

Lecture 5: Eigenvalues of A_G and the Laplacian

We continue our study of the adjacency matrix, and show that the multiplicity of the eigenvalue d is equal to the number of connected components. We then introduce the Laplacian of a graph. We will see how the second smallest eigenvalue of the Laplacian is related to the expansion of the graph.

Disclaimer: These lecture notes are informal in nature and are not thoroughly proofread. In case you find a serious error, please send email to the instructor pointing it out.

Spectrum of the Adjacency Matrix

Recall the definition of the adjacency matrix A_G of the a graph G . Once again, throughout this lecture, we will be dealing with graphs that are regular, i.e., all vertices have degree d .

We saw last time that any eigenvalue λ of A_G satisfies $|\lambda| \leq d$. Suppose we order the eigenvalues $\lambda_1 \leq \lambda_2 \leq \dots \leq \lambda_n$. Then we saw that $\lambda_n = d$. Since all the eigenvalues have magnitude $\leq d$, we have $\lambda_1 \geq -d$.

Exercise. For a connected graph G , $\lambda_1(A_G) = -d$ if and only if G is bipartite.

Today, we show a simple yet elegant connection between the eigenvalues and the number of connected components.

Theorem 1. *The multiplicity of the eigenvalue d is equal to the number of connected components of G .*

Before proving the theorem in general, let us work out a simpler case: suppose we know that the graph has precisely one connected component, i.e., the graph is connected. We saw that $\lambda_n = d$ (the eigenvector is the n dimensional vector with all entries being 1, which we shall denote by $\mathbf{1}_n$). The theorem states that in this case, $\lambda_{n-1} < d$. This is equivalent to saying that for any non-zero vector v orthogonal to $\mathbf{1}_n$, we cannot have $Av = dv$.

Let us start with any vector such that $Av = dv$. Let i^* be the index for which $|v_{i^*}|$ is maximized. Now, the equality $Av = dv$ implies that for every i ,

$$\sum_j A_{ij} v_j = dv_i.$$

In particular, this is true for $i = i^*$. Thus

$$\sum_j A_{i^*j} v_j = dv_{i^*}.$$

Now, the LHS has at most d nonzero terms (because the degree is d , precisely d of the A_{i^*j} terms are 1 and the rest are 0). Further, each of the terms is $\leq v_{i^*}$ in magnitude (by assumption, i^* maximizes $|v_i|$). Thus the only way the equality can hold is if we have d non-zero terms, and each term is precisely v_{i^*} .

This implies that for all j that are neighbors of i^* , we must have $v_j = v_{i^*}$. Now, we can apply the exact same argument with j instead of i^* . We would obtain that for all neighbors k of j , we have $v_k = v_j = v_{i^*}$. Now since the graph is connected, we can proceed this way and conclude that for *all* vertices i , we must have $v_i = v_{i^*}$!

Thus v has all its coordinates equal, i.e., it is parallel to $\mathbf{1}_n$ (and hence cannot be orthogonal unless it is zero). This establishes the simpler case. Let us now get to the general case.

Proof of Theorem 1. Suppose G has k connected components, say V_1, V_2, \dots, V_k . We need to show that the A_G has precisely k orthogonal eigenvectors of value d .

We show this by first exhibiting k orthogonal eigenvectors (which shows the multiplicity is $\geq k$), and then showing that any vector v satisfying $A_G v = dv$ is spanned by the vectors exhibited (which shows the multiplicity is $\leq k$).

The first is easy. Let $\mathbf{u}^{(i)}$ be a vector whose j th entry is 1 if $j \in V_i$ and 0 otherwise (i.e., it is the indicator vector for V_i). Then for every $i \leq k$, it is easy to see that

$$A_G \mathbf{u}^{(i)} = d \mathbf{u}^{(i)}.$$

Also, since the V_i form a partition of the vertex set, we have $\langle \mathbf{u}^{(i)}, \mathbf{u}^{(j)} \rangle = 0$ for all $i \neq j$. Thus there are at least k orthogonal eigenvectors of eigenvalue d .

