first. The transition matrix will have the canonical form as

$$\mathbf{P} = \begin{array}{c c} & \begin{array}{cc} \mathtt{TR} & \mathtt{ABS} \end{array} \\ \begin{array}{c} \mathtt{TR} \\ \\ \mathtt{ABS} \end{array} & \left(\begin{array}{c c} \mathbf{Q} & \vdots \mathbf{R} \\ \cdots\cdots\cdots & \cdots \\ 0 & \vdots \mathbf{I} \end{array} \right) \end{array}$$

- $\mathbf{I}$ is $r \times r$ identity matrix
- 0 is $r \times t$ zero matrix
- $\mathbf{R}$ is $t \times r$ non-zero matrix
- $\mathbf{Q}$ is $t \times t$ matrix

Example: Canonical form

$$\mathbf{P} = \begin{array}{c} \\ 1 \\ 2 \\ 3 \\ \\ 0 \\ 4 \end{array} \begin{array}{c} \begin{array}{ccc:cc} 1 & 2 & 3 & 0 & 4 \end{array} \\ \left(\begin{array}{ccc:cc} 0 & \frac{1}{2} & 0 & \frac{1}{2} & 0 \\ \frac{1}{2} & 0 & \frac{1}{2} & 0 & 0 \\ 0 & \frac{1}{2} & 0 & 0 & \frac{1}{2} \\ \hdashline 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{array} \right) \end{array}$$

Fundamental Matrix

- The fundamental matrix is given by $\mathbf{N} = (\mathbf{I} - \mathbf{Q})^{-1}$
- This is useful to calculate the absorption probabilites and the time to absorption.

$$\mathbf{N} = (\mathbf{I} - \mathbf{Q})^{-1} = \begin{pmatrix} \frac{3}{2} & 1 & \frac{1}{2} \\ 1 & 2 & 1 \\ \frac{1}{2} & 1 & \frac{3}{2} \end{pmatrix}$$

Absorption probability

- Absorption probability is given by $\mathbf{B} = \mathbf{NR}$

$$\mathbf{B} = \mathbf{NR} = \begin{pmatrix} \frac{3}{2} & 1 & \frac{1}{2} \\ 1 & 2 & 1 \\ \frac{1}{2} & 1 & \frac{3}{2} \end{pmatrix} \begin{pmatrix} \frac{1}{2} & 0 \\ 0 & 0 \\ 0 & \frac{1}{2} \end{pmatrix} = \begin{array}{c c} & \begin{array}{cc} 0 & 4 \end{array} \\ \begin{array}{c} 1 \\ 2 \\ 3 \end{array} & \begin{pmatrix} \frac{3}{4} & \frac{1}{4} \\ \frac{1}{2} & \frac{1}{2} \\ \frac{1}{4} & \frac{3}{4} \end{pmatrix} \end{array}$$

- Starting from state 1, the probability of getting absorbed in state 0 is $\frac{3}{4}$ while the probability of getting absorbed in state 4 is $\frac{1}{4}$

Time to absorption

- Expected number of steps before chain is absorbed.
- Given in state s_i, t_i is expected number of steps before chain is absorbed,
- The time to absorption is given by $\mathbf{t} = \mathbf{N}\mathbf{c}$.
- $\mathbf{c}$ is a column matrix of 1's.

Time to absorption

- $\mathbf{t} = \mathbf{Nc} =$

$$\begin{pmatrix} \frac{3}{2} & 1 & \frac{1}{2} \\ 1 & 2 & 1 \\ \frac{1}{2} & 1 & \frac{3}{2} \end{pmatrix} \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} = \begin{pmatrix} 3 \\ 4 \\ 3 \end{pmatrix}$$

- Starting from state 1, expected time of absorption is 3 steps
- Starting from state 2, expected time of absorption is 4 steps
- Starting from state 3, expected time of absorption is 3 steps

