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

Lecture 11: Fiedler's Nodal Domain Theorem

February 23, 2021

Lecturer: Matthew Hirn

In the last lecture we partially proved the following result.

Theorem 32. *Let $G = (V, E, w)$ be a weighted path graph on n vertices. Let $\boldsymbol{L}$ be its graph Laplacian with eigenvalues $0 = \lambda_1 < \lambda_2 < \cdots < \lambda_n$, and let $\boldsymbol{\psi}_k$ be the k^{th} eigenvector of $\boldsymbol{L}$ with eigenvalue λ_k. If $\boldsymbol{\psi}_k(a) \neq 0$ for all $a \in V$, then there are exactly $k-1$ edges $(a-1, a) \in E$ for which $\boldsymbol{\psi}_k(a-1)\boldsymbol{\psi}_k(a) < 0$.*

We now want to complete the proof. Recall we defined the matrix $\boldsymbol{A}$ as

$$\boldsymbol{A} := \boldsymbol{\Psi}_k(\boldsymbol{L} - \lambda_k \boldsymbol{I})\boldsymbol{\Psi}_k \, ,$$

where $\boldsymbol{\Psi}_k$ is the $n \times n$ matrix with $\boldsymbol{\psi}_k$ down its diagonal and zeros elsewhere. To complete the proof of Theorem 32, we need to show that

$$\boldsymbol{A} = \sum_{a=2}^{n} w(a-1, a)\boldsymbol{\psi}_k(a-1)\boldsymbol{\psi}_k(a)\boldsymbol{L}_{G_{a-1,a}} \, ,$$

which was labeled as equation (26) in the previous lecture. Let us do that now.

Proof of equation (26). We will prove the alternate formula for $\boldsymbol{A}$ in two steps. First, we will show (26) holds for the off diagonal entries; then, we will show it holds for the diagonal entries. For the off diagonal entries we first note that

$$\boldsymbol{A} = \boldsymbol{\Psi}_k(\boldsymbol{L} - \lambda_k \boldsymbol{I})\boldsymbol{\Psi}_k = \boldsymbol{\Psi}_k \boldsymbol{L} \boldsymbol{\Psi}_k - \lambda_k \boldsymbol{\Psi}_k^2 \, .$$

Since $\boldsymbol{\Psi}_k$ is a diagonal matrix, the off diagonal entries of $\boldsymbol{A}$ are contained solely in $\boldsymbol{\Psi}_k \boldsymbol{L} \boldsymbol{\Psi}_k$. So, we will compute the entries of this matrix. First we compute:

$$\boldsymbol{\Psi}_k \boldsymbol{L}(u, v) = \sum_{c \in V} \boldsymbol{\Psi}_k(u, c)\boldsymbol{L}(c, v) = \boldsymbol{\Psi}_k(u, u)\boldsymbol{L}(u, v) = \boldsymbol{\psi}_k(u)\boldsymbol{L}(u, v) \, .$$

Then we have:

$$\boldsymbol{\Psi}_k \boldsymbol{L} \boldsymbol{\Psi}_k(u, v) = \sum_{c \in V} \boldsymbol{\Psi}_k \boldsymbol{L}(u, c)\boldsymbol{\Psi}_k(c, v) = \boldsymbol{\Psi}_k \boldsymbol{L}(u, v)\boldsymbol{\Psi}_k(v, v) = \boldsymbol{\psi}_k(u)\boldsymbol{\psi}_k(v)\boldsymbol{L}(u, v) \, .$$

Therefore, since $\boldsymbol{L}(u, v) = -w(u, v)$ for $(u, v) \in E$ and $\boldsymbol{L}(u, v) = 0$ for $(u, v) \notin E$, we have,

$$\boldsymbol{A}(u, v) = \boldsymbol{\Psi}_k \boldsymbol{L} \boldsymbol{\Psi}_k(u, v) = \begin{cases} -w(u, v)\boldsymbol{\psi}_k(u)\boldsymbol{\psi}_k(v) & (u, v) \in E \\ 0 & (u, v) \notin E \end{cases} \, .$$

On the other hand, for $u \neq v$,

$$\boldsymbol{L}_{G_{a,b}}(u,v) = \begin{cases} -1 & (a,b) = (u,v) \\ 0 & (a,b) \neq (u,v) \end{cases} ,$$

and thus for $u \neq v$,

$$\begin{aligned} \sum_{(a,b)\in E} w(a,b)\boldsymbol{\psi}_k(a)\boldsymbol{\psi}_k(b)\boldsymbol{L}_{G_{a,b}}(u,v) &= \begin{cases} -w(u,v)\boldsymbol{\psi}_k(u)\boldsymbol{\psi}_k(v) & (u,v) \in E \\ 0 & (u,v) \notin E \end{cases} \\ &= \boldsymbol{A}(u,v)\,. \end{aligned}$$

