Columbia UGTCS: Algebraic and Spectral Graph Theory Fall 2023

Eigenvalues, Optimisation; Graph Laplacian

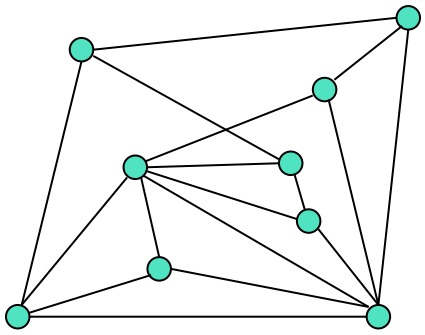
Ekene Ezeunala *Scribe: Ekene Ezeunala*

1 Recap

Let us start by recalling some of the things we introduced last week. First, one we all surely remember:

Definition 1. *A graph $G = (V, E)$ is a pair of sets V and E where $E \subseteq V \times V$. The elements of V are called vertices and the elements of E are called edges.*

We say that $e = (u, v) \in E$ is an edge between u and v. We say that u and v are adjacent if $(u, v) \in E$. We say that u and v are neighbours (or $u \sim v$) if they are adjacent. Here's a graph:



We will (mostly) consider cases where G is finite, has no isolated vertices (i.e. no $v \in V$ s.t. $\deg(v) = 0$), could have self-loops and parallel edges, is not necessarily connected, and is unweighted. However, from time to time we will consider graphs that do not satisfy some of these properties (this doesn't change things too much; we can deal with typical weights w by using parallel unweighted edges).

We are very interested in the matricial representations of graphs. One of the most natural ones is the adjacency matrix:

Definition 2 (Adjacency matrix)**.** *Let $\mathcal{G}$ be an unweighted graph with vertices $v_1, v_2, \ldots, v_n$. Then the adjacency matrix of $\mathcal{G}$ is the matrix $A \in \mathrm{Mat}(n \times n; \{0, 1\})$, whose (i, j) entry, denoted by $[A]_{i,j}$, is defined by*

$$[A]_{i,j} = \begin{cases} 1 & \text{if } v_i \text{ and } v_j \text{ are adjacent,} \\ 0 & \text{otherwise.} \end{cases}$$

This is merely a spreadsheet of the graph. It is quite natural, but we will see that it is not the most useful representation of a graph. Here's a more useful one:

Definition 3 (Laplacian matrix)**.** *Define the degree matrix* $D \in \mathrm{Mat}(n \times n; \mathbb{R})$ *as*

$$D_{i,j} = \begin{cases} \deg(v_i), & \text{if } i = j \\ 0, & \text{otherwise.} \end{cases}$$

The Laplacian matrix is then defined as $L = D - A$*.*

As we will see today, there is a close analogy between the Laplacian matrix and the Laplace operator in vector calculus, so we can (very) roughly think of the Laplacian matrix as a discrete "measure of the smoothness" of a graph.

We also saw last time that we can define a function $f : V \to \mathbb{R}$ on G. This function could represent, say, the temperature at each vertex or the electrical voltage in a circuit with resistance, or simply a 0/1 indicator for a subset $S \subset V$. f looks like:

$$f : V \to \mathbb{R} \equiv \begin{bmatrix} f(v_1) \\ f(v_2) \\ \vdots \\ f(v_n) \end{bmatrix}$$

Note that $(f + g)(x) \coloneqq f(x) + g(x)$ and $(\alpha \cdot f)(x) \coloneqq \alpha \cdot f(x)$ for $c \in \mathbb{R}$ are well-defined operations on these functions, so they behave like $\mathbb{R}^n$.

We also saw that we can speak of the *local variance* of f:

Definition 4 (Local variance)**.** *The local variance of* f *is defined as:*

$$\mathcal{E}(f) \coloneqq \mathbb{E}_{u \sim v}\left[(f(u) - f(v))^2\right],$$

where $u \sim v$ *is a probability distribution over the edges.*

Roughly speaking, $\mathcal{E}(f)$ is "small" if f is "smooth" over the edges E and "large" if f is "rough" over E.

