COVER TIMES

ALAN LEE

Abstract. This paper will discuss cover times and how they relate to various forms of hitting times. It will use this relation to show how certain random walks and other Markov chains' cover times can be bounded quite well with only the Matthews Method.

1. Introduction

In class, we have covered hitting times, mixing times, stationary times, and more. There are more than just those types of times, however. This paper will be cover the important topic of cover times by first introducing some common examples. An example of a cover time can be seen in many random walks, such as the following question.

Question 1.1. *A monkey controls a typewriter with only 4 keys: up, down, left and right. Ey randomly selects a key at each turn to press, which will move a dot on the screen initially located at* $(0, 0)$ *up, down, left, or right by 1 accordingly. On average, how many key presses will be made before every point* $(x, y)\{|x|, |y| \leq 10\}$ *is visited?*

While not directly addressed in this paper, this gives us an idea of a cover time. After solving for some cover times manually using elementary methods, the remainder of this paper will be used to bound cover times from both above and below in terms of hitting times. Finally, we will see some applications of this bound.

2. Definitions and Examples

In this paper, all Markov chains are assumed to be irreducible because we will be talking about cover times, which only make sense when we are working with irreducible Markov chains. We will focus on the following question, finding the necessary cover times to solve it.

Definition 2.1. Let $X_0, X_1, \ldots$ be a Markov chain on a state space Ω, and let $i \in \Omega$. The *cover time of the Markov chain starting from* i, denoted $C_i(X)$, is the expected amount of time taken to visit every state in Ω.

Definition 2.2. Let $X_0, X_1, \ldots$ be a Markov chain on a state space Ω, and let $i \in \Omega$. The *cover time* of the Markov chain, denoted $C(X)$, is defined as

$$C(X) = \max_{i \in \Omega}(C_i(X)).$$

We often like to talk about a random walk on a graph G instead of a specific Markov chain when discussing about cover times, as in the real world, more cover-time related problems apply to random walks than to specific-probability defined Markov chains. When we do so, we denote the cover time of a random walk on the graph G starting from a vertex i as $C_i(G)$.

Date: November 19, 2020.

$C(G)$ is similarly defined so that $C(G) = \max_{i \in \Omega}(C_i(G))$, where Ω is the set of points in G.

Example. One common example of a cover time can be found as a result of a random walk on a complete graph K_n, with the addition of a self-loop at each vertex. At each point during the walk, the walker moves from eir current vertex to any vertex since all vertices are connected by an edge. We see this is equivalent to the coupon collector's problem because visiting a new vertex is analogous to collecting a new coupon, with the chance of arriving at a new vertex decreasing as more unique vertices are visited. Hence, we have $C(K_n) \approx n\log(n)$.

Example. Another example that can easily be found using the method outlined by László Lovász in [L+93] is of an n-cycle.

The set of vertices visited by the walker must all be consecutive since the graph is a cycle. The first turn we have visited a total of $n-1$ vertices, the walker must have just arrived a new vertex and hence must be at one of the ends of eir set of vertices visited.

Now we equate this scenario to that of the gambler's ruin Markov chain with absorbing states 0 and n and with $p_{(i)(i+1)} = p_{(i)(i-1)} = \frac{1}{2}$ for all other states i. From a homework problem in week 3, we know that it takes an expected $k(n-k)$ moves to reach an absorbing state in this chain from state k. Since we are either at state 1 or $n-1$, it will take us an expected $(1)(n-1) = n-1$ more steps in order to reach a new vertex. If we denote c_m as the first time the walker has reached m unique vertices, we have

$$c_m = c_{m-1} + (m-1).$$

Thus, we can find that

$$c_n = c_{n-1} + (n-1) = c_{n-2} + (n-2) + (n-1) + \cdots = \sum_{k=1}^{n-1} k = \frac{n(n-1)}{2}$$

is the cover time of this graph.

Combining these two examples, we see that it takes slightly over $50\log(50) \approx 196$ moves to cover a complete graph with 50 vertices, while it takes $\frac{20 \cdot 19}{2} = 190$ moves to cover a 20-cycle. This close yet easily calculated comparison tells us that a random walk on $\mathbb{Z}/20\mathbb{Z}$ has a shorter cover time than one on K_{50} with self loops.