For the diagonal entries, we will show the row sums of $\boldsymbol{A}$ are the same as the row sums of the right hand side of (26). Since we know the off diagonal entries are the same, this will show the diagonal entries are the same too. We compute the row sums of $\boldsymbol{A}$ via:

$$\boldsymbol{A}\mathbf{1} = \boldsymbol{\Psi}_k(\boldsymbol{L} - \lambda_k \boldsymbol{I})\boldsymbol{\Psi}_k \mathbf{1} = \boldsymbol{\Psi}_k(\boldsymbol{L} - \lambda_k \boldsymbol{I})\boldsymbol{\psi}_k = \boldsymbol{\Psi}_k(\boldsymbol{L}\boldsymbol{\psi}_k - \lambda_k\boldsymbol{\psi}_k) = \boldsymbol{\Psi}_k(\lambda_k\boldsymbol{\psi}_k - \lambda_k\boldsymbol{\psi}_k) = \mathbf{0}\,.$$

On the other hand, $\boldsymbol{L}_{G_{a,b}}\mathbf{1} = \mathbf{0}$ for all $(a,b) \in E$, so the row sums of the right hand side of (26) are zeros too. The proof is complete. □

Remark 18. In Theorem 32, we assumed λ_k has multiplicity one for each $1 \leq k \leq n$, but in fact one can prove this must be true, and thus this assumption can be removed.

Remark 19. We also assume that $\boldsymbol{\psi}_k(a) \neq 0$ for all $a \in V$ in Theorem 32. We can also relax this assumption and show that the eigenvector changes sign $k-1$ times over the path.

21 Nodal domains

We are now going to prove Fiedler's Nodal Domain Theorem, which will generalize the result on the weighted path graph to any connected, weighted graph! In doing so, our analysis and interpretation of what the "frequency" of an eigenvector means will be complete.

Let $G = (V, E, w)$ be a weighted, connected graph, and let us first define a nodal domain. For $S \subseteq V$, recall that $G(S)$ is the subgraph induced by S; see equation (23) for the definition. For each eigenvector $\boldsymbol{\psi}_k$ of the graph Laplacian of G, $1 \leq k \leq n$, define the set $W_k \subseteq S$ as

$$W_k := \{a \in V : \boldsymbol{\psi}_k(a) \geq 0\}\,.$$

The subgraph induced by W_k, $G(W_k)$, is the k^{th} *nodal domain* for the graph G. Note, it will depend on the choice of $\boldsymbol{\psi}_k$, but we will be able to say something about the number of connected components of $G(W_k)$ irrespective the choice of $\boldsymbol{\psi}_k$.

To gain some intuition, let us go back to the weighted path graph. We know its eigenvector $\boldsymbol{\psi}_k$ will change sign exactly $k-1$ times. Thus $\boldsymbol{\psi}_2$ will change sign once, and $G(W_2)$ will have one connected component. $\boldsymbol{\psi}_3$ will change sign twice, and depending on the sign of $\boldsymbol{\psi}_3$, $G(W_3)$ will have ether one or two connected components. Similarly, $G(W_4)$ will have two connected components, and $G(W_5)$ will have two or three connected components. In other words, the number of connected components of $G(W_k)$ for the path graph will increase with the eigenvalue index. One may also wish to look back at Figures 9, 10, 22, and 23.

We see from the path graph example that measuring the number of connected components is a good proxy for the frequency of the eigenvector. Indeed, the more connected components $G(W_k)$ has, the more $\boldsymbol{\psi}_k$ must change its sign over the edges of G. For general weighted graphs, we will take the number of connected components of $G(W_k)$ as our measure of the frequency of $\boldsymbol{\psi}_k$. Fiedler's Nodal Domain Theorem will show that the number of these connected components, and hence the frequency of $\boldsymbol{\psi}_k$, is upper bounded by the index of the eigenvector. Its proof will leverage two of our previous big results, namely the Cauchy Eigenvalue Interlacing Theorem (Theorem 21) and the Perron-Frobenius Theorem (Theorem 26 and Corollary 28).

Theorem 33 (Fiedler's Nodal Domain Theorem)**.** *Let $G = (V, E, w)$ be a connected graph. Then for $k \geq 2$, $G(W_k)$ has at most $k - 1$ connected components.*

Proof. Since $\boldsymbol{\psi}_1 = \mathbf{1}$ and $\langle \boldsymbol{\psi}_1, \boldsymbol{\psi}_k \rangle = 0$ for all $k \geq 2$, we know that $\boldsymbol{\psi}_k$ must have positive and negative entries and $W_k \neq \emptyset$.

