MARKOV CHAINS WEEK 10: COVER TIMES

LORENZO WOLCZKO

1. Cover Times

Definition 1.1. The *cover time* of a Markov Chain $X_1, X_2, \ldots$ over a state space Ω is

$$C \equiv \max_j \tau_j(1)$$

We use the $\equiv$ sign to emphasize that C is random. This will also be referred to as τ_{cov} or just cov or cover. Also, $C(z)$ denotes the cover time starting at z.

This is essentially the time it takes to cover every state in a Markov chain. Obviously this is undefined if the Markov chain has multiple absorbing states. We also have:

Definition 1.2. The *"cover-and-return" time* of a Markov Chain $X_1, X_2, \ldots$ over a state space Ω is

$$C^+ \equiv \min\{t \geq C : X_t = X_0\}$$

This is essentially the time to hit every element of Ω and return to the initial state.

We now turn to an application of cover times on random walks.

2. Random Walks on Unweighted Graphs

Definition 2.1. A *random walk* on a graph is the Markov chain with the following structure. We start at a vertex, and for each step we travel along one of the edges connected to the current vertex. If the probabilities of going along any two of these edges are equal for any vertex, then the random walk is said to be *unweighted*. Otherwise, the random walk is *weighted*.

Example. If the graph K_n^* is the complete graph with n vertices (and also self loops at each node to make it aperiodic), then its cover time for an unweighted random walk is $\theta(n \log n)$. This is equivalent to the coupon collector problem which we covered in class.

Example. Where L_n is a line of n vertices, its cover time is $\theta(n^2)$. To prove this, let us call the vertices 1 through n. If h_{xy} is the hitting time of y from x, we have $h_{12} = 1$ and

$$h_{i,i+1} = \frac{1}{2} + \frac{1}{2}(1 + h_{i-1,i+1}) = 1 + \frac{1}{2}(h_{i-1,i+1})$$

$$h_{i,i+1} = 1 + \frac{1}{2}(h_{i-1,i} + h_{i,i+1}).$$

This gives us a recurrence

$$h_{i,i+1} = 2 + h_{i-1,i}$$

which yields

$$h_{i,i+1} = 2i - 1.$$

Date: November 22, 2020.

This means that

$$\begin{aligned} h_{1,n} &= \sum_{i=1}^{n-1} h_{i,i+1} = \sum_{i=1}^{n-1} (2i-1) \\ &= 2\sum_{i=1}^{n-1} i - \sum_{i=1}^{n-1} 1 \\ &= 2\frac{n(n-1)}{2} - (n-1) \\ &= (n-1)^2 \end{aligned}$$

This brings us to our first theorem, after we define a few more things.

Definition 2.2. A *spanning tree* of a graph is a version of the graph with as many edges removed as possible so that it is still connected.

Definition 2.3. A depth-first search is a method of searching a graph in which each branch is followed to the end before backtracking.

Lemma 2.4. *If two vertices x and y are connected by an edge, we have*

$$h_{xy} + h_{yx} \leq 2m$$

where m is the number of edges on the graph. The sum $h_{xy} + h_{yx}$ is also called a commute time.

We will prove this Lemma later on.

Theorem 2.5. *Let G be a connected graph with n vertices and m edges. The expected time for a random walk to cover all vertices of G is bounded above by $4m(n-1)$*

Proof. Consider a depth-first search of G starting on some vertex z, and let T be the spanning tree of that search. This covers every vertex. Now, consider the time to cover every vertex in the order of the depth-first search. Clearly this bounds the cover time of G starting from z. Also note that each edge is traversed exactly twice, once in each direction. This gives us

$$C(z) \leq \sum_{(x,y)\in T,(y,x)\in T} h_{xy}$$

where h denotes hitting time. By Lemma 2.4, this is less than or equal to

$$2m \cdot 2 \cdot (n-1)$$

since there are $n-1$ edges in the spanning tree. This holds for all vertices z. ■

3. Electrical Networks

We can use Markov chains to model certain electrical networks, but first we need to lay out some ground work.

