CS 860 Spectral graph theory, Spring 2019, Waterloo.

Lecture 2 : Graph spectrum

First, 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 show a characterization of eigenvalues using Rayleigh quotient.

Adjacency matrix

Given an undirected graph G with $V(G)=\{1,\dots,n\}$, the adjacency matrix $A(G)$ is an $n \times n$ matrix with $a_{ij}=a_{ji}=1$ if $ij \in E(G)$, otherwise $a_{ij}=a_{ji}=0$ if $ij \notin E(G)$.

The adjacency matrix of an undirected graph is symmetric.

So, by the spectral theorem for real symmetric matrices, the adjacency matrix has an orthonormal basis of eigenvectors with real eigenvalues.

Apriori, 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 start with 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 an example with the largest eigenvalue gap between the largest eigenvalue and the second largest.

Complete bipartite graph What is the spectrum of a complete bipartite graph $K_{m,n}$?

The adjacency matrix of $K_{m,n}$ looks like this: $\begin{array}{c} m\{ \\ n\{ \end{array} \left[\begin{array}{cc} 0 & \boxed{1} \\ \boxed{1} & 0 \end{array}\right]$
$\quad\;\;\; \underbrace{\;}_{m}\ \underbrace{\;}_{n}$.

It is of rank 2, so there are $n+m-2$ eigenvalues of 0, and two non-zero eigenvalues λ_1, λ_2.

As $\sum_{i=1}^{n+m} \lambda_i = \text{trace}(A) = 0$, we have $\lambda_1 = -\lambda_2$. Let this value be k.

As $\sum_{i=1}^{n+m} \lambda_i = \text{trace}(A) = 0$, we have $\lambda_1 = -\lambda_2$. Let this value be k.

Thus, $\det(\lambda I - A) = \lambda^{n+m} - k^2 \lambda^{n+m-2}$.

To determine k, we use the (expansion) definition of determinant of $\begin{pmatrix} \lambda & & \boxed{-1} & & \\ & \lambda & & & \\ \boxed{-1} & & \lambda & & \\ & & & \ddots & \\ & & & & \lambda \end{pmatrix}$.

Any term that contributes to λ^{n+m-2} must have $n+m-2$ diagonal entries,

and the remaining two entries must be $-a_{ij}$ and $-a_{ji}$ for some i, j.

There are totally mn such terms, where the sign of each term is -1.

So, $k^2 = mn$, and thus $k = \sqrt{mn}$.

To conclude, there are $n+m-2$ eigenvalues of 0, and one eigenvalue of $\sqrt{mn}$, and one of $-\sqrt{mn}$.

Bipartite graphs We can charactize bipartite graphs by the spectrums.

Claim If G is a bipartite graph and λ 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 $u = \begin{pmatrix} x \\ y \end{pmatrix}$ is an eigenvector of $A(G)$ with eigenvalue λ.

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 $\begin{pmatrix} 0 & B \\ B^T & 0 \end{pmatrix} \begin{pmatrix} x \\ -y \end{pmatrix} = \begin{pmatrix} -By \\ B^T x \end{pmatrix} = \begin{pmatrix} -\lambda x \\ \lambda y \end{pmatrix} = -\lambda \begin{pmatrix} x \\ -y \end{pmatrix}$,

and thus $\begin{pmatrix} x \\ -y \end{pmatrix}$ is an eigenvector of $A(G)$ with eigenvalue $-\lambda$.

k linearly independent eigenvectors with eigenvalue λ 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 λ_i, λ_j with $\lambda_i = -\lambda_j$, then G is bipartite.

Proof Let k be an odd number.

Then $\sum_{i=1}^{n} \lambda_i^k = 0$, by the symmetry of the spectrum.

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

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

Observe that $A^k_{i,j}$ 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 $\text{trace}(A^k) > 0$.

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

To conclude, a graph is bipartite if and only if the spectrum of the adjacency matrix is symmetric around the origin.

