Topics in Quantum Information	**U. Waterloo CO 781, Spring 2004**
Ashwin Nayak (Instructor)	**Kelly Itakura (Scribe)**
Classical hitting time. Szegedy quantization of Markov chains	**July 6, 2004**

1 Classical hitting time

In the last lecture, we derived an expression for the hitting time of a classical Markov chain in terms of the transition matrix and the initial distribution.

To recall, let P be an $N \times N$ symmetric stochastic matrix that defines a Markov chain (M.C.) on state space X. Suppose that the MC is also irreducible and aperiodic. Let $M \subseteq X$ be a subset of m marked elements, and P_M be the matrix obtained by deleting rows and columns of P corresponding to M.

Fact 1 *The hitting time of M from a given distribution s on X is $s_M^T(I - P_M)^{-1} \cdot \boldsymbol{1}$, where s_M^{T} is the vector s restricted to M.*

Note that P_M is orthogonally diagonalizable, and has spectral norm strictly smaller than 1, and so $I - P_M$ is invertible. If (λ_i, v_i) are a complete set of eigenvalues with the corresponding orthogonal eigenvectors for P_M, then $I - P_M$ has the same eigenvectors but with eigenvalues $\frac{1}{1-\lambda_i}$.

By choosing the initial state to be the stationary distribution (in this case, the uniform distribution), we get a good bound on hitting time.

Corollary 2 *If $s = (\frac{1}{N}, \frac{1}{N}, \ldots, \frac{1}{N})$, then hitting time $\mathrm{E}(T)$ is*

$$\mathrm{E}(T) \quad = \quad \frac{1}{N}\boldsymbol{1}^{\mathrm{T}} \cdot (I - P_M)^{-1} \cdot \boldsymbol{1}$$

If $\frac{\boldsymbol{1}}{\sqrt{N}} = \sum_{i=1}^{N-m} \nu_i v_i$ then,

$$\mathrm{E}(T) \quad = \quad \sum_{i=1}^{N-m} \nu_i^2 \frac{1}{1-\lambda_i} \quad \leq \quad \frac{1}{1 - \| P_M \|}. \tag{1}$$

Here, we have made use of the fact that $\| P_M \|$ is the size of the eigenvalue with the largest magnitude, and that $\sum_i \nu_i^2 = 1 - m/N \leq 1$.

If λ_1 is the eigenvalue with maximum magnitude, and $\nu_1 = \Omega(1)$, then $\mathrm{E}(T) = \Theta\left(\frac{1}{1-\| P_M \|}\right)$. We will see later that the hitting time of the Szegedy quantitization of this Markov chain is the square-root of this quantity, and hence gives a quadratic speed-up.

2 Analysis of the random walk algorithm for Element Distinctness

Using the results above, we can improve the crude upper bound we gave in the last lecture for the Element Distinctness algorithm based on random walk.

We will need the following lemma with which we can bound the eigenvalues of P_M, given the spectrum of P. Define $\sigma_i(P)$ to be ith largest singular value of P. Note that since P has a uniform stationary distribution, $\sigma_i(P) = \lambda_1 = 1$. So the spectral gap, which is the gap between the two largest singular values, is $\sigma_1(P) - \sigma_2(P) = 1 - \sigma_2(P)$.

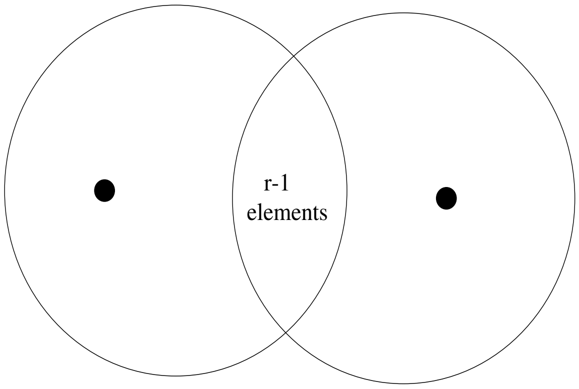

Figure 1: r-subsets x, y for which $p_{x,y} \neq 0$.

Lemma 3 *If the spectral gap $1 - \lambda_2(P)$ of P is $\geq \delta$, and if $m/N = \frac{|M|}{|X|} \geq \mu$, then $\| P_M \| \leq 1 - \frac{\delta\mu}{2}$.*

We omit the proof of the lemma.

