18.409 Applied Extremal Combinatorics	September 15, 1998

Lecture 2

Lecturer: Daniel A. Spielman | *Scribe: Leonid Reyzin*

First, we recall some basic concepts, related to graphs. For now, we will consider only undirected graphs.

Definition 1 *A graph G is a pair of sets (V, E), where E is a binary symmetric non-reflexive relation on V (i.e., $E \subset V \times V$ s.t. $\forall i, j \in V$, $(i,i) \notin E$ and if $(i,j) \in E$ then $(j,i) \in E$. The elements of V are known as nodes or vertices; the elements of E are known as edges. Two nodes i, j are called neighbors if the edge $(i,j) \in E$. The degree of a node is the number of its neighbors. A graph is called d-regular if every node has degree d. A graph is called connected if the transitive closure of E is $V \times V$.*

For each graph $G = (V, E)$ on n nodes (i.e., $|V| = n$), we can define an $n \times n$ adjacency matrix A as follows: $A_{i,j} = 1$ if $(i,j) \in E$ and $A_{i,j} = 0$ otherwise. Note that A is a real symmetric matrix. Therefore, there exists an orthonormal matrix X and a diagonal matrix D such that $X^T AX = D$. The eigenvalues of A are the diagonal elements of D, and the eigenvectors come from the corresponding rows of X and form an orthonormal basis. The collection of the eigenvalues of A is known as the *spectrum* of G (or of A).

The coordinates x_i of a vector $\overline{x} \in R^n$ naturally correspond to the nodes i of the graph. Consider now what happens when the adjacency matrix A is applied to a vector $\overline{x}$:

$$A(\overline{x})_i = \sum_{(j,i) \in E} x_j.$$

That is, the new value at each node is the sum of the old values at its neighbors in the graph. Therefore, if $\overline{x}$ is an eigenvector with eigenvalue λ, then

$$\sum_{(j,i) \in E} x_j = \lambda x_i.$$

Theorem 2 *Let G have the spectrum $\lambda_1 \geq \lambda_2 \geq \ldots \geq \lambda_n$. Then*

1. *$\lambda_1 \leq d_{max}$, where d_{max} is the maximal degree of a node in G*
2. *If G is connected, then $\lambda_1 = d_{max}$ if and only if G is d_{max}-regular.*

Proof

1. Let $\overline{x}$ be an eigenvector of λ_1. Assume without loss of generality that $x_1 > 0$ and $x_1 \geq x_i \forall 1 < i \leq n$. Note that because the degree of node 1 is at most d_{max},

$$d_{max} x_1 \geq \sum_{(j,1) \in E} x_j = A(\overline{x})_1 = \lambda_1 x_1.$$

   Therefore, $d_{max} \geq \lambda_1$.

2. If $\lambda_1 = d_{max}$, then $d_{max} x_1 = \sum_{(j,1) \in E} x_j$. Note that there are at most d_{max} elements in the sum, and that each of them is at most x_1, and that $x_1 > 0$. Therefore, each of them has to be equal to x_1 and there have to be exactly d_{max} of them. Thus, the degree of x_1 is d_{max}, and $x_j = x_1$ for all nodes j which are neighbors of the node 1. By induction, we can extend this

argument to all the neighbors of the node 1, and to all their neighbors, and so on. Since the graph is connected, the degree of every node is d_{max}.

The other direction follows from the fact that $\lambda_1 \leq d_{max}$ and that $(1, 1, \ldots, 1)$ is an eigenvector with the eigenvalue d_{max}. Therefore, $\lambda_1 = d_{max}$.

■

The following theorem is left as an exercise for the reader:

Theorem 3 *If G is connected and regular, then the following are equivalent:*

- $\lambda_1 = -\lambda_n$
- *G is bipartite*
- $\lambda_i = -\lambda_{n-i+1}$ *for all* $1 \leq i \leq n$

The following theorem (known as Perron's theorem) is stated here for the case of graphs, but has a more general form.

Theorem 4 *Let G be a connected graph with adjacency matrix A. Then there exists a vector $\overline{x} > 0$ (i.e., each coordinate is positive) and $\lambda > 0$ such that*

- $\lambda\overline{x} = A\overline{x}$
- $\lambda \geq |\mu|$ *for any eigenvalue μ of A*
- *λ has multiplicity 1*
- *if $\overline{y} > 0$ is an eigenvector of A, then $\overline{y}$ is a multiple of $\overline{x}$*

Proof In order to prove Perron's theorem, we will use the following fact from analysis without proof. It is known as "Brouwer's Fixed Point Theorem."