Laplacian Matrices

Given an undirected graph G, the Laplacian matrix $L(G)$ is defined as $D(G) - A(G)$, where $D(G) = \begin{pmatrix} d_1 & & & 0 \\ & d_2 & & \\ & & \ddots & \\ 0 & & & d_n \end{pmatrix}$ is a diagonal matrix with d_i = degree of vertex i in G.

When G is a regular graph, then $D = \begin{pmatrix} d & & 0 \\ & d & \\ 0 & & \ddots d \end{pmatrix}$ and $L = D - A$. Any eigenvector of A with eigenvalue λ 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 λ_i as eigenvalues of the Laplacian matrix, and use the name α_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 $\lambda_1 \leq \lambda_2 \leq \ldots \leq \lambda_n$ 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 $\vec{1}$ be the all-one vector. Then it can be easily checked that $L\vec{1} = 0$.

So L has 0 as an eigenvalue.

Can L have a smaller eigenvalue ?

Let $e = ij$ be an edge in G.

call this L_e

Then it can be verified that $L(G) = L(G-e) + \begin{array}{c|cc} & i & j \\ \hline i & 1 & -1 \\ j & -1 & 1 \end{array} = L(G-e) + L_e = L(G-e) + \begin{array}{c|c} i & 1 \\ j & -1 \end{array} \begin{array}{cc} i & j \\ 1 & -1 \end{array}$

Let $e = ij$. Let b_e be the column vector with the i-th position $= +1$ and the j-th position $= -1$, and 0 otherwise.

By induction, we can write $L(G) = \sum_{e \in E} L_e = \sum_{e = ij \in G} b_e b_e^T$.

Let $B = \begin{pmatrix} | & | & & | \\ b_1 & b_2 & \cdots & b_m \\ | & | & & | \end{pmatrix}$ be the matrix whose columns are $\{ b_e \mid e \in G \}$. Then $L = BB^T$.

This shows that L is positive semidefinite, and thus 0 is the smallest eigenvalue.

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 $L(G) = \begin{pmatrix} L(G_1) & 0 \\ 0 & L(G_2) \end{pmatrix}$ and so $\begin{pmatrix} 1 \\ \vdots \\ 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix}$ and $\begin{pmatrix} 0 \\ \vdots \\ 0 \\ 1 \\ \vdots \\ 1 \end{pmatrix}$ are both eigenvectors of $L(G)$ with eigenvalue 0, hence multiplicity ≥ 2.

If G is connected, consider $x^T L x = x^T \left(\sum_{e \in G} L_e \right) x = x^T \left(\sum_{e \in G} b_e b_e^T \right) x = \sum_{e = ij \in G} (x_i - x_j)^2 \geq 0$

If x is an eigenvector with eigenvalue 0, then $Lx = 0$ and thus $x^T L x = 0$.

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

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

Hence, the eigenvalue 0 has multiplicity one. □

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.

Robust generalizations

So far we have just used the graph spectrum to deduce some simple properties of the graph, such as 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 nontrivially:

- λ_2 is "small" iff the graph is "close" to disconnected (i.e. existence of a "sparse" cut).

- λ_k is "small" iff the graph is "close" to having k connected components (i.e. k disjoint "sparse" cuts).
- α_n is "close" to $-\alpha_1$ (adjacency matrix) iff the graph has a component "close" to bipartite.

We will discuss these results in the next lectures.

Rayleigh Quotient

The main tool in relating eigenvalues and eigenvectors to optimization problems is the Rayleigh quotient,

which is defined as $\frac{x^TAx}{x^Tx} = \frac{\sum_{i,j} a_{ij}x_ix_j}{\sum_i x_i^2}$.

