Eigenvalues, Expansion, and Flows

1 Characterizing Slow Mixing

Which graphs would have very slow mixing? Intuitively, graphs with bad bottlenecks. One example is the "barbell graph": two complete graphs K_n attached by a single edge. It is as "close" to disconnected as possible without actually being disconnected. In general, if graphs have low "expansion" – that is, they have large cuts with few edges across – then these bottlenecks will hurt mixing speed. If expansion is high, the walks mix rapidly.

To keep the analysis clean, we focus on undirected d-regular graphs. In most applications of this technique, we are building such graphs to sample objects from a uniform stationary distribution.

Let $\mathbf{A}$ be the (symmetric) adjacency matrix of a graph G. Then $\mathbf{P} = \frac{1}{d}\mathbf{A}$ is the transition matrix of the walk. Just in case $\mathbf{P}$ is not aperiodic, look at

$$\mathbf{Q} = \frac{1}{2}\mathbf{P} + \frac{1}{2}\mathbf{I},$$

which gives a self-loop with probability $1/2$ to each node. $\mathbf{Q}$ is symmetric, and it has therefore n real eigenvalues $\lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_n$. It is not difficult to show that $\lambda_1 = 1$ and $\lambda_n \geq 0$.

One of the main theorems of algebraic graph theory (not very difficult, but skipped here) is that the number of connected components of G is the largest i with $\lambda_i = 1$. By analogy, if λ_2 is "very close" to 1, it means there are "almost" two components (and similarly for larger i).

So $\lambda_1 - \lambda_2 = 1 - \lambda_2$ is a useful measure of connectivity. It is called the **spectral gap** of $\mathbf{Q}$.

2 Spectral Gaps and Mixing Speed

The stationary distribution $\boldsymbol{\pi}$ by definition satisfies $\boldsymbol{\pi}\mathbf{Q} = \boldsymbol{\pi}$. In linear algebra terms, $\boldsymbol{\pi}$ is a left eigenvector of $\mathbf{Q}$ with eigenvalue $1 = \lambda_1$. Let $\boldsymbol{\mu}$ be any starting distribution. After t steps, the L_2-error is

$$\|\boldsymbol{\mu}\mathbf{Q}^t - \boldsymbol{\pi}\|_2.$$

Because $\mathbf{Q}$ is symmetric, it has an orthonormal basis of eigenvectors $\boldsymbol{\omega}_1, \boldsymbol{\omega}_2, \ldots, \boldsymbol{\omega}_n$ with $\boldsymbol{\omega}_1 = \boldsymbol{\pi}$. These correspond to the eigenvalues $\lambda_1, \lambda_2, \ldots, \lambda_n$. Write $\boldsymbol{\mu}$ in the basis as

$$\boldsymbol{\mu} = \sum_{i=1}^{n} \alpha_i \boldsymbol{\omega}_i.$$

Then

$$\begin{aligned}
\boldsymbol{\mu}\mathbf{Q}^t &= \Big(\sum_{i=1}^{n} \alpha_i \boldsymbol{\omega}_i\Big)\mathbf{Q}^t \\
&= \sum_i \alpha_i (\boldsymbol{\omega}_i \mathbf{Q}^t) \\
&= \sum_i \alpha_i \lambda_i^t \boldsymbol{\omega}_i.
\end{aligned}$$

Because G has a single connected component, λ_1 is the only eigenvalue which equals 1. So as $t \to \infty$, this converges to

$$\alpha_1 \boldsymbol{\omega}_1 = \alpha_1 \boldsymbol{\pi}.$$

Since $\boldsymbol{\mu}\mathbf{Q}^t$ is known to converge to $\boldsymbol{\pi}$, we have $\alpha_1 = 1$. Going back to the L_2 error, we have

