Lecture 16

Lecturer: Ronitt Rubinfeld   *Scribe: Yinmeng Zhang*

1   Overview

1. Randomized Logspace and an algorithm for S,T-connectivity

2. Mixing Times

2   RL

RL is the class of problems which can be solved in randomized logspace. That is, there is a Turing Machine M with access to a read-only input tape of length n, a write-only output tape, and a work tape of size $O(\log n)$, which runs in polynomial time, such that

$$\text{If } x \in \text{L then } \Pr[M(x) \text{ accepts}] \geq \tfrac{2}{3}$$
$$\text{If } x \notin \text{L then } \Pr[M(x) \text{ accepts}] \leq \tfrac{1}{3}$$

Let's look at the problem USTCONN, or undirected S-T connectivity: given an undirected graph G, and two nodes S (source) and T (target), can we check if there is a path from S to T in randomized logspace?

Theorem 1 USTCONN $\in$ RL

Proof   The obvious algorithm turns out to work. Starting at node S, we take a random walk on G for $3 * 8n^3$ steps. If we ever see T, output "Yes", else no.

The only information that needs to be maintained is the name of the current node and how many steps we have taken so far, so the algorithm is logspace. Last time we saw that $8n^3$ is an upper bound for the cover time of a connected graph. This means it is also an upper bound on the expected time to see all nodes in the connected component of S, and in particular to see T. Thus if T is reachable from S, then let X be a random variable for the length of a random walk until we see T. By Markov's inequality,

$$\Pr[\text{accept}] = \Pr[X \geq 3 * 8n^3] \leq \Pr[X \geq 3 * E[X]] \leq 1/3.$$

While if T is not reachable from S, then clearly

$$\Pr[\text{accept}] = 0.$$

∎

One interesting implication of this theorem is the existence of Universal Traversal Sequence for graphs. Suppose there is a lexicographic order on the neighbors of any vertex in the graph; then we can specify a random walk by a sequence $\{a_i\}$, where at the i-th step we move to the a_i-th neighbor of the current vertex (modulo the degree of the current vertex). Now, if we pump up the probability of the above algorithm, we see that for a fixed connected graph on n nodes, overwhelmingly many random walks of some $O(n^3)$ length hit every vertex. Then, with an Adleman trick, there exist a particular (deterministic!) walk which works for every connected graph on n nodes. Such a walk is called a Universal Traversal Sequence. USTCONN is a great natural problem in RL $\subset$ RP. As a first step in our derandomization program, can we derandomize USTCONN? The answer is yes. So can we derandomize RL? The best result is RL $\subset$ L$^{3/2}$, due to Saks & Zhou. And what about RP? Well, let's finish looking at USTCONN first.

3   Linear Algebra Review

Definition 2 *A non-zero vector v is an* **eigenvector** *of a matrix A with corresponding* **eigenvalue** λ *if $vA = \lambda v$. (Notice that for us, vectors go on the right.)*

Definition 3 *The L_2 or* **Euclidean norm** *of a vector $v = (v_1, \cdots, v_n)$ is $|v|_2 = \sqrt{\sum_{i=1}^{n} v_i^2}$.*

We are used to seeing the L_1 norm, $|v|_1 = \sum_{i=1}^{n} |v_i|$, in the context of probability vectors, which must have L_1 norm equal to 1.

Fact 4 $L_2 \leq L_1 \leq \sqrt{n} L_2$

The first inequality follows by algebra ($\sum |v_i|^2 \leq (\sum |v_i|)^2)$), and the second by Cauchy-Schwartz $((\sum |v_i|^2)(\sum 1^2) \geq (\sum |v_i|)^2)$.

Definition 5 *The vectors $v^{(1)}, \cdots, v^{(n)}$ are* **orthonormal** *if*

$$v^{(i)} \cdot v^{(j)} = \sum_k v_k^{(i)} v_k^{(j)} = \begin{cases} 1 & \text{if } i = j \\ 0 & \text{otherwise} \end{cases}$$

Fact 6 *The norm is linear; in particular, if $v = \sum \alpha_i b_i$ for constants α_i and orthonormal basis vectors b_i, then $||v||_2 = \sqrt{\sum \alpha_i^2}$*

What does this have to do with connectivity? Let P be a transition matrix for a d-regular undirected graph. Then an eigenvector with eigenvalue 1, normalized so that its L_1 norm is 1, is by definition a stationary distribution for random walks on the graph. For example, if P is the transition for a d-regular undirected graph, then recall that its stationary distribution is the uniform distribution: it has eigenvector $v = (\frac{1}{n}, \cdots, \frac{1}{n})$. It will be convenient for us to normalize this to $v' = (\frac{1}{\sqrt{n}}, \cdots, \frac{1}{\sqrt{n}})$, which has L_2 norm 1.

