CS 798 - Algorithmic Spectral Graph Theory, Fall 2015, Waterloo

Lecture 2: Graph Spectrum

We start with studying the eigenvalues of the adjacency matrix of some simple graphs. Then, we introduce the Laplacian matrix of a graph and see some of its properties. Finally, we see a characterization of eigenvalues using Rayleigh quotient.

**Adjacency matrix**

Given an undirected graph $G$ with $V = [n]$, the adjacency matrix $A$ is an $n \times n$ matrix

where $a_{ij} = a_{ji} = 1$ if $ij$ is an edge of $G$, otherwise $a_{ij} = a_{ji} = 0$.

The adjacency matrix of an undirected graph is symmetric.

So, by the spectral theorem, it has an orthonormal basis of eigenvectors with real eigenvalues. It is not clear that these eigenvalues should provide any information about the graph properties. But they do, and surprisingly much information can be obtained from them.

Let's look at some examples and compute their spectrums.

**Complete graph**

What is the spectrum of a complete graph $K_n$?

If $G$ is a complete graph, then $A(G) = J - I$, where $J$ denotes the all-one matrix.

Any vector is an eigenvector of $I$ with eigenvalue 1.

Hence the eigenvalues of $A$ are one less than that of $J$.

Since $J$ is of rank 1, there are $n-1$ eigenvalues of 0.

The all-one vector is an eigenvector of $J$ with eigenvalue $n$.

So, $A$ has one eigenvalue of $n-1$, and $n-1$ eigenvalues of -1.

This is the example with largest possible spectral gap. It is an excellent expander, but it is too dense.

What we are interested in later in this course is to construct sparse graphs (linear number of edges) and yet with a large spectral gap.

**Bipartite graphs**

We can characterize bipartite graphs by the spectrums.

**Claim** If $G$ is a bipartite graph and $\lambda$ is an eigenvalue of $A(G)$ with multiplicity $k$, then $-\lambda$ is an eigenvalue of $A(G)$ with multiplicity $k$.

**Proof** If $G$ is a bipartite graph, then we can permute the rows and columns of $A(G)$ to
obtain the form $A(G) = \begin{pmatrix} 0 & B \\ B^T & 0 \end{pmatrix}$.

Suppose $u = (x \ y)$ is an eigenvector of $A(G)$ with eigenvalue $\lambda$.

Then $\begin{pmatrix} 0 & B \\ B^T & 0 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = \lambda \begin{pmatrix} x \\ y \end{pmatrix}$ which implies $B^T x = \lambda y$ and $By = \lambda x$.

This implies that $\begin{pmatrix} 0 & B \\ B^T & 0 \end{pmatrix} \begin{pmatrix} x \\ -y \end{pmatrix} = (-\lambda x) = (-\lambda) \begin{pmatrix} x \\ -y \end{pmatrix}$, and thus $\begin{pmatrix} x \\ -y \end{pmatrix}$ is an eigenvector of $A(G)$ with eigenvalue $-\lambda$.

$k$ linearly independent eigenvectors with eigenvalue $\lambda$ would give $k$ linearly independent with eigenvalue $-\lambda$, hence the claim.

The above result shows that the spectrum of a bipartite graph is symmetric around the origin. We now prove that the converse is also true.

Claim. If the nonzero eigenvalues occur in pairs $\lambda_i, -\lambda_i$ with $\lambda_i = -\lambda_i$, then $G$ is bipartite.

Proof. Let $k$ be an odd number.

Then $\sum_{i=1}^{k} \lambda_i^{k} = 0$.

Note that $\lambda_1^k, \lambda_2^k, \ldots, \lambda_n^k$ are the eigenvalues of $A^k$, because if $Au = \lambda u$ then $A Au = \lambda u$.

So, we have trace $(A^k) = \sum_{i=1}^{n} \lambda_i^k = 0$.

Observe that $A^k_{ij}$ is the number of length $k$ walks from $i$ to $j$ in $G$. (by induction)

If $G$ has an odd cycle of length $k$, then $A^k_{ii} > 0$ for some $i$ and trace $(A^k) > 0$.

So, since trace $(A^k) = 0$, $G$ must have no odd cycles and thus bipartite.

---

**Laplacian Matrices**

Given an undirected graph $G$, the **Laplacian matrix** $L(G)$ is defined as $D(G) - A(G)$, where

$$D(G) = \begin{pmatrix} d_1 & \ldots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \ldots & d_n \end{pmatrix}$$

is a diagonal matrix with $d_i =$ degree of vertex $i$ in $G$.