Theorem 5 *Let S be a closed convex bounded set in R^n and let $f : S \to S$ be a continuous function. Then $\exists x \in S$ such that $f(x) = x$.*

To apply Brouwer's Fixed Point Theorem, let

$$S = \{\overline{x} | x_i \geq 0 \; \forall 1 \leq i \leq n \text{ and } \sum_{i=1}^{n} x_i = 1\}$$

and let

$$f(\overline{x}) = \frac{A\overline{x}}{\sum_{i=1}^{n} (A\overline{x})_i}$$

(the denominator of f is there in order to normalize the result, to make sure it stays in S; the denominator is greater than 0 because both the matrix and the vector are nonnegative and at least one $x_i > 0$ and every column of the matrix contains a non-zero element if the graph is connected). Note that S is closed, convex and bounded, and $f : S \to S$ is continuous. Therefore, we can apply the fixed point theorem to get $\overline{x}$ such that

$$\frac{A\overline{x}}{\sum_{i=1}^{n} (A\overline{x})_i} = \overline{x}$$

or, equivalently,

$$A\overline{x} = \left(\sum_{i=1}^{n} (A\overline{x})_i \right) \overline{x}.$$

Let $\lambda = \sum_{i=1}^{n} (A\overline{x})_i$. As discussed above, $\lambda > 0$. We now need to show that $\overline{x} > 0$ (we already know that $\overline{x} \geq 0$ because $\overline{x} \in S$). Indeed, if $x_i = 0$ for some i, then $x_j = 0$ for all such nodes j that are neighbors of i and j (because $0 = \lambda x_i = (A\overline{x})_i = \sum_{(i,j)\in E} x_j$ and each $x_j \geq 0$). By induction, because the graph is connected, $\overline{x} = 0$, which is impossible because $\sum_{i=1}^{n} x_i = 1$.

We now need to show that $\lambda \geq |\mu|$ for any eigenvalue μ of A. Indeed, let $\overline{y} = (y_1, y_2, \ldots, y_n)$ be a corresponding to μ eigenvector. Let $\overline{y}^+ = (|y_1|, |y_2|, \ldots, |y_n|)$. Then

$$(A\overline{y}^+)_i = \sum_{j=1}^{n} a_{ij}|y_i| \geq \left|\sum_{j=1}^{n} a_{ij}y_i\right| = |\mu y_i|,$$

so $A\overline{y}^+$ is componentwise greater than or equal to $|\mu|\overline{y}^+$. Therefore, $\overline{x}^T A\overline{y}^+ \geq x^T|\mu|\overline{y}^+$ (because $\overline{x}^T > 0$), so $\lambda\overline{x}^T\overline{y}^+ \geq |\mu|\overline{x}^T\overline{y}^+$. So $\lambda \geq |\mu|$, because $\overline{x}^T\overline{y}^+ > 0$ because $\overline{y}^+ \geq 0$ and $\overline{x}^T > 0$.

In order to show that λ has multiplicity 1, let $\overline{y} \neq 0$ be such that $A\overline{y} = \lambda\overline{y}$. Then there exists an α such that $\overline{z} = \overline{x} - \alpha\overline{y}$ is component-wise greater than or equal to 0, but has a 0 component (because $\overline{x}$ has all positive components, we can always find a small enough α). Then $\overline{z}$ is a nonnegative eigenvector with eigenvalue λ, but it has a 0 component. By the same proof as that for $\overline{x} > 0$ above, we can show that $\overline{z}$ has to be 0. Therefore, $\overline{x}$ and $\overline{y}$ are linearly dependent. Thus, λ has multiplicity 1.

All that remains to show is that every eigenvector $\overline{y}$ corresponding to an eigenvalue $\mu \neq \lambda$ has both positive and negative components. Indeed, suppose $\overline{y} \geq 0$. Then

$$\overline{x}^T A\overline{y} = \lambda\overline{x}^T\overline{y} = \mu\overline{x}^T\overline{y}.$$

Because $\overline{x} > 0$ and $\overline{y} \geq 0$ and $\overline{y} \neq 0$, we have that $\overline{x}^T\overline{y} \neq 0$, so $\lambda = \mu$. ■

We will now prove the following theorem, known as "Courant-Fischer Theorem" (or "inequality"), which will be a very useful tool later.