Theorem 7 *Let P be a real symmetric matrix. Then it has an orthonormal eigenbasis $v^{(1)}, \cdots, v^{(n)}$. If P is the transition matrix for a connected undirected graph, then the corresponding eigenvalues are $1 = \lambda_1 > |\lambda_2| \geq \cdots \geq |\lambda_n|$.*

Fact 8 *Assume P has eigenvectors $v^{(1)}, \cdots, v^{(n)}$ with corresponding eigenvalues $\lambda_1, \cdots, \lambda_n$ in nonincreasing order.*

1. *αP has the same eigenvectors $v^{(1)}, \cdots, v^{(n)}$ with corresponding eigenvalues $\alpha\lambda_1, \cdots, \alpha\lambda_n$.*

2. *$P + I$ has the same eigenvectors $v^{(1)}, \cdots, v^{(n)}$ with corresponding eigenvalues $\lambda_1 + 1, \cdots, \lambda_n + 1$.*

3. *P^k has the same eigenvectors $v^{(1)}, \cdots, v^{(n)}$ with corresponding eigenvalues $\lambda_1^k, \cdots, \lambda_n^k$.*

4. *If P is stochastic, then for all i, $|\lambda_i| \leq 1$.*

5. *If P is a transition matrix of an undirected graph, then $w = \sum \alpha_i v^{(i)}$. The L_2 norm $|w| = \sum \alpha_i^2$.*

Our trick for making graphs aperiodic — choosing with $1/2$ probability to stay at the same node, and with $1/2$ probability to move to a random neighbor — corresponds to studying the matrix $(P + I)/2$. The first two facts imply that this matrix has the same eigenvectors, and related eigenvalues: $(\lambda_i + 1)/2$.

4 Mixing Times and Eigenstuff

Last time we saw that if a graph is aperiodic and connected, then there exists a unique stationary or limiting distribution. How quickly does a random walk approach the stationary distribution?

Definition 9 *The **mixing time** (parameterized by ϵ) of Markov chain A is*

$$T(\epsilon) = \min\{t : \|\pi - \pi^{(0)}A^t\|_1 < \epsilon\},$$

*where π is the stationary distribution and $\pi^{(0)}$ is the initial distribution. We say A is **rapidly mixing** if $T(\epsilon) = \mathrm{poly}(\log |V|, \log \frac{1}{\epsilon})$, where V is the set of possible states.*

Question: What is the relationship between mixing time and cover time?

It is tempting to say that the mixing time must be greater than the cover time, but in fact it is not necessary to be sure to have visited every node for the probabilities to have mixed. Consider the complete graph: regardless of the starting distribution, after one step the distribution will be uniform. Thus, we see that rapidly mixing graphs, which have the property that after visiting very few of the nodes of the graph, the distribution will be very close to uniform, are not ipso facto impossible.

Theorem 10 *Suppose P is the transition matrix for a d-regular undirected ergodic graph. It has a unique stationary distribution $\bar{\pi}$; let the initial distribution be π_0. Then*

$$\|\pi_0 P^t - \bar{\pi}\|_2 \leq |\lambda_2|^t$$

Proof By a previous theorem, P has an orthonormal basis of eigenvectors $v^{(1)}, \cdots, v^{(n)}$, so we can write $\pi_0 = \sum_{i=1}^n \alpha_i v_i$.

$$\pi_0 P^t = \sum \alpha_i \lambda_i^t v^{(i)} = \alpha_1 v^{(1)} + \sum_{i=2}^n \alpha_i \lambda_i^t v^{(i)}$$

where the summation is disappearing, since the λ_i are strictly less than 1.

$$
\begin{aligned}
\|\pi_0 P^t - \alpha_1 v^{(1)}\|_2 &= \left\|\sum_{i=2}^n \alpha_i \lambda_i^t v^{(i)}\right\|_2 \\
&= \sqrt{\sum_{i=2}^n \alpha_i^2 \lambda_i^{2t}} \qquad\qquad \text{(see Fact 6)} \\
&\leq |\lambda_2^t| \sqrt{\sum_{i=2}^n \alpha_i^2} \\
&\leq |\lambda_2^t| \cdot \|\pi_0\|_2 \qquad\qquad \text{(see Fact 6)} \\
&\leq |\lambda_2^t| \cdot \|\pi_0\|_1 \qquad\qquad \text{(see Fact 4)} \\
&\leq |\lambda_2^t| \qquad\qquad \text{(Because } \pi_0 \text{ is a probability vector)}
\end{aligned}
$$

∎

Next time we'll see how to use random walks on rapidly mixing graphs to save randomness.