1 Random Walks and Electrical Networks

Random walks are widely used tools in algorithm design and probabilistic analysis and they have numerous applications. Given a graph and a starting vertex, select a neighbor of it uniformly at random, and move to this neighbor; then select a neighbor of this point at random, and move to it etc. The random sequence of vertices selected this way is a random walk on the graph.

1.1 Basic Definitions

More formally, consider a random walk on $G(V, E)$, starting from $v_0 \in V$: we start at a node v_0; if at the t-th step we are at a node v_t, we move to a neighbor of v_t with probability $\frac{1}{d(v_t)}$, where $d(v)$ is the degree of node v. The node v_0 may be fixed, but may itself be drawn from some initial distribution P_0. We denote by P_t the distribution of v_t:

$$P_t(i) = Pr[v_t = i].$$

We denote by $M = (p_{ij})_{i,j \in V}$ the matrix of transition probabilities where $p_{ij} = \frac{1}{d(i)}$ and 0 otherwise. In matrix form, we can write M as DA where A is the adjacency matrix of G and D denotes the diagonal matrix with $(D)_{ii} = \frac{1}{d(i)}$. The evolution of probabilities can be expressed by the simple equation $P_{t+1} = M^T P_t$, and hence

$$P_t = (M^T)^t P_0.$$

We say that the distribution P_0 is stationary (or steady-state) for the graph G if $P_1 = P_0$. In this case, of course, $P_t = P_0$ for all $t > 0$. For every graph G, the distribution

$$\pi(i) = \frac{d(i)}{2m},$$

is stationary.

1.2 Markov Chains

Random walks on graphs are special cases of Markov chains, from the general theory of which we will state a few relevant results. A **Markov chain** $\{X_0, X_1, ...\}$ is a discrete-time random process defined over a set of states S with a matrix P of transition probabilities. At each time step t the Markov chain is in some state $X_t \in S$; its state X_{t+1} at the next time step depends only on its current state, not the current time or any previous states. Thus the only information needed to specify the random variables $\{X_0, X_1, ...\}$ is a distribution over the starting state X_0 and the transition probabilities p_{ij}, the probability that the next state will be j, given that the current state is i. The correspondence between random walks and Markov chains is made clear by taking $S = V$ and $P = M$.

A **stationary distribution** for a Markov chain is a probability distribution π over the states such that $\pi = P^T \pi$.

A Markov chain is **irreducible** if for any two states i and j, there is a path of positive probability transitions from i to j and vice versa.

The periodicity of a state i is the maximum integer T for which there exists an initial distribution q_0 and positive integer a such that if t does not belong to the arithmetic progression $\{a + Ti \mid i \geq 0\}$ then $q_t(i) = 0$. An **aperiodic** Markov chain is one in which all states have periodicity 1.

Theorem 1 *(Fundamental Theorem of Markov Chains) Any irreducible, finite, and aperiodic Markov chain has the following properties:*

1. *There is a unique stationary distribution* π *such that* $\pi(i) > 0$ *for all states* i.
2. *Let* $N(i,t)$ *be the number of times the Markov chain visits state* i *in* t *steps. Then*

$$\lim_{t\to\infty} \frac{N(i,t)}{t} = \pi(i).$$

In applying this theorem to random walks on simple connected graphs, we first note that the periodicity of every state (vertex) is the greatest common divisor of the length of all closed walks in G. Thus the random walk is aperiodic iff G is not bipartite. G connected implies the random walk is irreducible. We may conclude that if G is connected and non-bipartite, then the distribution of v_t tends to the unique stationary distribution $\pi(i) = d(i)/2m$, as $t \to \infty$, for any initial distribution P_0. Additionally we may see that the expected number of steps between two visits of a vertex i is $2m/d(i)$.

One interesting property of random walks on undirected graphs is their time-reversibility: a random walk considered backwards is also a random walk. This can be verified by observing $p_{ij}\pi(i) = p_{ji}\pi(j)$. A heuristic motivation for the above Markov chain-based result is stated as follows. Time-reversibility implies that a stationary walk passes through an edge in every direction with the same frequency of on average once every $1/2m$ times. By the same argument, the expected number of steps between two visits of a vertex i is also $2m/d(i)$.

