CS 294-2 Fall 2004 | Quantum Walks | 10/21/04 Lecture 13

Quantum Walks

1 Classical Random Walks

Random walks on graphs are used in designing algorithms for many sampling and counting problems. For example, assume that we are given a graph G as input, and we want to compute the (approximate) number of spanning trees of G. It can be shown that this problem is essentially equivalent to the problem of generating a uniformly random spanning tree. We can uniformly sample spanning trees by a random walk on another graph H. The vertex set of H consists of all spanning trees of H. Two spanning trees T and T' are adjacent in H if and only if T' is obtained from T by removing one edge and adding a new one.

In order to ensure that random walks on a graph H converge to a unique stationary distribution, the underlying graph H should satisfy the following two properties:

- H should be connected.
- H should not be bipartite, so that the random walk is aperiodic.

It is easy to check that the stationary distribution of a random walk on a simple graph is uniform on edges, i.e. the probability of traversing each edge is the same.

Sometimes we consider random walks on weighted undirected graphs, where each edge (x,y) has some weight w_{xy}. For each vertex x, let $w_x = \sum_y w_{xy}$ denote the sum of weights of all edges incident to x. If $P = (p_{xy})$ denotes the transition probability matrix of the Markov chain associated to the random walk, i.e. p_{xy} be the probability of moving to vertex y when we are at x, then $p_{xy} = w_{xy}/w_x$. The stationary distribution over vertices is given by $\pi_x = w_x/\sum_y w_y$. For the most part of this lecture, we assume that we walk on regular graphs, i.e. $w(x) = w(y)$ for every x and y. For regular graphs, the matrix P is symmetric, and the uniform distribution $\pi_x = 1/N$ is stationary. (N is the number of vertices.)

Consider a random walk on a regular graph. If the vector $|v\rangle$ is a distribution over vertices, then $P|v\rangle$ is the distribution of the vertices after one step of the random walk. Since P is symmetric, it has N real eigenvectors $|v_1\rangle, \ldots, |v_N\rangle$ that form an orthogonal basis for R^N. Let $\lambda_1 \geq \ldots \geq \lambda_N$ be the corresponding eigenvalues in ascending order. Moreover, P is stochastic, i.e. it has nonnegative entries and its columns sum to 1. (In fact P is doubly stochastic; both its rows and columns sum to 1.) Thus, the eigenvectors of P are between 1 and -1. For the stationary distribution $|\pi\rangle$, we have $P|\pi\rangle = |\pi\rangle$. Hence, $\lambda_1 = 1$.

Remark. We can assume that all eigenvectors of P are nonnegative. If P has negative eigenvectors, we can modify the random walk by adding self-loops with probability 1/2. The eigenvalues of the new transition matrix $(P+I)/2$ are then $(\lambda_1+1)/2, \ldots, (\lambda_N+1)/2$.

It turns out that the rate of convergence of the random walk is governed by the gap $1-\lambda_2$ between the first and the second largest eigenvalues. To see why, let $|v\rangle$ be a distribution over vertices. $|v\rangle$ is a linear combination $\sum_i \alpha_i |v_i\rangle$ of eigenvectors, where $|v_1\rangle$ is the stationary distribution. The distribution after t steps of the random walk would be $\sum_i \alpha_i \lambda_i^t |v_i\rangle$. Therefore, $1-\lambda_2$ determines how fast non-stationary eigenvectors diminish.

Let M be a subset of vertices of the graph that are marked. The hitting time T of the random walk is the number of steps needed to encounter a marked vertex. Let P_M denote the matrix obtained from P by removing rows and columns corresponding to vertices in M. Since the sum of elements of each column of P is at most 1, the eigenvalues of P_M are at most 1.

Fact 1. The expected hitting time $E[T]$ is $\leq \frac{1}{1-\lambda}$, where λ is the largest eigenvalue of P_M.

Fact 2. If $|M| = \varepsilon N$ and the eigenvalues of P are $1, \lambda_2, \ldots$, then $1-\lambda \geq \frac{\varepsilon(1-\lambda_2)}{2}$.

We will see that quantum walks achieve a quadratic speed-up in terms of hitting time.

1.1 Example: Hypercubes.

Let us now look at random walks on hypercubes as an example. The n-dimensional hypercube is a graph with vertex set $\{0,1\}^n$. Two vertices are connected if they differ in exactly one position.

