Advanced Algorithms 2012A
Lecture 6 – Spectral graph theory*

Robert Krauthgamer

1 Basic spectral graph theory

Today we will see how combinatorial properties of the graph are manifested by eigenvalues and eigenvectors of matrices related to the graph.

1.1 Adjacency and Laplacian matrices

Let $G = (V, E)$ be an undirected graph, with edge weights $w_e \geq 0$, where $w_{ij} = 0$ effectively means that $ij \notin E$. As usual, it is illustrative to think of the unit-weight case, and in fact even regular graphs. The analog of the degree of vertex i is defined as $d_i = \sum_{j:ij\in E} w_i$, and it is useful to put these values in a diagonal matrix $D = \mathrm{diag}(\vec{d})$.

The graph can be described by its *adjacency matrix* $A = A_G$ given by:

$$A_{ij} = \begin{cases} w_i & \text{if } ij \in E, \\ 0 & \text{otherwise.} \end{cases}$$

It is often more convenient to work with the graph's *Laplacian matrix* $L = L_G$ given by:

$$L_{ij} = \begin{cases} -w_{ij} & \text{if } ij \in E, \\ d_i & \text{if } i = j, \\ 0 & \text{otherwise.} \end{cases}$$

Fact 1. $L = D - A$.

1.2 Recall (eigenvalues and eigenvectors)

Let M be a square $n \times n$ matrix whose entries are real values. Then $\lambda \in \mathbb{R}$ is an eigenvalue associated with nonzero eigenvector $\vec{x} \in \mathbb{R}^n$ if $M\vec{x} = \lambda\vec{x}$. Note that scaling $\vec{x}$ preserves this condition.

*These notes summarize the material covered in class, usually skipping proofs, details, examples and so forth, and possibly adding some remarks, or pointers. The exercises are for self-practice and need not be handed in. In the interest of brevity, most references and credits were omitted.

The eigenvalues are exactly the roots of characteristic polynomial $\det(A - \lambda I) = 0$, hence A has at most n eigenvalues, possibly with multiplicities. Since A (and similarly L) is symmetric, it has n pairs (eigenvalue λ ,eigenvector v), all with real values, such that the n eigenvectors are orthogonal meaning $\langle x, y\rangle = x^T y = \sum_i x_i y = 0$.

Fact 2. If G is r-regular then $D = rI$. Thus, $A\vec{v} = \lambda\vec{v}$ iff $L\vec{v} = (r - \lambda)v$, meaning that A and L have the same eigenvectors (and the eigenvalues are "reflected").

1.3 Recall (from variational characterization):

The eigenvalues of a symmetric matrix M can be found by minimizing/maximizing the Rayleigh quotient:

$$\lambda_{\max}(M) = \max_x \frac{x^T M x}{x^T x}; \qquad \lambda_{\min}(M) = \min_x \frac{x^T M x}{x^T x}.$$

Exer: Prove that $\lambda_{\max}(A)$ is between the average degree $\frac{1}{n}\sum_i d_i$ and the maximum degree $\max_i d_i$.

Fact 3. For every $\vec{x} \in R^V$,

$$x^T L x = \sum_{ij\in E} w_{ij}(x_i - x_j)^2.$$

Proof: Write L as summation of $|E|$ matrices, each corresponding to one edge and is "effectively" a 2×2 matrix $\left(\begin{smallmatrix} w_{ij} & -w_{ij} \\ -w_{ij} & w_{ij} \end{smallmatrix}\right)$, which contributes $w_{ij}(x_i - x_j)$.

Exer: Prove that we can write $L = B^T B$ where $B \in \mathbb{R}^{E\times V}$ is a (signed and weighted) incidence matrix. Verify that $x^T L x = \|Bx\|^2$ and use it to give a different proof for Fact 3.

Fact 4. Denote the eigenvalues of L by $\lambda_1 \le \cdots \le \lambda_n$. Then $\lambda_1 = 0$ (in particular, L is PSD).

Proof: $\lambda_1 \ge 0$ follows from Fact 3. Plugging in the all-ones vector, we further get $\lambda_1 \le 0$.