Finally, we stated and proved the spectral theorem for symmetric matrices:

Theorem 5 (Spectral Theorem)**.** *If* $M \in \mathrm{Mat}(n \times n; \mathbb{R})$ *is an* n*-by-*n*, real, symmetric matrix, then there exist real numbers* $\lambda_1, \ldots, \lambda_n$ *and* n *mutually orthogonal unit vectors* $\nu_1, \ldots, \nu_n$ *and such that* ν_i *is an eigenvector of* M *of eigenvalue* λ_i*, for each* i*.*

2 Eigenvalues and Optimisation

We'll start today by talking about symmetric matrices, matrices $A \in \mathrm{Mat}(n \times n; \mathbb{R})$ satisfying $A^\top = A$. These matrices have many nice properties:

1. They have real eigenvalues.
2. They are orthogonally diagonalisable, i.e. there exists an orthogonal matrix Q and a diagonal matrix Λ such that $A = Q\Lambda Q^\top$. This is called the *spectral decomposition* of A (indeed it is a consequence of the spectral theorem). Furthermore, Λ has the eigenvalues of A on its diagonal, and Q has the corresponding orthonormal eigenvectors of A as its columns.

One useful property of these symmetric matrices is that they let us define the so-called *quadratic form*:

Definition 6 (Quadratic form)**.** *Let $A \in \mathrm{Mat}(n \times n; \mathbb{R})$ be a symmetric matrix. Then the quadratic form of A is defined as:*

$$Q_A(x) \coloneqq x^\top Ax, \qquad \text{for all } x \in R^n.$$

One thing we can immediately note is that every quadratic form is a popularity with terms of degree at most 2. This is because:

$$Q_A(x) = x^\top Ax = \sum_{i=1}^{n}\sum_{j=1}^{n} A_{i,j}x_i x_j,$$

so that if we pick, say, $A = \begin{bmatrix} 1 & 5 \\ 5 & 4 \end{bmatrix}$ and $x = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$, then

$$Q_A(x) = x^\top Ax = x_1^2 + 4x_2^2 + 20x_1x_2.$$

(If $A = L$, we can describe the *Laplacian* quadratic form of a graph with weights $w_{i,j}$ as:

$$Q_L(f) = f^\top Lf = \sum_{(i,j)\in E} w_{i,j}(f(i) - f(j))^2,$$

where $f \in \mathbb{R}^V$ is a function on the vertices of the graph G. This quadratic form measures the "smoothness" of f; it is small if f does not "jump too much" over any edge.)
Now, we can ask the following question: given a symmetric matrix A, what is the maximum value of $Q_A(x)$ over all $x \in \mathbb{R}^n$ with $\|x\| = 1$? More concretely, what is the solution to the optimisation problem

$$\max_{x\in\mathbb{R}^n} x^\top Ax \text{ subject to } \|x\|^2 = 1?$$

The following is an unconstrained version of this problem:

$$\max_{\substack{x\in\mathbb{R}^n \\ x\neq 0}} \frac{x^\top Ax}{x^\top x}.$$

The "scaling-invariant" quantity $\frac{x^\top Ax}{x^\top x}$ is called the *Rayleigh quotient* of A at x. This is an important quantity; it appears in quite a few popular applications:

1. principal component analysis (PCA): Wish to solve $\max_{\substack{v\in\mathbb{R}^n \\ v\neq 0}} \frac{v^\top \Sigma v}{v^\top v}$, where Σ is the covariance matrix of a random variable $X \in \mathbb{R}^n$—here, we consider the *normalised* Rayleigh quotient.

2. latent Dirichlet allocation (LDA): Wish to solve $\max_{\substack{v\in\mathbb{R}^n \\ v\neq 0}} \frac{v^\top S_b v}{v^\top S_w v}$, where $S_b, S_w \in \mathrm{Mat}(n \times n; \mathbb{R})$ are the between-class and within-class scatter matrices, respectively—here, we consider the *generalised* Rayleigh quotient.

3. spectral clustering: Wish to solve $\min_{\substack{v\in\mathbb{R}^n \\ v\neq 0}} \frac{v^\top Lv}{v^\top Dv}$, where L is the Laplacian matrix of a graph G and D is its degree matrix—here, we again consider the *generalised* Rayleigh quotient.

