Lecture 2

In which we review linear algebra and introduce spectral graph theory.

1 Eigenvalues and Eigenvectors

Spectral graph theory studies how the eigenvalues of the adjacency matrix of a graph, which are purely algebraic quantities, relate to combinatorial properties of the graph.

We begin with a brief review of linear algebra.

If $x = a + ib$ is a complex number, then we let $x^* = a - ib$ denote its *conjugate*.

If $M \in \mathbb{C}^{n\times n}$ is a square matrix, $\lambda \in \mathbb{C}$ is a scalar, $\mathbf{v} \in \mathbb{C}^n - \{\mathbf{0}\}$ is a non-zero vector and we have

$$M\mathbf{v} = \lambda \mathbf{v} \qquad (1)$$

then we say that λ is an *eigenvalue* of M and that $\mathbf{v}$ is *eigenvector* of M corresponding to the eigenvalue λ.

When (1) is satisfied, then we equivalently have

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

for a non-zero vector $\mathbf{v}$, which is equivalent to

$$\det(M - \lambda I) = 0 \qquad (2)$$

For a fixed matrix M, the function $\lambda \to \det(M - \lambda I)$ is a univariate polynomial of degree n in λ and so, over the complex numbers, the equation (2) has exactly n solutions, counting multiplicities.

If $G = (V, E)$ is a graph, then we will be interested in the adjacency matrix A of G, that is the matrix such that $A_{ij} = 1$ if $(i, j) \in E$ and $A_{ij} = 0$ otherwise. If G is a multigraph or a weighted graph, then A_{ij} is equal to the number of edges between (i, j), or the weight of the edge (i, j), respectively.

The adjacency matrix of an undirected graph is symmetric, and this implies that its eigenvalues are all real.

Definition 1 *A matrix $M \in \mathbb{C}^{n \times n}$ is Hermitian if $M_{ij} = M^*_{ji}$ for every i, j.*

Note that a real symmetric matrix is always Hermitian.

Lemma 2 *If M is Hermitian, then all the eigenvalues of M are real.*

PROOF: Let M be an Hermitian matrix and let λ be a scalar and $\mathbf{x}$ be a non-zero vector such that $M\mathbf{x} = \lambda\mathbf{x}$. We will show that $\lambda = \lambda^*$, which implies that λ is a real number. We define the following *inner product* operation over vectors in $\mathbb{C}^n$:

$$\langle \mathbf{v}, \mathbf{w} \rangle := \sum_i v_i^* \cdot w_i$$

Notice that, by definition, we have $\langle \mathbf{v}, \mathbf{w} \rangle = (\langle \mathbf{w}, \mathbf{v} \rangle)^*$ and $\langle \mathbf{v}, \mathbf{v} \rangle = ||\mathbf{v}||^2$. The lemma follows by observing that

$$\begin{aligned} &\langle M\mathbf{x}, \mathbf{x} \rangle \\ &= \sum_i \sum_j M^*_{ij} x^*_j x_i \\ &= \sum_i \sum_j M_{ji} x_i x^*_j \\ &= \langle \mathbf{x}, M\mathbf{x} \rangle \end{aligned}$$

where we use the fact that M is Hermitian, and that

$$\langle M\mathbf{x}, \mathbf{x} \rangle = \langle \lambda\mathbf{x}, \mathbf{x} \rangle = \lambda^* ||x||^2$$

and

$$\langle \mathbf{x}, M\mathbf{x} \rangle = \langle \mathbf{x}, \lambda\mathbf{x} \rangle = \lambda ||x||^2$$

so that $\lambda = \lambda^*$. $\square$

From the discussion so far, we have that if A is the adjacency matrix of an undirected graph then it has n real eigenvalues, counting multiplicities of the number of solutions to $\det(A - \lambda I) = 0$.

If G is a d-regular graph, then instead of working with the adjacency matrix of G it is somewhat more convenient to work with the normalized matrix $M := \frac{1}{d} \cdot A$.

In the rest of this section we shall prove the following relations between the eigenvalues of M and certain purely combinatorial properties of G.

