Notes for Lecture 22

Notes on Reingold's Theorem, Part I

Today we begin a proof that the *undirected* (s,t)-connectivity problem can be solved in *deterministic* logarithmic space. We give a number of preliminary definitions and results. We will get to the actual algorithm next time.

1 Review of Linear Algebra

We think of vectors as row vectors, or $1 \times n$ matrices. If $\mathbf{v} = (v_1, \ldots, v_n) \in \mathbb{R}^n$ is a real vector, then its length is defined as $||\mathbf{v}|| = \sqrt{v_1^2 + \cdots v_n^2}$. If $\mathbf{v}_1, \ldots, \mathbf{v}_n \in \mathbb{R}^n$ are mutually orthogonal vectors of length one, then they are called an orthonormal basis of $\mathbb{R}^n$. Every vector $\mathbf{v}$ can then be written in a unique way as $\mathbf{v} = \alpha_1 \mathbf{v}_1 + \cdots + \alpha_n \mathbf{v}_n$, and we have $||\mathbf{v}|| = \sqrt{\alpha_1^2 + \cdots + \alpha_n^2}$.

Let $A \in \mathbb{R}^{n \times n}$ be a matrix. A (left) *eigenvalue* of A is a value $\lambda \in \mathbb{C}$ such that, for some non-zero vector $\mathbf{v} \in \mathbb{C}^n$,

$$\mathbf{v} \cdot A = \lambda \mathbf{v}$$

and, if so, the vector $\mathbf{v}$ is called an *eigenvector*. If λ is an *eigenvalue* of A, then we have

$$\mathbf{v} \cdot (A - \lambda I) = \mathbf{0}$$

where I is the identity matrix and $\mathbf{0}$ is the all-zero vector. This implies that the columns of the matrix $(A - \lambda I)$ are not linearly independent, and so

$$\det(A - \lambda I) = 0$$

where det denotes the determinant. The value $\det(A - \lambda I)$, as a function of λ, is a polynomial of degree n, and so A can have at most n eigenvalues. Note also that if $\det(A - \lambda I) = 0$, then the columns of A cannot all be linearly independent, and so there must be a nonzero vector $\mathbf{v}$ such that $\mathbf{v} \cdot (A - \lambda I) = 0$. It follows that λ is an eigenvalue of A if and only if it is a root of the polynomial $\det(A - \lambda I)$. If we count multiplicities, then there are precisely n eigenvalues for A

If $A \in \mathbb{R}^{n \times n}$ is a *symmetric* matrix, that is, a matrix such that $A_{i,j} = A_{j,i}$ for every i, j, then there are precisely n real eigenvalues of A, counting multiplicities. Furthermore, several additional results hold (all the results from now on assume that A is symmetric).

Let λ be an eigenvalue of A, and let $\mathbf{v}$ and $\mathbf{w}$ be eigenvectors of λ. Then if $a, b \in \mathbb{R}$ are reals, we have that

$$(a\mathbf{v} + b\mathbf{w}) \cdot A = a\mathbf{v}A + b\mathbf{w}A = a\lambda\mathbf{v} + b\lambda\mathbf{w} = \lambda(a\mathbf{v} + b\mathbf{w})$$

and so $(a\mathbf{v} + b\mathbf{w})$ is an eigenvector. This means that set of eigenvectors of λ (plus the all-zero vector) form a linear subspace of $\mathbb{R}^n$. We have the following theorems:

1. The dimension of the space of eigenvectors of λ is equal to the multiplicity of λ as a root of $\det(A - \lambda I)$;

2. If $\lambda \neq \lambda'$ are two eigenvalues, then the space of eigenvectors of λ is orthogonal to the space of eigenvectors of λ'.

This means that if $\lambda_1, \ldots, \lambda_n$ are the eigenvalues of A, with multiplicities, then we can find mutually orthogonal unit vectors $\mathbf{v}_1, \ldots, \mathbf{v}_n$ such that each $\mathbf{v}_i$ is an eigenvector of λ_i. Such eigenvectors form a basis for $\mathbb{R}^n$.

We finally come to an observation that will be important later. Consider the matrix $A^2 = A \cdot A$. If λ is an eigenvalue of A with eigenvector $\mathbf{v}$, then

$$\mathbf{v} \cdot A^2 = (\mathbf{v} \cdot A) \cdot A = \lambda \mathbf{v} \cdot A = \lambda^2 \mathbf{v}$$

so that λ^2 is an eigenvalue of A^2 with eigenvector $\mathbf{v}$.