$$\begin{aligned}
\|\boldsymbol{\mu}\mathbf{Q}^t - \boldsymbol{\pi}\|_2 &= \|\sum_i \alpha_i \lambda_i^t \boldsymbol{\omega}_i - \boldsymbol{\pi}\|_2 \\
&= \|\sum_{i=2}^{n} \alpha_i \lambda_i^t \boldsymbol{\omega}_i\|_2 \\
&= \sqrt{\sum_{i=2}^{n} \alpha_i^2 \lambda_i^{2t}} \\
&\leq \lambda_2^t \sqrt{\sum_{i=1}^{n} \alpha_i^2} \\
&\leq \lambda_2^t \|\boldsymbol{\mu}\|_2 \\
&\leq \lambda_2^t \\
&= (1 - (1 - \lambda_2))^t \\
&\leq e^{-(1-\lambda_2)t}.
\end{aligned}$$

So to get error δ, it suffices to have $t \geq \frac{\ln 1/\delta}{1-\lambda_2}$. In particular, if the spectral gap $1 - \lambda_2$ is large enough – $\Omega(1)$ or $\Omega(1/polylog(n))$ – the walk mixes rapidly. Thus, our goal is to bound the spectral gap.

3 Spectral Gaps and Expansion Measures

For a cut $(S, \bar{S})$, define the:

- Edge expansion as
$$\alpha(S) := \frac{|e(S, \bar{S})|}{|S|}.$$
- Vertex expansion as
$$\frac{|\mathcal{N}(S) \setminus S|}{|S|}.$$
- Conductance as
$$\frac{1}{\sum_{u \in S} \boldsymbol{\pi}_u} \sum_{e=(u,v), u \in S, v \notin S} \boldsymbol{\pi}_u p_{uv}.$$

For undirected d-regular graphs, all $\boldsymbol{\pi}_u = {}^1\!/_n$, so the conductance is

$$\frac{1}{|S|} \sum_{e=(u,v), u \in S, v \notin S} \frac{1}{d} = \frac{e(S, \bar{S})}{d|S|} = \frac{\alpha(S)}{d}.$$

All of these in a sense are a measure of surface-to-volume ratio of S. Define the edge expansion of a graph G as the worst expansion of any set:

$$\alpha(G) := \min_{S, |S| \leq n/2} \alpha(S).$$

Theorem 1. *[Cheeger's Inequality].* Let G be a d-regular undirected graph. Then,

$$1 - \frac{\alpha(G)}{d} \leq \lambda_2(G) \leq 1 - \frac{1}{4}(\frac{\alpha(G)}{d})^2.$$

Rewritten to see the spectral gap,

$$\frac{1}{4}(\frac{\alpha(G)}{d})^2 \leq 1 - \lambda_2(G) \leq \frac{\alpha(G)}{d}.$$

Proof. The proof is nontrivial, in particular the upper bound on $\lambda_2(G)$. The key idea is to interpret the eigenvector $\boldsymbol{\omega}_2$ as embedding the nodes into a line and show that one of the "line cuts" must have corresponding expansion. □

In particular, if $\alpha(G)$ is $\Omega(1)$ or $\Omega(1/polylog(n))$, then so is the spectral gap – so the random walk mixes rapidly. Our new goal is to prove lower bounds on the expansion $\alpha(G)$ for graphs we care about.

4 Expansion and Flows

The intuition is that cuts with few edges are bottlenecks for flows. In particular, if we can send a lot of flow, we cannot have small cuts. Here, different types of flows correspond to different types of cuts.

Specifically, we focus on all-to-all multicommodity flow. Give each edge a capacity of 1 and find the largest Ψ such that for each node pair (s,t) simultaneously, we can send Ψ units of flow. Denote this quantity by $\Psi^*(G)$.

Equivalently, find the smallest congestion $c(G)$ such that we can route one unit of flow for each pair (s,t) simultaneously with congestion no more than $c(G)$ on any edge. Then $\Psi^*(G) = 1/c(G)$.