Perhaps the biggest value the Rayleigh quotient has for us is that it lets us find the eigenvalues of a matrix. Indeed, we have a following min-max characterisation of eigenvalues via the Courant-Fischer theorem:

Theorem 7 (Courant-Fischer)**.** *Let $M \in \mathrm{Mat}(n\times n;\mathbb{C})$ be Hermitian. Then for each $1 \le k \le n$, let $\{S_k^\alpha\}_{\alpha\in I_k}$ be the set of all k-dimensional linear subspaces of $\mathbb{C}^n$. Also, enumerate the n eigenvalues $\lambda_1,\dots,\lambda_n$ of M (counting multiplicities) in increasing order, i.e. $\lambda_1 \le \cdots \le \lambda_n$. Then we have*

$$\lambda_k \overset{(i)}{=} \min_{\alpha\in I_k} \max_{x\in S_k^\alpha\setminus\{0\}} \frac{x^\top Mx}{x^\top x} \overset{(ii)}{=} \max_{\alpha\in I_{n-k+1}} \min_{x\in S_{n-k+1}^\alpha\setminus\{0\}} \frac{\nu^\top M\nu}{\nu^\top\nu},$$

where λ_k is the kth largest eigenvalue of M and S,T are subspaces of $\mathbb{C}^n$.

Proof. Denote an orthonormal basis for the eigenvectors as $u_1,\dots,u_n$ corresponding to eigenvalues $\lambda_1,\dots,\lambda_n$. We will first proved equation (i), and then (ii):

(i) First let $W = \mathsf{span}\{u_1,\dots,u_n\}$, so that $\dim(W) = n-k+1$. Thus for any k-dimensional subspace S_k^α, we should have $\dim(S_k^\alpha \cap W) \ge 1$; this is because $\dim(S_k^\alpha + W) = \dim(S_k^\alpha) + \dim(W) - \dim(S_k^\alpha \cap W)$, and of course $\dim(S_k^\alpha + W) \le n$.
Now choose $x \in (S_k^\alpha \cap W)\setminus\{0\}$, and note that $x = \sum_{j=k}^n \langle x,u_j\rangle u_j$ and $Au_j = u_j$, so that

$$\begin{aligned}\frac{x^\top Mx}{x^\top x} &= \frac{\langle Mx,x\rangle}{\|x\|^2} = \frac{\langle \sum_{j=k}^n \lambda_j\langle x,u_j\rangle u_j, x\rangle}{\|x\|^2}\\ &= \frac{\sum_{j=k}^n \lambda_j|\langle x,u_j\rangle|^2}{\|x\|^2} \ge \lambda_k \frac{\sum_{j=k}^n \lambda_j|\langle x,u_j\rangle|^2}{\|x\|^2}\\ &= \lambda_k.\end{aligned}$$

Note that we have used the fact that $\lambda_k \le \lambda_{k+1} \le \ldots \le \lambda_n$, and $\|x\|^2 = \sum_{j=k}^n |\langle x,u_j\rangle|^2$. Thus for any S_k^α,

$$\sup_{x\in S_k^\alpha\setminus\{0\}} \frac{x^\top Mx}{x^\top x} \ge \lambda_k.$$

Indeed for any $x \in S_k^\alpha$, we have

$$\sup_{x\in S_k^\alpha\setminus\{0\}} \frac{x^\top Mx}{x^\top x} = \sup_{x\in S_k^\alpha,\|x\|=1} x^\top Mx,$$

and since $\{x \in S_k^\alpha : \|x\| = 1\}$ is compact, the supremum is attained at some $x \in S_k^\alpha$ with $\|x\| = 1$. Thus we have

$$\max_{x\in S_k^\alpha\setminus\{0\}} \frac{x^\top Mx}{x^\top x} = \sup_{x\in S_k^\alpha\setminus\{0\}} \frac{x^\top Mx}{x^\top x} \ge \lambda_k.$$

On the other hand, consider a particular k-dimensional subspace $S_k^\alpha = \mathsf{span}(u_1,\dots,u_k)$, so that

$$\frac{x^\top Mx}{x^\top x} = \frac{\sum_{j=k}^n |\langle x,u_j\rangle|^2}{\|x\|^2} \le \lambda_k.$$

Choosing $x = u_k$, we see that the inequality is tight, so that we have

$$\max_{x \in S_k^\alpha \setminus \{0\}} \frac{x^\top M x}{x^\top x} = \lambda_k.$$

This also implies that the minimum of $\max_{x \in S_k^\alpha \setminus \{0\}} \frac{x^\top M x}{x^\top x}$ over all the $\alpha \in I_k$ is also attained, and we conclude that

$$\lambda_k = \min_{\alpha \in I_k} \max_{x \in S_k^\alpha \setminus \{0\}} \frac{x^\top M x}{x^\top x}$$

.