3. Bounding Cover Times

Now that we have a general sense of methods of calculating cover times, let us bound them in terms of hitting times. We also slightly modify the definition of hitting time to account for the entire Markov chain, not just a specific ordered pair of states.

Definition 3.1. Let $X_0, X_1, \ldots$ be a Markov chain on a state space Ω. The hitting time of the Markov chain, denoted $H(X)$, is defined as

$$H(X) = \max_{i,j \in \Omega} \mathbb{E}_i(\tau_j).$$

Our first theorem can easily be deduced by noting the definitions of cover times and hitting times.

Theorem 3.2. *For any irreducible Markov chain $X_0, X_1, \ldots$ on a state space Ω with n states, we have*

$$H(X) \leq C(X).$$

Proof. Selecting states $i, j \in \Omega$ such that $H(X) = \mathbb{E}_i(\tau_j)$, we have

$$H(X) = \mathbb{E}_i(\tau_j) \leq C_i(X) \leq C(X).$$

We must have $\mathbb{E}_i(\tau_j) \leq C_i(X)$ because by the time all states have been visited starting from state i, state j must have also been visited, causing the process to take at least the hitting time of j when starting from i.

□

This bound is often because weak there are more than 2 states in the majority of Markov chains. We shall now focus the rest of this paper on proving the upper bound, which is found in a more complicated manner.

Theorem 3.3 (Matthews Method)**.** *For any irreducible Markov chain $X_0, X_1, \ldots$ on a state space Ω with n states, we have*

$$C(X) \leq H(X)\left(1 + \frac{1}{2} + \frac{1}{3} + \cdots + \frac{1}{n-1}\right).$$

Proof. We start by assuming that without loss of generality our state space is $\Omega = \{1, 2, \ldots, n\}$ and that $X_0 = n$. We also let σ be a uniform random permutation of $\{1, 2, \ldots, n-1\}$. We denote T_k as the first time that the states $\sigma(1), \ldots, \sigma(k)$ have been visited, and $L_k = X_{T_k}$ as the last state among $\sigma(1), \ldots, \sigma(k)$ to have been visited.

Since $H(X)$ is at least the time it takes to get from any one state to any other, we have

$$\mathbb{E}_n(T_1 | \sigma(1) = s) \leq H(X)$$

for all $1 \leq s \leq n-1$. Taking into account all possible s, we have $\mathbb{E}_n(T_1) \leq H(X)$.

Now, we choose $1 \leq r \neq s \leq n-1$ to prove some new inequalities. We have

$$\mathbb{E}_n(T_k - T_{k-1} | L_{k-1} = r, \sigma(k) = s = L_k) \leq H(X)$$

because the left side of the inequality is less at most the time it takes to go from state r to s, while the right side is at least the time it takes. Averaging over all r and all s results in the simplified inequality

$$\mathbb{E}_n(T_k - T_{k-1} | L_k = \sigma(k)) \leq H(X).$$

We also have, for any set S, we have

$$\mathbb{P}_n\{L_k = \sigma(k) | \{\sigma(1), \ldots, \sigma(k)\} = S\} = \frac{1}{k}$$

since every $1 \leq k \leq n-1$ is equally likely to be the last state to be visited. Additionally, if we know that $L_k \neq \sigma(k)$, $T_k = T_{k-1}$ because k will already have been visited by the time the states $1, \ldots, k-1$ have been visited as well. Thus we have

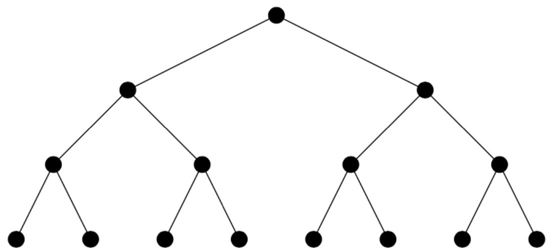

Figure 1. A rooted binary tree of depth 3.

$$\mathbb{E}_n(T_k - T_{k-1}|L_k \neq \sigma(k)) = 0.$$

Taking

$$\mathbb{E}_n(T_k - T_{k-1}|L_k = \sigma(k)) \leq H(X)$$

from above, we can rearrange to get

$$\mathbb{E}_n(T_k - T_{k-1}) \leq \mathbb{P}_n\{L_k = \sigma(k)\} \cdot H(X) = \frac{H(X)}{k}.$$

Finally, we sum over all $1 \leq k \leq n-1$ to get

$$C(X) = \mathbb{E}_n(T_1) + \sum_{k=2}^{n-1} \mathbb{E}_n(T_k - T_{k-1}) \leq H(X) + \frac{H(X)}{2} + \cdots + \frac{H(X)}{n-1},$$

as desired.