1.4 Graph connectivity and λ_2

It turns out that λ_2 represents the connectivity of G, and is thus called the algebraic connectivity.

Lemma 5. Denote the eigenvalues of L by $\lambda_1 \le \cdots \le \lambda_n$. Then G is disconnected iff $\lambda_2 = 0$.

The proof was seen in class. One direction follows by using the vectors $x = 1_S$ and $y = 1_{\bar{S}}$, or a suitable linear combination of them (which is orthogonal to all ones vector). For the other direction, take an eigenvector that is orthogonal to the all ones vector and letting $S \subset V$ be all coordinates of the same (say maximum) value.

Exer: Prove that the multiplicity of eigenvalue 0 equals the number of connected components in G.

Exer: The analogous claim for A would be that G is disconnected iff the two largest eigenvalues of A are equal. Is it true?

2 Cheeger's inequalities

As we will now, Lemma 5 above has an approximate version: if λ_2 is close to 0 then the graph is "almost" disconnected. The connectivity will be in terms of a variant of edge-expansion/sparse-cut usually called conductance (the names are sometimes interchanged).

We will do a version that does not depend on the maximum degree $d_{\max} = \max_{i \in V} d_i$.

2.1 Conductance and sparsity

Let us extend the weights w and d to sets by defining $w(S_1, S_2) = \sum_{ij \in E \cap (S_1 \times S_2)} w_{ij}$ for $S_1 \cap S_2 = \emptyset$ and $d(S) = \sum_{i \in S} d_i$.

The *sparsity* of a set $S \subset V$ is defined as

$$\mathrm{sp}_G(S) := \tfrac{w(S,\bar{S})}{d(S)d(\bar{S})/d(V)},$$

and the sparsity of a graph is

$$\mathrm{sp}(G) := \min_{S \subset V} \mathrm{sp}_G(S).$$

Exer: Prove that $\mathrm{sp}(G)$ is an instance of sparse-cut from last week.

We can define the *conductance* of a set $S \subset V$ to be

$$\phi_G(S) := \tfrac{w(S,\bar{S})}{\min\{d(S),d(\bar{S})\}}, \quad \text{thus} \quad \phi_G(S) \le \mathrm{sp}_G(S) \le 2\phi_G(S).$$

and similarly

$$\phi(G) := \min_{S \subset V} \phi_G(S), \quad \text{thus} \quad \phi(G) \le \mathrm{sp}(G) \le 2\phi(G).$$

Notice that both $\phi_G(S) = \phi_G(\bar{S})$ and similarly for sparsity, i.e., both have symmetry between S and $\bar{S}$. It is thus useful to think of S as the "smaller" one according to $d(.)$, and then $d(\bar{S})/d(V) \in [1/2, 1]$.

Interpretation: $\phi_G(S)$ measures what fraction of the edges incident to S actually leave S (i.e., go out to $\bar{S}$).

Example: Suppose G is a 2d-grid of size $\sqrt{n} \times \sqrt{n}$. Let $S_j \subset V$ contain the j leftmost columns, thus $|S_j| = j\sqrt{n}$. To compute $\phi(G)$ we need to consider all subsets of V, but for the sake of example let us consider here only the subsets S_j (without proving that one of these sets gives the minimum). Observe that $w(S_j, \bar{S}_j) = \sqrt{n}$. Since almost all vertices have degree 4 (except for the boundary), $d(S_j) \approx 4|S_j|$ and since we want assume $d(S_j) \le d(\bar{S}_j)$ we are constrained to $j \le \sqrt{n}/2$. Then

$$\phi(G) = \min_{j \le \sqrt{n}/2} \frac{\sqrt{n}}{4j\sqrt{n}} = \frac{\Theta(1)}{\sqrt{n}}$$

2.2 The normalized Laplacian

To get a more general bound (more sensitive to degrees, which is important for graphs that are not regular or bounded degree), define the graph's *Normalized Laplacian* to be the matrix $\hat{L} = \hat{L}_G$ given by:

$$\hat{L}_{ij} = \begin{cases} -w_{ij}/\sqrt{d_i d_j} & \text{if } ij \in E, \\ 1 & \text{if } i = j, \\ 0 & \text{otherwise.} \end{cases}$$

Exer: Prove that $\hat{L} = I - D^{-1/2} A D^{-1/2} = D^{-1/2} L D^{-1/2}$, and that for all $x \in R$

$$x^T \hat{L} x = \sum_{ij \in E} w_{ij} \cdot (x_i/\sqrt{d_i} - x_j/\sqrt{d_j})^2.$$

Furthermore, the smallest eigenvalue of $\hat{L}$ is 0, and that $\vec{d}^{1/2}$ is always an associated eigenvector.