When $G$ is a regular graph, then $D = \begin{pmatrix} d & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & d \end{pmatrix}$ and $L = D - A$. Any eigenvector of $A$ with eigenvalue $\lambda$ is an eigenvector of $L$ with eigenvalue with eigenvalue $d - \lambda$, and vice versa.

So in this case the spectrums of the adjacency matrix and the Laplacian matrix are basically
equivalent, but when \( G \) is non-regular it may not be easy to relate their eigenvalues.

As a convention, I try to reserve the name \( \lambda_i \) for eigenvalues of the Laplacian matrix, and use the name \( \alpha_i \) for the \( i \)-th eigenvalue for the adjacency matrix.

Also, as the largest eigenvalue of the adjacency matrix corresponds to the smallest eigenvalue of the Laplacian matrix (in d-regular graphs). It is natural to order the eigenvalues of the adjacency matrix as \( \alpha_1 \geq \alpha_2 \geq \ldots \geq \alpha_n \), while using the reverse order \( \lambda_1 \geq \lambda_2 \geq \ldots \geq \lambda_n \) for the Laplacian matrix. So, later on, when I say the second (or \( k \)-th) eigenvalue of a graph, I mean the second largest eigenvalue of the adjacency matrix or the second smallest eigenvalue of the Laplacian matrix.

Let’s try to understand more about the spectrum of the Laplacian matrices.

Let \( \mathbf{1} \) be the all-one vector. Then it can be easily checked that \( \mathbf{L} \mathbf{1} = 0 \).

So \( \mathbf{L} \) has 0 as an eigenvalue.

Can \( \mathbf{L} \) have a smaller eigenvalue?

Let \( e = ij \) be an edge in \( G \),

Then it can be verified that \( \mathbf{L}(\mathbf{G}) = \mathbf{L}(\mathbf{G} - e) + [\begin{bmatrix} 1 & -1 \\ -1 & 1 \end{bmatrix} + j \begin{bmatrix} 1 \\ 1 \end{bmatrix}] \).

Let \( \mathbf{e} = ij \). Let \( \mathbf{b} \) be the column vector with the \( i \)-th position \( = 1 \) and the \( j \)-th position \( = -1 \), and 0 elsewhere.

By induction, we can write \( \mathbf{L}(\mathbf{G}) = \sum_{e \in \mathbf{G}} \mathbf{L}_e = \sum_{e \in \{ij \in \mathbf{G} \}} \mathbf{b} \mathbf{b}^T \).

Let \( \mathbf{B} = \begin{bmatrix} \mathbf{b}_1 & \mathbf{b}_2 & \ldots & \mathbf{b}_n \end{bmatrix} \) be the matrix whose columns are \( \{ \mathbf{b} | \mathbf{b} \in \{e \in \mathbf{G} \} \} \). Then \( \mathbf{L} = \mathbf{B} \mathbf{B}^T \).

This shows that \( \mathbf{L} \) is positive semidefinite, and thus 0 is the smallest eigenvalue.

One advantage of the Laplacian matrix is that the first eigenvalue is zero, and the first eigenvector is the all-one vector, and so they are easy to deal with.

Another way to see that \( \mathbf{L} \geq 0 \) is to show that \( \mathbf{x}^T \mathbf{L} \mathbf{x} \geq 0 \) for all \( \mathbf{x} \in \mathbb{R}^n \).

Another advantage of the Laplacian matrix is that it has a nice quadratic form.

Note that \( \mathbf{x}^T \mathbf{L} \mathbf{x} = \mathbf{x}^T \left( \sum_{e \in \mathbf{E}} \mathbf{L}_e \right) \mathbf{x} = \sum_{e \in \mathbf{E}} \mathbf{x}^T \mathbf{L}_e \mathbf{x} = \sum_{e \in \mathbf{E}} (\mathbf{x}_i - \mathbf{x}_j)^2 \) (assuming \( e = ij \)).

This immediately implies that \( \mathbf{x}^T \mathbf{L} \mathbf{x} \geq 0 \) for all \( \mathbf{x} \in \mathbb{R}^n \), proving that \( \mathbf{L} \geq 0 \).
**Connectedness**

**Claim.** A graph is connected if and only if 0 is an eigenvalue of \( L(G) \) with multiplicity 1.

**Proof.** If \( G \) is disconnected, then the vertex set can be partitioned into two sets \( S_1 \) and \( S_2 \) such that there are no edges between them. Then \( L(G) = \begin{pmatrix} 0 & 0 \\ 0 & L_1 \end{pmatrix} \) and so \( \begin{pmatrix} 1 \\ 0 \end{pmatrix} \) and \( \begin{pmatrix} 0 \\ 1 \end{pmatrix} \) are both eigenvectors of \( L(G) \) with eigenvalue 0, hence multiplicity \( \geq 2 \).

