Spectral Graph Theory	Lecture 3
Laplacian and the Adjacency Matrices	
Daniel A. Spielman	September 9, 2009

3.1 Weighted Path Graphs

We will now prove the following theorem of Fiedler:

Theorem 3.1.1. *Let P be a weighted path graph on n vertices, let L_P have eigenvalues $0 = \lambda_1 < \lambda_2 \cdots \leq \lambda_n$, and let $\boldsymbol{v}_k$ be an eigenvector of λ_k. Then, $\boldsymbol{v}_k$ changes sign $k-1$ times.*

The main ingredient in the proof will be Sylvester's law of intertia, which I will first recall.

Theorem 3.1.2 (Sylvester's Law of Intertia)**.** *Let A be any symmetric matrix and let B be any non-singular matrix. Then, the matrix BAB^T has the same number of positive, negative and zero eigenvalues as A.*

I will not prove this in class today. Instead, I'll give a proof next lecture. I include the following proof in the lecture notes, which is different from the proof I will give next lecture.

While you might not know Sylvester's law of intertia, we will prove it by using two things that you do know. First, recall that if A is a square matrix and B is a non-singular matrix, then

$$BAB^{-1}$$

has the same eigenvalues as A. This holds because

$$A\boldsymbol{v} = \lambda \boldsymbol{v} \quad \text{if and only if} \quad \left(BAB^{-1}\right)(B\boldsymbol{v}) = \lambda(B\boldsymbol{v}).$$

We also know that the rank of A and BAB are the same.

Proof. We first recall that every non-singular matrix B can be written $B = QR$, where Q is an orthonormal matrix Q and R is upper-triangular matrix R with positive diagonals[1] We will use a slight variation of this fact, writing $B = RQ$. Now, since $Q^T = Q^{-1}$, QAQ^T has exactly the same eigenvalues as A. Let R_t be the matrix $t * R + (1-t)I$, and consider the family of matrices

$$M_t = R_t Q A Q^T R_t^T,$$

as t goes from 0 to 1. At $t = 0$, the matrix has the same eigenvalues as A. At $t = 1$, we get $B^T AB$. All of these matrices are symmetric, so they all have Real eigenvalues. As the eigenvalues of a

[1] This is called the QR-factorization. It follows from Gram-Schmidt orthonormalization.

symmetric matrix are continuous functions of the matrix coefficients[2], if the number of positive, negative or zero eigenvalues of $B^T AB$ differs from that of A, then there must be some t for which M_t has more zero eigenvalues than does A. But, as the matrices R_t are upper-triangular with positive diagonal entries, they are all non-singular. So, the rank of M_t must equal the rank of A, which means this cannot happen. $\square$

Fiedler's Theorem will follow from an analysis of the eigenvalues of tri-diagonal matrices with zero row-sums. These may be viewed as Laplacians of weighted path graphs in which some edges are allowed to have negative weights.

Proposition 3.1.3. *Let M be a symmetric matrix such that*

$$M\mathbf{1} = \mathbf{0}.$$

Then,

$$M = \sum_{i \neq j} -M(i,j) L_{(i,j)}. \tag{3.1}$$

Proof. The expression on the right-hand side of (3.1) clearly agrees with M in all off-diagonal entries. Given all the off-diagonal entries, the diagonal entries are determined by the constraint $M\mathbf{1} = \mathbf{0}$, which the right-hand side of (3.1) satisfies as well because $L_{(i,j)}\mathbf{1} = \mathbf{0}$ for all $i \neq j$. $\square$

Lemma 3.1.4. *Let M be a symmetric tri-diagonal matrix with $2p$ positive off-diagonal entries such that*

$$M\mathbf{1} = \mathbf{0}. \tag{3.2}$$

Then, M has p negative eigenvalues.