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

Claim $\quad \lambda_1 = \max_x \frac{x^TAx}{x^Tx}$

proof $\quad$ Let $x = c_1v_1 + c_2v_2 + \ldots + c_nv_n$, as $v_1, \ldots, v_n$ form a basis

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

$$= (c_1v_1 + \ldots + c_nv_n)^T (c_1\lambda_1v_1 + c_2\lambda_2v_2 + \ldots + c_n\lambda_nv_n)$$

$$= \sum_{i=1}^n c_i^2\lambda_i \quad \text{(because } v_1, \ldots, v_n \text{ are orthonormal)}$$

Similarly, $x^Tx = (c_1v_1 + \ldots + c_nv_n)^T(c_1v_1 + \ldots + c_nv_n) = \sum_{i=1}^n c_i^2$.

So, $\frac{x^TAx}{x^Tx} = \frac{\sum_{i=1}^n c_i^2\lambda_i}{\sum_{i=1}^n c_i^2} \leq \frac{\lambda_1 \sum_{i=1}^n c_i^2}{\sum_{i=1}^n c_i^2} = \lambda_1$.

Since v_1 attains the maximum, the claim follows. □

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 $\quad \lambda_k = \max_{x \in T_k} \frac{x^TAx}{x^Tx}$

proof $\quad$ 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, $\frac{x^TAx}{x^Tx} = \frac{\sum_{i=k}^n c_i^2\lambda_i}{\sum_{i=k}^n c_i^2} \leq \frac{\lambda_k \sum_{i=k}^n c_i^2}{\sum_{i=k}^n c_i^2} = \lambda_k$.

Since $v_k \in T_k$ and $\frac{v_k^TAv_k}{v_k^Tv_k} = \lambda_k$, the claim follows. □

The above result gives a characterization of λ_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.

Courant-Fischer Theorem $$\lambda_k = \max_{\substack{S\subseteq R^n \\ \dim(S)=k}} \min_{x\in S} \frac{x^TAx}{x^Tx} = \min_{\substack{S\subseteq R^n \\ \dim(S)=n-k+1}} \max_{x\in S} \frac{x^TAx}{x^Tx}$$

Proof We first consider the max-min term.

Let S_k be the k-dimensional subspace spanned by $v_1,\dots,v_k$, i.e. $\{ x \mid x = c_1v_1+\dots+c_kv_k \text{ for some } c_1,\dots,c_k \}$.

For any $x\in S_k$, $$\frac{x^TAx}{x^Tx} = \frac{(c_1v_1+\dots+c_kv_k)^T A (c_1v_1+\dots+c_kv_k)}{(c_1v_1+\dots+c_kv_k)^T(c_1v_1+\dots+c_kv_k)} = \frac{\sum_{i=1}^k c_i^2\lambda_i}{\sum_{i=1}^k c_i^2} \geq \frac{\lambda_k\sum_{i=1}^k c_i^2}{\sum_{i=1}^k c_i^2} = \lambda_k .$$

So, $$\max_{\substack{S\subseteq R \\ \dim(S)=k}} \min_{x\in S} \frac{x^TAx}{x^Tx} \geq \min_{x\in S_k} \frac{x^TAx}{x^Tx} \geq \lambda_k .$$

To prove that the maximum cannot be greater that λ_k, observe that any k-dimensional subspace must intersect the n-k+1 dimensional subspace T_k spanned by $\{v_k, v_{k+1}, \dots, v_n\}$.

For any $x\in T_k$, $$\frac{x^TAx}{x^Tx} = \frac{\sum_{i=k}^n c_i^2\lambda_i}{\sum_{i=k}^n c_i^2} \leq \lambda_k .$$

So, $$\max_{\substack{S\subseteq R \\ \dim(S)=k}} \min_{x\in S} \frac{x^TAx}{x^Tx} \leq \max_{\substack{S\subseteq R \\ \dim(S)=k}} \min_{x\in S\cap T_k} \frac{x^TAx}{x^Tx} \leq \lambda_k . \quad \square$$

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 A be an n×n symmetric matrix and let B be a principle submatrix of dimension n-1 (i.e. B is obtained from A by deleting the same row and column from A). Then

$$\alpha_1 \geq \beta_1 \geq \alpha_2 \geq \beta_2 \geq \dots \geq \alpha_{n-1} \geq \beta_{n-1} \geq \alpha_n ,$$

when $\alpha_1 \geq \alpha_2 \geq \dots \geq \alpha_n$ are the eigenvalues of A and $\beta_1 \geq \beta_2 \geq \dots \geq \beta_{n-1}$ are the eigenvalues of B.