If \( G \) is connected, consider \( x^T L x = x^T \left( \sum_{k \in V} L_k \right) x = x^T \left( \sum_{e \in E} b_e e_e \right) x = \sum_{e \in E} (x_i - x_j)^2 \geq 0 \).

If \( x \) is an eigenvector with eigenvalue 0, then \( L x = 0 \) and thus \( x^T L x = 0 \).

For \( x^T L x = \sum_{e \in E} (x_i - x_j)^2 \geq 0 \), we must have \( x_i = x_j \) for every edge \( e \).

Since \( G \) is connected, it implies that \( x = c \cdot \bar{1} \) for some \( c \), i.e. a multiple of \( \bar{1} \).

Hence the eigenvalue 0 has multiplicity one, as required. \( \square \)

Actually, the proof can be used to prove the following (exercise).

**Claim.** The Laplacian matrix \( L(G) \) has 0 as its eigenvalue with multiplicity \( k \) if and only if the graph \( G \) has \( k \) connected components.

**Bipartiteness.** Like for adjacency matrix, it is also possible to characterize bipartite graphs by looking at its Laplacian spectrum.

For example, one can prove that for a connected \( d \)-regular graph \( G \), \( G \) is bipartite if and only if the maximum eigenvalue of \( L(G) \) is \( 2d \).

We leave the proof of this fact as a homework problem.

**Robust generalizations**

So far we have just used the graph spectrum to deduce some simple properties of the graph, like bipartiteness or connectedness, which are easy to deduce by other methods (e.g. BFS).

But the nice thing about these spectral characterizations is that they can be generalized in a robust way:

- \( \lambda_2 \) is "small" iff the graph is "close" to disconnected (i.e. existence of a "sparse" cut).
- \( \lambda_k \) is "small" iff the graph is "close" to having \( k \) connected components (i.e. \( k \) disjoint "sparse" cuts).
- \( \lambda_n \) is "close" to 2\( d \) iff the graph has a component "close" to being bipartite.
We will prove the first item (Cheeger's inequality) next time, and discuss the last two items after that.

**Rayleigh Quotient**

The main tool in relating eigenvalues and eigenvectors to optimization problem is the Rayleigh quotient, which is defined as \( \frac{x^TAx}{x^Tx} \).

Let \( A \) be a real symmetric matrix with eigenvalues \( \lambda_1, \lambda_2, \ldots, \lambda_n \), and orthonormal eigenvectors \( v_1, v_2, \ldots, v_n \).

Claim \( \lambda_i = \max_x \frac{x^TAx}{x^Tx} \)

Proof. Let \( x = c_1v_1 + c_2v_2 + \ldots + c_nv_n \), as \( v_1, v_2, \ldots, v_n \) form a basis.

Then \( x^TAx = (c_1v_1 + \ldots + c_nv_n)^T A (c_1v_1 + \ldots + c_nv_n) \)

\[ = (c_1v_1 + \ldots + c_nv_n)^T \left( \begin{array}{ccc} \lambda_1v_1 & & \\ & \ddots & \\ & & \lambda_nv_n \end{array} \right) \left( \begin{array}{c} c_1v_1 \\ \vdots \\ c_nv_n \end{array} \right) \]

\[ = \sum_i c_i^2 \lambda_i \quad \text{(because } v_1, \ldots, v_n \text{ are orthonormal)} \]

Similarly, \( x^Tx = (c_1v_1 + \ldots + c_nv_n)^T (c_1v_1 + \ldots + c_nv_n) = \sum_i c_i^2 \).

So, \( \frac{x^TAx}{x^Tx} = \frac{\sum_i c_i^2 \lambda_i}{\sum_i c_i^2} \leq \frac{\lambda_1}{\sum_i c_i^2} = \lambda_i \).

Since \( v_i \) attains the maximum, the claim follows. \( \square \)

This can be extended to characterize other eigenvalues.

Let \( T_k \) be the set of vectors that are orthogonal to \( v_1, v_2, \ldots, v_{k-1} \).

Claim \( \lambda_k = \max_{x \in T_k} \frac{x^TAx}{x^Tx} \)

Proof. Let \( x \in T_k \). Write \( x = c_1v_1 + \ldots + c_nv_n \).

Recall that \( c_i = \langle x, v_i \rangle \). Since \( x \in T_k \), we have \( c_1, c_2, \ldots, c_{k-1} = 0 \).

Then, \( \frac{x^TAx}{x^Tx} = \frac{\sum_i c_i^2 \lambda_i}{\sum_i c_i^2} \leq \frac{\lambda_k}{\sum_i c_i^2} = \lambda_k \).