Suppose that $G(W_k)$ has t connected components. Let us order the vertices of G so that the vertices in one connected component of $G(W_k)$ come first, the vertices in another connected component of $G(W_k)$ come second, and so on so forth, and the vertices $a \in V$ for which $\boldsymbol{\psi}_k(a) < 0$ come last. Then the graph Laplacian of G can be written as:

$$\boldsymbol{L} = \begin{pmatrix} \boldsymbol{B}_1 & \mathbf{0} & \mathbf{0} & \cdots & \mathbf{0} & \boldsymbol{C}_1 \\ \mathbf{0} & \boldsymbol{B}_2 & \mathbf{0} & \cdots & \mathbf{0} & \boldsymbol{C}_2 \\ \mathbf{0} & \mathbf{0} & \boldsymbol{B}_3 & \cdots & \mathbf{0} & \boldsymbol{C}_3 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & \cdots & \boldsymbol{B}_t & \boldsymbol{C}_t \\ \boldsymbol{C}_1^T & \boldsymbol{C}_2^T & \boldsymbol{C}_3^T & \cdots & \boldsymbol{C}_t^T & \boldsymbol{A} \end{pmatrix},$$

where:

- $\boldsymbol{B}_i$ encodes the vertices and edges in the i^{th} connected component of $G(W_k)$. Note that each $\boldsymbol{B}_i$ is symmetric with non-positive entries on the off-diagonal, and hence the Perron-Frobenius Theorem (specifically Corollary 28) applies to each one.

- $\boldsymbol{C}_i$ encodes the edges between the vertices of the i^{th} connected component of of $G(W_k)$ and those vertices $a \in V$ for which $\boldsymbol{\psi}_k(a) < 0$. Note that every entry of $\boldsymbol{C}_i$ is non-positive and at least one entry of $\boldsymbol{C}_i$ must be non-zero (thus negative) since otherwise G would not be connected.

- $\boldsymbol{A}$ encodes the edges between the vertices $a \in V$ for which $\boldsymbol{\psi}_k(a) < 0$.

We can also decompose $\boldsymbol{\psi}_k$ using the same blocks of vertices:

$$\boldsymbol{\psi}_k = \begin{pmatrix} \boldsymbol{x}_1 \\ \boldsymbol{x}_2 \\ \boldsymbol{x}_3 \\ \vdots \\ \boldsymbol{x}_t \\ \boldsymbol{y} \end{pmatrix},$$

where

- $\boldsymbol{x}_i$ is a vector of length the number of vertices in the i^{th} connected component of $G(W_k)$, and the entries of $\boldsymbol{x}_i$ are non-negative.
- $\boldsymbol{y}$ is a vector of length the number of vertices $a \in V$ for which $\boldsymbol{\psi}_k(a) < 0$, and every entry of $\boldsymbol{y}$ is negative.

We will now show that the smallest eigenvalue of each $\boldsymbol{B}_i$ is strictly less than λ_k, the eigenvalue of $\boldsymbol{\psi}_k$. First note that we have:

$$\begin{pmatrix} \boldsymbol{B}_1 & \boldsymbol{0} & \boldsymbol{0} & \cdots & \boldsymbol{0} & \boldsymbol{C}_1 \\ \boldsymbol{0} & \boldsymbol{B}_2 & \boldsymbol{0} & \cdots & \boldsymbol{0} & \boldsymbol{C}_2 \\ \boldsymbol{0} & \boldsymbol{0} & \boldsymbol{B}_3 & \cdots & \boldsymbol{0} & \boldsymbol{C}_3 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ \boldsymbol{0} & \boldsymbol{0} & \boldsymbol{0} & \cdots & \boldsymbol{B}_t & \boldsymbol{C}_t \\ \boldsymbol{C}_1^T & \boldsymbol{C}_2^T & \boldsymbol{C}_3^T & \cdots & \boldsymbol{C}_t^T & \boldsymbol{A} \end{pmatrix} \begin{pmatrix} \boldsymbol{x}_1 \\ \boldsymbol{x}_2 \\ \boldsymbol{x}_3 \\ \vdots \\ \boldsymbol{x}_t \\ \boldsymbol{y} \end{pmatrix} = \begin{pmatrix} \boldsymbol{B}_1\boldsymbol{x}_1 + \boldsymbol{C}_1\boldsymbol{y} \\ \boldsymbol{B}_2\boldsymbol{x}_2 + \boldsymbol{C}_2\boldsymbol{y} \\ \boldsymbol{B}_3\boldsymbol{x}_3 + \boldsymbol{C}_3\boldsymbol{y} \\ \vdots \\ \boldsymbol{B}_t\boldsymbol{x}_t + \boldsymbol{C}_t\boldsymbol{y} \\ \sum_{i=1}^T \boldsymbol{C}_i^t\boldsymbol{x}_i + \boldsymbol{A}\boldsymbol{y} \end{pmatrix} = \lambda_k \begin{pmatrix} \boldsymbol{x}_1 \\ \boldsymbol{x}_2 \\ \boldsymbol{x}_3 \\ \vdots \\ \boldsymbol{x}_t \\ \boldsymbol{y} \end{pmatrix},$$