If $\lambda_1, \ldots, \lambda_n$ are eigenvalues of A and $\mathbf{v}_1, \ldots, \mathbf{v}_n$ are the respective eigenvectors, then $\lambda_1^t, \ldots, \lambda_n^t$ are eigenvalues of A^t and $\mathbf{v}_1, \ldots, \mathbf{v}_n$ are the respective eigenvectors.

If A is a symmetric matrix, $\lambda_1, \ldots, \lambda_n$ are its eigenvalues, and $\mathbf{v}_1, \ldots, \mathbf{v}_n$ are their respective eigenvectors, then for every vector $\mathbf{v}$ we can write in a unique way $\mathbf{v} = \alpha_1 \mathbf{v}_1 + \ldots + \alpha_n \mathbf{v}_n$. Now, note that

$$\mathbf{v} \cdot A^t = (\alpha_1 \mathbf{v}_1 + \cdots + \alpha_n \mathbf{v}_n) = \alpha_1 \lambda_1^t \mathbf{v}_1 + \ldots + \alpha_n \lambda_n^t \mathbf{v}_n$$

that is, we can compute $\mathbf{v} \cdot A^t$ very easily without having to actually compute the matrix A^t.

2 Graphs and Eigenvalues

Let $G = (V, E)$ be an undirected regular graph of degree d. We allow G to have multiple edges and self-loops. We identify V with $\{1, \ldots, n\}$.

We define the "random walk" matrix (normally it is called the *normalized* adjacency matrix) of G to be the matrix $A \in \mathbb{R}^{V \times V}$ obtained by dividing each entry of the adjacency matrix of G by d. Another way to look at it is as follows: let i and j be vertices of G, and consider the probabilistic process of picking at random one of the edges that are incident on i; then $A_{i,j}$ is the probability that j is the other endpoint of the selected edge.

Consider now the matrix A^2. It should be clear that $A^2_{i,j}$ is the probability that, starting from i, a two-step *random walk* will land on j. In general, $A^t_{i,j}$ is the probability of going from i to j in a t-step random walk.

Let $\mathbf{p} = (p_1, \ldots, p_n) \in \mathbb{R}^V$ be a probability distribution, that is $p_i \geq 0$ for all i, and $\sum_i p_i = 1$. Then the vector $\mathbf{p} \cdot A$ is also a probability distribution, and it is the distribution that we get in the following way: pick a vertex i at random according to probability $\mathbf{p}$, then move to a random neighbor. Similarly, $\mathbf{p}A^t$ is the probability distribution that we get by picking a start vertex according to distribution $\mathbf{p}$ and then performing t steps of a random walk.

Since, as we saw above, it is easy to compute a product of the form $\mathbf{p} \cdot A^t$ given the eigenvalues and eigenvectors of A, it is no wonder that eigenvalues and eigenvectors play an important role in the study of random walks.

First, note that A is a symmetric matrix (because we assumed that G was an undirected graph), and so all its eigenvalues are real numbers.

It is easy to see that 1 is an eigenvalue of A, with eigenvector the uniform distribution $\mathbf{u} := (1/n, \ldots, 1/n)$. (Recall that we assumed that G was a regular graph, otherwise this would not be true.) It is also a theorem that for every other eigenvalue λ we have $|\lambda| \leq 1$.

If G is not connected, then the multiplicity of 1 as an eigenvalue is more than 1, that is, there are at least two linearly indpendent eigenvectors of 1. (For example, think of the distribution that is uniform in one connected component, and zero elsewhere.) If G is connected, then the eigenvalue 1 has multiplicity 1. It is however possible that -1 could be an eigenvalue, and, for example, this happens if the graph is bipartite: if $(S, V - S)$ a bipartition of the vertices such that all the edges go between S and $V - S$, then we can construct an eigenvector of -1 by defining a vector $\mathbf{v}$ such that $v_i = 1$ if $i \in S$ and $v_i = -1$ otherwise. Fortunately, it is a theorem that these are the only cases.

Theorem 1 *Let G be an undirected, regular, connected, non-bipartite, graph of degree d with n vertices, and let A be the random walk matrix of G. Then if $\lambda_1 \geq \lambda_2 \geq \lambda_n$ are the eigenvalues of A, with multiplicities, in sorted order, we have $\lambda_1 = 1$ and $|\lambda_i| \leq 1 - 1/dn^2$ for $i = 2, \ldots, n$.*

Finally, we are ready to study random walks in G, where G is, as in the theorem above, an undirected, regular, connected, non-bipartite, graph of degree d with n vertices.