Proof. By Proposition 3.1.3, we may write

$$M = \sum_{i=1}^{n-1} -M(i,i+1) L_{(i,i+1)}.$$

Thus,

$$\boldsymbol{x}^T M \boldsymbol{x} = \sum_{i=1}^{n-1} -M(i,i+1)(\boldsymbol{x}(i) - \boldsymbol{x}(i+1))^2.$$

We now perform a change of variables that will diagonalize the matrix M. Let $\boldsymbol{\delta}(1) = \boldsymbol{x}(1)$, and $\boldsymbol{\delta}(i) = \boldsymbol{x}(i) - \boldsymbol{x}(i-1)$ for $i \geq 2$. So,

$$\boldsymbol{x}(i) = \boldsymbol{\delta}(1) + \boldsymbol{\delta}(2) + \cdots + \boldsymbol{\delta}(i).$$

This change of variables is realized by the lower-triangular matrix L which has 1's on and below the diagonal:

$$\boldsymbol{x} = L\boldsymbol{\delta}.$$

[2] You might not know this yet, but you will see why in the next few lectures

By Sylvester's law of intertia, we know that

$$L^T M L$$

has the same number of positive, negative, and zero eigenvalues as M. On the other hand,

$$\boldsymbol{\delta}^T L^T M L \boldsymbol{\delta} = \sum_{i=2}^{n} -M(i, i+1)\delta(i)^2,$$

so this matrix clearly has one zero eigenvalue, and as many negative eigenvalues as there are negative $M(i, i+1)$. $\square$

Proof of Theorem 3.1.1. We will just consider the case in which $\boldsymbol{v}_k$ has no zero entries. The proof for the general case may be obtained by splitting the graph by removing the vertices with zero entries. For simplicity, we will also assume that λ_k has multiplicity 1.

We wish to show that the number of i for which $\boldsymbol{v}_k(i)\boldsymbol{v}_k(i+1) < 0$ equals $k-1$.

Let V_k denote the diagonal matrix with v_k on the diagonal, and let λ_k be the corresponding k-th eigenvalue of L_P. Consider the matrix

$$M = V_k^T (L_P - \lambda_k I) V_k.$$

The inner matrix obviously has one zero eigenvalue and $k-1$ negative eigenvalues. So, by Sylvester's law of intertia, M has $k-1$ negative eigenvalues, one zero eigenvalue, and $n-k$ positive eigenvalues. The matrix M satisfies the conditions of Lemma 3.1.4 because

$$M\mathbf{1} = V_k^T (L - \lambda_k I) V_k \mathbf{1} = V_k^T (L - \lambda_k I)\boldsymbol{v}_k = V_k^T \mathbf{0} = \mathbf{0}.$$

Morevover,

$$M(i, i+1) = -w(i, i+1)\boldsymbol{v}_k(i)\boldsymbol{v}_k(i+1)$$

is positive precisely when $\boldsymbol{v}_k(i)\boldsymbol{v}_k(i+1) < 0$, Thus, by Lemma 3.1.4 there are exactly $k-1$ such indices i. $\square$

3.2 Adjacency Matrices

It is sometimes convenient to consider adjacency matrices of graphs. The adjacency matrices provide somewhat natural operators. In the unweighted case,

$$(A_G \boldsymbol{v})(i) = \sum_{j:(i,j)\in E} \boldsymbol{v}(j).$$

As the notation $\sum_{j:(i,j)\in E}$ is somewhat cumbersome, we will instead write $\sum_{j\sim i}$ when E is clear from context.

Adjacency matrices can behave somewhat differently from Laplacian matrices. But, let's begin by seeing how they can be similar.

For now, let $G = (V, E)$ be an unweighted graph. The degree of a vertex $i \in V$, usually written $d(i)$, is the number of edges attached to it[3]. The graph G is said to be d-regular if every vertex has degree d. In this case,

$$L_G = D_G - A_G = dI - A_G.$$

So, there is a very clean relationship between the spectra of A_G and L_G. Let

$$0 = \lambda_1 \leq \lambda_2 \leq \cdots \leq \lambda_n$$

be the eigenvalues of L_G, with corresponding eigenvectors $\boldsymbol{v}_1, \ldots, \boldsymbol{v}_n$. Similarly, let

$$\alpha_1 \geq \alpha_2 \geq \cdots \geq \alpha_n$$

be the eigenvalues of A_G. We have

$$\alpha_i = d - \lambda_i,$$

and

$$A_G \boldsymbol{v}_i = (dI - L_G)\boldsymbol{v}_i = (d - \lambda_i)\boldsymbol{v}_i.$$

In particular, we see that the constant vectors are eigenvectors of eigenvalue d.