and in particular,

$$\boldsymbol{B}_i\boldsymbol{x}_i + \boldsymbol{C}_i\boldsymbol{y} = \lambda_k\boldsymbol{x}_i\,, \quad 1 \le i \le t\,. \tag{27}$$

Now, every entry of $\boldsymbol{C}_i$ is non-positive, at least one entry of $\boldsymbol{C}_i$ is negative, and every entry of $\boldsymbol{y}$ is negative; therefore, every entry of $\boldsymbol{C}_i\boldsymbol{y}$ is non-negative and at least one entry of $\boldsymbol{C}_i\boldsymbol{y}$ is positive. It follows that $\boldsymbol{x}_i \neq \boldsymbol{0}$, since if this were the case we would have $\boldsymbol{C}_i\boldsymbol{y} = \boldsymbol{0}$, which contradicts the fact that at least one entry of $\boldsymbol{C}_i\boldsymbol{y}$ is positive.

Furthermore, using the fact that every entry of $\boldsymbol{x}_i$ is non-negative, plus our conclusions regarding $\boldsymbol{C}_i\boldsymbol{y}$, as well as (27), we have

$$\lambda_k\boldsymbol{x}_i(a) - \boldsymbol{C}_i\boldsymbol{y}(a) \le \lambda_k\boldsymbol{x}_i(a) \quad \Longrightarrow \quad \boldsymbol{B}_i\boldsymbol{x}_i(a) \le \lambda_k\boldsymbol{x}_i(a)\,, \quad \forall\, a\,.$$

We deduce that

$$\boldsymbol{x}_i^T\boldsymbol{B}_i\boldsymbol{x}_i \le \lambda_k\boldsymbol{x}_i^T\boldsymbol{x}_i\,. \tag{28}$$

Now, by the Courant-Fischer Theorem (Theorem 3) and (28), we know that

$$\lambda_1(\boldsymbol{B}_i) = \min_{\boldsymbol{x}\neq\boldsymbol{0}} \frac{\boldsymbol{x}^T\boldsymbol{B}_i\boldsymbol{x}}{\boldsymbol{x}^T\boldsymbol{x}} \le \frac{\boldsymbol{x}_i^T\boldsymbol{B}_i\boldsymbol{x}_i}{\boldsymbol{x}_i^T\boldsymbol{x}_i} \le \lambda_k\,.$$

We have two cases:

1. Suppose $\boldsymbol{x}_i$ has at least one zero entry and $\boldsymbol{x}_i^T \boldsymbol{B}_i \boldsymbol{x}_i = \lambda_k \boldsymbol{x}_i^T \boldsymbol{x}_i$ (if it is a strict inequality we done). Applying the Perron-Frobenius Theorem (Corollary 28), though, we conclude that λ_k cannot be $\lambda_1(\boldsymbol{B}_i)$ since the eigenvector associated to $\lambda_1(\boldsymbol{B}_i)$ cannot have any zero entries. Thus $\lambda_1(\boldsymbol{B}_i) < \lambda_k$.

2. Suppose $\boldsymbol{x}_i$ has all positive entries. Since $\boldsymbol{C}_i \boldsymbol{y}$ is non-negative with at least one positive entry, we have that $\boldsymbol{x}_i^T \boldsymbol{C}_i \boldsymbol{y} > 0$. Therefore, using (27),
$$\boldsymbol{x}_i^T \boldsymbol{B}_i \boldsymbol{x}_i = \lambda_k \boldsymbol{x}_i^T \boldsymbol{x}_i - \boldsymbol{x}_i^T \boldsymbol{C}_i \boldsymbol{y} < \lambda_k \boldsymbol{x}_i^T \boldsymbol{x}_i \,,$$
and thus $\lambda_1(\boldsymbol{B}_i) < \lambda_k$.