Theorem 3 *Let G be a d-regular undirected graph, and $M = \frac{1}{d} \cdot A$ be its normalized adjacency matrix. Let $\lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_n$ be the real eigenvalues of M with multiplicities. Then*

1. $\lambda_1 = 1$ *and* $\lambda_n \geq -1$.
2. $\lambda_2 = 1$ *if and only if G is disconnected.*
3. $\lambda_n = -1$ *if and only if at least one of the connected components of G is bipartite.*

In the next lecture we will begin to explore an "approximate" version of the second claim, and to show that λ_2 is close to 1 if and only if G has a sparse cut.

1.1 More on Eigenvalues and Eigenvectors

In order to relate the eigenvalues of the adjacency matrix of a graph to combinatorial properties of the graph, we need to first express the eigenvalues and eigenvectors as solutions to *optimization problems*, rather than solutions to algebraic equations.

First, we observe that if M is a real symmetric matrix and λ is a real eigenvalue of M, then λ admits a real eigenvector. This is because if $M\mathbf{x} = \lambda\mathbf{x}$ for some $\mathbf{x} \in \mathbb{C}^n$, then we also have $M\mathbf{x}' = \lambda\mathbf{x}'$, where $\mathbf{x}' \in \mathbb{R}^n$ is the vector whose i-th coordinate is the real part of the i-th coordinate of $\mathbf{x}$. Now, if λ is a (real) eigenvalue of a symmetric real matrix M, then the set $\{\mathbf{x} \in \mathbb{R}^n : M\mathbf{x} = \lambda\mathbf{x}\}$ is a vector subspace of $\mathbb{R}^n$, called the *eigenspace* of λ.

Fact 4 *If $\lambda \neq \lambda'$ are two distinct eigenvalues of a symmetric real matrix M, then the eigenspaces of λ and λ' are orthogonal.*

Proof: Let $\mathbf{x}$ be an eigenvector of λ and $\mathbf{y}$ be an eigenvector of λ'. From the symmetry of M and the fact that M, $\mathbf{x}$ and $\mathbf{y}$ all have real entries we get

$$\langle M\mathbf{x}, \mathbf{y} \rangle = \langle \mathbf{x}, M\mathbf{y} \rangle$$

but

$$\langle M\mathbf{x}, \mathbf{y} \rangle = \lambda \cdot \langle \mathbf{x}, \mathbf{y} \rangle$$

and

$$\langle \mathbf{x}, M\mathbf{y} \rangle = \lambda' \cdot \langle \mathbf{x}, \mathbf{y} \rangle$$

so that

$$(\lambda - \lambda') \cdot \langle \mathbf{x}, \mathbf{y} \rangle = 0$$

which implies that $\langle \mathbf{x}, \mathbf{y} \rangle = 0$, that is, that $\mathbf{x}$ and $\mathbf{y}$ are orthogonal. $\Box$

Definition 5 *The algebraic multiplicity of an eigenvalue λ of a matrix M is the multiplicity of λ as a root of the polynomial* $\det(M - \lambda I)$*. The geometric multiplicity of λ is the dimension of its eigenspace.*

The following is the only result of this section that we state without proof.

Fact 6 *If M is a symmetric real matrix and λ is an eigenvalue of M, then the geometric multiplicity and the algebraic multiplicity of λ are the same.*

This gives us the following "normal form" for the eigenvectors of a symmetric real matrix.

Fact 7 *If $M \in \mathbb{R}^{n \times n}$ is a symmetric real matrix, and $\lambda_1, \ldots, \lambda_n$ are its eigenvalues with multiplicities, and $\mathbf{v}_1$ is a length-1 eigenvector of λ_1, then there are vectors $\mathbf{v}_2, \ldots, \mathbf{v}_n$ such that $\mathbf{v}_i$ is an eigenvector of λ_i and $\mathbf{v}_1, \ldots, \mathbf{v}_n$ are orthonormal.*

PROOF: For each eigenvalue, choose an orthonormal basis for its eigenspace. For λ_1, choose the basis so that it includes $\mathbf{v}_1$. $\Box$

Finally, we get to our goal of seeing eigenvalue and eigenvectors as solutions to continuous optimization problems.

