Markov Chains

Discrete time Markov Chains (DTMC)

- A discrete time stochastic process $\{X_n, n \in \mathbb{Z}_+\}$ is essentially a collection of random variables.These random variables could be dependent and need not have identical distribution.

- A stochastic process $\{X_n, n \in \mathbb{Z}_+\}$ is a discrete time Markov chain if for any $n_1 < n_2 < \ldots < n_k < n$,

$$P(X_n = j | X_{n_1} = x_1, .., X_{n_k} = i) = P(X_n = j | X_{n_k} = i)$$

- This is called as the Markov property.

- $P(next\ state | past\ states,\ present\ state) = P(next\ state |\ present\ state)$

- We will throughout assume that the state space $\mathcal{S}$ is countable.

Example: Coin with memory!

- In a Markovian coin with memory, the outcome of the next toss depends on the current toss.
- $X_n = 1$ for heads and $X_n = -1$ otherwise. $\mathcal{S} = \{+1, -1\}$.
- Sticky coin : $P(X_{n+1} = 1 | X_n = 1) = 0.9$ and $P(X_{n+1} = -1 | X_n = -1) = 0.8$ for all n.
- Flippy Coin: $P(X_{n+1} = 1 | X_n = 1) = 0.1$ while $P(X_{n+1} = -1 | X_n = -1) = 0.3$ for all n.
- This can be represented by a transition diagram (see board)
- The transition probability matrix P for the two cases is $P_s = \begin{bmatrix} 0.9 & .1 \\ 0.2 & 0.8 \end{bmatrix}$ and $P_f = \begin{bmatrix} 0.1 & 0.9 \\ 0.7 & 0.3 \end{bmatrix}$
- The row corresponds to present state and the column corresponds to next state.

Running example: Dice with memory!

- In a markovian dice with memory, the outcome of the next roll depends on the current roll.
- $X_n = i$ for $i \in \mathcal{S}$ where $\mathcal{S} = \{1, \dots, 6\}$.
- Example transition probability matrix

$$P = \begin{bmatrix} 0.9 & .1 & 0 & 0 & 0 & 0 \\ 0 & .9 & .1 & 0 & 0 & 0 \\ 0 & 0 & 0.9 & 0.1 & 0 & 0 \\ 0 & 0 & 0 & 0.9 & 0.1 & 0 \\ 0 & 0 & 0 & 0 & 0.9 & 0.1 \\ 0.1 & 0 & 0 & 0 & 0 & 0.9 \end{bmatrix}$$

- State transition diagram on board

Finite dimensional distributions

- Consider a Markov dice with transition probability P.
- What is $P(X_0 = 4, X_1 = 5, X_2 = 6)$?
- $= P(X_2 = 6 | X_1 = 5, X_0 = 4) P(X_1 = 5 | X_0 = 4) P(X_0 = 4)$
- $= p_{65} p_{54} P(X_0 = 4)$.
- What is $P(X_0 = 4)$?
- This probability of starting in a particular state is called initial distribution of the markov chain.
- Let $\bar{\mu} = (\mu_1, \ldots, \mu_M)$ denote the initial distribution, i.e., $P(X_0 = x_0) = \mu_{x_0}$.
- $P(X_0 = x_0, X_1 = x_1, \ldots X_k = x_k) = p_{x_{k-1}, x_k} \times \ldots \times p_{x_0, x_1} \mu_{x_0}$

Chapman Kolmogorov Equations for DTMC

- Consider a Markov coin and its transition probability matrix (tpm) $P = \begin{bmatrix} p_{1,1} & p_{1,-1} \\ p_{-1,1} & p_{-1,-1} \end{bmatrix}$.

- Given $X_0 = 1$, what is $P(X_2 = 1)$?

$$\begin{aligned} P(X_2 = 1|X_0 = 1) &= P(X_2 = 1|X_1 = 1, X_0 = 1)P(X_1 = 1|X_0 = 1) \\ &+ P(X_2 = 1|X_1 = -1, X_0 = 1)P(X_1 = -1|X_0 = 1) \\ &= p_{1,1}^2 + p_{-1,1}p_{1,-1} \end{aligned}$$

- Here the first inequality follow from the fact that

$$P(C|A) = P(C|BA)P(B|A) + P(C|B^cA)P(B^c|A) \text{ HW: Verify}$$

- Similarly, $P(X_2 = -1|X_0 = 1), P(X_2 = 1|X_0 = -1), P(X_2 = -1|X_0 = -1)$ can be obtained and these are elements of a two-step transition matrix $P^{(2)}$.

Chapman Kolmogorov Equations for DTMC

- The two step transition probability matrix $P^{(2)}$ is given by
$$P^{(2)} = \begin{bmatrix} p_{1,1}^2 + p_{1,-1}p_{-1,1} & p_{1,1}p_{1,-1} + p_{1,-1}p_{-1,-1} \\ p_{-1,1}p_{1,1} + p_{-1,-1}p_{-1,1} & p_{-1,1}p_{1,-1} + p_{-1,-1}^2 \end{bmatrix}.$$

- This implies that $P^{(2)} = P \times P = P^2$.

- In general, $P^{(n)} = P^n$.

- Chapman-Kolmogorov equations are a further generalization of this.

$$\boxed{P^{(n+l)} = P^{(n)}P^{(l)}}$$

- We wont see the proof of this.