One more related expansion measure:

$$\sigma(G) = \min_S \frac{|e(S,\bar{S})|}{|S| \cdot |\bar{S}|}.$$

Theorem 2. *[Leighton & Rao].*

$$\Psi^*(G) \leq \sigma(G) \leq \Psi^*(G) \cdot \mathcal{O}(\log n).$$

For any set S with $|S| \leq n/2$,

$$\alpha(S) = \sigma(S) \cdot |\bar{S}|$$

lies between $\frac{n}{2}\sigma(S)$ and $n\sigma(S)$. Thus,

$$\frac{n}{2}\sigma(G) \leq \alpha(G) \leq n\sigma(G).$$

So

$$\frac{n}{2}\Psi^*(G) \leq \alpha(G) \leq \mathcal{O}(n \log n)\Psi^*(G).$$

If we can prove a lower bound of $\frac{1}{n \cdot polylog(n)}$ on $\Psi^*(G)$, we get a $\frac{1}{polylog(n)}$ lower bound on $\alpha(G)$ and thus on the spectral gap (which means our chain mixes rapidly).

Theorem 3. If it is possible in G to route one unit of flow simultaneously for each node pair (s,t) such that the congestion of each edge e is at most $\mathcal{O}(n \cdot polylog(n))$, then the random walk on G mixes in time $\mathcal{O}(d^2 polylog(n))$.

This is called the **canonical flows** technique, and is most often applied by having the flow for each pair (s,t) along a single path. Then it is called the **canonical paths** technique. In general, analysis using canonical flows often loses several log factors compared to the best possible (but that doesn't matter if you just want to prove rapid mixing).

5 Application: Sampling Matchings in Dense Bipartite Graphs

Given a bipartite graph G, count the number of perfect matchings in G. Equivalently, compute $\mathrm{perm}(A(G))$. This is $\#P$-complete (even though finding a perfect matching is polytime). We will count them approximately using sampling. Define $\mathcal{M}_k$ as the set of all matchings of G with exactly k edges. We want to compute $|\mathcal{M}_n|$.

Write

$$|\mathcal{M}_n| = \frac{|\mathcal{M}_n|}{|\mathcal{M}_{n-1}|} \cdots \frac{|\mathcal{M}_2|}{|\mathcal{M}_1|}|\mathcal{M}_1|.$$

Clearly $|\mathcal{M}_1| = m$. Our goal is to estimate all of these ratios sufficiently accurately, so that even multiplying them leads to an error of at most $1 \pm \varepsilon$.

Why estimate the ratios instead? Any $|\mathcal{M}_i|$ could be very large or very small, so estimating its size by sampling is error-prone. We will show that each ratio is bounded between n^{-2} and n^2 for dense graphs. (A graph is considered dense if each node has degree at least $n/2$).

To estimate the ratio $|\mathcal{M}_i|/|\mathcal{M}_{i-1}|$, we sample nearly uniformly from the union of the two and count the fraction of samples that are from $\mathcal{M}_i$. Because the sizes are similar enough, this gives a $1 \pm \varepsilon$ approximation of the ratio.

Our new problem is to sample nearly uniformly from $\mathcal{M}_k \cup \mathcal{M}_{k-1}$. This can be reduced fairly easily to sampling from $\mathcal{M}_n \cup \mathcal{M}_{n-1}$ (by adding a bunch of new nodes). We want to sample perfect or near-perfect matchings using random walks.

We define a suitable random walk on $\mathcal{M}_n \cup \mathcal{M}_{n-1}$. Suppose we are at a matching M. We choose a uniformly random edge $e = (u, v) \in G$.

- If M is perfect and $e \in M$, remove e from M.
- If M is not perfect and both u and v are unmatched, add e to M.
- If M is not perfect and u is matched to some w and v is unmatched, replace (u, w) with (u, v) – symmetrically if v is matched and u is unmatched.
- Otherwise, do nothing.

Is the Markov graph undirected? Yes, addition and removal go both ways, and edge rotation can be undone. Is the Markov graph regular? Yes, each matching M has degree m (one edge for each possible edge pick, including self-loops). Once we show connectedness, we will have shown the stationary distribution is uniform.