We have proved that
$$\lambda_1(\boldsymbol{B}_i) < \lambda_k \,, \quad 1 \le i \le t \,.$$
It follows that the matrix
$$\boldsymbol{B} = \begin{pmatrix} \boldsymbol{B}_1 & \boldsymbol{0} & \cdots & \boldsymbol{0} \\ \boldsymbol{0} & \boldsymbol{B}_2 & \cdots & \boldsymbol{0} \\ \vdots & \vdots & \ddots & \vdots \\ \boldsymbol{0} & \boldsymbol{0} & \cdots & \boldsymbol{B}_t \end{pmatrix} ,$$
has at least t eigenvalues less than λ_k. We now want to apply the Cauchy Interlacing Theorem to $\boldsymbol{L}$ and the sub-matrix $\boldsymbol{B}$. However, $\boldsymbol{B}$ is not a principal sub-matrix of $\boldsymbol{L}$ so we need to verify that we can.

First, let $\widetilde{\boldsymbol{A}}$ be an $n \times n$ matrix and let $\widetilde{\boldsymbol{B}}$ be an $(n-1) \times (n-1)$ principal sub-matrix of $\widetilde{\boldsymbol{A}}$. This time, let us order the eigenvalues of $\widetilde{\boldsymbol{A}}$ and $\widetilde{\boldsymbol{B}}$ in increasing order:
$$\lambda_1(\widetilde{\boldsymbol{A}}) \le \lambda_2(\widetilde{\boldsymbol{A}}) \le \cdots \le \lambda_n(\widetilde{\boldsymbol{A}}) \quad \text{and} \quad \lambda_1(\widetilde{\boldsymbol{B}}) \le \lambda_2(\widetilde{\boldsymbol{B}}) \le \cdots \le \lambda_{n-1}(\widetilde{\boldsymbol{B}}) \,.$$
Applying the Cauchy Interlacing Theorem (Theorem 21), we know that
$$\lambda_1(\widetilde{\boldsymbol{A}}) \le \lambda_1(\widetilde{\boldsymbol{B}}) \le \lambda_2(\widetilde{\boldsymbol{A}}) \le \lambda_2(\widetilde{\boldsymbol{B}}) \le \cdots \le \lambda_{n-1}(\widetilde{\boldsymbol{A}}) \le \lambda_{n-1}(\widetilde{\boldsymbol{B}}) \le \lambda_n(\widetilde{\boldsymbol{A}}) \,.$$
In fact, even if $\widetilde{\boldsymbol{B}}$ is an $(n-\ell) \times (n-\ell)$ sub-matrix of $\widetilde{\boldsymbol{A}}$ obtained by removing the same rows and columns from $\widetilde{\boldsymbol{A}}$, we can prove that
$$\lambda_i(\widetilde{\boldsymbol{A}}) \le \lambda_i(\widetilde{\boldsymbol{B}}) \,, \quad 1 \le i \le n - \ell \,. \tag{29}$$
The proof is nearly identical to the proof of Corollary 22, and in particular uses the Cauchy Interlacing Theorem.

Let us apply (29) with $\widetilde{\boldsymbol{A}} = \boldsymbol{L}$ and $\widetilde{\boldsymbol{B}} = \boldsymbol{B}$. We conclude that
$$\lambda_i(\boldsymbol{L}) \le \lambda_i(\boldsymbol{B}) < \lambda_k(\boldsymbol{L}) \,, \quad 1 \le i \le t \,.$$
Since it is only possible for $\lambda_i(\boldsymbol{L}) < \lambda_k(\boldsymbol{L})$ for $1 \le i \le k-1$, we conclude that $t \le k-1$. $\square$

Remark 20. The theorem is sharp. Indeed, consider the star graph $S_5 = (V, E)$ on five vertices such that $V = \{1, \ldots, 5\}$ with $a = 1$ being the hub vertex. One can verify that $\boldsymbol{\psi}$ defined as $\boldsymbol{\psi}(1) = 0$, $\boldsymbol{\psi}(2) = \boldsymbol{\psi}(3) = \boldsymbol{\psi}(4) = 1$, and $\boldsymbol{\psi}(5) = -3$ is an eigenvector with eigenvalue $\lambda_2 = 1$ (use Lemma 10). On the other hand, if we consider the set $\{a \in V : \boldsymbol{\psi}(a) > 0\}$, we get three connected components.

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.

[3] David I. Shuman, Sunil K. Narang, Pascal Frossard, Antonio Ortega, and Pierre Vandergheynst. The emerging field of signal processing on graphs: Extending high-dimensional data analysis to networks and other irregular domains. *IEEE Signal Processing Magazine*, 30(3):83–98, 2013.