(ii) The proof of (ii) is similar to that of (i)—we will omit similar details. Choose again $W = \mathsf{span}\{u_1, \ldots, u_k\}$, so that $\dim(W) = k$. Then for any subspace S_{n-k+1}^α, we should have $\dim(S_k^\alpha \cap W) \geq 1$.

Next, choose any $x \in (S_{n-k+1}^\alpha \cap W) \setminus \{0\}$, such that $\frac{x^\top M x}{x^\top x} \leq \lambda_k$, and therefore

$$\min_{x \in S_{n-k+1}^\alpha \setminus \{0\}} \frac{x^\top M x}{x^\top x} \leq \lambda_k.$$

We again choose a particular $S_{n-k+1}^\alpha = \mathsf{span}(u_{k+1}, \ldots, u_n)$, so that

$$\min_{x \in S_{n-k+1}^\alpha \setminus \{0\}} \frac{x^\top M x}{x^\top x} = \lambda_k.$$

We can finally say that:

$$\max_{\alpha \in I_{n-k+1}} \min_{x \in S_{n-k+1}^\alpha \setminus \{0\}} \frac{x^\top M x}{x^\top x} = \lambda_k,$$

and we are done.

Thus,

$$\lambda_k \overset{(i)}{=} \min_{\alpha \in I_k} \max_{x \in S_k^\alpha \setminus \{0\}} \frac{x^\top M x}{x^\top x} \overset{(ii)}{=} \max_{\alpha \in I_{n-k+1}} \min_{x \in S_{n-k+1}^\alpha \setminus \{0\}} \frac{\nu^\top M \nu}{\nu^\top \nu},$$

as desired. ■

In the above proof, we considered Hermitian matrices, but the theorem is true for real symmetric matrices as well. The proof is similar; we only need to restrict the subspaces S_k^α to be real subspaces.

We will now discuss a few theorems and corollaries that can be considered as consequences of the Courant-Fischer theorem. We will start with the following theorem by Weyl, which allos us to obtain bounds on the kth largest eigenvalue of a matrix $M = A + B$:

Theorem 8 (Weyl)**.** *Let $A, B \in \mathrm{Mat}(n \times n; \mathbb{C})$ be both Hermitian, and let $\{\lambda_j(A)\}_{j=1}^n$, $\{\lambda_j(B)\}_{j=1}^n$ and $\{\lambda_j(M)\}_{j=1}^n$ be the eigenvalues of A, B and $M = A + B$, respectively, in increasing order. Then for each $1 \leq k \leq n$, we have*

$$\lambda_k(A) + \lambda_1(B) \leq \lambda_k(M) \leq \lambda_k(A) + \lambda_n(B).$$

By symmetry, we also have

$$\lambda_k(B) + \lambda_1(A) \leq \lambda_k(M) \leq \lambda_k(B) + \lambda_n(A).$$

Proof. From last time, we know that

$$\lambda_1(B) \leq \frac{x^\top B x}{x^\top x} \leq \lambda_n(B),$$

for $x \neq 0$. Then, by Courant-Fischer's theorem, we have for any $1 \leq k \leq n$,

$$\begin{aligned}
\lambda_k(A+B) &= \min_{\alpha \in I_k} \max_{x \in S_k^\alpha \setminus \{0\}} \frac{x^\top (A+B) x}{x^\top x} \\
&= \min_{\alpha \in I_k} \max_{x \in S_k^\alpha \setminus \{0\}} \frac{x^\top A x}{x^\top x} + \min_{\alpha \in I_k} \max_{x \in S_k^\alpha \setminus \{0\}} \frac{x^\top B x}{x^\top x} \\
&\geq \min_{\alpha \in I_k} \max_{x \in S_k^\alpha \setminus \{0\}} \frac{x^\top A x}{x^\top x} + \lambda_1(B) \\
&= \lambda_1(B) + \min_{\alpha \in I_k} \max_{x \in S_k^\alpha \setminus \{0\}} \frac{x^\top A x}{x^\top x} \\
&= \lambda_1(B) + \lambda_k(A).
\end{aligned}$$