Proof It should be clear that $\alpha_i \geq \beta_i$, because $$\alpha_i = \max_{\substack{S\subseteq R^n \\ \dim(S)=i}} \min_{x\in S} \frac{x^TAx}{x^Tx} \geq \max_{\substack{S\subseteq R^{n-1} \\ \dim(S)=i}} \min_{x\in S} \frac{x^TAx}{x^Tx}$$
$$= \max_{\substack{S\subseteq R^{n-1} \\ \dim(S)=i}} \min_{x\in S} \frac{x^TBx}{x^Tx} = \beta_i .$$

Simply put, because the search space for A is larger than that for B.

Next, we prove $\beta_i \geq \alpha_{i+1}$. For any $S\subseteq R^n$ with $\dim(S)=i+1$, its restriction to the first n-1 coordinates (i.e. $S\cap R^{n-1}$) 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 β_i can do as well as α_{i+1}.

More formally, let S^* be the set that attains maximum for α_{i+1}.

Then, $\alpha_{i+1} = \min_{x \in S^*} \frac{x^T A x}{x^T x} \leq \min_{x \in S^* \cap R^{n-1}} \frac{x^T A x}{x^T x} \leq \max_{\substack{S \subseteq R^{n-1} \\ \dim(S)=i}} \min_{x \in S} \frac{x^T A x}{x^T x} = \max_{\substack{S \subseteq R^{n-1} \\ \dim(S)=i}} \min_{x \in S} \frac{x^T B x}{x^T x} = \beta_i$. □

First Eigenvalue

Let A be the adjacency matrix of an undirected graph. Let α_1 be its largest eigenvalue.

Claim $\alpha_1 \leq d_{max}$, where d_{max} denotes the maximum degree in G.

Proof Let v_1 be an eigenvector with eigenvalue λ_1.

Let j be the vertex with $v_1(j) \geq v_1(i)$ for all i.

$$\alpha_1 v_1(j) = (Av_1)(j) = \sum_{i: ij \in E(G)} v_1(i) \leq \sum_{i: ij \in E(G)} v_1(j) = \deg(j)\, v_1(j) \leq d_{max}\, v_1(j).$$

Therefore, $\alpha_1 \leq d_{max}$. □

In fact, if $\lambda_1 = d_{max}$, then the above inequalities must hold as equalities, i.e. $v_1(i) = v_1(j)$ for every neighbor i of j and $\deg(j) = d_{max}$. It implies that when G is connected and $\lambda_1 = d_{max}$, then G must be d_{max}-regular and the eigenvalue λ_1 is of multiplicity 1, since the eigenvectors for λ_1 must be of the form $c \cdot \vec{1}$ for some constant c.

Exercise: Prove that $\alpha_1 \geq d_{avg}$, where d_{avg} denotes the average degree of the graph.

More generally, prove that λ_1 is at least the average degree of the densest induced subgraph.

Homework: Prove that the maximum eigenvalue of the adjacency matrix of a tree with maximum degree d is at most $2\sqrt{d-1}$.

The Perron-Frobenius theorem for non-negative matrices tell 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 here. See Chapter 8.6 of [GR] for proofs.

Summary Let me highlight some key points for future references.

- A graph is bipartite iff the spectrum of the adjacency matrix is symmetric around the origin.

  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 : $L \succcurlyeq 0$, $L = \sum_{e \in E} L_e$, and $x^T L x = \sum_{ij \in E} (x_i - x_j)^2$.
- The Rayleigh quotient and its uses in characterizing eigenvalues.
- The Perron-Frobenius theorem will be used when we study random walks.

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