Since \( v_k \in T_k \) and \( \frac{v_k^TAv_k}{v_k^Tv_k} = \lambda_k \), the claim follows. \( \square \)

The above result gives a characterization of \( \lambda_k \), but it requires the knowledge of the previous eigenvectors.

The result below gives a characterization without knowing the eigenvectors, and is more useful in giving bounds on eigenvalues.
The result below gives a characterization without knowing the eigenvectors, and is more useful in giving bounds on eigenvalues.

**Courant-Fischer Theorem**  
\[
\lambda_k = \max_{S \in \mathcal{S}} \min_{\mathbf{x} \in S} \frac{\mathbf{x}^T \mathbf{A} \mathbf{x}}{\mathbf{x}^T \mathbf{x}} = \min_{S \in \mathcal{S}} \max_{\mathbf{x} \in S} \frac{\mathbf{x}^T \mathbf{A} \mathbf{x}}{\mathbf{x}^T \mathbf{x}}
\]

**Proof.** We first consider the max-min term.

Let \( S_k \) be the \( k \)-dimensional subspace spanned by \( v_1, \ldots, v_k \), i.e., \( \{ \mathbf{x} \mid \mathbf{x} = c_1 v_1 + \cdots + c_k v_k \} \) for some \( c_1, \ldots, c_k \).

For any \( \mathbf{x} \in S_k \), \[
\frac{\mathbf{x}^T \mathbf{A} \mathbf{x}}{\mathbf{x}^T \mathbf{x}} = \frac{(c_1 v_1 + \cdots + c_k v_k)^T (c_1 v_1 + \cdots + c_k v_k)}{(c_1 v_1 + \cdots + c_k v_k)^T (c_1 v_1 + \cdots + c_k v_k)} = \sum_{i=1}^k c_i^2 \lambda_i \leq \frac{\lambda_k}{c_i^2} \leq \lambda_k.
\]

So, \[
\max_{S \in \mathcal{S}} \min_{\mathbf{x} \in S} \frac{\mathbf{x}^T \mathbf{A} \mathbf{x}}{\mathbf{x}^T \mathbf{x}} \geq \min_{\mathbf{x} \in S_k} \frac{\mathbf{x}^T \mathbf{A} \mathbf{x}}{\mathbf{x}^T \mathbf{x}} \geq \lambda_k.
\]

To prove that the maximum cannot be greater than \( \lambda_k \), observe that any \( k \)-dimensional subspace must intersect the \( k+1 \) dimensional subspace \( T_k \) spanned by \( \{ v_k, v_{k+1}, \ldots, v_n \} \).

For any \( \mathbf{x} \in T_k \), \[
\frac{\mathbf{x}^T \mathbf{A} \mathbf{x}}{\mathbf{x}^T \mathbf{x}} = \frac{\sum_{i=1}^k c_i^2 \lambda_i}{\sum_{i=1}^k c_i^2} \leq \lambda_k.
\]

So, \[
\min_{S \in \mathcal{S}} \max_{\mathbf{x} \in S} \frac{\mathbf{x}^T \mathbf{A} \mathbf{x}}{\mathbf{x}^T \mathbf{x}} \leq \max_{S \in \mathcal{S}} \min_{\mathbf{x} \in S} \frac{\mathbf{x}^T \mathbf{A} \mathbf{x}}{\mathbf{x}^T \mathbf{x}} \leq \lambda_k.
\]

The min-max characterization can be proved by a similar argument (i.e., intersection of two subspaces), and we leave the details to the reader.

One consequence of the Courant-Fischer theorem is the eigenvalue interlacing theorem.

We will study eigenvalues interlacing intensively in the last part of the course.

**Eigenvalue Interlacing Theorem**  
Let \( \mathbf{A} \) be an \( mn \) symmetric matrix and let \( \mathbf{B} \) be a principle submatrix of dimension \( n-1 \) (i.e., \( \mathbf{B} \) is obtained from \( \mathbf{A} \) by deleting the same row and column from \( \mathbf{A} \)). Then \[
\lambda_1, \lambda_2, \lambda_3, \lambda_4, \ldots, \lambda_{n-1} \leq \beta_1, \beta_2, \beta_3, \beta_4, \ldots, \beta_{n-1}, \]

when \( \lambda_1, \lambda_2, \ldots, \lambda_n \) are the eigenvalues of \( \mathbf{A} \) and \( \beta_1, \beta_2, \ldots, \beta_{n-1} \) are the eigenvalues of \( \mathbf{B} \).