The exact same argument can be applied for the other inequality to get $\lambda_k(A)+\lambda_n(B) \geq \lambda_k(A+B)$, and we are done. ■

The *monotonicity corollary* follows immediately from Weyl's theorem:

Corollary 9 (Monotonicity)**.** *Adopt all assumptions and notations from the above Weyl's theorem. Then for each $1 \leq k \leq n$, we have $\lambda_k(A) \leq \lambda_k(M)$.*

Proof. As B is PSD, $\lambda_j \geq 0$ for all j, and so $\lambda_k(A) + \lambda_1(B) \leq \lambda_k(M)$. Thus $\lambda_k(A) \leq \lambda_k(M)$. ■

The *interlacing corollary* is another application of Courant-Fischer. It offers a relationship between the eigenvalues of a Hermitian matrix A and those of its rank one perturbation. We will not prove it, but a statement is given below:

Corollary 10 (Interlacing)**.** *Adopt all assumptions and notations from the above Weyl's theorem. Then for each $1 \leq k \leq n$, we have*

$$\lambda_1(A) \leq \lambda_1(M) \leq \lambda_2(A) \leq \cdots \leq \lambda_k(A) \leq \lambda_k(M) \leq \lambda_{k+1}(A) \leq \cdots \leq \lambda_n(M) \leq \lambda_n(A).$$

Problems

1. Show that if $A \in \mathrm{Mat}(n \times n; \{0,1\})$ is the adjacency matrix of a graph, then

$$\lambda_1 = \max_{\nu \neq 0} \frac{\nu^\top A \nu}{\nu^\top \nu} \geq \overline{d},$$

   where $\overline{d}$ is the average degree of the graph and λ_1 is the largest eigenvalue of A.

2. Show that if $A \in \mathrm{Mat}(n \times n; \{0,1\})$ is the adjacency matrix of a graph G, then

$$\lambda_1 = \max_{\nu \neq 0} \frac{\nu^\top A \nu}{\nu^\top \nu} \leq M(G),$$

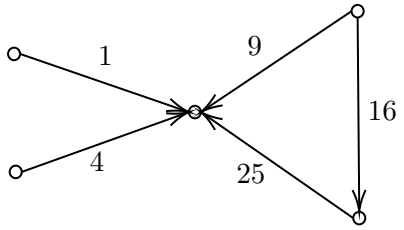
   where $M(G)$ is the maximum degree of all the vertices.

3 Graph Laplacian

Recall that the Laplacian matrix of a (weighted) graph G is defined as $L = D - A$, where A is its (weighted) adjacency matrix and $D = \sum_j (\mathrm{col}_i(A))_j$ is its degree matrix. We will now describe the *incidence matrix* $\mathscr{E}$ of a graph G. It is the matrix whose entities are given by

$$\mathscr{E}_{e,v} = \begin{cases} -\sqrt{A_e} & \text{if } v \text{ is the initial vertex of edge } e \\ +\sqrt{A_e} & \text{if } v \text{ is the terminal vertex of edge } e \\ 0 & \text{otherwise.} \end{cases}$$

As a quick example, consider the following graph:



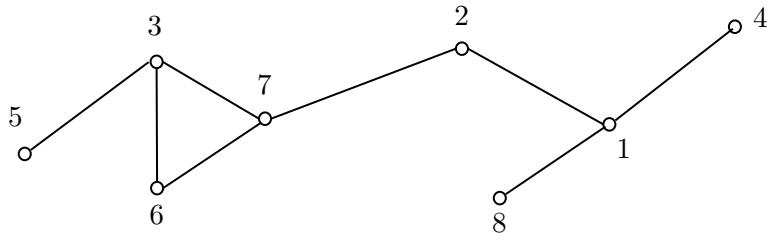
We can readily see that the incidence matrix of this graph is:

$$\mathscr{E} = \begin{bmatrix} -1 & 0 & 0 & 0 & 0 \\ 0 & -2 & 0 & 0 & 0 \\ 1 & 2 & 3 & 5 & 0 \\ 0 & 0 & -3 & 0 & -4 \\ 0 & 0 & 0 & -5 & 4 \end{bmatrix}.$$

This matrix[1] offers us an alternative definition for L: $L = \mathscr{E}\mathscr{E}^\top$.