Now we would like an explicit description of the matrix P so that we can compute the the spectral gap for the matrix. Recall from the last lecture, that the state space X of the walk is the collection of r-subsets of $[n]$. Given an r-subset x, there are $r(n-r)$ other r-subsets y reachable by swapping one of the r elements in x with one of the $n-r$ elements not in the subset. Each of these $r(n-r)$ subsets y is equally likely to be chosen as the next state of the Markov chain. Then

$$N \;=\; |X| \;=\; \binom{n}{r},$$

and

$$\begin{aligned} p_{x,y} &= \begin{cases} \frac{1}{r(n-r)} & \text{if } |x \cap y| = r-1 \\ 0 & \text{otherwise} \end{cases} \\ &= \frac{J_{n,r,r-1}}{r(n-r)}, \end{aligned}$$

where $J_{n,r,r-1}$ is a boolean matrix indexed by r-subsets with the (x,y)th entry $= 1$ iff the intersection $x \cap y$ is of size $r-1$.

We have encountered the collection of matrices $J_{n,k,t}$ several times in class before—once in the communication protocol for sampling disjoint subsets, and once in the Razborov lower bound for the communication complexity of Set Disjointness. The spectrum of the matrices is well known. The matrix $J_{n,k,t}$ has exactly $k+1$ distinct eigenvalues, and for $k = r$ and $t = r-1$, these are:

$$\lambda_j \;=\; (r-j)(n-r) - j(r-j+1), \quad 0 \leq j \leq r.$$

Note that λ_j is a decreasing function of j, and ranges from $r(n-r)$ to $-r$. The eigenvalues of P are $\lambda_j/r(n-r)$. We may assume that $r \in o(n)$. Otherwise, the algorithm may as well query f at all n points and solve the problem. Thus, the eigenvalues of P lie in the range $[1, -1/n - o(1/n)]$. From this, we can compute the spectral gap $1 - \frac{\lambda_1}{r(n-r)} = \frac{n}{r(n-r)} > \frac{1}{r}$.

The marked states in M are the r-subsets that contain a colliding pair of elements. This collection is smallest when we have exactly one pair of colliding elements. Then,

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

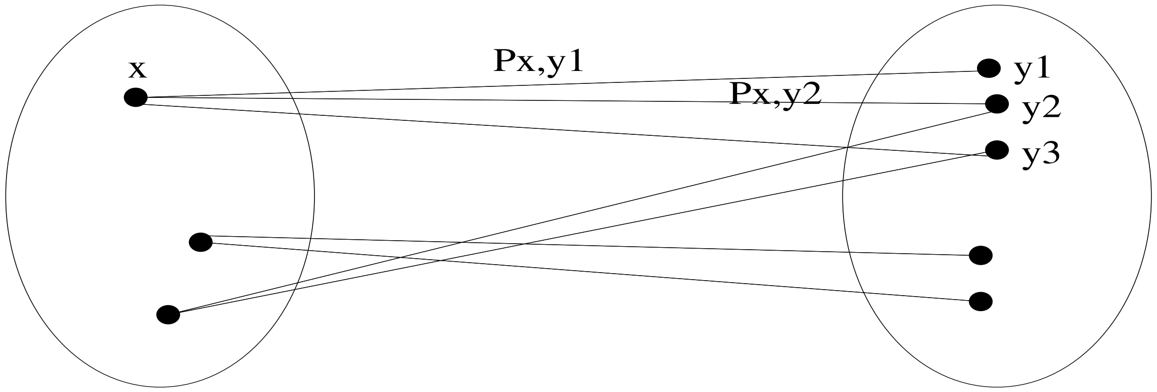

Figure 2: A bipartite version of the classical walk P.

Invoking the Lemma 3 above, we have $\| P_M \| \leq 1 - \frac{r}{4n^2}$. So $\mathrm{E}(T) \leq \frac{1}{1 - \| P_m \|} \leq \frac{4n^2}{r}$. This is an upper bound on the hitting time of the random walk.

The query complexity of the algorithm is calculated as follows. There are r initial queries for the value of f at each point in the starting r-subset. In each of the $\frac{n^2}{r}$ steps of the walk, there is a query for the value of f at the new point we swap into the subset. Thus, the query complexity is $O(r + \frac{n^2}{r})$.

