CMSE 890-001: Spectral Graph Theory and Related Topics, MSU, Spring 2021

Lecture 04: The Complete Graph and Drawing Graphs

January 28, 2021

Lecturer: Matthew Hirn

8 The complete graph

Throughout the course, but especially in the first part of it, we will compute the eigenvalues (and eigenvectors) of particular graphs. We begin doing so in this section, and we start with the complete graph.

The *complete graph* on n vertices, often denoted by $K_n = (V, E)$, is defined as:

$$V = \{1, \ldots, n\},$$
$$E = \{(a, b) : a, b \in V \text{ and } a \neq b\}.$$

In other words, the complete graph contains every possible edge; see Figure 14 for a picture of K_5.

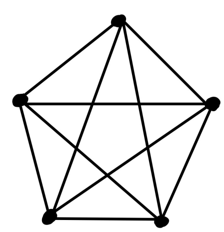

Figure 14: The complete graph on $n = 5$ vertices, K_5.

Theorem 7. *The graph Laplacian of K_n has eigenvalue 0 with multiplicity 1 and eigenvalue n with multiplicity $n - 1$.*

Proof. Since K_n is connected, we know from Theorem 5 that $\lambda_1 = 0$ and $\lambda_2 > 0$; therefore, the eigenvalue 0 has multiplicity 1, and we know its eigenvector is $\mathbf{1}$.

To compute the non-zero eigenvalues, let $\boldsymbol{\psi} \neq \mathbf{0}$ be any vector orthogonal to $\mathbf{1}$. Thus:

$$\langle \boldsymbol{\psi}, \mathbf{1} \rangle = \sum_{a \in V} \boldsymbol{\psi}(a) = 0 \quad \Longrightarrow \quad \boldsymbol{\psi}(a) = -\sum_{\substack{b \in V \\ b \neq a}} \boldsymbol{\psi}(b).$$

Let us now compute $\boldsymbol{L}\boldsymbol{\psi}(a)$ for an arbitrary vertex $a \in V$. Using (7) and the above equation, we get:

$$\boldsymbol{L}\boldsymbol{\psi}(a) = \sum_{\substack{b \in V \\ b \neq a}} (\boldsymbol{\psi}(a) - \boldsymbol{\psi}(b)) = (n-1)\boldsymbol{\psi}(a) - \sum_{\substack{b \in V \\ b \neq a}} \boldsymbol{\psi}(b) = n\boldsymbol{\psi}(a).$$

Since a was arbitrary we have $\boldsymbol{L}\boldsymbol{\psi} = n\boldsymbol{\psi}$ for any vector orthogonal to $\mathbf{1}$. In other words, every vector orthogonal $\mathbf{1}$ is an eigenvector of $\boldsymbol{L}$ with eigenvalue n, and so the eigenvalue n has multiplicity $n-1$. □

9 Drawing with Laplacian eigenvectors

In this section we give mathematical justification to embedding a connected graph into $\mathbb{R}^2$ using the first two non-trivial eigenvectors of its graph Laplacian. Recall the map was given by $a \mapsto (\boldsymbol{\psi}_2(a), \boldsymbol{\psi}_3(a)) \in \mathbb{R}^2$, and an example was given in Figure 12.

Let us begin with the simpler problem of embedding a graph into $\mathbb{R}$. Let $\boldsymbol{x} : V \to \mathbb{R}$ be a candidate function on the vertices that we are going to use to map G into $\mathbb{R}$ via $a \mapsto \boldsymbol{x}(a)$. Our modeling assumption is that we would like vertices that are neighbors to be close to one another in the embedding. So a natural candidate is:

$$\boldsymbol{x} = \arg\inf_{\boldsymbol{z} \in \mathbb{R}^n} \boldsymbol{z}^T \boldsymbol{L} \boldsymbol{z} = \sum_{(a,b) \in E} (\boldsymbol{z}(a) - \boldsymbol{z}(b))^2 \,.$$