Remark 11. *The definition $L = D - A$ gives $L_{ij} = D_{ij} - A_{ij} = D_{ii}\delta ij - A_{ij}$, where δ_{ij} is the Kronecker delta on the vertices. Recalling thar $D_{ii} = \sum_{j=1}^{n} A_{ij}$, we can see that $L_{ij} = \sum_{j=1}^{n} A_{ij}\delta ij - A_{ij} = \deg(v_i)\delta_{ij} - A_{ij}$.*

Now let's try to motivate why we call L a *Laplacian*. Consider the following graph, and assume you applied heat to vertex 1 and that all of the edges can conduct heat:



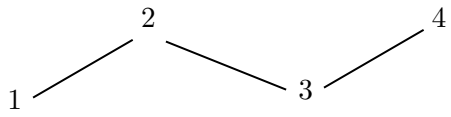
[1] We didn't discuss this, but we can reconstruct $G = (V, E)$ from $\mathscr{E}$ in time $\mathcal{O}(|V||E|)$, regardless of whether G is connected or not.

Newton's law of cooling states that the heat gets transferred like $dT/dt = -k(T - T_{\text{env}})$, where T is the temperature of the vertex and k is some constant. So, in the example above, we have $dT_1/dt = k(T_4 - T_1) + k(T_2 - T_1) + k(T_8 - T_1)$, and so on. We can speak about this more generally:

$$\begin{aligned}
\frac{dT_i}{dt} &= -k \sum_j A_{ji}(T_i - T_j) = -k \left(\sum_j A_{ij} T_i - \sum_j A_{ij} T_j \right) \\
&= -k \left(\sum_j A_{ij} \sum_\ell \delta_{\ell i} T_\ell - \sum_j A_{ij} T_j \right) = -k \left(\sum_\ell \sum_j A_{ij} \delta_{\ell i} T_\ell - \sum_\ell A_{i\ell} T_\ell \right) \\
&= -k \left(\sum_\ell \left[\sum_j A_{ij} \delta_{\ell i} - A_{i\ell} \right] T_\ell \right) = -k \sum_\ell L_{i\ell} T_\ell \\
&= -kLT
\end{aligned}$$

We can "think" of this expression as a discrete version of the heat equation $dT/dt = -k\nabla^2 T$, where L is the discrete version of the Laplace operator ∇^2—hence, the *Laplacian*. We can also think about how the Laplacian characterises the smoothness of a function on a graph via the same analogy.

We'll quickly discuss some examples. Consider the path graph



which has

$$L = \begin{bmatrix} 1 & -1 & 0 & 0 \\ -1 & 2 & -1 & 0 \\ 0 & -1 & 2 & -1 \\ 0 & 0 & -1 & 1 \end{bmatrix}$$

We can find that the eigenbasis a and spectrum b of L are

$$a = \begin{bmatrix} -0.5 & -0.6538 & 0.5 & -0.2706 \\ -0.5 & 0.2706 & -0.5 & 0.6538 \\ -0.5 & -0.2706 & 0.5 & -0.6538 \\ -0.5 & -0.6538 & 0.5 & 0.2706 \end{bmatrix}, \quad b = \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 0.5838 & 0 & 0 \\ 0 & 0 & 2 & 0 \\ 0 & 0 & 0 & 3.4142 \end{bmatrix}.$$

Observe that $a^\top a = \mathbb{I}$ and $a^\top L a = b$.