Exer: Does Lemma 5 hold for the normalized Laplacian $\hat{L}$?

2.3 Cheeger's inequality

Theorem 6: [Alon, Alon-Milman, Sinclair-Jerrum, Mihail, after Cheeger] Let λ_2 be the second smallest eigenvalue of the normalized Laplacian $\hat{L}_G$. Then

$$\tfrac{1}{2}\lambda_2 \leq \phi(G) \leq \sqrt{2\lambda_2}.$$

We may assume $\lambda_2 > 0$ (otherwise we're done similarly to Lemma 5).

Observation 7: For every $x \in \mathbb{R}^V$, we can set $y := D^{-1/2} x$ and then

$$\frac{x^T \hat{L} x}{x^T x} = \frac{x^T D^{-1/2} L D^{-1/2} x}{x^T x} = \frac{y^T L y}{(D^{1/2} y)^T D^{1/2} y} = \frac{\sum_{ij \in E} w_{ij} (y_i - y_j)^2}{\sum_{i \in V} d_i y_i^2}.$$

This is "almost" like the Rayleigh quotient of y with respect to L but with "weights" in the denominator.

2.4 The easy direction

We want to show that the eigenvector problem is a relaxation of the cut problem, hence its value can be only smaller. Specifically, for every cut $(S, \bar{S})$ (including the optimal one) we look at the Rayleigh quotient with respect to a vector that is roughly like 1_S, but of course λ_2 is the "minimum" Rayleigh quotient.

The proof was seen in class. Here we only outline the main idea. Recall:

$$\lambda_2 = \min_{x \perp \vec{d}^{1/2}} \frac{x^T \hat{L} x}{x^T x}.$$

Using Observation 7, we see that λ_2 is the minimizer of $\frac{x^T \hat{L} x}{x^T x} = \frac{y^T L y}{y^T D y}$ under the condition $0 = x^T d^{1/2} = y^T D^{1/2} d^{1/2} = \langle y, \vec{d} \rangle$.

First attempt. Fix $S \subset V$. Intuitively, it should be any set with small sparsity $\mathrm{sp}_G(S)$, perhaps even the minimizer of $\mathrm{sp}(G)$. Building on Observation 7, it makes sense to choose $x = D^{1/2} y$ for $y = 1_S$. Then

$$\frac{x^T \hat{L} x}{x^T x} = \frac{y^T L y}{y^T D y} = \frac{w(S, \bar{S})}{d(S)}.$$

But $y^T \vec{d} = d(S) \neq 0$.

Second attempt. We use a single positive value for all the coordinates $i \in S$, and a single negative value for all coordinates $i \in \bar{S}$. An appropriate weighting (similarly to Lemma 5) is to choose y to be $\frac{1}{d(S)} 1_S - \frac{1}{d(\bar{S})} 1_{\bar{S}}$. We need to verify that $y^T \vec{d} = 0$ and $\frac{x^T \hat{L} x}{x^T x} = \mathrm{sp}_G(S)$.

Finally, we let S be a minimizer of $\phi_G(S)$ to prove that $\lambda_2 \leq \frac{x^T \hat{L} x}{x^T x} \leq \mathrm{sp}(G) \leq 2\phi(G)$.

Exer: Prove a statement similar to Theorem 6 for $\lambda_2(L)$ and the isoperimetric number $\alpha_G = \min_{S \subset V} \frac{w(S, \bar{S})}{minn|S|, |\bar{S}|}$. Note that now the inequalities might involve the maximum degree $d_{\max} = \max_{i \in V} d_i$.