However, there are some issues. The first is that there is nothing to prevent us from picking $\boldsymbol{x} = \mathbf{0}$, the all zeros vector, which maps every vertex to 0. We can remedy this by enforcing that $\boldsymbol{x}$ have unit norm, i.e.,

$$\boldsymbol{x} = \arg\inf_{\substack{\boldsymbol{z} \in \mathbb{R}^n \\ \|\boldsymbol{z}\| = 1}} \boldsymbol{z}^T \boldsymbol{L} \boldsymbol{z} \,.$$

This new version still does not work, though, as we can pick $\boldsymbol{x} = (1/\sqrt{n})\mathbf{1}$; that is the constant vector. This embedding still maps every vertex to the same point in $\mathbb{R}$, which is not a very interesting or useful embedding. We can fix this problem by adding the constraint that $\boldsymbol{x}$ be orthogonal to $\mathbf{1}$:

$$\boldsymbol{x} = \arg\inf_{\substack{\boldsymbol{z} \in \mathbb{R}^n \\ \|\boldsymbol{z}\| = 1 \\ \langle \boldsymbol{z}, \mathbf{1} \rangle = 0}} \boldsymbol{z}^T \boldsymbol{L} \boldsymbol{z} \,. \tag{8}$$

This final formulation will give us something interesting. In fact, we already know what it will give us! Using the Courant-Fischer Theorem (Theorem 3), we know that the value of the argument of (8) is λ_2 and that $\boldsymbol{x} = \boldsymbol{\psi}_2$ will be the solution.

But our original goal was to embed G into $\mathbb{R}^2$. To do so, we need two coordinate functions, $\boldsymbol{x} : V \to \mathbb{R}$ and $\boldsymbol{y} : V \to \mathbb{R}$, that will give us an embedding $a \mapsto (\boldsymbol{x}(a), \boldsymbol{y}(a)) \in \mathbb{R}^2$. Using again the modeling assumption that we want neighboring vertices to be close to one another, we seek to minimize the sum of the squares of the lengths between neighboring vertices in the embedding:

$$\boldsymbol{x}, \boldsymbol{y} = \arg\inf_{\boldsymbol{z}_1, \boldsymbol{z}_2 \in \mathbb{R}^n} \sum_{(a,b) \in E} \left\| \begin{pmatrix} \boldsymbol{z}_1(a) \\ \boldsymbol{z}_2(a) \end{pmatrix} - \begin{pmatrix} \boldsymbol{z}_1(b) \\ \boldsymbol{z}_2(b) \end{pmatrix} \right\|^2 \,.$$

We first notice that this minimization problem can also be written in terms of the graph Laplacian, as:

$$\sum_{(a,b)\in E} \left\| \begin{pmatrix} \boldsymbol{z}_1(a) \\ \boldsymbol{z}_2(a) \end{pmatrix} - \begin{pmatrix} \boldsymbol{z}_1(b) \\ \boldsymbol{z}_2(b) \end{pmatrix} \right\|^2 = \sum_{(a,b)\in E} (\boldsymbol{z}_1(a) - \boldsymbol{z}_1(b))^2 + (\boldsymbol{z}_2(a) - \boldsymbol{z}_2(b))^2 \,,$$
$$= \boldsymbol{z}_1^T \boldsymbol{L} \boldsymbol{z}_1 + \boldsymbol{z}_2^T \boldsymbol{L} \boldsymbol{z}_2 \,.$$

As before we need to impose constraints on $\boldsymbol{x}$ and $\boldsymbol{y}$ to avoid degenerate solutions. We we impose the unit norm constraint,

$$\|\boldsymbol{x}\| = \|\boldsymbol{y}\| = 1 \,,$$

and the orthogonality constraint with respect to the constant vector,

$$\langle \boldsymbol{x}, \mathbf{1} \rangle = \langle \boldsymbol{y}, \mathbf{1} \rangle = 0 \,.$$