Lemma 8 *If M is a symmetric matrix and λ_1 is its largest eigenvalue, then*

$$\lambda_1 = \sup_{\mathbf{x} \in \mathbb{R}^n : ||\mathbf{x}|| = 1} \mathbf{x}^T M \mathbf{x}$$

Furthermore, the sup *is achieved, and the vectors achieving it are precisely the eigenvectors of λ_1.*

PROOF: That the sup is achieved follows from the fact that the set $\{\mathbf{x} \in \mathbb{R}^n : ||\mathbf{x}|| = 1\}$ is compact and that $\mathbf{x} \to \mathbf{x}^T M \mathbf{x}$ is a continuous function.

If $\mathbf{v}_1$ is a length-1 eigenvector of λ_1, then

$$\sup_{\mathbf{x}\in\mathbb{R}^n:||\mathbf{x}||=1} \mathbf{x}^T M\mathbf{x} \geq \mathbf{v}_1^T M\mathbf{v}_1 = \lambda_1$$

If $\mathbf{y}$ is a length-1 vector that achieves the sup, then let $\mathbf{v}_1, \ldots, \mathbf{v}_n$ be as in Fact 7 and write

$$\mathbf{y} = \alpha_1\mathbf{v}_1 + \ldots \alpha_n\mathbf{v}_n$$

Then

$$\sup_{\mathbf{x}\in\mathbb{R}^n:||\mathbf{x}||=1} \mathbf{x}^T M\mathbf{x} = \mathbf{y}^T M\mathbf{y} = \sum_i \alpha_i^2\lambda_i$$

Since $\sum_i \alpha_i^2 = ||\mathbf{y}||^2 = 1$, we have

$$\sup_{\mathbf{x}\in\mathbb{R}^n:||\mathbf{x}||=1} \mathbf{x}^T M\mathbf{x} = \sum_i \alpha_i^2\lambda_i \leq \lambda_1 \cdot \sum_i \alpha_i^2 = \lambda_1$$

Finally, we see that we have $\mathbf{y}^T M\mathbf{y} = \lambda_1$ precisely when, for every i such that $\alpha_i \neq 0$ we have $\lambda_i = \lambda_1$, that is, precisely when $\mathbf{y}$ is in the eigenspace of λ_1. □

Similarly we can prove

Lemma 9 *If M is a symmetric matrix, λ_1 is its largest eigenvalue, and $\mathbf{v}_1$ is an eigenvector of λ_1, then*

$$\lambda_2 = \sup_{\mathbf{x}\in\mathbb{R}^n:||\mathbf{x}||=1,\mathbf{x}\perp\mathbf{v}_1} \mathbf{x}^T M\mathbf{x}$$

Furthermore, the sup *is achieved, and the vectors achieving it are precisely the eigenvectors of λ_2. (If $\lambda_1 = \lambda_2$, then the vectors achieving the* sup *are the eigenvalues of $\lambda_1 = \lambda_2$ which are orthogonal to $\mathbf{v}_1$.)*

And

Lemma 10 *If M is a symmetric matrix and λ_n is its largest eigenvalue, then*

$$\lambda_n = \inf_{\mathbf{x}\in\mathbb{R}^n:||\mathbf{x}||=1} \mathbf{x}^T M\mathbf{x}$$

Furthermore, the inf *is achieved, and the vectors achieving it are precisely the eigenvectors of λ_n.*

1.2 Proof of Theorem 3

We will make repeated use of the following identity, whose proof is immediate: if M is the normalized adjacency matrix of a regular graph, and $\mathbf{x}$ is any vector, then

$$\sum_{i,j} M_{i,j}(x_i - x_j)^2 = 2\mathbf{x}^T\mathbf{x} - 2\mathbf{x}^T M\mathbf{x} \tag{3}$$

That is,

$$\mathbf{x}^T M\mathbf{x} = \mathbf{x}^T\mathbf{x} - \frac{1}{2}\sum_{i,j} M_{i,j}(x_i - x_j)^2 \le \mathbf{x}^T\mathbf{x}$$

And so

$$\lambda_1 = \max_{\mathbf{x}\in\mathbb{R}^n:||\mathbf{x}||=1} \mathbf{x}^T M\mathbf{x} \le 1$$

If we take $\mathbf{1} = (1,\ldots,1)$ to be the all-one vector, we see that $\mathbf{1}^T M\mathbf{1} = 1$, and so 1 is the largest eigenvalue of M, with $\mathbf{1}$ being one of the vectors in the eigenspace of 1.