Classification of states

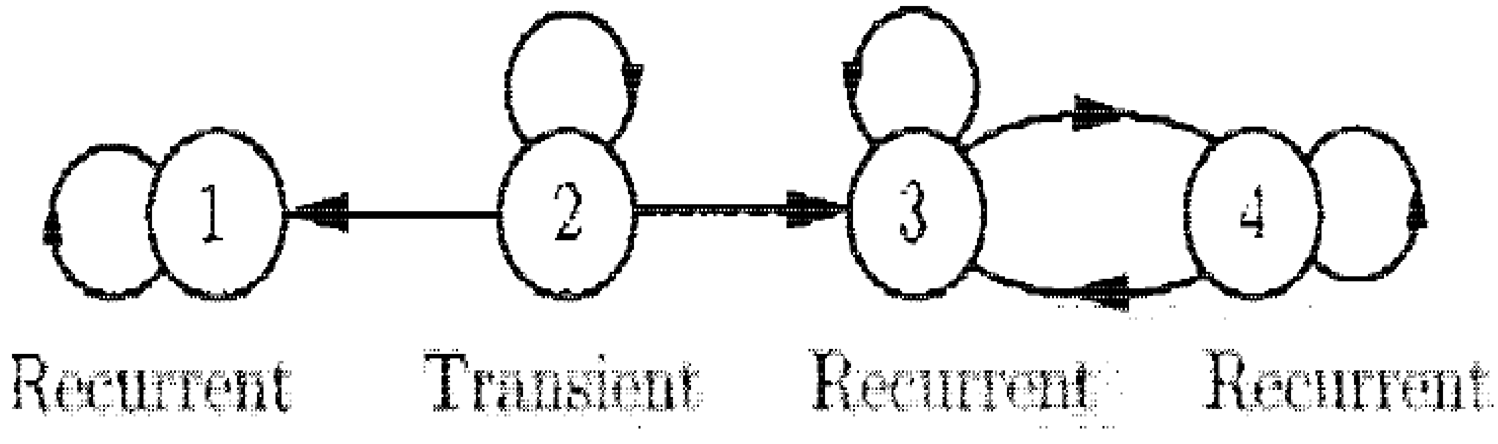
- Consider a Markov process with state space $\mathcal{S}$
- We say that j is accessible from i if $p_{ij}^{n} > 0$ for some n.
- This is denoted by $i \to j$.
- if $i \to j$ and $j \to i$ then we say that i and j communicate. This is denoted by $i \leftrightarrow j$.

  A chain is said to be irreducible if $i \leftrightarrow j$ for all $i, j \in \mathcal{S}$.

- Are the examples of Markovian coin and dice we have considered till now irreducible? check!

Recurrent and Transient states

- We say that a state i is recurrent if
  $F_{ii} = P($ ever returning to i having started in i $) = 1$.
- F_{ii} is not easy to calculate. (Not part of this course)
- If a state is not recurrent, it is transient.
- For a transient state i, $F_{ii} < 1$.
- If $i \leftrightarrow j$ and i is recurrent, then j is recurrent.



Limiting probabilities

- $P = \begin{bmatrix} 0 & 0 & 1 \\ 0 & 0.6 & 0.4 \\ 0.6 & 0.4 & 0 \end{bmatrix}$ $P^5 = \begin{bmatrix} .06 & .3 & .64 \\ .18 & .38 & .44 \\ .38 & .44 & .18 \end{bmatrix}$ $P^{30} = \begin{bmatrix} .23 & .385 & .385 \\ .23 & .385 & .385 \\ .23 & .385 & .385 \end{bmatrix}$

- $P = \begin{bmatrix} 1-a & a \\ b & 1-b \end{bmatrix}$ $\lim_{n\to\infty} P^n = \begin{bmatrix} \frac{b}{a+b} & \frac{a}{a+b} \\ \frac{b}{a+b} & \frac{a}{a+b} \end{bmatrix}$

- What is the interpretation of $\lim_{n\to\infty} p_{ij}^{(n)} = [\lim_{n\to\infty} P^n]_{ij}$?

- $\pi_j = \lim_{n\to\infty} p_{ij}^{(n)}$ denotes the limiting probability of being in state j when starting in state i.

- For an M state DTMC, $\pi = (\pi_1, \ldots, \pi_M)$ denotes the limiting distribution.

- How do we obtain the limiting distribution π? Does it always exist?

Stationary distribution

The **stationary distribution** of a Markov chain is defined as a solution to the equation $\pi = \pi P$.

- πP is essentially the p.m.f of X_1 having picked X_0 according to π.
- $\pi = \pi P$ says that, if the initial distribution is π, then the distribution of X_1 is also π.
- Continuing this argument, the p.m.f of X_n for any n is π and there is no dependence on the starting state.
- MCMC algorithms use this idea (at stationarity successive states of the Markov chain have p.m.f π) to sample from target distribution π.

Limiting vs Stationary distribution

- Obtain stationary distribution for the Markov Chain with transition probability $P = \begin{bmatrix} 0 & 0 & 1 \\ 0 & 0.6 & 0.4 \\ 0.6 & 0.4 & 0 \end{bmatrix}$
- The limiting distribution need not exist for some Markov chains, but the stationary distribution exists. For example for $P = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}$.
- The limiting distribution if it exists, is same as the stationary distribution.