In Ambainis' algorithm, r is chosen to be $n^{2/3}$. This gives us a classical hitting time of $O(n^{4/3})$, with query complexity $O(n^{4/3})$. This is much worse than the naive (and optimal) classical algorithm. However, the hitting time may be speeded up quadratically by a suitable quantization of the walk. This gives us a quantum query complexity of $O(r + \frac{n}{\sqrt{r}})$. This expression is minimized when $r = n^{2/3}$ and leads to Ambainis' bound of $O(n^{2/3})$.

Finally, we turn to the quantization of the classical walks that leads to the quadratic speedup.

3 The Szegedy quantization of classical walks

We will focus on a quantization of a Markov chain specified by a classical transition matrix P due to [Sze04]. For previous work on quantum walks and their algorithmic applications, we refer the reader to a survey article by Ambainis (listed in the bibliography web page for the course).

The quantum walk is defined over a bipartite state space, corresponding to a random walk on a bipartite graph (see Figure 2). Each side of the graph contains r-subsets as vertices. A pair of vertices in the left and the right hand partition of the graph are joined by an edge iff one subset can be changed into the other by removing one of the elements in the subset and adding one that is not in the set. The probability of moving from a subset x on the left side of the graph to a subset y in the right side of the graph is given by $p_{x,y}$.

For each vertex $x \in X$ on the left, we consider the superposition:

$$|\phi_x\rangle \quad = \quad |x\rangle \sum_y \sqrt{p_{x,y}} |y\rangle,$$

and for each vertex $y \in X$ on the right, the superposition

$$|\psi_y\rangle \quad = \quad \left(\sum_x \sqrt{p_{y,x}} |x\rangle \right) |y\rangle.$$

Note that $\{\phi_x\}_x$ and $\{\psi_y\}_y$ are both orthonormal sets. Let $E_1 = \mathrm{span}_x\{\phi_x\}$ and $E_2 = \mathrm{span}_y\{\psi_y\}$, and Π_1, Π_2 be orthonormal projections onto E_1, E_2 respectively. We define two operators R_1 and R_2 as $R_1 = 2\Pi_1 - I$

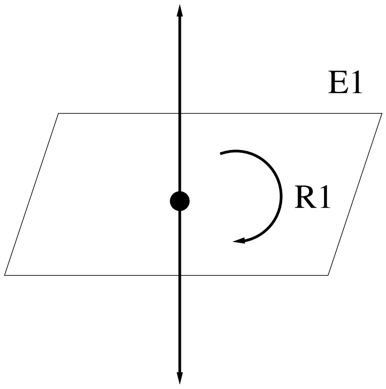

Figure 3: The action of the unitary operator R_1 is a reflection.

and $R_2 = 2\Pi_2 - I$. Then R_1 and R_2 are reflections about E_1 and E_2 respectively, and therefore unitary. For example, R_1 preserves the space E_1 respectively, and inverts any vector orthogonal to E_1 (see Figure 3).

The quantitization W_P of a Markov Chain P is defined as the product of the two reflections:

$$W_P = R_2 R_1.$$

Why is this a natural definition? The naïve way to define a step of the walk is:

$$|x\rangle|\bar{0}\rangle \to |x\rangle \sum_y \sqrt{p_{x,y}}|y\rangle.$$

The next step would similarly consist of diffusing over the neighbours of the vertex in the second register using a fresh set of ancilla qubits. However this would be merely a simulation of a classical walk. Instead of doing this, Szegedy's quantization keeps memory of only the previous step. To do so, it uses an operator that diffuses x into all its neighbors y's (and vice versa) in a manner analogous to the Grover diffusion operator. This idea also occurs in [Wat01]. The difference here is that the diffusion is guided by the transition probabilities in the classical walk.

In the next lecture, we will analyze the quantum walk W_P.

References

[Sze04] Mario Szegedy. Spectra of quantized walks and a $\sqrt{\delta\epsilon}$ rule. LANL Quantum Physics preprint quant-ph/0403120, January 2004.

[Wat01] John Watrous. Quantum simulations of classical random walks and undirected graph connectivity. *Journal of Computer and System Sciences*, 62(2):376–391, 2001.