Definition 3.1. Each edge has a value for *conductivity*, which determines the probability of crossing that edge in a random walk. For an edge (x, y) this is denoted as c_{xy}. For a vertex x, its conductivity is

$$\sum_y c_{xy}.$$

We also have

$$p_{xy} = \frac{c_{xy}}{c_x}.$$

Note that the stationary distribution in this situation is defined as

$$\pi_x = \frac{c_x}{c_0}$$

where $c_0 = \sum_x c_x$. Additionally,

Definition 3.2. The *resistance* of an edge (x, y) is defined as

$$r_{xy} = \frac{1}{c_{xy}}$$

Definition 3.3. The *effective resistance* of a graph G is

$$r_{\text{eff}}(G) = \max_{x,y}(r_{xy})$$

This brings us to our next theorem

Theorem 3.4. *Let G be an undirected graph with m edges. Then*

$$mr_{\mathit{eff}}(G) \leq C(G) \leq 2e^3 mr_{\mathit{eff}} \log n + n$$

Before we prove this, we need another useful theorem.

Theorem 3.5. *Given an undirected graph, consider the electrical network where each edge is replaced with a 1-ohm resistor. Given vertices x and y, the commute time commute(x, y), equals $2mr_{xy}$, where $r_{\mathit{eff}}(x, y)$ is the effective resistance from x to y and m is the number of edges in the graph.*

However, before proving this, we need to learn about voltage and current. *Voltage* (v) is a property of every vertex in an electrical network. Normally we fix the voltages of two vertices and calculate the rest using equations which will be shown shortly. Also, current (i) is a property of any path connecting two vertices. We have the two following equations to help us determine the meaning of both:

$$i_{xy} = \frac{v_x - v_y}{r_{xy}} = (v_x - v_y)c_{xy}$$

(Ohm's law) and

$$\sum_y i_{xy} = 0.$$

(Kirchhoff's Law). Therefore,

$$\sum_y c_{xy}(v_x - v_y) = 0$$

$$\sum_y c_{xy}(v_x) = \sum_y c_{xy}(v_y)$$

$$c_x v_x = \sum_y c_{xy}(v_y)$$

$$v_x = \sum_y c_{xy}(v_y)\frac{1}{c_x}$$

$$v_x = \sum_y p_{xy}(v_y).$$

Thus, the voltage of a vertex is a weighted average of of the voltages of the adjacent vertices. This leads to the following result

Theorem 3.6. *If the voltage of vertex a is 1 and the voltage of vertex b is 0, then the voltage of a vertex x, v_x, is the probability that a random walk starting on x reaches a before b.*

Proof. Let us consider p_x, the probability that a random walk starting on x will reach a before b. It is obvious that $p_a = 1$ and $p_b = 0$. Furthermore, the probability of the random walk from x reaching a before b is equal to the sum over all y adjacent to x of $p_{xy}p_y$, because this is equivalent to starting on x, going to an adjacent y, and doing the random walk from there. Since p_x and v_x have equivalent definitions, they must be equal. ■

And finally,

Definition 3.7. The *effective resistance* from x to y r_{eff} is equal to

$$\frac{v_x - v_y}{i_{xy}}.$$

We are finally ready to prove theorem 3.5.

Proof. of theorem 3.5 Insert on each vertex i a current of d_i (the degree of i.) This means the total current is $2m$. Extract all of the current from a single vertex j. Now, let v_{ij} be the voltage difference from a vertex i to j. Now let k be a vertex adjacent to i. The current through the resistor between i and k is $v_{ij} - v_{jk}$, because the resistor has resistance 1. Thus, we must have

$$d_i = \sum_{k \text{ adj to } i} (v_{ij} - v_{jk}) = d_i v_{ij} - \sum_{k \text{ adj to } i} v_{kj}.$$

Solving for v_{ij},

$$v_{ij} = 1 + \sum_{k \text{ adj to } i} \frac{1}{d_i} v_{kj} = \sum_{k \text{ adj to } i} \frac{1}{d_i}(1 + v_{jk}).$$