□

4. Covering Binary Trees

Using Theorem 3.3, we can now bound cover times of some common Markov chains such as in binary trees. We will focus primarily on the upper bound as that brings more useful and strictly bounded results. The following examples are from [LP17].

Definition 4.1. A *rooted binary tree* of depth k is a graph with $2^{k+1} - 1$ vertices with 1 vertex on the top level (called the root), 2 vertices on the next, 4 on the next, and so on. Each vertex (except those on the bottom row, which are called roots) is connected to exactly two unique vertices on the next level so that there are no cycles.

See Figure 1 for an example of the rooted binary tree of depth 3. For sake of simplicity, we will also denote $2^{k+1} - 1$ as simply n.

We can find the hitting time of a random walk on this rooted binary tree by first noting that the commute time between the root vertex and any specific leaf vertex is

$$t_{p \leftrightarrow a} = 2(n-1)k,$$

where commute time is defined by the time it takes to go from one vertex to another and then back to the first vertex.

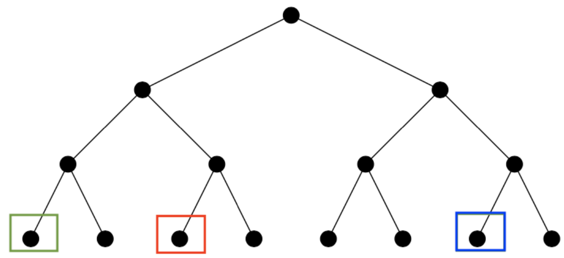

Figure 2. $H(X)$ is equal to the hitting time from the green vertex to the blue vertex, but not between the red and green vertices as their most common ancestor is not the root vertex.

The maximal hitting time in this random walk will be from one leaf to another so that no common ancestors (except the root vertex) are shared between them (see Figure 2 for an example of such a pair of vertices). However, this is the same as the time it should take to commute between the root and any leaf, which is $2(n-1)k$. Hence $H(X) = 2(n-1)k$, where $X_0, X_1, \ldots$ denotes our Markov chain. As a result, we have

$$C(X) \leq H(X)\left(1 + \frac{1}{2} + \cdots + \frac{1}{n-1}\right) = 2(n-1)k\left(1 + \frac{1}{2} + \cdots + \frac{1}{n-1}\right).$$

Note that we can approximate the nth harmonic number as $\log(n)$, so we have

$$C(X) \leq 2(n-1)k\left(1 + \frac{1}{2} + \cdots + \frac{1}{n-1}\right) = (2 + o(1))(\log 2)nk^2$$

as our final result.

5. Covering Coin-Flip Sequences

In class, we examined how to determine if a specific sequence of coin flips will appear before another specific sequence, as well as the expected time needed to arrive at a particular sequence. Now, the amount of time needed to reach *all* the sequences, where the method used to generate new sequences is as follows.

Definition 5.1. Let $X_0, X_1, \ldots$ be a Markov chain on the state space $\{H, T\}^k$ with a uniform initial state. At each step, the first letter of the k-tuple is removed and the either a T or an H is appended to the end of the result, each with probability $\frac{1}{2}$. This chain is referred to as the *shift chain on binary k-tuples.*

Example. A sequence for the shift chain on binary quintuples could be $HTTTT$, $TTTTH$, $TTTHH$. However, $HTTTT, TTTHH, TTHTH$ would not be a valid sequence because the last four letters of the preceding term must be the first four letters of the following term.

We also need to define what a waiting time is to help us grasp the upper bound of the cover time.

Definition 5.2. The *waiting time* w_x for a term $x \in \{H, T\}^k$ is the number of steps required for a x to appear in the chain without any overlap from the current state's bits.