This happens only if the graph is d-regular. Let d_{max} denote the maximum degree of any vertex in G.

Lemma 3.2.1.

$$\alpha_1 \leq d_{max}.$$

Proof. Let $\boldsymbol{v}_1$ be an eigenvector of eigenvalue α_1. Let j be the vertex on which it takes its maximum value, so $\boldsymbol{v}_1(j) \geq \boldsymbol{v}_1(i)$ for all i, and assume without loss of generality that $\boldsymbol{v}_1(j) \neq 0$. We have

$$\alpha_1 = \frac{(A\boldsymbol{v}_1)(j)}{\boldsymbol{v}_1(j)} = \frac{\sum_{i \sim j} \boldsymbol{v}_i(i)}{\boldsymbol{v}_1(j)} = \sum_{i \sim j} \frac{\boldsymbol{v}_i(i)}{\boldsymbol{v}_1(j)} \leq \sum_{i \sim j} 1 \leq d(j) \leq d_{max}. \tag{3.3}$$

□

Lemma 3.2.2. *If G is connected and $\alpha_1 = d_{max}$, then G is d_{max}-regular.*

Proof. If we have equality in (3.3), then it must be the case that $d(j) = d_{max}$ and $\boldsymbol{v}_1(i) = \boldsymbol{v}_1(j)$ for all $i \sim j$. Thus, we may apply the same argument to every neighbor of j. As the graph is connected, we may keep applying this argument to neighbors of vertices to which it has already been applied to show that $\boldsymbol{v}_1(k) = \boldsymbol{v}_1(j)$ and $d(k) = d_{max}$ for all $k \in V$. □

[3]In the weighted case, all this goes through if we sum the weights of the attached edges

3.3 The Perron-Frobenius Theorem

We now address what happens when G is not regular. We already know that $\alpha_1 < \delta_{max}$, but what does $\boldsymbol{v}_1$ look like? Actually, I do not know any convenient description of $\boldsymbol{v}_1$. But, we can show that it is a positive vector. This is a consequence of the Perron-Frobenius Theorem, of which we will prove a special case.

Before proving it, we will see some of its consequences.

Lemma 3.3.1. *Let $G = (V, E, w)$ be a connected weighted graph. Assume there is a positive vector $\boldsymbol{v}$ such that*

$$A_G \boldsymbol{v} = \alpha \boldsymbol{v}.$$

Then,

(a) There is a non-negative, non-singular diagonal matrix S such that

$$S^{-1} A_G S \mathbf{1} = \alpha \mathbf{1},$$

(b) For every eigenvalue α_i of A_G,

$$|\alpha_i| \leq \alpha.$$

(c) The eigenvalue α has multiplicity 1.

Point (a) is an example of a matrix-scaling theorem. In this case, it says that we can scale the matrix so that all of its row-sums are the same. This sort of statement is useful in many applications, and so it will receive much of our attention.

Proof. Let S be the diagonal matrix with $\boldsymbol{v}$ on the diagonal. Part (a) follows from

$$S^{-1} A_G S \mathbf{1} = S^{-1} A_G \boldsymbol{v} = S^{-1} \alpha \boldsymbol{v} = \alpha S^{-1} \boldsymbol{v} = \alpha \mathbf{1}.$$

To prove part (b), let

$$B = S^{-1} A_G S,$$

and recall that B has the same eigenvalues of A_G. Let

$$B \boldsymbol{u} = \beta \boldsymbol{u}.$$

Let j be the index of the largest entry of $\boldsymbol{u}$, in absolute value. We then have

$$\beta = \frac{(B\boldsymbol{u})(j)}{\boldsymbol{u}(j)} = \sum_i B(j,i) \frac{\boldsymbol{u}(i)}{\boldsymbol{u}(j)}.$$