Not that the hitting time h_{ij} can be expressed as

$$\sum_{k \text{ adj to } i} \frac{1}{d_i}(1 + h_{kj})$$

so we can subtract these to get

$$v_{ij} - h_{ij} = \sum_{k \text{ adj to } i} \frac{1}{d_i}(v_{jk} - h_{jk})$$

thus $v_{ij} - h_{ij}$ is harmonic. Let us set j as our boundary vertex, meaning we define

$$v_{jj} - h_{jj} = 0$$

(because this is obviously true). This means that $v_{ij} - h_{ij}$ is zero everywhere, meaning the voltage is equal to the hitting time. Now imagine that the current is extracted from i. This means that, following the same line of reasoning, $v_{ji} = h_{ji}$. Now reverse the current, and we get $-v_{ji} = h_{ji}$. Since $-v_{ji} = v_{ij}$, we have $v_{ij} = h_{ji}$.
Thus, when we apply a current of d_i at each i, we have $v_{ij} = h_{ij}$. When we reverse the

currents and extract the current from i, we have $v_{ij} = h_{ji}$. Now, superpose these situations. All current cancels except for $2m$ amps that flow from i to j. Thus,

$$2mr_{ij} = v_{ij} = h_{ij} + h_{ji}.$$

This also proves lemma 2.4. ■

Proof. of Theorem 3.4 By definition $r_{\text{eff}}(G) = \max_{x,y} r_{xy}$. Let u and v be the vertices for which r_{xy} is maximal. Then, $r_{\text{eff}}(G) = r_{uv}$. By theorem 3.5, we have $\text{commute}(u, v) = 2mr_{uv}$, so $mr_{uv} = \frac{1}{2}\text{commute}(u, v)$. The commute time from u to v and back to u is clearly less than twice $\max(h_{uv}, h_{vu})$, which is clearly less than $\text{cover}(G)$. Putting this together, we have the first half of the inequality

$$mr_{\text{eff}}(G) \leq \text{cover}(G).$$

Now for the second half of the inequality. By theorem 3.5, we know that for any x, y,$\text{commute}(x, y) = 2mr_{xy}$, which is less than or equal to $2mr_{\text{eff}}(G)$, meaning $h_{xy} \leq 2mr_{\text{eff}}(G)$. By Markov's inequality, since the expected time to reach y starting at any x is less than $2mr_{\text{eff}}(G)$ the probability that y is not reached from x in less than $2mr_{\text{eff}}(G)e^3$ steps is at most $\frac{1}{e^3}$. Thus, the probability that a vertex y has not been reached in $2e^3mr_{\text{eff}}(G)\log n$ steps is at most $\frac{1}{e^3}^{\log n} = \frac{1}{n^3}$ because a random walk of length $2e^3mr_{\text{eff}}(G)\log n$ steps is a sequence of $\log n$ independent random walks of length $2e^3mr_{\text{eff}}(G)$. Suppose after a walk of length $2e^3mr_{\text{eff}}(G)$ vertices $v_1, v_2, \ldots, v_l$ had not been reached. Walk until v_1 has been reached, then v_2 and so on. Each of these has expected time at most n^3 but each only happens with probability $\frac{1}{n^3}$, so essentially we add $O(1)$ for each v_i, a total at most n. So we have

$$\text{cover}(G) \leq 2e^3mr_{\text{eff}}(G)\log n + \sum_v \frac{1}{n^3}n^3 \leq 2e^3mr_{\text{eff}}(G)\log n + n.$$

I did not go into the reason why it is n^3 but it's fairly simple to show that this is greater than or equal to than the number in theorem 3.5. ■

4. Sources

(1) https://www.stat.berkeley.edu/~aldous/RWG/Chap6.pdf
(2) https://people.csail.mit.edu/ronitt/COURSE/S14/Handouts/Lec_6_scribe.pdf
(3) https://www.win.tue.nl/~jkomjath/SPBlecturenotes.pdf
(4) https://www.cs.cmu.edu/~avrim/598/chap5only.pdf

E-mail address: lorenzo.w@sonic.net