1.3 The Gambler's Ruin on a Graph

A gambler enters a casino with a plan to play the following game. At each turn he will bet 1 dollar to win 2 dollars with probability 1/2 and lose his money with probability 1/2. He is determined to leave either when he is "ruined" (i.e. he has no money left) or as soon as he collects N dollars. What is ϕ_k the probability that he leaves with N dollars if he starts with $k \le N$ dollars?

The intuitive answer is k/N. Since this is a fair game, the gambler should leave with the same amount of money on average as when he started. Thus, starting with k dollars, his final expected value is $\phi_k N + (1-\phi_k)0$ and hence $\phi_k = k/N$. One can make this argument formal with the help of the optional stopping theorem.

We can also see this process as a random walk on a path graph. We can write a system of equations determining the value of every node

$$\phi_0 = 0,\ \phi_N = 1,\ \text{and}\ \phi_x = 1/2(\phi_{x-1} + \phi_{x+1})$$

which can easily be verified to have the solution $v_k = k/N$.

Now, consider a generalization. Suppose we have a random walk in a graph like the one introduced at the start of the lecture. We are placed on some starting vertex and may walk through the maze until we reach either vertex a or vertex b. If we hit vertex b we win 1 dollar and we get nothing if we reach vertex a. What is the probability ϕ_x we win 1 dollar starting at vertex x? Similar to the above example we can write the system of equations

$$\phi_a = 0,\ \phi_b = 1,\ \phi_x = \frac{1}{d(x)} \sum_{y::y\sim x} \phi_y \tag{1}$$

where $d(x)$ is the degree of vertex x.

1.4 Graphs as Electrical Circuits

We may view a graph $G(V,E)$ as an electrical network (or circuit). This network consists of wires such that each $(x,y) \in E$ is a resistor with resistance r_{xy} across which flows a current i_{xy}. The intersections of the wires at vertices of V has a voltage v_x measured at $x \in V$. For any such electrical network the following properties hold.

1. Kirchoff's Law: For any vertex, the current flowing in equals the current flowing out.
$$\sum_{y\,:\,x\sim y} i_{xy} = 0 \qquad \forall\ x \in V$$
2. Ohm's Law: The voltage difference across any edge is equal to the current times the resistance.
$$v_x - v_y = r_{xy} i_{xy} \qquad \forall\ (x, y) \in E$$

Let us consider a network in which each resistor has a resistance of 1 Ohm. Furthermore, put a voltage or charge of 1 volt at vertex a and ground vertex b. Then $v_a = 0$, $v_b = 1$ and by Ohm's law we have that $i_{xy} = v_x - v_y$ for all $x \neq y \neq b$ with $(x, y) \in E$. Applying Kirchoff's law yields that $\sum_{y\,:\,x\sim y}(v_x - v_y) = 0$, which yields the following system of equations:

$$v_a = 0,\ v_b = 1,\ v_x = \frac{1}{d(x)} \sum_{y\,:\,x\sim y} v_y.$$

Observe that the voltages and probabilities in equation (1) follow the same law. Both ϕ and v are harmonic functions: the value of each function at a node besides a and b is the average of its values of its neighbors. They have the same value at at the boundary (nodes a and b). So by uniqueness principle[1] they have the same value everywhere.

The above argument provides a connection which can be exploited. In particular, various methods and results from the theory of electricity often motivated by physics, can be applied to provide results about random walks. We will explore a few of these.

Escape probability. Let us start by computing P_{esc}, the probability that the random walk starting at a reaches b before returning to a.

$$P_{esc} = \frac{1}{d(a)} \sum_{x\,:\,x\sim a} (v_x - v_a) = \frac{1}{d(a)} \sum_{x\,:\,x\sim a} v_x\ .$$

Observe that $\sum_{x\,:\,x\sim a} v_x = \sum_{x\,:\,x\sim a} i_{ax} = i_a$ which is the total current leaving vertex a with $v_a = 0$ and $v_b = 1$. The total current between two points when one unit of voltage is applied between them is referred to as the *effective conductance* between a and b. The *effective resistance* between a and b is the inverse quantity.