So,

$$|\beta| = \left| \sum_i B(j,i) \frac{\boldsymbol{u}(i)}{\boldsymbol{u}(j)} \right| \leq \sum_i B(j,i) \left| \frac{\boldsymbol{u}(i)}{\boldsymbol{u}(j)} \right| \leq \sum_i B(j,i) = \alpha.$$

To prove part (c), we similarly note that

$$\beta = \frac{(B\boldsymbol{u})(j)}{\boldsymbol{u}(j)} = \sum_i B(j,i)\frac{\boldsymbol{u}(i)}{\boldsymbol{u}(j)} \leq \sum_i B(j,i) = \alpha,$$

and that we can only have equality if

$$\frac{\boldsymbol{u}(i)}{\boldsymbol{u}(j)} = 1$$

for all i for which $B(j,i)$ is non-zero. As in the proof of Lemma 3.2.2, we may exploit the connectivity of the graph to prove that $\boldsymbol{u}$ must be the constant vector. □

We will now prove that the vector $\boldsymbol{v}$ exists. Our proof will go through matrix scaling. We exploit the fact that if $S = \mathtt{diag}(\boldsymbol{v})$, by which I mean that S is the diagonal matrix with $\boldsymbol{v}$ on the diagonal, then

$$S^{-1}A_G S\mathbf{1} = \alpha\mathbf{1} \quad \text{if and only if} \quad A_G\boldsymbol{v} = \alpha\boldsymbol{v}.$$

Theorem 3.3.2. *Let $G = (V,E)$ be a connected graph, and let A be a non-negative matrix such that $A(i,j) > 0$ for all $(i,j) \in E$. Then, there exists a positive vector $\boldsymbol{v}$ and an $\alpha > 0$ such that*

$$S^{-1}AS\mathbf{1} = \alpha\mathbf{1},$$

where $S = \mathtt{diag}(\boldsymbol{v})$.

We will actually prove this theorem in a special case. I will leave the derivation of the theorem from the special case to the first problem set.

Lemma 3.3.3. *Let A be a matrix such that $A(i,j) > 0$ for all i and j. Then, there exists a positive vector $\boldsymbol{v}$ and an $\alpha > 0$ such that*

$$S^{-1}AS\mathbf{1} = \alpha\mathbf{1},$$

where $S = \mathtt{diag}(\boldsymbol{v})$.

We will prove Lemma 3.3.3 by providing an algorithm for computing $\boldsymbol{v}$. As a first attempt, let $\boldsymbol{s} = A\mathbf{1}$ be the vector of row-sums in A, and let $S = \mathtt{diag}(\boldsymbol{s})$. All of the row-sums in the matrix $S^{-1}A$ are the same, which would help us except that we need to multiply by S on the right. Let's do it anyway. For a matrix A, define

$$f(A) = S^{-1}AS, \quad \text{where } S = \mathtt{diag}(A\mathbf{1}).$$

On the problem set, we will show that by iteratively applying f, one can balance all the row sums of A. In the meantime, we define

$$\phi(A) = \max(A\mathbf{1}) - \min(A\mathbf{1}).$$

Lemma 3.3.4. *Let A be a matrix such that $A(i,j) > 0$ for all i and j. Then,*

$$\phi(f(A)) < \phi(A).$$

Proof. We will show that this operation decreases the gap between the largest and smallest row-sum, provided that it is non-zero. Let $\boldsymbol{s} = A\mathbf{1}$. Let $\boldsymbol{t} = S^{-1}A_G S\mathbf{1}$, so $\phi(f(A)) = \max(\boldsymbol{t}) - \min(\boldsymbol{t})$. We will prove that $\max(\boldsymbol{t}) \leq \max(\boldsymbol{s})$. In fact, for every i For every i, we have