It is easily seen that $w_x \geq k$ because all of the current state's bits (H's or T's) must have rotated out, and a completely new set of k bits must have appeared. Additionally, we have $w_x \geq \tau_x$ for all x, where τ_x denotes the hitting time of the particular term x. Note also that w_x is independent of the initial state and that our chain has a uniform stationary distribution. Hence

$$\mathbb{E}(w_x) \geq \mathbb{E}(\max(\tau_x)) = 2^k.$$

We also define a value H_k that can be equated to $H(X)$ while simultaneously finding its exact value.

Proposition 5.3. *Fix $k \geq 1$. For the shift chain on binary k-tuples, H_k defined by*

$$H_k = \max_{x \in \{0,1\}^k} \mathbb{E}(w_x)$$

is equal to $2^{k+1} - 2$.

Proof. We proceed by induction. For $k = 1$, there is a $\frac{1}{2}$ chance that it will take 1 step, $\frac{1}{4}$ that it will take 2, and so on. Thus, we have

$$\begin{aligned}
H_1 &= \frac{1}{2} + \frac{2}{4} + \frac{3}{8} + \frac{4}{16} + \cdots \\
&= \left(\frac{1}{2} + \frac{1}{4} + \frac{1}{8} + \cdots\right) + \left(\frac{1}{4} + \frac{1}{8} + \cdots\right) + \left(\frac{1}{8} + \cdots\right) \\
&= 1 + \frac{1}{2} + \frac{1}{4} + \cdots \\
&= 2.
\end{aligned}$$

Now fix a term x of length $k = 1$ and let x^- be the pattern consisting of the first k bits (H's or T's) of x. We must arrive at the term x^- first before arriving at x, and we have a $\frac{1}{2}$ chance of going from x^- to x. The other $\frac{1}{2}$ of the time, we start over, now requiring as much time as we needed at the start to arrive at x. Thus we have the recurrence

$$\mathbb{E}(w_x) \leq \mathbb{E}(w_{x^-}) + 1 + \frac{1}{2}\mathbb{E}(w_x),$$

where the additional 1 represents the flip where we go from x^- to either the start or x. Now, we bound H_{k+1} in terms of H_k by choosing an x with $H_{k+1} = \mathbb{E}(w_x)$. Then we have $\frac{1}{2}\mathbb{E}(w_x) = \frac{1}{2}H_{k+1}$ and $\mathbb{E}(w_{x^-}) \leq H_k$. Replacing these into the inequality above, we have

$$H_{k+1} \leq H_k + 1 + \frac{1}{2}H_{k+1},$$

which we rewrite as

$$H_{k+1} \leq 2H_k + 2.$$

Given $H_1 = 2$, we have, as a result, $H_k \leq 2^{k+1} - 2$. Note that when we want to find the hitting time from a term that ends in H to a term with only T's, it will take on average $2^{k+1} - 2$ coin flips, using the martingale method. Thus, we have $H_k = 2^{k+1} - 2$. □

As we did before, we can approximate the nth harmonic number as $\log(n)$, so we have

$$1 + \frac{1}{2} + \frac{1}{3} + \cdots + \frac{1}{2^k - 1} \approx \log(2^k) = k \log 2.$$

Combining the results we have found, we can obtain, for the shift chain on binary k-tuples,

$$\mathbb{E}(C_i(X)) = C(X) \leq H_k \left(1 + \frac{1}{2} + \frac{1}{3} + \cdots + \frac{1}{2^k - 1}\right) \approx (2^{k+1} - 2)(k \log 2),$$

or

$$C(X) \leq (k \log 2) 2^{k+1}(1 + o(1)).$$

As we have seen, Theorem 3.3 is useful in determining various cover times in terms of the easier-to-find hitting times. We have also seen the usefulness of cover times in real world examples, such as the coupon collector problem. The Matthews Method seems quite elementary given the involvement of harmonic numbers, but the upper bound has proven to be a very strong one. As more research is done into bounding cover times, mixing times, hitting times, and more, we may be able to find an even better bound in the future.

References

[L+93] László Lovász et al. Random walks on graphs: A survey. *Combinatorics, Paul erdos is eighty*, 2(1):1–46, 1993.

[LP17] David A Levin and Yuval Peres. *Markov chains and mixing times*, volume 107. American Mathematical Soc., 2017.