Now consider any v such that $A_G \mathbf{u} = d \mathbf{u}$. We can now use exactly the same reasoning we used in the case we had a single connected component, to conclude that in any connected component V_i , the values of $\mathbf{u}_j : j \in V_i$ for j are all equal! (Formally, we can look at the $j^* \in V_i$ that has the largest magnitude of \mathbf{u}_j , and argue similarly).

Thus when restricted to connected component V_i , \mathbf{u} is a scaling of $\mathbf{u}^{(i)}$. This implies that we can write $\mathbf{u} = \sum_i \alpha_i \mathbf{u}^{(i)}$, for some constants α_i . This implies that any vector with $A_G \mathbf{u} = d \mathbf{u}$ is a linear combination of the $\mathbf{u}^{(i)}$, implying that the multiplicity is at most k .

This completes the proof. □

Note. For graphs that are not regular, such a clean connection between the multiplicity of the top eigenvalue and connected components does not hold. However, many of the other results we will show will turn out to hold.

Let us now go back to the question of finding sparse cuts in a graph. We recall a little bit of notation from the previous lecture.

Partitioning objectives

Recall the definitions of the sparsity $\sigma(S)$ and the expansion $\phi(S)$, for a set $S \subseteq V$:

$$\sigma(S) := \frac{E(S, \bar{S})}{|S||\bar{S}|} \quad ; \quad \phi(S) := \frac{E(S, \bar{S})}{d|S|}.$$

From the definitions, it is easy to see that for any S of size $\leq n/2$, we have

$$\frac{n}{2d} \sigma(S) \leq \phi(S) \leq \frac{n}{d} \sigma(S).$$

In the last lecture, we also saw a *relaxation* for the problem of finding $\min_{S \subseteq V} \sigma(S)$,

$$\frac{1}{2n} \cdot \min_{x \in \mathbb{R}^n, \sum_i x_i = 0} \frac{\sum_{i \sim j} (x_i - x_j)^2}{\sum_i x_i^2},$$

which we said can be formulated as an eigenvalue problem. Let us now see how.

The Graph Laplacian. The Laplacian of a graph, denoted L_G , is the matrix $dI - A_G$.^a The key property of the Laplacian matrix is that for any vector x ,

$$x^T L_G x = \sum_{ij \in E} (x_i - x_j)^2. \quad (1)$$

^aFor graphs that are not regular, the Laplacian is defined to be $D - A_G$, where D is a diagonal matrix whose i th entry is the degree of the i th vertex.

To see Eq. (1), simply expand the RHS: $\sum_{ij \in E} x_i^2 + x_j^2 - x_i x_j - x_j x_i$. This is equal to $\sum_i d x_i^2 - x^T A_G x$, because every x_i^2 appears precisely d times. We can now write $\sum_i d x_i^2$ as $x^T (dI)x$, and thus Eq. (1) follows.

Eigenvalues of the Laplacian. By definition, the eigenvalues of the Laplacian are related to the eigenvalues of A_G in a simple way. To see this, note that any eigenvector \mathbf{u} of A_G with $A_G \mathbf{u} = \lambda \mathbf{u}$ is also an eigenvector of $dI - A_G$ – in fact $(dI - A_G) \mathbf{u} = (d - \lambda) \mathbf{u}$. Thus if the eigenvalues of A_G are $\lambda_1, \lambda_2, \dots, \lambda_n$, then the eigenvalues of L_G are $d - \lambda_1, d - \lambda_2, \dots, d - \lambda_n$.

The eigenvalues of A_G lie in the range $[-d, d]$. Thus the eigenvalues of L_G lie in $[0, 2d]$. Theorem 1 is equivalent to saying that the multiplicity of the eigenvalue 0 of L_G is the number of connected components of G . Also, for a connected graph, the eigenvector of L_G corresponding to the eigenvalue 0 is $\mathbf{1}_n$.