However, since have two coordinates now, we have another type of degenerate solution in which $\boldsymbol{x} = \boldsymbol{y} = \boldsymbol{\psi}_2$. In order to avoid this solution, we impose that $\boldsymbol{x}$ be orthogonal to $\boldsymbol{y}$, i.e., $\langle \boldsymbol{x}, \boldsymbol{y} \rangle = 0$, which leads to the optimization problem:

$$\boldsymbol{x}, \boldsymbol{y} = \underset{\substack{\boldsymbol{z}_1, \boldsymbol{z}_2 \in \mathbb{R}^n \\ \|\boldsymbol{z}_1\| = \|\boldsymbol{z}_2\| = 1 \\ \langle \boldsymbol{z}_1, \mathbf{1} \rangle = \langle \boldsymbol{z}_2, \mathbf{1} \rangle = 0 \\ \langle \boldsymbol{z}_1, \boldsymbol{z}_2 \rangle = 0}}{\arg\inf} \sum_{(a,b)\in E} \left\| \begin{pmatrix} \boldsymbol{z}_1(a) \\ \boldsymbol{z}_2(a) \end{pmatrix} - \begin{pmatrix} \boldsymbol{z}_1(b) \\ \boldsymbol{z}_2(b) \end{pmatrix} \right\|^2 . \tag{9}$$

A natural candidate for the solution of (9) is $\boldsymbol{x} = \boldsymbol{\psi}_2$ and $\boldsymbol{y} = \boldsymbol{\psi}_3$.

More generally, suppose we want to embed G into $\mathbb{R}^k$ with k coordinate functions $\boldsymbol{x}_1, \ldots, \boldsymbol{x}_k$ that are orthonormal and are orthogonal to $\mathbf{1}$ and that minimize $\sum_{i=1}^k \boldsymbol{z}_i^T \boldsymbol{L} \boldsymbol{z}_i$, i.e.,

$$\boldsymbol{x}_1, \ldots, \boldsymbol{x}_k = \underset{\boldsymbol{z}_1, \ldots, \boldsymbol{z}_k}{\arg\inf} \sum_{i=1}^k \boldsymbol{z}_i^T \boldsymbol{L} \boldsymbol{z}_i \quad \text{subject to} \quad \langle \boldsymbol{z}_i, \boldsymbol{z}_j \rangle = \delta(i-j) \text{ and } \langle \boldsymbol{z}_i, \mathbf{1} \rangle = 0 \,. \tag{10}$$

As in the two-dimensional case, a natural candidate is to take $\boldsymbol{x}_i = \boldsymbol{\psi}_{i+1}$. Since $\boldsymbol{\psi}_i^T \boldsymbol{L} \boldsymbol{\psi}_i = \lambda_i$, we see that the value of the argument of (10) will be $\sum_{i=2}^{k+1} \lambda_i$. The following theorem says this is the best we can do.

Theorem 8. *Let $G = (V, E)$ be a graph and let $0 = \lambda_1 \leq \lambda_2 \leq \cdots \leq \lambda_n$ be the eigenvalues of its graph Laplacian $\boldsymbol{L}$ with associated orthonormal eigenvectors $\boldsymbol{\psi}_1, \ldots, \boldsymbol{\psi}_n$. Let $\boldsymbol{x}_1, \ldots, \boldsymbol{x}_k \in \mathbb{R}^n$ be orthonormal vectors that are all orthogonal to $\mathbf{1}$. Then*

$$\sum_{i=1}^k \boldsymbol{x}_i^T \boldsymbol{L} \boldsymbol{x}_i \geq \sum_{i=2}^{k+1} \lambda_i \,,$$

and furthermore, one achieves equality if and only if

$$\langle \boldsymbol{x}_i, \boldsymbol{\psi}_j \rangle = 0 \text{ for all } j \text{ such that } \lambda_j > \lambda_{k+1} \,.$$

Proof. It is a standard fact from Linear Algebra that a set of orthonormal vectors $\boldsymbol{x}_1, \dots, \boldsymbol{x}_k \in \mathbb{R}^n$ can be completed to form an orthonormal basis $\boldsymbol{x}_1, \dots, \boldsymbol{x}_n$. We now have two ONBs, $\boldsymbol{x}_1, \dots, \boldsymbol{x}_n$ and $\boldsymbol{\psi}_1, \dots, \boldsymbol{\psi}_n$. Using (6),

$$\sum_{i=1}^{n} |\langle \boldsymbol{x}_i, \boldsymbol{\psi}_j \rangle|^2 = 1 \quad \text{and} \quad \sum_{j=1}^{n} |\langle \boldsymbol{x}_i, \boldsymbol{\psi}_j \rangle|^2 = 1\,. \tag{11}$$

Also, let $\boldsymbol{\psi}_1$ be the constant vector.

