Diameter, Probability, and Concentration of Measure

Daniel A. Spielman November 4, 2009

19.1 A first bound on the Diameter

The diameter of a graph is the maximum over vertices u and v of the distance between u and v. It is easy to show that graphs with high conductance must have low diameter. If the set of edges leaving a set must be large relative to the size of the set, then after taking neighborhoods a few times, one must encounter most of the edges. We now make this intuition precise.

Recall from Lecture 7 the definion of the sparsity of a set $S \subseteq V$:

$$\mathsf{sp}(S) \stackrel{\text{def}}{=} \frac{|\partial(S)|}{\min\left(d(S), d(\bar{S})\right)},$$

where we recall that $d(S)$ is the sum of the degrees of vertices in S. We now bound the diameter of G in terms of

$$\mathsf{sp}_G \stackrel{\text{def}}{=} \min_S \mathsf{sp}(S).$$

For a vertex u, let U^k denote the set of vertices at distance at most k from u. In particular, $U^0 = \{u\}$. For every k such that $d(U^k) \leq |E|/2$, we have

$$d(U^k) \leq d(\bar{U^k}),$$

so

$$\left|\partial(U^k)\right| \geq \mathsf{sp}_G d(U^k).$$

On the other hand,

$$d(U^{k+1}) \geq d(U^k) + \left|\partial(U^k)\right|.$$

So, when $d(U^k) \leq |E|/2$,

$$d(U^{k+1}) \geq (1 + \mathsf{sp}_G)\, d(U^k).$$

As $d(u) \geq 1$, this implies

$$d(U^k) \geq (1 + \mathsf{sp}_G)^k.$$

So, if we choose a k such that

$$(1 + \mathsf{sp}_G)^k \geq m, \tag{19.1}$$

we will guarantee that $d(U^k)$ is at least $m = d(V)/2$. Define V^k analogously to U^k. As both U^k and V^k touch at least half of the edges, there must be an edge between U^k and V^k. So, the distance between u and v is at most $2k + 1$.

When sp_G is small, $1 + \mathsf{sp}_G$ is approximately e^{sp_G}, so up to first order (19.1) will be satisfied when

$$k \geq \frac{\ln m}{\mathsf{sp}_G}.$$

So, we obtain an upper bound on the diameter that is essentially

$$\frac{2 \ln m}{\mathsf{sp}_G}.$$

We can easily relate this bound to the spectrum of a graph. In Lecture 7, we proved that

$$\mathsf{sp}_G \geq \nu_2 / 2,$$

where ν_2 is the second-smallest eigenvalue of the normalized Laplacian matrix of a graph. We thus find that the diameter is essentially[1] at most

$$\frac{4 \ln m}{\nu_2}.$$

For regular graphs, Chung [Chu89] obtained the cleaner upper bound of

$$\frac{\ln n}{2\nu_2}.$$

But, one can obtain a quadratically better bound in terms of the eigenvalues.

Before we do that, let me quickly explain how the bound we derived could be made precise and stronger. To make it stronger, note that

$$d(U^{k+1}) \geq d(U^k) + \left|\partial(U^k)\right| + \left|\partial(U^{k+1})\right|.$$

This saves us a factor of approximately 2. To make the bound precise, I point out that an examination of the Taylor series of $\ln(1+x)$ reveals that

$$\frac{1}{\ln(1+x)} \leq \frac{1}{x} + \frac{1}{2}. \tag{19.2}$$

19.2 A better bound on the Diameter

Intuitively, if we can quickly solve linear equations in the Laplacian matrix of a graph by an iterative method, then the graph should have small diameter. We now make that intuition precise.

Let L be the Laplacian of a connected graph G. Recall that iterative methods work well when there is a low-degree polynomial in L that is approximately the inverse of L. For non-singular matrices, we measured this by comparing the product of L with the approximate inverse to the identity. As

[1] This bound is off by a multiplicative factor of approximately $1 + 2\nu_2$.

L is singular, we only want to approximate the identity on its span. So, we should compare to the projector onto its span, which we denote Π and which we recall equals $\frac{1}{n}L_{K_n}$.

Assume that we have a polynomial p of degree $k-1$ such that

$$\|p(L)L - \Pi\| \leq \epsilon. \tag{19.3}$$

We know from the analysis of the conjugate gradient that we can find ϵ-approximate solutions to linear equations in L by performing k multiplications by L. In fact, recall that we got an approximate solution to $L\boldsymbol{x} = \boldsymbol{b}$, for $\boldsymbol{b}$ in the span of L, by setting

$$\boldsymbol{x}^k = p(L)\boldsymbol{b}.$$

We then found

$$\left\| L\boldsymbol{x}^k - \boldsymbol{b} \right\| = \|(Lp(L) - \Pi)\boldsymbol{b}\| = \|(p(L)L - \Pi)\boldsymbol{b}\| \leq \|(p(L)L - \Pi)\| \, \|\boldsymbol{b}\| \leq \epsilon \, \|\boldsymbol{b}\| \, .$$