Let us write down what the *second smallest* eigenvalue of L_G is, for a connected graph G . From the characterization of eigenvalues we saw last class,

$$\lambda_2(L_G) = \min_{x \perp \mathbf{1}_n} \frac{x^T L_G x}{x^T x}.$$

The condition $x \perp \mathbf{1}_n$ is exactly the same as $\sum_i x_i = 0$, and using Eq. (1), we have

$$\lambda_2(L_G) = \min_{\sum_i x_i = 0} \frac{\sum_{ij \in E} (x_i - x_j)^2}{\sum_i x_i^2}.$$

Note that this is precisely the relaxation for $\min_S \sigma(S)$ (without the $(1/2n)$ factor)! Thus, we can relate the sparsest cut value to $\lambda_2(L_G)$ as

$$\min_S \sigma_S \leq \frac{1}{2n} \lambda_2(L_G).$$

Using the relation between $\min_S \sigma(S)$ and $\min_{|S| \leq n/2} \phi(S)$ (which we denoted by $\Phi(G)$ in the previous lecture), we have

$$\Phi(G) \geq \frac{1}{4d} \lambda_2(L_G).$$

This means that by finding the value $\lambda_2(L_G)$ (which is a quantity we can efficiently compute) we obtain a *lower bound* for the minimum expansion of a cut in G . In particular, if $\lambda_2(L_G)$ happens to be large, it means that every cut in G has many edges going across! (which, as we will see, is an important property

in graphs). However, this inequality does not rule out the possibility that $\lambda_2(L_G)$ is always *tiny*. We would ideally like to say that $\lambda_2(L_G)/d$ is *not much smaller* than $\Phi(G)$.

This is precisely what the so-called Cheeger's inequality talks about. In the next lecture we will show the following:

Theorem 2. *Let G be a d -regular graph. Then we have*

$$\Phi(G) \leq \sqrt{\frac{2\lambda_2(L_G)}{d}}.$$

Note that for graph that is not connected, $\lambda_2(L_G)$ is zero. In fact, if V_1 is a connected component of size $\leq n/2$, we have $\phi(V_1) = 0$, thus the inequality above indeed holds in this case. We can also view Cheeger's inequality in general as a *robust* form of the above statement. (*Robust* in the following sense: we know it holds when $\lambda_2 = 0$; does it hold when λ_2 is nearly zero?)

Non-regular graphs

For graphs that are not regular, the right matrix to look at is $A'_G := D^{-1/2}A_GD^{-1/2}$. (Here $D^{-1/2}$ is simply the diagonal matrix whose (i, i) th entry is $(\deg(i))^{-1/2}$ – we assume there are no isolated vertices, so none of the degrees is zero) This matrix is sometimes called the *normalized adjacency matrix* of a graph. Note that it also symmetric. Now consider the vector \mathbf{u} , whose i th entry is $\deg(i)^{1/2}$. Clearly, we have

$$D^{-1/2}\mathbf{u} = \mathbf{1}_n \implies D^{-1/2}A_GD^{-1/2}\mathbf{u} = D^{-1/2}A_G\mathbf{1}_n = \mathbf{u}.$$

The last equality is because $A_G\mathbf{1}_n$ is a vector whose i th entry is $\deg(i)$. Thus \mathbf{u} is an eigenvector with eigenvalue 1. It turns out $\lambda_{\max}(A'_G) = 1$. This is not entirely trivial. From the characterization of λ_{\max} , we have

$$\lambda_{\max}(A'_G) = \max_x \frac{x^T A'_G x}{x^T x} = \max_x \frac{\sum_{ij \in E} \frac{2x_i x_j}{\sqrt{\deg(i)\deg(j)}}}{\sum_i x_i^2}.$$

(The factor 2 is because the sum includes ij and ji .) Now, since $2ab \leq (a^2 + b^2)$ for any real numbers a, b , we have

$$\sum_{ij \in E} \frac{2x_i x_j}{\sqrt{\deg(i)\deg(j)}} \leq \sum_{ij \in E} \frac{x_i^2}{\deg(i)} + \frac{x_j^2}{\deg(j)} = \sum_i x_i^2.$$

The last inequality holds because $\frac{x_i^2}{\deg(i)}$ appears precisely $\deg(i)$ times. Thus $\lambda_{\max}(A'_G) \leq 1$.

An analog of Theorem 1 can be proved for the matrix A'_G . For general graphs, we can define the Normalized Laplacian as $L'_G := I - A'_G$. Its eigenvalues are also $1 -$ eigenvalues of A'_G , and lie in $(0, 2)$. It turns out that Cheeger's inequality also holds in terms of the second smallest eigenvalue of L'_G (without the factor d in the denominator).