EIGENVALUES OF SYMMETRIC MATRICES, AND GRAPH THEORY

Last week we saw how to use the eigenvalues of a matrix to study the properties of a graph. If our graph is undirected, then the adjacency matrix is symmetric. There are many special properties of eigenvalues of symmetric matrices, as we will now discuss.

Let \( A \) be a symmetric matrix. Let \( \lambda \) and \( \mu \) be eigenvalues of \( A \), with corresponding eigenvectors \( u \) and \( v \). We claim that, if \( \lambda \) and \( \mu \) are distinct, then \( u \) and \( v \) are orthogonal. Proof: We have \( u^T Av = \lambda (u^T v) \). But, also, \( u^T Av = (Au)^T v = \mu u^T v \). So \( \lambda u^T v = \mu u^T v \) and we deduce that \( u^T v = 0 \).

Thus, if \( A \) has \( n \) distinct eigenvalues, with \( n \) real eigenvectors \( v_i \), then the \( v_i \)'s are orthogonal and can be normalized to be orthonormal. In fact, more is true. As you should have learned in your linear algebra class, we have

**The Spectral Theorem:** If \( A \) is a symmetric real matrix, then the eigenvalues of \( A \) are real and \( \mathbb{R}^n \) has an orthonormal basis of eigenvectors for \( A \).

Let \( v_1, v_2, \ldots, v_n \) be the promised orthogonal basis of eigenvectors for \( A \). Let \( S \) be the matrix which takes the standard basis vector \( e_i \) to \( v_i \); explicitly, the columns of \( S \) are the \( v_i \). As we learned before, we have \( A = S^{-1} \text{diag}(\lambda_1, \ldots, \lambda_n) S \). However, in this case, things get even better. Since the \( v_i \) are orthonormal, the matrix \( S \) is orthogonal and we have \( S^{-1} = S^T \). In other words, the rows of \( S \) are again the \( v_i \):

\[
A = \begin{pmatrix}
  v_1^T \\
v_2^T \\
  \vdots \\
v_n^T
\end{pmatrix}
\begin{pmatrix}
  \lambda_1 \\
  \lambda_2 \\
  \vdots \\
  \lambda_n
\end{pmatrix}
\begin{pmatrix}
  v_1 \\
v_2 \\
  \vdots \\
v_n
\end{pmatrix}.
\]

This result is so important that we write it in several equivalent ways:

\[
A = \sum_{i=1}^{n} \lambda_i v_i v_i^T.
\]

\[
A(u) = \sum_{i=1}^{n} \lambda_i v_i \langle v_i, u \rangle.
\]

And, most importantly for our current purposes,

\[
\langle u, Au \rangle = \sum_{i=1}^{n} \lambda_i \langle v_i, u \rangle^2 \quad (*).
\]

Write \( \lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_n \). From equation (*), we have the following consequences:

For any vector \( u \neq 0 \), we have

\[
\lambda_1 \geq \frac{\langle Au, u \rangle}{\langle u, u \rangle} \geq \lambda_n.
\]

For any vector \( u \) which is orthogonal to \( v_1 \), we have

\[
\lambda_2 \geq \frac{\langle Au, u \rangle}{\langle u, u \rangle} \geq \lambda_n.
\]
The proof of the first equation is simple enough: if \( u = \sum c_i v_i \) then \( \langle Au, u \rangle / \langle u, u \rangle = (\sum \lambda_i c_i^2) / (\sum c_i^2) \), which is between \( \lambda_1 \) and \( \lambda_n \). The second result is similar, just noting that the condition that \( u \) is orthogonal to \( v_1 \) means that \( \lambda_1 = 0 \).

\* \* \*

Let \( G \) be a \( d \) regular graph.\(^1\) Let \( A \) be the adjacency matrix, and let \( \lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_n \) be the eigenvalues of \( A \). As a first application of these ideas, we show that all the \( \lambda_i \)'s lie between \(-d \) and \( d \). By the spectral theorem, we know that the \( \lambda_i \) are real.

Proof: Let \( L \) be the Laplacian matrix, \( L = d \cdot \Id - A \). So the eigenvalues of \( L \) are \( d - \lambda_1, \ldots, d - \lambda_n \). Note that, for any vector \( u \), we have \( \langle u, Lu \rangle = \sum_{(i,j) \in \text{Edge}(G)} (u(i) - u(j))^2 \geq 0 \). In particular, if \( v_i \) are the orthonormal eigenvectors of \( A \), then we have \( \langle v_i, Lv_i \rangle = (d - \lambda_i)\langle v_i, v_i \rangle = d - \lambda_i \). So \( d - \lambda_i \geq 0 \) and we see that \( d \geq \lambda_i \). A similar argument, using \( d \cdot \Id + A \), shows that \( -d \leq \lambda_i \).

**Exercise** Show that \( \lambda_2 = d \) if and only if \( G \) is disconnected. Show that \( \lambda_n = -d \) if and only if \( G \) is bipartite.

\* \* \*

We now return to our study of how well random walks mix on \( G \). Let \( \ell \) be the largest absolute value of any of \( \lambda_2, \lambda_3, \ldots, \lambda_n \), so \( \ell \) is either \( \lambda_2 \) or \( -\lambda_n \).