Since the hypercube is bipartite, we modify the random walk on the hypercube by adding self-loops of probability $1/2$ to it. That is, at each vertex x, we go to a neighbor vertex with probability $1/2$ and we stay at x with probability $1/2$. But this is equivalent to choosing a random position $1 \leq i \leq n$ and then setting x_i, the ith position of x, to a random value. We see that the distribution of the random walk is uniform after t steps if each of the n positions has been chosen at least once after the t steps. Therefore, the mixing time of the random walk is determined by the coupon collector's problem, which is $O(n \log n)$.

We can also get the same result by looking at the eigenvalues $\lambda_1, \ldots, \lambda_{2^n}$ of the transition matrix. (We consider the original transition matrix, the one without self-loops; note that eigenvalues of the modified transition matrix are $\{(\lambda_i+1)/2)\}$). For every $s \in \{0,1\}^n$, the 2^n-dimensional vector $\chi_s(x) = (-1)^{s.x}$ is an eigenvector with eigenvalue $1-2|s|/n$, where $|s|$ is the number of nonzero components of s. Notice that the gap between the first and second largest eigenvector is $2/n$.

2 Quantization of Markov Chains

Let $P = (p_{xy})$ be the transition matrix of a classical Markov chain with state space X consisting of N states. We define a quantum walk that corresponds to P.

The quantum walk operates on the Hilbert space $C^N \otimes C^N$ with basis states $\{|x\rangle\,|y\rangle : x, y \in X\}$. Define

$$\begin{aligned} |\phi_x\rangle &= \sum_{y \in Y} \sqrt{p_{xy}}|x\rangle\,|y\rangle, \text{ for } x \in X, \\ |\psi_y\rangle &= \sum_{x \in X} \sqrt{p_{yx}}|x\rangle\,|y\rangle, \text{ for } y \in X. \end{aligned}$$

Let E_1 and E_2 be the subspaces spanned by $\{|\phi_x\rangle\}$ and $\{|\psi_y\rangle\}$ respectively. If we had operators $T_1 : C^N \otimes C^N \to E_1$ and $T_2 : C^N \otimes C^N \to E_2$ such that for arbitrary $|r\rangle$, T_1 mapped $|x\rangle\,|r\rangle$ to $|\phi_x\rangle$, and T_2 mapped $|r\rangle\,|y\rangle$ to $|\psi_y\rangle$, then by applying T_1 and T_2 alternately, we could get an analogue of the Markov chain P. However, T_1 and T_2 are not realizable in quantum mechanics since they are not reversible.

Instead, we will define unitary operators R_1 and R_2 that serve as quantum analogues of T_1 and T_2: Let π_1 and π_2 denote orthogonal projections on E_1 and E_2. We define $R_1 = 2\pi_1 - 1$ and $R_2 = 2\pi_2 - 1$ to be reflections with respect to E_1 and E_2 respectively. Note that R_i keeps E_i pointwise fixed (as does T_i), and tosses up the rest of the space as much as possible.

Each step of the quantum walk is given by $W_P = R_2 R_1$. Running the quantum walk for t steps corresponds to applying W_P^t.

When P is symmetric, the state

$$\left|\Phi_0\right\rangle = \frac{1}{\sqrt{N}} \sum_{x,y} \sqrt{p_{x,y}} \left|x\right\rangle \left|y\right\rangle$$

lies in $E_1 \bigcap E_2$. Thus $W_P \left|\Phi_0\right\rangle = \left|\Phi_0\right\rangle$.

3 Quantum Hitting Time

Assume that some subset M of states of a Markov chain P are marked. The hitting time of M is defined as the number of iterations necessary to encounter an element of M. For the purpose of analyzing the hitting time, we can modify the random walk such that as soon as we reach some vertex $x \in M$, we never leave x. The transition matrix corresponding to the modified random walk equals

$$\tilde{P} = \begin{pmatrix} P_M & P' \\ 0 & I \end{pmatrix},$$

where P_M is the submatrix of P obtained by deleting rows and columns of M.

Given a symmetric matrix P and subset M of marked states, we will describe how to check whether M is empty or not by running the quantum walk $W_{\tilde{P}}$. We start by setting the initial state of the walk to $\left|\Phi_0\right\rangle$, the stationary state for W_P. Next, we run the quantum walk for t steps to get $\left|\Phi_t\right\rangle = W_{\tilde{P}}^t \left|\Phi_0\right\rangle$. If M is empty, then $P = \tilde{P}$ and $\left|\Phi_t\right\rangle = \left|\Phi_0\right\rangle$. If M is not empty, then we will show that for large enough T, if the number of steps t is chosen at random from $\{1, \ldots, T\}$, then $E_t[\langle \Phi_t \left| \Phi_0 \right\rangle^2] = 1 - \Omega(1)$. So if we measure $\left|\Phi_t\right\rangle$ along $\left|\Phi_0\right\rangle$, with probability $\Omega(1)$, the result of the measurement is not $\left|\Phi_0\right\rangle$. Therefore we can distinguish between the case where M is empty and the case where it is not. Notice that the above algorithm is not a search algorithm, that is, we only check the existence of a marked state, and cannot necessarily find one.