Since $\langle \boldsymbol{x}_i, \boldsymbol{\psi}_1 \rangle = 0$ for $1 \le i \le k$, we have using Lemma 4 and (11):

$$\begin{aligned}
\forall\, 1 \le i \le k, \quad \boldsymbol{x}_i^T \boldsymbol{L} \boldsymbol{x}_i &= \sum_{j=2}^{n} \lambda_j |\langle \boldsymbol{x}_i, \boldsymbol{\psi}_j \rangle|^2 \\
&= \lambda_{k+1} + \sum_{j=2}^{n} (\lambda_j - \lambda_{k+1}) |\langle \boldsymbol{x}_i, \boldsymbol{\psi}_j \rangle|^2 \\
&= \lambda_{k+1} + \sum_{j=2}^{k+1} \underbrace{(\lambda_j - \lambda_{k+1})}_{\le 0} |\langle \boldsymbol{x}_i, \boldsymbol{\psi}_j \rangle|^2 + \sum_{j=k+2}^{n} \underbrace{(\lambda_j - \lambda_{k+1})}_{\ge 0} |\langle \boldsymbol{x}_i, \boldsymbol{\psi}_j \rangle|^2 \\
&\ge \lambda_{k+1} + \sum_{j=2}^{k+1} (\lambda_j - \lambda_{k+1}) |\langle \boldsymbol{x}_i, \boldsymbol{\psi}_j \rangle|^2\,.
\end{aligned}$$

Notice that the inequality will be an equality if and only if $\langle \boldsymbol{x}_i, \boldsymbol{\psi}_j \rangle = 0$ for all j such that $\lambda_j > \lambda_{k+1}$.

Now let us sum over $1 \le i \le k$ and use the previous calculation along with (11):

$$\begin{aligned}
\sum_{i=1}^{k} \boldsymbol{x}_i^T \boldsymbol{L} \boldsymbol{x}_i &\ge k\lambda_{k+1} + \sum_{j=2}^{k+1} (\lambda_j - \lambda_{k+1}) \sum_{i=1}^{k} |\langle \boldsymbol{x}_i, \boldsymbol{\psi}_j \rangle|^2 \\
&= k\lambda_{k+1} + \sum_{j=2}^{k+1} \underbrace{(\lambda_j - \lambda_{k+1})}_{\le 0} \underbrace{\left(1 - \sum_{i=k+1}^{n} |\langle \boldsymbol{x}_i, \boldsymbol{\psi}_j \rangle|^2 \right)}_{0 \le \cdots \le 1} \\
&\ge k\lambda_{k+1} + \sum_{j=2}^{k+1} (\lambda_j - \lambda_{k+1}) \qquad (12) \\
&= \sum_{j=2}^{k+1} \lambda_j\,.
\end{aligned}$$

The last step is to show that the final inequality, (12), is an equality if and only if $\langle \boldsymbol{x}_i, \boldsymbol{\psi}_j \rangle = 0$ for all j such that $\lambda_j > \lambda_{k+1}$. I leave that for you to verify on your own. $\square$

Theorem 8 shows that the eigenvector embedding of a graph G is optimal in the sense that we defined in this section. It thus gives a mathematical justification for this embedding method. However, one must be careful to remember that the embedding will only be good if the quantity we are minimizing makes sense for the particular graph you are studying. There are lots of other graph embedding/data visualization algorithms that minimize (or maximize) other quantities.

References

[1] Daniel A. Spielman. Spectral and algebraic graph theory. Book draft, available at: http://cs-www.cs.yale.edu/homes/spielman/sagt/, 2019.

[2] Michael Perlmutter, Feng Gao, Guy Wolf, and Matthew Hirn. Geometric scattering networks on compact Riemannian manifolds. In *Proceedings of The First Mathematical and Scientific Machine Learning Conference, Proceedings of Machine Learning Research*, volume 107, pages 570–604, 2020.