We consider now the complete graph K_n on n vertices. We can find (either by using $L = D - A$ or by using the incidence matrix) that

$$L_{K_n} = \begin{bmatrix} n-1 & -1 & \cdots & -1 \\ -1 & n-1 & \cdots & -1 \\ \vdots & \vdots & \ddots & \vdots \\ -1 & -1 & \cdots & n-1 \end{bmatrix} = n\mathbb{I} - \mathbf{1}\mathbf{1}^\top,$$

where $\mathbf{1}$ is the all-ones vector. We can immediately see that $L\mathbf{1} = 0$, and so 0 is an eigenvalue of L_{K_n} with corresponding eigenvector $\nu_1 = \mathbf{1}$. We can compute the spectrum of the graph by symmetry (left as an exercise). Regardless, the following bit from linear algebra can help us find other eigenvectors given ν_1:

Theorem 12. *For a symmetric matrix M with eigenvalues $\lambda_i \neq \lambda_j$, $\mathcal{E}_{\lambda_i} \perp \mathcal{E}_{\lambda_j}$. Furthermore, in the case of just two distinct eigenvalues λ_1 and λ_2, we have $\mathcal{E}_{\lambda_1} \oplus \mathcal{E}_{\lambda_2} = \mathcal{V} \subseteq \mathbb{R}^n$.*

Proof. The first part is straightforward: assume ν_i and ν_j are two eigenvectors of M corresponding to λ_i and λ_j. Then $\nu_j^\top M \nu_i = \lambda_i \nu_j^\top \nu_i = \lambda_j \nu_j^\top \nu_i$, and so $(\lambda_i - \lambda_j)\nu_j^\top \nu_i = 0$, which implies $\nu_j^\top \nu_i = 0$ (since $\lambda_i \neq \lambda_j$). Thus $\mathcal{E}_{\lambda_i} \perp \mathcal{E}_{\lambda_j}$.

We'll prove a general version of the second part. let $\nu_1, \ldots, \nu_n$ be n linearly independent eigenvectors; for each $i \in \{1, 2, \ldots, n\}$, let $\lambda_i \in \mathbb{R}$ be such that $A\nu_i = \lambda_i \nu_i$. We can assume without loss of generality that $\lambda_1 \leqslant \lambda_2 \leqslant \cdots \leqslant \lambda_n$. Suppose that $\lambda_1 = \lambda_2 = \cdots = \lambda_k$, for some $k \in \{1, 2, \ldots, n\}$. Then the eigenspace $\mathcal{E}_{\lambda_1}$ associated with λ_1 is $\bigoplus_{i=1}^{k} \mathbb{R}\nu_i$. And if $\lambda_{k+1} = \lambda_{k+2} = \cdots = \lambda_l$ for some $l \in \{k+1, \cdots, n\}$, $\mathcal{E}_{\lambda_{k+1}} = \bigoplus_{i=k+1}^{l} \mathbb{R}\nu_i$. And so on. So, if $k_1, \ldots, k_m \in \{1, 2, \ldots, n\}$ are such that $\lambda_{k_1} < \cdots < \lambda_{k_m}$ and that $\{\lambda_{k_1}, \ldots, \lambda_{k_m}\} = \{\lambda_1, \ldots, \lambda_n\}$, then

$$\mathbb{R}^n = \bigoplus_{j=1}^{m} \mathcal{E}_{\lambda_j},$$

and each partial sum is a subspace of $\mathbb{R}^n$. ■

This theorem is useful for constructing the eigenvectors of L given ν_1: each ν_ℓ is an eigenvector so long as it is orthogonal to L_{K_n} for some n. Thus, we can construct $\nu_2, \ldots, \nu_n$ by taking the $n-1$ vectors

$$\begin{bmatrix} 1 \\ -1 \\ 0 \\ \vdots \\ 0 \end{bmatrix}, \ldots, \begin{bmatrix} 1 \\ 0 \\ \vdots \\ 0 \\ -1 \end{bmatrix}.$$

We conclude our discussion for today with a theorem relating the eigenvalues of its L to the components of any graph G.