Example 2

- Calculate the fundamental matrix for the given transition matrix

$$\mathbf{P} = \begin{array}{c c} & \begin{array}{cccccc} \texttt{GGGG} & \texttt{GGGg} & \texttt{GGgg} & \texttt{GgGg} & \texttt{Gggg} & \texttt{gggg} \end{array} \\ \begin{array}{c} \texttt{GGGG} \\ \texttt{GGGg} \\ \texttt{GGgg} \\ \texttt{GgGg} \\ \texttt{Gggg} \\ \texttt{gggg} \end{array} & \begin{pmatrix} 1 & 0 & 0 & 0 & 0 & 0 \\ 0.25 & 0.5 & 0 & 0.25 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 \\ 0.062 & 0.25 & 0.125 & 0.25 & 0.25 & 0.062 \\ 0 & 0 & 0 & 0.25 & 0.5 & 0.25 \\ 0 & 0 & 0 & 0 & 0 & 1 \end{pmatrix} \end{array}$$

Mean First Passage time

- If an ergodic Markov chain started in s_i. The expected number of steps it takes to reach s_j for the first time is called as mean first passage time denoted by m_{ij}. By convention $m_{ii} = 0$.
- These become the individual entries of the matrix $\mathbf{M}$. Lets see how to calculate that.

Mean First Passage Matrix

- Represented as $\mathbf{M}$.
- Calculated from $\mathbf{Z}$ where $\mathbf{Z} = (\mathbf{I} - \mathbf{P} + \mathbf{W})^{-1}$

$$m_{ij} = \frac{z_{jj} - z_{ij}}{w_j}$$

- Calculate the mean recurrence time for

$$P = \begin{pmatrix} 0.5 & 0.25 & 0.25 \\ 0.5 & 0 & 0.5 \\ 0.25 & 0.25 & 0.5 \end{pmatrix}$$

Mean First Passage Matrix

- $\mathbf{Z} = (\mathbf{I} - \mathbf{P} + \mathbf{W})^{-1}$ where

$$\mathbf{W} = \begin{pmatrix} 0.4 & 0.2 & 0.4 \\ 0.4 & 0.2 & 0.4 \\ 0.4 & 0.2 & 0.4 \end{pmatrix}$$

- Therefore $\mathbf{Z}$ is

$$\mathbf{Z} = \begin{pmatrix} 9/10 & -1/20 & 3/20 \\ -1/10 & 6/5 & -1/10 \\ 3/20 & -1/20 & 9/10 \end{pmatrix}^{-1}$$

Mean First Passage Matrix

- Get **M** using

$$m_{ij} = \frac{z_{jj} - z_{ij}}{w_j}$$

- The diagonal elements of **M** will be zero.

$$\mathbf{M} = \begin{pmatrix} 0 & 4 & 10/3 \\ 8/3 & 0 & 8/3 \\ 10/3 & 4 & 0 \end{pmatrix}$$

Mean Recurrence time

- If an ergodic Markov chain started in s_i. The expected number of steps it takes to return to s_i for the first time is called as mean recurrence time for s_i.
- For a ergodic Markov chain, the mean recurrence time for a state s_i is $r_i = \frac{1}{w_i}$, w_i is the i^{th} component of stationary vector.
- For the stationary vector $\mathbf{w} = [0.4, 0.2, 0.4]$, the recurrence time is $\mathbf{r} = [2.5, 5, 2.5]$. If the chain began in state 2 it takes on an average 5 steps to return back to the same state.

Reversible Markov chains

- Let the original Markov chain be

$$X_{n-1}, X_n, X_{n+1}, X_{n+2}$$

- Given present, past and future are independent
- When reversed it will look like

$$X_{n+1}, X_n, X_{n-1}, X_{n-2}$$

- Given present, future and past are independent.
- Reverse process is also a Markov chain.

Reverse transition probability

- The reverse transition probability is given by

$$P(X_n = j | X_{n+1} = i) = \frac{P(X_n = j, X_{n+1} = i)}{P(X_{n+1} = i)}$$

- Using Bayes rule

$$\begin{aligned} P(X_n = j | X_{n+1} = i) &= \frac{P(X_n = j) P(X_{n+1} = i | X_n = j)}{P(X_{n+1} = i)} \\ &= \frac{P(X_{n-1} = j) p_{ji}}{P(X_n = i)} \\ p_{ij} &= \frac{w_j p_{ji}}{w_i} \end{aligned}$$

- A Markov chain is said to be reversible if there is a stationary vector **w** such that

$$w_i p_{ij} = w_j p_{ji}$$