# Notes for Lecture 23

Let $G = (V, E)$ be a $d$-regular graph, $|V| = n$. Last time we defined the *edge expansion* of $G$,

$$h(G) = \min_{S, |S| \leq n/2} \frac{\#\text{ edges}(S, \overline{S})}{d|S|}.$$

Today we will look at a different approach to defining expanders via random walks.

Let $M$ denote the transition matrix of a ramdom walk on $G$, defined by

$$M(u, v) = \frac{\#\text{ edges from } u \text{ to } v}{d}$$

Let $\pi$ be a distribution over the vertices of $G$, $\pi \in \mathbb{R}^V$. Then $\pi M$ is the distribution of the final position, after starting at $v \sim \pi$ and taking one step of the random walk. $\pi M^t$ is the distribution of the final position, after starting at $v \sim \pi$ and taking $t$ steps of the random walk. Let $\mathcal{U}_n = (\frac{1}{n}, \ldots, \frac{1}{n})$ be the uniform distribution on the vertices.

We say $G$ is an expander if for every distribution $\pi$, $||\mathcal{U}_n - \pi M||_2$ is smaller by a constant factor than $||\mathcal{U}_n - \pi||_2$.

We also define the following notion of expansion:

$$\lambda(G) = \max_{\text{distributions } \pi} \frac{||\pi M - \mathcal{U}_n||_2}{||\pi - \mathcal{U}_n||_2}$$

Observe that

$$\begin{aligned}
||\mathcal{U}_n - \pi M^t||_{\text{SD}} &= \frac{1}{2}||\mathcal{U}_n - \pi M^t||_1 \\
&\leq \frac{\sqrt{n}}{2}||\mathcal{U}_n - \pi M^t||_2 && \text{by Cauchy-Schwarz} \\
&= \frac{\sqrt{n}}{2}||\mathcal{U}_n - \pi M^{t-1} M||_2 \\
&\leq \frac{\sqrt{n}}{2}\lambda(G)||\mathcal{U}_n - \pi M^{t-1}||_2 \\
&\leq \ldots \\
&\leq \frac{\sqrt{n}}{2}\lambda^t(G)||\mathcal{U}_n - \pi||_2 && \text{by induction} \\
&\leq \frac{\sqrt{n}}{2}\lambda^t(G) && \text{since } ||\mathcal{U}_n - \pi||_2 \leq 1 \\
&\leq \varepsilon && \text{provided } t \geq \log_{\lambda(G)} \frac{\sqrt{n}}{2\varepsilon}
\end{aligned}$$

Thus, if the number of steps $t$ is $\Omega\left(\frac{1}{1-\lambda(G)}(\log n + \log \frac{1}{\varepsilon})\right)$ then the resulting distribution is $\varepsilon$ close to uniform. In particular this means that $\text{Diam}(G) = O(\frac{1}{1-\lambda(G)} \log n)$. Also $G$ is an expander if and only if $\lambda(G)$ is constant and bounded away from 1.

Our goal is to prove

**Theorem 1 (Cheeger's inequality)** *For all $d$-regular graphs $G$,*

$$\frac{1-\lambda(G)}{2} \leq h(G) \leq \sqrt{2(1-\lambda(G))}$$

We remark that these bounds are tight up to constant factors. The usefulness of Cheeger's inequality lies in the fact that $\lambda$ is efficiently computable, whereas $h$ is NP-hard to compute exactly.

From the definition of $\lambda$ we see that it may be equivalently written as

$$\lambda(G) = \max_x \frac{||xM||_2}{||x||_2}$$

where the maximum is taken over all $x \in \mathbb{R}^V$ such that $\mathcal{U}_n + x$ is a distribution. Now, $\mathcal{U}_n + x$ is a distribution means that $\sum_v (\frac{1}{n} + x(v)) = 1$ so that $\sum_v x(v) = 0$ and also that $\forall v, x(v) \geq -\frac{1}{n}$. The latter condition is just a normalization. Using the former condition in our equivalent formulation, we get

$$\lambda(G) = \max_{x \in \mathbb{R}^V, \sum x(v) = 0} \frac{||xM||_2}{||x||_2}$$

This maximum is achieved by the vector orthogonal to the all-ones vector that is least shortened under the action of $M$.