So we have the following formula for λ_2:

$$\lambda_2 = \sup_{\mathbf{x}\in\mathbb{R}^n:||\mathbf{x}||=1,\sum_i x_i=0} \mathbf{x}^T M\mathbf{x}$$

where we equivalently expressed the condition $\mathbf{x} \perp \mathbf{1}$ as $\sum_i x_i = 0$.

Using (3), we have

$$\lambda_2 = 1 - \inf_{\mathbf{x}\in\mathbb{R}^n:||\mathbf{x}||=1,\sum_i x_i=0} \frac{1}{2}\sum_{ij} M_{ij}(x_i - x_j)^2$$

So, if $\lambda_2 = 1$, there must exist a non-zero $\mathbf{v} \in \mathbb{R}^n$ such that $\sum_i v_i = 0$ and $\sum_{ij} M_{ij}(v_i - v_j)^2 = 0$, but this means that, for every edge $(i,j) \in E$ of positive weight we have $v_i = v_j$, and so $v_i = v_j$ for every i,j which are in the same connected component. The conditions $\sum_i v_i = 0$ and $\mathbf{v} \neq \mathbf{0}$ imply that $\mathbf{v}$ has both positive and negative coordinates, and so the sets $A := \{i : v_i > 0\}$ and $B := \{i : v_i < 0\}$ are non-empty and disconnected, and so G is not connected.

Conversely, if G is disconnected, and S and $V - S$ are non-empty sets such that $E(S, V - S) = 0$, then we can define $\mathbf{v}$ so that $v_i = |S|/(|V - S|)$ if $i \notin S$, and $v_i = -|V - S|/|S|$ if $\in S$, so that $\sum_i v_i = 0$. This gives us a non-zero vector such that $\sum_{ij} M_{ij}(v_i - v_j)^2 = 0$ and, after dividing every coordinate by $||\mathbf{v}||$, a length-1 vector proving that $\lambda_2 \ge 1$.

Finally, to study λ_n we observe that for every vector $\mathbf{x} \in \mathbb{R}^n$ we have

$$\sum_{i,j} M_{i,j}(x_i + x_j)^2 = 2\mathbf{x}^T\mathbf{x} + 2\mathbf{x}^T M\mathbf{x}$$

and so

$$\begin{aligned}\lambda_n &= \min_{\mathbf{x}\in\mathbb{R}^n:||\mathbf{x}||=1} \mathbf{x}^T M\mathbf{x} \\ &= \min_{\mathbf{x}\in\mathbb{R}^n:||\mathbf{x}||=1} -\mathbf{x}^T\mathbf{x} + \frac{1}{2}\sum_{i,j} M_{i,j}(x_i + x_j)^2 \\ &= -1 + \min_{\mathbf{x}\in\mathbb{R}^n:||\mathbf{x}||=1} \frac{1}{2}\sum_{i,j} M_{i,j}(x_i + x_j)^2\end{aligned}$$

From which we see that it is always $\lambda_n \geq -1$, and that if $\lambda_n = -1$ then there must be a non-zero vector $\mathbf{x}$ such that $x_i = -x_j$ for every edge $(i, j) \in E$. Let i be a vertex such that $x_i = a \neq 0$, and define the sets $A := \{j : x_j = a\}$, $B := \{j : x_j = -a\}$ and $R = \{j : x_j \neq \pm a\}$. The set $A \cup B$ is disconnected from the rest of the graph, because otherwise an edge with an endpoint in $A \cup B$ and an endpoint in R would give a positive contribution to $\sum_{i,j} M_{i,j}(x_i + x_j)^2$; furthermore, every edge incident on a vertex on A must have the other endpoint in B, and vice versa. Thus, $A \cup B$ is a connected component, or a collection of connected components, of G which is bipartite, with the bipartition A, B.