Theorem 13. *If the unweighted graph G has n vertices and λ is an eigenvalue of its Laplacian L_G, then $0 \leq \lambda \leq n$. Furthermore, the multiplicity of $\lambda = 0$ is equal to the number of connected components of G, and the multiplicity of $\lambda = n$ is one less than the number of connected components of $\overline{G}$[2].*

Proof. Suppose ν has $\|\nu\| = \sqrt{\langle \nu, \nu \rangle} = 1$ and $L_G \nu = \lambda \nu$. Then since $\nu \cdot \nu = 1$, we have

$$\begin{aligned} \lambda &= (\lambda\nu) \cdot \nu = (L\nu) \cdot \nu = (\mathscr{E}\mathscr{E}\nu) \cdot \nu \\ &= \nu^\top (\mathscr{E}^\top)^\top \mathscr{E}^\top \nu = (\mathscr{E}^\top \nu)^\top (\mathscr{E}^\top \nu) \\ &= \left\| \mathscr{E}^\top \nu \right\|^2 \geq 0, \end{aligned}$$

[2]This is the complement of G, specified by deleting all of G's edges

where $\mathscr{E}$ is the incidence matrix of G. Thus the spectrum of L_G is non-negative. Let $v_1, \ldots, v_k$ be the vertices of a connected component of G. Then the sum of the corresponding sums of $\mathscr{E}$ is 0 (verify this for yourself) and any $k-1$ of these sums are linearly independent. It follows that the nullity of $\mathscr{E}$ (and hence $L = \mathscr{E}\mathscr{E}^\top$) is equal to the number of components of G. For the second inequality, we use the fact that if G has n vertices, then $L_G + L_{\overline{G}} = L_{K_n} = n\mathbb{I} - \mathbf{1}\mathbf{1}^\top$. Consider now some $\nu^* = (1, \ldots, 1)$, so that $L_G\nu^* = L_{\overline{G}}\nu^* = L_{K_n}\nu^* = 0$. Suppose $L_G\nu = \lambda\nu$ for some $\nu \perp \nu^*$. Since the spectrum of K_n is $[0, n, \ldots, n]$, we find that $L_{K_n}\nu = n\nu$. Furthermore,

$$L_{\overline{G}}\nu = (L_{K_n} - L_G)\,\nu = n\nu - \lambda\nu = (n-\lambda)\nu.$$

Hence ν is an eigenvector for $L_{\overline{G}}$ with eigenvalue $n - \lambda$. Therefore $n - \lambda \geq 0 \implies n \geq \lambda$. So $0 \leq \lambda \leq n$.

By the calculation above, if $[\lambda_1, \ldots, \lambda_n]$ is the spectrum of L_G then $[n - \lambda_1, \ldots, n - \lambda_n]$ is the spectrum of $L_{\overline{G}}$. Thus the multiplicity of $\lambda = n$ in L_G is equal to the multiplicity of $n - \lambda = 0$ which indicates the number of connected components in $\overline{G}$. On the other hand, the multiplicity of $\lambda = 0$ is the same as the multiplicity of $n - \lambda = n$ which indicates the number of connected components in G. ■

Corollary 14. *If G has n vertices and $\lambda = n$ is an eigenvalue of L_G, then G is connected.*

Proof. If G was not connected then $\overline{G}$ would have more than one connected component and so $L_{\overline{G}}$ would be connected hence the ν satisfying $L_G\nu = n\nu$ and $\nu^* = (1, \ldots, 1)$ would be null to $L_{\overline{G}}$ and hence $\nu \perp \nu^*$, which contradicts the fact that ν^* is an eigenvector of $L_{\overline{G}}$. ■

Problems

1. Show that the Laplacian is always positive semidefinite.
2. Show that the nonzero eigenvalues of K_n are n (with algebraic multiplicity $n-1$) and 0 (with algebraic multiplicity 1).
3. Show that the Laplacian can be written as $L = \mathscr{E}\mathscr{E}^\top$.

References

[HJ12] R. Horn and C. Johnson. *Matrix Analysis*. Cambridge University Press, 2012.

[Fiedler72] M. Fiedler. Algebraic connectivity of graphs. *Czechoslovak Mathematical Journal*, 23(2):298–305, 1972.

[Chung97] F. Chung. *Spectral Graph Theory*. American Mathematical Society, 1997.

[Abreu06] N. M. M. de Abreu. *Old and new results on algebraic connectivity of graphs*. Federal University of Rio de Janeiro, 2006.

[AM85] W. Anderson and T. Morley. Eigenvalues of the Laplacian of a graph. *Linear and Multilinear Algebra*, 18(2):141–145, 1985.