**Proof.** It should be clear that \( \lambda_i \geq \beta_i \), because \( \lambda_i = \max_{S \in \mathcal{S}} \min_{\mathbf{x} \in S} \frac{\mathbf{x}^T \mathbf{A} \mathbf{x}}{\mathbf{x}^T \mathbf{x}} = \max_{S \in \mathcal{S}} \min_{\mathbf{x} \in S} \frac{\mathbf{x}^T \mathbf{B} \mathbf{x}}{\mathbf{x}^T \mathbf{x}} \)

Simply put, because the search space for \( \mathbf{A} \) is larger than that for \( \mathbf{B} \).

Next we prove \( \beta_i \leq \lambda_i \). For any \( S \in \mathcal{S} \) with \( \dim(S) = i+1 \), its restriction to the first \( i+1 \) coordinates (i.e., \( S \cap \mathbb{R}^{i+1} \)) is of dimension at least \( i \).
So, informally, if there is a good \((i+1)\)-dimensional subspace for \(A\), then there is a good \(i\)-dimensional subspace for \(B\), and so \(\beta_i\) can do as well as \(\alpha_i\).

More formally, let \(S_i\) be the set that attains maximum for \(\alpha_{i+1}\).

Then, \(\alpha_{i+1} = \min_{x \in S_i} \frac{x^T A x}{x^T x} \leq \min_{x \in S_i} \frac{x^T A x}{x^T x} \leq \max_{S \subseteq \mathbb{R}^m} \min_{x \in S} \frac{x^T A x}{x^T x} = \max_{S \subseteq \mathbb{R}^m} \min_{x \in S} \frac{x^T B x}{x^T x} = \beta_i.\)

**First Eigenvalue**

Let \(A\) be the adjacency matrix of an undirected graph. Let \(\lambda\) be its largest eigenvalue.

**Claim** \(\lambda \leq \Delta_{\max}\), where \(\Delta_{\max}\) denotes the maximum degree in \(G\).

**Proof** Let \(v_1\) be an eigenvector with eigenvalue \(\lambda_1\).

Let \(\bar{j}\) be the vertex with \(v_1(j) \geq v_1(i)\) for all \(i\).

\[\lambda_1 v_1(j) = (Av_1)(j) = \sum_{i:j \in E(G)} v_1(i) = \Delta_{\max} v_1(j) \leq \deg(j) v_1(j) \leq \Delta_{\max} v_1(j).\]

Therefore, \(\lambda_1 \leq \Delta_{\max}\) \(\Box\)

In fact, if \(\lambda_1 = \Delta_{\max}\), then the above inequalities must hold as equalities, i.e., \(v_1(i) = v_1(j)\) for every neighbor \(i\) of \(\bar{j}\) and \(\deg(j) = \Delta_{\max}\). It implies that when \(G\) is connected and \(\lambda_1 = \Delta_{\max}\), then \(G\) must be \(\Delta_{\max}\)-regular and the eigenvalue \(\lambda_1\) is of multiplicity 1, since the eigenvectors for \(\lambda_1\) must be of the form \(c \bar{j}\) for some constant \(c\).

**Claim** \(\lambda_1 \geq \Delta_{\text{avg}}\), where \(\Delta_{\text{avg}}\) denotes the average degree of \(G\).

**Proof** \(\lambda_1 = \max_x \frac{x^T A x}{x^T x} \geq \frac{1}{n} \sum_{i,j} a_{i,j} = \frac{\Delta_{\text{avg}}}{n} = \frac{2m}{n} = \Delta_{\text{avg}}\) \(\Box\)

More generally, \(\lambda_i\) is at least the average degree of the densest induced subgraph.

The Perron-Frobenius theorem for non-negative matrices tells us more about the first eigenvalue.

**Theorem** Let \(G\) be a connected undirected graph. Then

1. the first eigenvalue is of multiplicity one
2. \(|\lambda_i| \leq \lambda_1\) for all \(i\).
3. all entries of the first eigenvector are non-zero and have the same sign.
We will not prove it in class. See Chapter 8.6 of [GR] for proofs.

**Summary**  Let me highlight some key points for future reference:

- A graph is bipartite iff the spectrum of the adjacency matrix is symmetric.
  - We use an argument about the trace to prove it, and we will see this argument again.

- Remember the following properties of the Laplacian matrix: $L \leq 0$, $L = \sum_{e \in E} L_e$, and $x^T L x = \sum_{i \neq j} (x_i - x_j)^2$.

- The Rayleigh quotient and its uses in characterizing eigenvalues.

- The Perron-Frobenius theorem will be used when we study random walks and also in the homework!

**References**

[GR] Algebraic Graph Theory, by Godsil and Royle.