$$\boldsymbol{t}(i) = \frac{1}{\boldsymbol{s}(i)} \sum_j A(i,j)\boldsymbol{s}(j) \leq \frac{1}{\boldsymbol{s}(i)} \sum_j A(i,j) \max(\boldsymbol{s}) = \max(\boldsymbol{s}) \frac{1}{\boldsymbol{s}(i)} \sum_j A(i,j) = \max(\boldsymbol{s}),$$

so $\max(\boldsymbol{t}) \leq \max(\boldsymbol{s})$. We may similarly show that $\boldsymbol{t}(i) \geq \min(\boldsymbol{s})$. So, we know that $\phi(f(A)) \leq \phi(A)$. To obtain a strict inequality, let k be an index for which $\boldsymbol{s}(k)$ is minimized, and assume that $\boldsymbol{s}(k) < \max(\boldsymbol{s})$. Being more careful with the derivation, we find

$$\begin{aligned}
\boldsymbol{t}(i) &= \frac{1}{\boldsymbol{s}(i)} \sum_j A(i,j)\boldsymbol{s}(j) \\
&= \frac{1}{\boldsymbol{s}(i)} A(i,k)\boldsymbol{s}(k) \frac{1}{\boldsymbol{s}(i)} \sum_{j \neq k} A(i,j)\boldsymbol{s}(j) \\
&< \frac{1}{\boldsymbol{s}(i)} A(i,k) \max(\boldsymbol{s}) \frac{1}{\boldsymbol{s}(i)} \sum_{j \neq k} A(i,j) \max(\boldsymbol{s}) \\
&= \max(\boldsymbol{s}) \frac{1}{\boldsymbol{s}(i)} \sum_j A(i,j) \\
&= \max(\boldsymbol{s}).
\end{aligned}$$

$\square$

Proof of Lemma 3.3.3. Let us abuse notation by defining

$$\phi(\boldsymbol{v}) = \phi\left((\texttt{diag}(\boldsymbol{v}))^{-1} \ A \ (\texttt{diag}(\boldsymbol{v})) \right).$$

We have shown that for every vector $\boldsymbol{v}$ for which $\phi(\boldsymbol{v}) > 0$, there exists a vector $\boldsymbol{w}$ for which $\phi(\boldsymbol{w}) < \phi(\boldsymbol{v})$. We can either prove that the limit goes to zero by analysis (as we will do in a moment), or concretely (as we will do on the homework).

To prove that the matrices can actually be made uniform, we use a little analysis. We would like to let $\boldsymbol{v}$ be a vector at which $\phi(\boldsymbol{v})$ is minimized. But, we must first show that such a vector exists. To simplify this task, note that $\phi(\boldsymbol{v}) = \phi(c\boldsymbol{v})$ for every $c > 0$. So, we can now just consider vectors $\boldsymbol{v}$ beloning to the set

$$U_0 \stackrel{\text{def}}{=} \left\{ \boldsymbol{v} : \boldsymbol{v}^T \mathbf{1} = 1 \text{ and } \boldsymbol{v} > 0 \right\}.$$

It is easy to show that if some entry of $\boldsymbol{v}$ approaches zero, then $\phi(\boldsymbol{v})$ goes to infinity. So, there is some $\epsilon > 0$ for which we can restrict our attention to

$$U_\epsilon \stackrel{\text{def}}{=} \left\{ \boldsymbol{v} : \boldsymbol{v}^T \mathbf{1} = 1 \text{ and } \boldsymbol{v} \geq \epsilon \right\}.$$

The set U_ϵ is closed and compact, and so every function on U_ϵ achieves its infimum. We have shown that for every $\boldsymbol{v}$ for which $\phi(\boldsymbol{v}) > 0$ there exists a $\boldsymbol{w}$ for which $\phi(\boldsymbol{w}) < \phi(\boldsymbol{v})$, so this infimum must be zero, and there must be some vector $\boldsymbol{v} \in U_\epsilon$ for which $\phi(\boldsymbol{v}) = 0$. $\square$