The effective resistance comes up quite often in the context of random walks. Here is another useful example:

Proposition 1 *Let a and b be two adjacent vertices in G. The probability that a random walk starting at a visits b for the first time using the edge $\{a, b\}$ is $R_{eff}(a, b)$ in the corresponding electrical network.*

Proof: First note that since a and b are adjacent, $R_{eff}(a, b) \leq 1$. Let us denote the desired probability by q. Every time the random walk leaves a, it has a probability $\frac{1}{d(a)}$ of taking $\{a, b\}$ before returning to either a or b. It also has a probability $1 - P_{esc}$ of not visiting b before returning to a. In that case, the probability of visiting b using the edge $\{a, b\}$ is again q. Therefore, $q = \frac{1}{d(a)} + (1 - P_{esc})q$ and hence $q = \frac{1}{d(a)P_{esc}} = R_{eff}(a, b)$. ■

[1] See `http://en.wikipedia.org/wiki/Potential_theory`

Commute time. The expected **hitting time** $H(x, y)$ between two nodes x and y is the expected number of steps a random walk starting at vertex x takes before it visits vertex y for the first time. The commute time $\kappa(x, y) = H(x, y) + H(y, x)$ is the expected time to visit y and return to x.

Suppose that for each $x \in V$, we apply (externally) $d(x)$ units of current and at some node y we remove $\sum_{x \in V} d(x) = 2m$ units of current. Here $m = |E|$ the number of edges (wires). Then using Ohm's and Kirchoff's laws we can show that

$$v_x - v_y = 1 + \frac{1}{d(x)} \sum_{z\,:\,z \sim x} (v_z - v_y)$$

Similarly, for a random walk, the expected time it takes to reach node y from x is equal to 1 (for the current step) plus the expected time from each of the neighbor nodes times the probability of going to those neighbors. In other words,

$$H(x, y) = 1 + \sum_{z\,:\,z \sim x} H(z, y)/d(x)$$

The laws are identical so we conclude that $H(x, y) = v_x - v_y$.

Now consider repeating the above argument, exchanging the roles of x and y. We obtain $H(y, x) = v_y - v_x$ where v_y, v_x are determined by pushing $d(v)$ units of current into each node $v \in V$ and removing $2m$ from x. If instead we pull $d(v)$ units of current from each v and push $2m$ units into x, only the signs reverse and we get $H(y, x) = v_x - v_y$. We then super-pose this reversed current flow on top of the original. All current flows cancel except for $2m$ into x and $2m$ out of y, and the potentials simply add so that the potential difference between x and y is $H(x, y) + H(y, x) = \kappa(x, y)$. Thus we obtain the following well-known result.

Theorem 2 *For all pairs of vertices x and y the* commute time *$\kappa(x, y)$ between x and y is given by*

$$\kappa(x, y) = H(x, y) + H(y, x) = 2mR_{eff}(x, y),$$

where $R_{eff}(x, y)$ is the effective resistance between x and y.

Cover time. The cover time C starting from a given distribution is the expected number of steps to reach every node. If no starting distribution is specified, we mean to start the walk from the stationary distribution. Cover time is particular to the graph the random walk is taking place over; we encourage the reader to try to calculate the cover time for a few special classes of graphs such as complete graphs, paths, and a path connected to a complete graph (also known as the lollipop graph!). For the sake of brevity, we will mention only the following bound for cover time known as Matthew's bound:

Theorem 3 *The cover time starting from any node of a graph with n nodes is at most* $\log n$ *times the maximum hitting time between any two nodes. In other words,*

$$\max_{i,j \in V} H(i, j) \leq C \leq \max_{i,j \in V} H(i, j) \log n.$$

The lower bound is obvious. We will sketch the upper bound with $\log n$ replaced with $2 \log n$.

Proof: Let b be the expected number of steps before a random walk visits more than half of the nodes, and let h be the maximum hitting time between any two nodes. We will show that $b \leq 2h$. The theorem is easily followed afterwards. In $2h$ steps we have seen more than half of all nodes. By a similar argument, in another $2h$ steps we have seen more than half of the rest etc.

Now to prove the statement assume, for simplicity, that $n = 2k + 1$ is odd. Let t_v be the time when node v is first visited. Then the time β when we reach more than half of the nodes is the $(k + 1)$-st largest of the t_v. Hence

$$\sum_v t_v \geq (k+1)\beta$$

and so

$$b = E(\beta) \leq \frac{1}{k+1} \sum_v E(t_v) \leq 2h.$$

■