Theorem 6 *Let A be a real symmetric $n \times n$ matrix with the spectrum $\lambda_1 \geq \lambda_2 \ldots \geq \lambda_n$. Then*

$$\lambda_i = \max_{U} \min_{\overline{x}\in U, \overline{x}\neq 0} \frac{\overline{x}^T A\overline{x}}{\overline{x}^T\overline{x}},$$

where the maximum is taken over all i-dimensional subspaces $U \subseteq R^n$. (The quotient $\frac{\overline{x}^T A\overline{x}}{\overline{x}^T\overline{x}}$ is known as "Rayleigh quotient.")

Proof Let $\overline{v_1}, \overline{v_2}, \ldots, \overline{v_n}$ be the eigenvectors corresponding to the eigenvalues $\lambda_1, \lambda_2, \ldots, \lambda_n$, respectively. Let U be the space generated by $\overline{v_1}, \overline{v_2}, \ldots, \overline{v_i}$. The

$$\min_{\overline{x}\in U, \overline{x}\neq 0} \frac{\overline{x}^T A\overline{x}}{\overline{x}^T\overline{x}}$$

occurs at $\overline{x} = \overline{v_i}$. (To see that, write $\overline{x} = \sum_{j=1}^{i} \alpha_j\overline{v_j}$; then

$$\frac{\overline{x}^T A\overline{x}}{\overline{x}^T\overline{x}} = \frac{\sum_{j=1}^{i} {\alpha_j}^2\lambda_j}{\sum_{j=1}^{i} {\alpha_j}^2};$$

the minimum of that is clearly if $\alpha_j = 0$ for $j < i$ and $\alpha_i \neq 0$.) Therefore,

$$\max_{U} \min_{\overline{x}\in U, \overline{x}\neq 0} \frac{\overline{x}^T A\overline{x}}{\overline{x}^T\overline{x}} \geq \lambda_i.$$

To get the other direction, let U be some i-dimentional subspace of R^n. We need to show that there exists an $\overline{x} \neq 0 \in U$ such that $\frac{\overline{x}^T A \overline{x}}{\overline{x}^T \overline{x}} \leq \lambda_i$. Let V be the $(n-i+1)$-dimensional subspace generated by $\overline{v_i}, \overline{v_{i+1}}, \ldots, \overline{v_n}$. $U \cap V \neq \{0\}$ because $\dim(V) + \dim(U) > n$. So, let $0 \neq \overline{x} \in U \cap V$ and let $\overline{x} = \sum_{j=i}^{n} \alpha_j \overline{v_j}$. Then

$$\frac{\overline{x}^T A \overline{x}}{\overline{x}^T \overline{x}} = \frac{\sum_{j=i}^{n} {\alpha_j}^2 \lambda_j}{\sum_{j=i}^{n} {\alpha_j}^2} \leq \lambda_i.$$

■

By the same proof we can get another variant of the Courant-Fischer Theorem:

$$\lambda_i = \min_{U} \max_{\overline{x} \in U, \overline{x} \neq 0} \frac{\overline{x}^T A \overline{x}}{\overline{x}^T \overline{x}},$$

where the maximum is taken over all $n-i+1$-dimensional subspaces $U \subseteq R^n$.

We can now use the Courant-Fischer Theorem to prove the Interlacing Theorem.

Theorem 7 *Let G be a graph on n nodes with the adjacency matrix A and spectrum $\lambda_1 \geq \lambda_2 \geq \ldots \geq \lambda_n$. Let H be the same graph with one vertex (and all the edges containing it) removed. If H has the adjacency matrix B and spectrum $\mu_1 \geq \mu_2 \geq \ldots \geq \mu_{n-1}$, then $\lambda_i \geq \mu_i \geq \lambda_{i+1}$ for all i such that $1 \leq i < n$.*

Proof Assume, without loss of generality, that the first node of G was removed to get H. Then B is simply A without the first row and column.

$$\mu_i = \max_{U \subseteq R^{n-1}, \dim(U)=i} \min_{\overline{x} \in U, \overline{x} \neq 0} \frac{\overline{x}^T B \overline{x}}{\overline{x}^T \overline{x}} = \max_{U \subseteq R^n, \dim(U)=i, x_1=0 \forall \overline{x} \in U} \min_{\overline{x} \in U,\ \overline{x} \neq 0} \frac{\overline{x}^T A \overline{x}}{\overline{x}^T \overline{x}} \leq$$

$$\max_{U \subseteq R^n, \dim(U)=i} \min_{\overline{x} \in U, \overline{x} \neq 0} \frac{\overline{x}^T A \overline{x}}{\overline{x}^T \overline{x}} = \lambda_i.$$

(The $\leq$ above is because we simply removed a restriction on the max).

The other half of the interlacing theorem follows indentically from the other variant of the Courant-Fischer Theorem. ■