The minimum T that can be used in the above algorithm is called the quantum hitting time of M.

Theorem (Szegedy). The quantum hitting time of M is $O(\frac{1}{\sqrt{1-\|P_M\||}})$ where $\|P_M\|$ is the operator norm of P_M. ($\|P_M\|$ is the largest eigenvector of P_M too).

Proof. Let $n = N - |M|$. Let $\left|v_1\right\rangle, \ldots, \left|v_n\right\rangle \in R^n$ be an orthonormal basis of real eigenvectors for P_M with eigenvalues $\lambda_1, \ldots, \lambda_n$.

Let $1 \leq i \leq n$. Suppose $\left|v_i\right\rangle = \sum_x \alpha_{xi} \left|x\right\rangle$, where we assume $\alpha_{xi} = 0$ for all $x \in M$. Define

$$\left|e_{1i}\right\rangle = \sum_x \alpha_{xi} \left|\phi_x\right\rangle = \sum_{x,y} \alpha_{xi} \sqrt{\tilde{p}_{xy}} \left|x\right\rangle \left|y\right\rangle \in E_1,$$
$$\left|e_{2i}\right\rangle = \sum_y \alpha_{yi} \left|\psi_y\right\rangle = \sum_{x,y} \alpha_{yi} \sqrt{\tilde{p}_{yx}} \left|x\right\rangle \left|y\right\rangle \in E_2.$$

We have $\langle e_{1i} \left| e_{2i} \right\rangle = \langle v_i | P_M \left| v_i \right\rangle = \lambda_i$. Moreover, we have $\pi_2 \left|e_{1i}\right\rangle = \lambda_i \left|e_{2i}\right\rangle$ since $\left|e_{1i}\right\rangle - \lambda_i \left|e_{2i}\right\rangle$ is orthogonal to every $\left|\psi_y\right\rangle$. Similarly, $\pi_1 \left|e_{2i}\right\rangle = \lambda_i \left|e_{1i}\right\rangle$.

Let V_i denote the subspace generated by $\left|e_{1i}\right\rangle$ and $\left|e_{2i}\right\rangle$. It is easy to see that $V_1, \ldots, V_n$ are invariant subspaces of R_1 and R_2. In fact, R_j reflects every vector in V_i with respect to $\left|e_{ji}\right\rangle$. Thus, under $W_{\tilde{P}} = R_2 R_1$, every vector in subspace V_i is rotated by angle $2\theta_i$, where $\theta_i = \cos^{-1} \lambda_i$ is the angle between $\left|e_{1i}\right\rangle$ and $\left|e_{2i}\right\rangle$.

It is not hard to see that $V_1, \ldots, V_n$ are orthogonal. Indeed, for $i \neq j$, the orthogonality of $\left|v_i\right\rangle$ and $\left|v_j\right\rangle$ implies $\langle e_{1i} \left| e_{1j} \right\rangle = \langle e_{2i} \left| e_{2j} \right\rangle = 0$, and $\pi_2 \left|e_{1i}\right\rangle = \lambda_i \left|e_{2i}\right\rangle$ implies $\langle e_{1i} \left| e_{2j} \right\rangle = 0$.

Let $V^\perp$ denote the orthogonal complement of the subspace $V = V_1 + \cdots + V_n$ in the Hilbert space of the quantum walk. We will show that $V^\perp$ is another invariant subspace of R_1 and R_2. Let $\left|\phi\right\rangle$ and $\left|\phi^\perp\right\rangle$ be two arbitrary vectors in V and $V^\perp$ respectively. We will show that $\left|\phi\right\rangle$ and $\pi_1\left|\phi^\perp\right\rangle$ are orthogonal. Since $\langle\phi|\pi_1\left|\phi^\perp\right\rangle = \langle\phi|\pi_1\pi_1\left|\phi^\perp\right\rangle$, we can instead show that $\pi_1\left|\phi\right\rangle$ and $\pi_1\left|\phi^\perp\right\rangle$ are orthogonal. But this is true because $\pi_1\left|\phi\right\rangle$ is a linear combination of $\{\left|e_{1i}\right\rangle\}$, and $\left|\phi^\perp\right\rangle$, being in $V^\perp$, is orthongonal to all $\left|e_{1i}\right\rangle$. Thus, $V^\perp$ is invariant under π_1, and hence under R_1, and similarly under R_2.