To use (19.3) to prove an upper bound on the diameter of a graph, let S and T be two sets of vertices at distance $k+1$ from each other. That is, for every vertex $s \in S$ and $t \in T$, the distance from s to t is at least $k+1$. The vector $L\boldsymbol{\chi}_T$ is supported on the vertices of distance at most 1 from T, and generally for any polynomial $Lp(L)$ of degree k and zero constant[2] term, $Lp(P)$ is supported on the set of vertices of distance at most k from T. So,

$$\boldsymbol{\chi}_S^T p(L) L \boldsymbol{\chi}_T = 0.$$

On the other hand, as S and T are disjoint, we have

$$\boldsymbol{\chi}_S^T \Pi \boldsymbol{\chi}_T = \boldsymbol{\chi}_S^T \frac{1}{n} L_{K_n} \boldsymbol{\chi}_T = \frac{1}{n} |S| \, |T| \, .$$

Combining these facts with (19.3) gives

$$\frac{1}{n} |S| \, |T| = \boldsymbol{\chi}_S^T \left(p(L)L - \Pi \right) \boldsymbol{\chi}_T \leq \|\boldsymbol{\chi}_S\| \, \|(p(L)L - \Pi)\| \, \|\boldsymbol{\chi}_T\| \leq \epsilon \, \|\boldsymbol{\chi}_S\| \, \|\boldsymbol{\chi}_T\| \, . \tag{19.4}$$

Recall that in Lecture 16 we exploited the following polynomials, built from Chebychev polynomials.

Theorem 19.2.1. *For $t \geq 1$ and $0 < \alpha < \beta$, there exists a polynomial $q(X)$ of degree t such that $q(0) = 1$ and*

$$|q(\lambda)| \leq 2 \left(\frac{\sqrt{\kappa} - 1}{\sqrt{\kappa} + 1} \right)^t ,$$

for all $\alpha \leq \lambda \leq \beta$ and $\kappa = \beta/\alpha$.

We now use this to obtain a better bound on the diameter of a graph in terms of λ_n/λ_2.

[2] Be very careful here. You might at first think that a non-zero constant term would be safe, because it looks like it corresponds to adding a multiple I. But, it really corresponds to adding a multiple of Π.

Theorem 19.2.2. *Let $G = (V, E)$ be a connected graph, and let $\lambda_2 \leq \cdots \leq \lambda_n$ be its Laplacian eigenvalues. Then, the diameter of G is at most*

$$\left(\frac{1}{2}\sqrt{\frac{\lambda_n}{\lambda_2}} + 1 \right) \ln 2n.$$

Proof. Let s and t be any two vertices in V that are distance at least $k+1$ from each other. Setting $S = \{s\}$ and $T = \{t\}$, inequality (19.4) tells us that for every polynomial q of degree k such that $q(0) = 1$,

$$\frac{1}{n} \leq \max_{i \geq 2} q(\lambda_i).$$

Applying the polynomials from Theorem 19.2.1, we find

$$\frac{1}{n} \leq 2 \left(\frac{\sqrt{\kappa} - 1}{\sqrt{\kappa} + 1} \right)^k.$$

Taking logs and rearraning, this yields the following upper bound on k.

$$k \leq \frac{\ln 2n}{\ln \left(1 + 2\sqrt{\frac{\lambda_2}{\lambda_n}} \right)},$$

$$\frac{1}{2}\left(\sqrt{\frac{\lambda_n}{\lambda_2}} + 1 \right),$$

where the last inequality follows from (19.2). $\square$

This later bound is from [CFM94].

19.3 The Hypercube

It is instructive to see how these bounds apply to the hypercube on $\{0, 1\}^d$. We have $n = 2^d$, $\lambda_2 = 2$ and $\lambda_n = 2d$. So, Theorem 19.2.2 provides an upper bound on the diameter of the hypercube of approximately

$$(\sqrt{d}/2) \ln n \sim d^{3/2}.$$

As the diameter is in fact d, this is a very poor bound. However, we can use the technique from the previous section to obtain the correct bound on the diameter of the hypercube. Recall that the eigenvalues of the hypercube are precisely the integers $2i$ for $1 \leq i \leq d$. So, we may use the polynomial

$$q(x) = \prod_{i=1}^{d} (1 - x/2i).$$

While it is comforting to get the correct bound on the diameter of the hypercube, it is not the easiest way to do so.

19.4 Probability and Concentration of Measure

Consider d random variables, $x_1, \ldots, x_d$, each of which is chosen independently and uniformly from $\{\pm 1\}$. Taken as a vector, $(x_1, \ldots, x_d)$ is naturally associated with a vertex of a d-dimensional hypercube. We can learn a lot about functions of these random variables through a spectral analysis of the hypercube.

