6.1220 LECTURE 19

BRENDON JIANG

Contents

1. Markov Chains/Random Walks

Definition 1.1: Markov Chain

A markov chain is a sequence of random variables $\{X_t : t \in \mathbb{N}\}$ which is memory less:

$$\forall t \in \mathbb{N}, X_{t+1} \text{ is independent of history } X_0, \dots, X_{t-1} \text{ given } X_t$$

Example 1.2

Consider a simple gambling game, where you win/lose 5 dollars based on a coin flip. If we let X_t denote your balance at each time, note that X_{t+1} is only dependent upon the previous time X_t.

Definition 1.3: Time Homogeneous

The Markov chain $\{X_t : t \in \mathbb{N}\}$ is time-homogenous if there is a **transition probability matrix W** such that $\forall t \in \mathbb{N}$ and pairs of states u, v,

$$\Pr[X_{t+1} = v \mid X_t = u] = W(u, v)$$

We can perform random walks on unweighted directed graphs and weighted directed graphs.

Definition 1.4: Graphical Representation

Given W, we can build a weighted directed graph $G_W = (V, E)$ by adding a directed edge between every pair of states u, v and setting $wt(u, v) = W(u, v)$. Then X_t evolves as a random walk on G_W.

Date: November 2024

1.1. Markov Chains for Decryption

The simplest cipher one can use is a substitution cipher, which is a one to mapping of symbols to letters. Our goal is to find this cipher F, which has $26! = 4 \cdot 10^{26}$ probabilities. The first step to crack this cipher is to come up with some measure of reasonableness $\varphi : \{mappings F\} \to \mathbb{R}$ so that φ is larger if the message looks more like english.

1.2. Evolution of X_t

How can we describe the distribution π_t of X_t given W and the initial distribution π_0 of X_0? We can first start with the distribution of X_1.

$$\Pr[X_1 = x_1] = \sum_{x_0} \Pr[X_0 = x_0, X_1 = x_1] = \sum_{x_0} \pi_0(x_0) \cdot W(x_0, x_1)$$

We can do this similarly for X_t.

$$\Pr[X_t = x_t] = \sum_{x_{t-1}} \Pr[X_{t-1} = x_{t-1}, X_t = x_t] = \sum_{x_{t-1}} \pi_{t-1}(x_{t-1}) \cdot W(x_{t-1}, x_t)$$

We propose the lemma here then:

Lemma 1.5

The distribution π_t of X_t is given by $\pi_t = \pi_0 W^t$.

In the simulations, we observe that in most of our examples, W^t converges to a matrix with all rows equal to some distribution π_* as $t \to \infty$.

Definition 1.6: Stationary

We say a distribution π_* is stationary with respect to W if $\pi_* = \pi_* W$.

Can we construct a graph on which random walks has multiple stationary distributions? We can, if the graph has two isolated vertices. Can we construct a connected directed graph on which random walk never converges? We also c an, by constructing a directed cycle.

Definition 1.7: Connectivity

We say two vertices u, v communicate with each other if there is a directed path from u to v and vice versa. We write this as $u \leftrightsquigarrow v$.

We observe that $\leftrightsquigarrow$ is an equivalence relation (satisfies reflexive, transitive, symmetric), hence V can be partitioned into equivalence classes.

Definition 1.8: SCC

We say the equivalence classes formed by this equivalence relation are called strongly connected components.

Lemma 1.9

The "induced" directed graph on strongly connected components is acyclic. We say a component is transient if it has an outgoing edge. Otherwise, it is recurrent.

Definition 1.10: Cycles

The period of a vertex is the greatest common denominator of the lengths of all directed cycles of the lengths of all directed cycles in its strongly connected component. We say the vertex is periodic if its period is ¿ 1. Otherwise it is aperiodic.

Note that if a vertex has a self-loop, then it is automatically aperiodic. We say a markov chain is lazy is all vertices have self-loops.

Theorem 1.11: The Fundamental Theorem of Markov Chains

Every time-homogeneous Markov chain W has a stationary distribution. Furthermore, if the underlying directed graph is strongly connected, then the stationary distribution is unique. If all vertices in the recurrent components are aperiodic, then for every choice of initial distribution π_0, the distribution $\pi_t = \pi_0 W^t$ of X_t converges to some stationary distribution as $t \to \infty$.

Back in the scenario of gambling now (1.2). Note that all nodes that are not "making it big" or "losing it all" are transient, so as $t \to \infty$, we either make it big or lose it all. If making twice our buy in where our buy in is 20 is making it big, then $\pi_*(\$X) = 1 - p$ if $X = 0$ and p if $X = 40$. Intuitively, $p = \dfrac{1}{2}$ by symmetry. We can also prove this rigorously.

Define $\alpha_k = \Pr[\text{make it big} \mid \text{start with } \$(5 \cdot k)]$. Note that $\alpha_k = \dfrac{1}{2}\alpha_{k-1} + \dfrac{1}{2}\alpha_{k+1}$, and solving this gives us $\alpha_4 = \dfrac{1}{2}$.