Let A be the random walk matrix of G, let $1 > \lambda_2 \geq \cdots \geq \lambda_n$ be the eigenvalues of A and $\mathbf{v}_1, \ldots, \mathbf{v}_n$ the respective eigenvectors, chosen to be orthogonal and of length 1. Note that $\mathbf{v}_1$ is a multiple of the uniform distribution. Let $\lambda(G) = \max_{j=2}^{n} |\lambda_j|$ be the second largest eigenvalue in absolute value. Recall that $\lambda < 1 - 1/dn^2$.

Let $\mathbf{p} \in \mathbb{R}^n$ be a distribution. We can write $\mathbf{p} = \mathbf{u} + (\mathbf{p} - \mathbf{u})$, where $\mathbf{u} := (1/n, \cdots, 1/n)$ is the uniform distribution and where $(\mathbf{p} - \mathbf{u})$ is orthogonal to the uniform distribution. (It is easy to verify this last claim.) The vectors $\mathbf{v}_2, \cdots, \mathbf{v}_n$ are a basis for the set of vectors orthogonal to $\mathbf{u}$, and so we can then write, for some coefficients $\alpha_2, \cdots, \alpha_n$,

$$\mathbf{p} = \mathbf{u} + \alpha_2 \mathbf{v}_2 + \cdots \alpha_n \mathbf{v}_n \ .$$

Consider now $\mathbf{p} \cdot A$, that is the distribution obtained by picking a vertex according to $\mathbf{p}$ and then doing a one-step random walk. We have

$$\mathbf{p} \cdot A = \mathbf{u} \cdot A + \sum_{j=2}^{n} \alpha_j \mathbf{v}_j A = \mathbf{u} + \sum_{j=2}^{n} \alpha_j \lambda_j \mathbf{v}_j \ .$$

in other words, the distribution that we obtain is uniform, up to an error term, and the

error term is $\sum_{j=2}^{n} \alpha_j \lambda_j \mathbf{v}_j$. As a vector, how *long* is this error term? We can compute

$$\begin{aligned} ||\mathbf{p} \cdot A - \mathbf{u}|| &= ||\sum_{j=2}^{n} \alpha_j \lambda_j \mathbf{v}_j|| \\ &= \sqrt{\sum_{j=2}^{n} \alpha_j^2 \lambda_j^2} \\ &\leq \max_{j=2}^{n} |\lambda_j| \cdot \sqrt{\sum_{j=2}^{n} \alpha_j^2} \\ &= \lambda \cdot ||\mathbf{p} - \mathbf{u}|| \end{aligned}$$

That is, the new distribution is *closer by a factor of* λ to the uniform distribution. After t steps we clearly have $||\mathbf{p}A^t - \mathbf{u}|| \leq \lambda^t ||\mathbf{p} - \mathbf{u}||$.

For a vertex $i \in V$, define $\mathbf{p}_i = 1$ and $\mathbf{p}_j = 0$ for $j \neq 1$. This corresponds to the distribution that gives probability one to vertex i. As before, we have $||\mathbf{p}A^t - \mathbf{u}|| \leq \lambda^t ||\mathbf{p} - \mathbf{u}||$ and it can be seen that $||\mathbf{p} - \mathbf{u}|| = \sqrt{1 - 1/n} < 1$. So we get

$$||\mathbf{p}A^t - \mathbf{u}|| \leq \lambda^t$$

Pick a value of t such that $\lambda^t \leq 1/2n$, for example $t = O(\frac{1}{1-\lambda} \log n)$ is enough. Now we have $||\mathbf{p}A^t - \mathbf{u}|| < 1/2n$, from which it follows that every entry of $\mathbf{p}A^t$ must be at least $1/2n$. This also implies that every vertex of G is reachable from i in $O(\frac{1}{1-\lambda} \log n)$ steps, and since i was arbitrary we have actually proved that the diameter of G is at most $O(\frac{1}{1-\lambda} \log n)$.

For general graphs, we only know $\lambda \leq 1 - 1/dn^2$, so the above proof only shows that the diameter is at most $O(dn^2 \log n)$, which is not very interesting. The argument does, however, show that if λ is a constant, for example .9, then the diameter is logarithmic.

An *expander* is a graph such that $\lambda(G)$ is a constant bounded away from 1.[1] Expanders have a lot of applications in computer science in general and in complexity theory in particular. For this lecture, the only property we use is that the diameter is logarithmic.

It is known that there is a constant d such that for every n there is a d-regular expander with n vertices and $\lambda(G) \leq 1/2$. In fact, it is known how to get $\lambda(G) \leq \sqrt{2d-1}/d$ for certain values of d.

[1] This, of course, does not make sense as a definition because for fixed graph λ is always a "constant bounded away from one." In the formal definition, we think of an infinite family of graphs $\{G_n\}$, where G_n has n vertices, such that there is a fixed constant λ such that $\lambda(G_n) \leq \lambda$ for every n.