Markov Chains I

Lec.26

August 5, 2020

Markov Chains: Fundamental Idea

We now wish to model sequences of random variables $X_0, X_1, X_2, \ldots$. You can think of X_n as the state of a system at time n.

We will be working on the setting where time is discrete, and each X_i can only take on a finite set of values. This finite set of values is denoted $\mathcal{X}$ and is called the state space.

Markov Property

Think of X_n be the present/current state, and X_{n+1} as the future state. The Markov Property is:

$$\Pr(X_{n+1} = j | X_n = i, ..., X_0 = i_0) = \Pr(X_{n+1} = j | X_n = i)$$

This property is not saying the future is independent of the past. This property is saying that the past and future are *conditionally independent* given the present.

We call $\Pr(X_{n+1} = j | X_n = i) = P(i, j)$ the transition probability from state i to state j.
In this class, we will only deal with time homogeneous Markov chains.

Transition Probability Matrix

Let the state space $\mathcal{X}$ be $\{1, ..., k\}$. The transition probability matrix for a Markov chain P is a k by k matrix such that the entry in the ith row and j column is $P(i, j)$, and:

$$P(i,j) \geq 0 \quad \forall i,j \in \mathcal{X} \tag{1}$$

$$\sum_{j=1}^{k} P(i,j) = 1 \quad \forall i \in \mathcal{X} \tag{2}$$

Markov Chain Example

Distribution Over States

We use π_i to represent the distrution of our random variable X_i over the states in $\mathcal{X}$. The entries in π_i must be probabilities that sum up to 1.

π_0 is called our initial distribution. π_0, in conjunction with P and $\mathcal{X}$ fully specifies our Markov chain.

Moving in Time

Suppose at time n, X_n has distribution π_n. Then, by the Law of Total Probability,

$$\Pr(X_{n+1} = j) = \sum_i \Pr(X_{n+1} = j | X_n = i) \Pr(X_n = i) \tag{3}$$
$$= \sum_i P(i,j)\pi_n(i) \tag{4}$$

But $\sum_i P(i,j)\pi_n(i)$ is the jth entry of $\pi_n P$. So,

$$\pi_{n+1} = \pi_n P \tag{5}$$
$$\pi_{n+2} = \pi_{n+1} P = \pi_n P P = \pi_n P^2 \tag{6}$$

P^k is the probability transition matrix where the entry in the ith row jth column is the probability of going from state i to state j in k steps.

Hitting Time Example

Probability of A before B Example

Invariant Distribution Definition

A distribution π is *invariant* for the transition probability matrix P if it satisfies the following *balance equations*:

$$\pi = \pi P \tag{7}$$

Stationary Distribution Existence

Let P be the probability transition matrix for a Markov chain.

The rows of P add up to 1. Let $\mathbf{1}$ be a column vector of ones. This means $P\mathbf{1} = \mathbf{1} = 1 \cdot \mathbf{1}$.

This means P has a right eigenvector corresponding to eigenvalue 1. Since the right and left eigenvalues of a square matrix are the same, this means there exists some left eigenvector π such that $\pi P = 1 \cdot \pi$.

Note that this does not say anything about the uniqueness of the stationary distribution.

Stationary Distribution Example