Therefore, the operation of the quantum walk, $W_{\tilde{P}}$, can be decomposed through the direct sum decomposition $V_1 + \cdots + V_n + V^\perp$. Consider the initial state

$$\left|\Phi_0\right\rangle = \frac{1}{\sqrt{N}}\sum_{x,y}\sqrt{p_{xy}}\left|x\right\rangle\left|y\right\rangle = \frac{1}{\sqrt{N}}\sum_{x\notin M}\left|\phi_x\right\rangle + \frac{1}{\sqrt{N}}\sum_{x\in M}\sum_{y}\sqrt{p_{xy}}\left|x\right\rangle\left|y\right\rangle.$$

For $x \notin M$, the vector $\left|\phi_x\right\rangle$ is a linear combination of $\{\left|e_{1i}\right\rangle\}$. For $x \in M$, the vector $\left|x\right\rangle\left|y\right\rangle$ is orthogonal to $\{\left|e_{1i}\right\rangle\}$ and $\{\left|e_{2i}\right\rangle\}$. Thus, if $\left|u_i\right\rangle$ denotes the projection of $\left|\Phi_0\right\rangle$ onto V_i, then $\sum_i\langle u_i\left|u_i\right\rangle = n/N$.

We may assume $n/N \geq 1/2$, since otherwise at least half of the vertices are marked and we can solve the problem by random sampling. Since the quantum walk $W_{\tilde{P}}$ rotates vectors in subspace V_i by angle $2\theta_i$, if we run the quantum walk for t steps, where t is chosen at random from $\{1,\ldots,\Omega(1/|\theta_i|)\}$, then $E_t[|\langle u_i|W_{\tilde{P}}^t\left|u_i\right\rangle|] = 1-\Omega(1)$. Therefore, for some $T = O(1/\theta)$ where $\theta = \min_i|\theta_i|$, when t is chosen from $\{1,\ldots,T\}$, we have $E_t[|\langle\Phi_t\left|\Phi_0\right\rangle|] \leq 1-\Omega(1)\sum_i\langle u_i\left|u_i\right\rangle = 1-\Omega(1)$. This implies that there exists positive constant c such that $\Pr[|\langle\Phi_t\left|\Phi_0\right\rangle| < 1-c] > c$. Hence $E_t[\langle\Phi_t\left|\Phi_0\right\rangle^2] = 1-\Omega(1)$. Since $\theta \geq \sin\theta \geq \sqrt{1-\cos\theta} = \sqrt{1-||P_M||}$, we have proved that the hitting time is $O(\frac{1}{\sqrt{1-||P_M||}})$.

4 Element Distinctness

In this section, as an application of quantum random walks and their hitting time, we solve the element distinctness problem.

Definition. In the element distinctness problem, we are given a function $f : \{1,\ldots,n\} \to \{1,\ldots,m\}$, where $n \leq m$, and we want to check whether f is one-to-one.

Classically, any algorithm for this problem requires $\theta(n)$ queries of f. Quantumly, Ambainis has shown that $\theta(n^{2/3})$ queries is necessary and sufficient.

Here is an algorithm for the problem:

1. Start with some subset $S \subseteq \{1,\ldots,n\}$ of size r.
2. Check if there are two elements $x,y \in S$ such that $f(x) = f(y)$.
3. Remove a random element of S, add a random new element to S, and repeat step 2.

The algorithm is essentially a random walk on a graph whose vertices are subsets of size r of $\{1,\ldots,n\}$. There is an edge between two subsets if and only if they differ in exactly two elements. The marked vertices M are those subsets S that contain elements $x,y \in S$ such that $f(x) = f(y)$. The total number of vertices is $N = \binom{n}{r}$. If the function is not one-to-one, then the number of marked states is at least $\binom{n-2}{r-2}$. Thus,

$$\frac{|M|}{N} \geq \frac{\binom{n-2}{r-2}}{\binom{n}{r}} = \frac{r(r-1)}{(n-r+2)(n-r+1)} \approx \frac{r^2}{n^2}.$$

It is known that the second eigenvalue of the above graph is approximately $1-1/r$. Therefore, $1-||P_M|| = \Omega(r/n^2)$.

The number of queries that the algorithm makes is r (for the first step) plus the hitting time of M (for the second step). Classically, this is $O(r+n^2/r)$, which is at best $O(n)$. Quantumly, the number of queries is $O(r+n/\sqrt{r})$, which is $O(n^{2/3})$ when $r = n^{2/3}$.