Let us recall the following facts from linear algebra. Let $M$ be an $n \times n$ matrix. We say (a complex number) $\lambda$ is an eigenvalue for $M$ if $\exists x \in \mathbb{R}^n$, $x \neq 0$ such that $xM = \lambda x$. Equivalently, $x(M - \lambda I) = 0$. Thus $\lambda$ is an eigenvalue for $M$ exactly when $\det(M - \lambda I) = 0$. The stipulated vector $x$ is called an eigenvector corresponding to the eigenvalue $\lambda$.

If $M$ is a symmetric matrix, then its eigenvalues are real. If $\lambda$ has multiplicity $k$ (*i.e.* $(z-\lambda)^k$ divides the characteristic polynomial, $\det(M - zI)$) then $\{x \mid xM = \lambda x\}$ is a linear subspace of dimension $k$. If $\lambda$ and $\lambda'$ are eigenvalues and $x$ and $x'$ are the corresponding eigenvectors, then

$$\lambda x y^T = (xM)y^T = x(My^T) = x(yM)^T = \lambda' x y^T.$$

It follows that the eigenvectors corresponding to *distinct* eigenvalues are orthogonal. Since any multiple of an eigenvector is also an eigenvector, it follows that there is an orthonormal basis of eigenvectors.

Now, if $M$ is the transition matrix of a random walk on $d$-regular graph $G$, then

(a) $\mathcal{U}_n$ is an eigenvector for eigenvalue 1.

(b) If $\lambda$ is an eigenvalue of $M$ and $x$ the corresponding eigenvector, then $\lambda \leq 1$. Moreover, if $\lambda < 1$ then $x$ is orthogonal to $\mathcal{U}_n$.

The proof of (a) is immediate. To see (b), note that

$$
\begin{aligned}
0 &\leq \sum_{u,v \in V} M(u,v)(x(u)-x(v))^2 \\
&= \sum_{u,v} M(u,v)x^2(u) + \sum_{u,v} M(u,v)x^2(v) - 2\sum_{u,v} M(u,v)x(u)x(v) \\
&= \sum_{u} x^2(u) + \sum_{v} x^2(v) - 2\sum_{v} (xM)(v)x(v) && \text{since } M \text{ is a stochastic matrix} \\
&= 2\sum_{v} x^2(v) - 2\sum_{v} \lambda x^2(v) && \text{since } xM = \lambda x \\
&= 2(1-\lambda)\sum_{v} x^2(v)
\end{aligned}
$$

It follows that $1-\lambda \geq 0$, *i.e.* $\lambda \leq 1$. Moreover, if $\lambda$ is an eigenvalue of $M$ the $\lambda^2$ is an eigenvalue of $M^2$, which is the transition matrix of a random walk on $G^2$. ($G^2$ is the graph which has an edge between $u$ and $v$ exactly when there is a path of length 2 between them in $G$.) By the above argument, $\lambda^2 \leq 1$.

Now suppose $1 = \lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_n \geq -1$ are the eigenvalues of $M$. Let $x_1, x_2, \ldots, x_n$ be an orthonormal basis of eigenvectors, $x_i$ corresponding to eigenvalue $\lambda_i$ and $x_1 = (1/\sqrt{n}, \ldots, 1/\sqrt{n})$. Then $x \perp (1, \ldots, 1)$ implies that there are real numbers $a_2, \ldots, a_n$ such that $x = \sum_{i=2}^n a_i x_i$. We have