As we noted last time, the number of walks from \( r \) to \( s \) of length \( k \) is of the form \( c_1 d^k + c_2 \lambda_2^k + \cdots \) for various constants \( c_2, \ldots, c_n \) which we were not able to compute at that time. We now remedy this. We have

\[
A^k = S^T \begin{pmatrix} \lambda_1^k \\ \lambda_2^k \\ \vdots \\ \lambda_n^k \end{pmatrix} S.
\]

So the \((r,s)\) entry of \( A^k \) is

\[
\sum_{i=1}^n \langle v_i, v_i \rangle \lambda_i^k = \sum_{i=1}^n (v_i)_r (v_i)_s \lambda_i^k.
\]

Let’s start with the \( i = 1 \) term, which will give us the coefficient of \( d^n \). The eigenvector with eigenvalue \( d \) is \((1,1,\ldots,1)\). That’s before normalizing to become orthonormal. The all ones vector has length \( \sqrt{n} \), so \( v_1 = (1/\sqrt{n}, 1/\sqrt{n}, \ldots, 1/\sqrt{n}) \). Thus, the leading term of our sum is \( 1/\sqrt{n} d^n = d^n / n \). Notice that, if the \( d^n \) paths from \( r \) were distributed at random, we’d expect \( d^n / n \) of them to land at \( s \).

For a crude bound for the other terms, since \( \langle v_i, v_i \rangle = 1 \), every coordinate of \( v_i \) is at most \( 1 \). So we deduce that

\[
|\#(\text{paths from } r \text{ to } s \text{ of length } k) - d^n / n| \leq \ell^n + \ell^n + \cdots + \ell^n = (n-1)\ell^k.
\]

We can do a little better if we think about the fact that the \( v_i \) are orthonormal. Although any individual component in a \( v_i \) might be near \( 1 \), we can’t have many of them near \( 1 \) all at once. More specifically, we have the equality of matrices

\[
\sum v_i v_i^T = \Id
\]

so

\[
\sum_{i=1}^n (v_i)_r^2 = 1.
\]

\(^1\)Almost all of these ideas can be generalized to non-regular graphs, but the notation gets worse.
By Cauchy-Schwartz, $\sum |(v_i)_r||(v_i)_s| \leq 1$ and we have

$$|\#(\text{paths from } r \text{ to } s \text{ of length } k) - d^r/n| \leq \ell^k \left( \sum_{i=2}^{n} |(v_i)_r||(v_i)_s| \right) \leq \ell^k.$$ 

Let’s see what this implies about the diameter of $G$. Let $K + 1$ be the greatest distance between any two vertices of $G$. We must have that $d^K/n \leq \ell^K$, so that it is possible there can be no paths of length $\leq K$ from $r$ to $s$. In other words,

$$K \leq \frac{\log n}{\log d - \log \ell}.$$ 

Notice that $\log n/\log d$ is an obvious lower bound for $K$. So, if $\ell$ is significantly less than $d$, then we come within a constant factor of this lower bound.

This seems like a good point to tell the definition of an expander sequence: For fixed $d$, an expander sequence is a sequence of graphs $G_n$, all $d$ regular, such that the size of $G_n$ goes to $\infty$ there is a constant $R < d$ so that $\ell(G_n) < R$ for all $n$.

In particular, for an expander sequence, the diameter is close to the minimum possible, and we see that signals sent through an expander graph distribute very fast.

***

As a second, longer, application, we will show that, if $\lambda_2$ is significantly less than $d$, then any separation of $G$ into two pieces must cut many edges.

Let’s partition the vertices of $G$ into two sets, $X$ and $Y$. We will build a vector $u$ which is orthogonal to $v_1$. Namely, $u(w) = 1/|X|$ if $w \in X$ and $u(w) = -1/|Y|$ if $w \in Y$. Notice that $\langle u, v_1 \rangle$ is nothing but $\sum u(w) = |X|/|X| - |Y|/|Y| = 0$.

We thus know that

$$\frac{\langle u, Au \rangle}{\langle u, u \rangle} \leq \lambda_2.$$ 

Although it isn’t too bad to compute $\langle u, Au \rangle$, it is even easy to work with the Laplacian. The eigenvalues of $L$ are $0, d - \lambda_2, \ldots, d - \lambda_n$, and the eigenvectors are the same as for $A$. So

$$\frac{\langle u, Lu \rangle}{\langle u, u \rangle} \geq d - \lambda_2.$$ 

The numerator is $\sum_{(i,j) \in \text{Edge}(G)} (u(i) - u(j))^2$ This is

$$\#(\text{Edges between } X \text{ and } Y)(1/|X| + 1/|Y|)^2.$$ 

The denominator, meanwhile, is $\sum u(w)^2 = |X|/|X|^2 + |Y|/|Y|^2 = 1/|X| + 1/|Y|$.

Putting it all together,

$$\frac{\#(\text{Edges between } X \text{ and } Y)(1/|X| + 1/|Y|)^2}{1/|X| + 1/|Y|} \geq d - \lambda_2.$$ 

We have $1/|X| + 1/|Y| = |X| \cdot |Y|/(|X| + |Y|) = |X| \cdot |Y|/n$ so we have

$$\#(\text{Edges between } X \text{ and } Y)(1/|X| + 1/|Y|)^2 \geq \frac{d - \lambda_2}{n} |X| \cdot |Y|.$$ 

In other words, of the $|X| \cdot |Y|$ possible edges between $X$ and $Y$, at least $(d - \lambda_2)/n$ of them are present. For a random pair of vertices, the odds that there is an edge between them is $d/n$. So, when $\lambda_2$ is significantly less than $d$, the number of edges between $X$ and $Y$ is always within a constant ratio of what you would expect at random.