We will consider functions f of these random variables that satisfy a *Lipschitz* condition:

$$|\{i : x_i \neq y_i\}| = 1 \quad \Longrightarrow \quad |f(x_1, \ldots, x_d) - f(y_1, \ldots, y_d)| \leq 1.$$

For example, consider

$$f(x_1, \ldots, x_d) = \frac{1}{2} \sum_i x_i. \tag{19.5}$$

The *Concentration of Measure Phenomenon* is the observation that such functions are often tightly concentrated around their *median*. This may be understood as a strong generalization of the Chernoff/Hoeffding bounds, which say that the sum (19.5) is tightly concentrated around its *mean*. Statements that only rely on the Lipschitz condition are very powerful, as such a condition can often be proved, even for functions that we understand very poorly.

Let's see why a function that satisfies the Lipschitz condition is concentrated around its median, let f be any such function and let $V = \{\pm 1\}^d$ be the vertex set of the d-dimensional hypercube. Let E be the natural set of edges. The Lipschitz condition says that for every edge $(\boldsymbol{x}, \boldsymbol{y})$,

$$|f(\boldsymbol{x}) - f(\boldsymbol{y})| \leq 1.$$

Let μ be the median of f over V, and let

$$S = \{\boldsymbol{v} \in V : f(\boldsymbol{v}) \leq \mu\}.$$

Also, let T be the set of vertices on which f exceeds μ by at least k:

$$T = \{\boldsymbol{v} \in V : f(\boldsymbol{v}) > \mu + k\}.$$

The Lipschitz condition tells us that the distance between S and T is at least k. As $|S| = n/2$, a spectral analysis will tell us that T must be small if k is big.

We will quantify this by emulating the proof of Theorem 19.2.2. However, we first slightly tighten up (19.4). Observe that

$$\boldsymbol{\chi}_S^T \left(p(L)L - \Pi\right) \boldsymbol{\chi}_T = \boldsymbol{\chi}_S^T \Pi \left(p(L)L - \Pi\right) \Pi \boldsymbol{\chi}_T.$$

So,

$$\frac{1}{n} |S|\, |T| \leq \epsilon \, \|\Pi \boldsymbol{\chi}_S\| \, \|\Pi \boldsymbol{\chi}_T\|. \tag{19.6}$$

We have $|T| = \tau n$ and $\|\boldsymbol{\chi}_T\| = \sqrt{n(\tau - \tau^2)}$. So, if S and T are at distance at least $k+1$,

$$\sqrt{\frac{\sigma\tau}{(1-\sigma)(1-\tau)}} \leq \max_{i \geq 2} q(\lambda_i),$$

for every degree k polynomial q for which $q(0) = 1$. Again applying the polynomials of Theorem 19.2.1, we obtain

$$\sqrt{\frac{\sigma\tau}{(1-\sigma)(1-\tau)}} \leq 2\left(\frac{\sqrt{\kappa}-1}{\sqrt{\kappa}+1}\right)^k,$$

where $\kappa = \lambda_n/\lambda_2$. We now set $\sigma = 1/2$, and use this inequality to find an upper bound on τ in terms of k. We find

$$\sqrt{\tau} \leq 2\left(\frac{\sqrt{\kappa}-1}{\sqrt{\kappa}+1}\right)^k = 2\left(1 - \frac{2}{\sqrt{\kappa}+1}\right)^k \leq 2e^{-\frac{2k}{\sqrt{\kappa}+1}}.$$

So,

$$\tau \leq e^{-\frac{4k}{\sqrt{\kappa}+1}}.$$

In our case, $\kappa = d$. So, we obtain the probability bound

$$\mathrm{Pr}_{\boldsymbol{x}}\left[f(\boldsymbol{x}) > \mu + k\right] = \mathrm{Pr}_{\boldsymbol{x}}\left[\boldsymbol{x} \in T\right] = \tau \leq e^{-\frac{4k}{\sqrt{d}+1}}.$$

So, as soon as k exceeds $\sqrt{d}$, the probability that f exceeds the median by more than k becomes very small. This is a very reasonable concentration bound. However, it is possible to prove even strong bounds.

For example one can show that

$$\mathrm{Pr}_{\boldsymbol{x}}\left[f(\boldsymbol{x}) > \mu + k\right] \leq e^{-\frac{k^2}{2d}}$$

(see [AS00, Theorem 7.4.2]). This latter bound is much stronger when k is big, say around a constant times d. It is possible to improve the bound obtained by the spectral method in this case as well by using a stronger analysis of the convergence of the Conjugate Gradient (in particular, case ii in [AL86]). This analysis would also provide a probability exponentially small in d when k is a constant times d, but would still not be quite as strong.

References

[AL86] Owe Axelsson and Gunhild Lindskog. On the rate of convergence of the preconditioned conjugate gradient method. *Numerische Mathematik*, 48(5):499–523, 1986.

[AS00] Noga Alon and Joel Spencer. *The Probabilistic Method.* John Wiley & Sons, 2000.

[CFM94] F. R. K. Chung, V. Faber, and T. A. Manteuffel. On the diameter of a graph from eigenvalues associated with its laplacian. *SIAM Journal on Discrete Mathematics*, 7:443–457, 1994.

[Chu89] F. R. K. Chung. Diameters and eigenvalues. *Journal of the American Mathematical Society*, 2(2):187–196, 1989.