$$
\begin{aligned}
\lambda(G) &= \max_{x \perp (1,\ldots,1)} \frac{||xM||_2}{||x||_2} \\
&= \max_{a_2,\ldots a_n} \frac{||(\sum_2^n a_i x_i)M||_2}{||\sum_2^n a_i x_i||_2} \\
&= \max_{a_2,\ldots a_n} \frac{||\sum_2^n a_i \lambda_i x_i||_2}{\sqrt{\sum_2^n a_i^2}} \\
&= \max_{a_2,\ldots a_n} \frac{\sqrt{\sum_2^n a_i^2 \lambda_i^2}}{\sqrt{\sum_2^n a_i^2}} \\
&= \max_{a_2,\ldots a_n} \frac{\max_{i \in \{2,\ldots,n\}} |\lambda_i| \sqrt{\sum_2^n a_i^2}}{\sqrt{\sum_2^n a_i^2}} \\
&= \max_{i \in \{2,\ldots,n\}} |\lambda_i| \\
&= \max\{\lambda_2, -\lambda_n\}
\end{aligned}
$$

We would like to say that $\lambda(G) = \lambda_2$, but there may be negative eigenvalues. However negative eigenvalues only occur in bipartite graphs. If $G$ is bipartite, then the random walk on $G$ is periodic, and can therefore never mix (*i.e.* get close to uniform). To fix this problem, we can add self-loops to $G$ and do a *lazy* random walk, staying at the same vertex with probability 1/2. This random walk has transition matrix $\frac{1}{2}(M+I)$, whose eigenvalues are $\frac{1}{2}(1+\lambda_i)$. Thus the eigenvalues are non-negative.

In fact, Cheeger's inequality actually involves $\lambda_2$, rather than $\lambda$.

Another problem occurs if $\lambda_2 = 1$. In this case the graph is disconnected, and any random walk is confined to one component, and therefore cannot mix. To see that the graph is disconnected, observe that when $\lambda_2 = 1$, there is a non-zero vector $x \perp (1, \ldots, 1)$, with $xM = x$. For this $x$, we have $0 \leq \sum_{u,v \in V} M(u,v)(x(u) - x(v))^2 = 2(1 - \lambda_2) \sum_v x^2(v) = 0$. Thus for each $u, v \in V$ either $M(u,v) = 0$ or $x(u) = x(v)$. Fix $u$ and look at the all the vertices that are reachable from $u$. An easy induction shows that if $u'$ is reachable from $u$ then $x(u') = x(u)$. But since $x$ is not a multiple of $(1, \ldots, 1)$, there is a vertex $v$ such that $x(v) \neq x(u)$. Thus $v$ is not reachable from $u$ and therefore $G$ is disconnected.

**Theorem 2 (Expander Mixing Lemma)** *For all $A$, $B \subseteq V$, $A \cap B = \emptyset$,*

$$\left| \#\ edges(A, B) - \frac{d|A||B|}{n} \right| \leq \lambda_2 d \sqrt{|A||B|}$$

Here note that $\frac{d|A||B|}{n}$ is the expected number of edges from $A$ to $B$ in a random $d$-regular graph $G$.

As a consequence of the expander mixing lemma, if $|A| = \alpha n$ and $|B| = \beta n$, then

$$|\#\ \text{edges}(A, B) - d\alpha\beta n| \leq \lambda_2 d n \sqrt{\alpha\beta} \leq \varepsilon d\alpha\beta n$$

provided $\lambda_2 \leq \varepsilon\sqrt{\alpha\beta}$.

Another consequence is that the largest independent set in $G$ has size $\lambda_2 n$. To see this, note that $\#\ \text{edges}(A, B) = \sum_{u \in A, v \in B} \#\ \text{edges}\ (u, v)$ in $G$. If $A$ is an independent set in $G$ and $B = A$ then $\#\ \text{edges}(A, B) = 0$. According to the theorem, $\left| \frac{d|A|^2}{n} \right| \leq \lambda_2 d|A|$. In other words, $|A| \leq \lambda_2 n$.

It also follows that the chromatic number of $G$ is at least $1/\lambda_2$.