20 Introduction

In this lecture we continue our discussion of algorithms for bandwidth minimization, and show how *volume respecting embeddings* can be used for this purpose. Recall that, the MINIMUM BANDWIDTH problem is to find a one-to-one map $f : V(G) \xhookrightarrow{\text{1-1}} [n]$ to minimize

$$\mathsf{bw}(f) = \max_{\{i,j\}\in E} |f(i) - f(j)|.$$

Recall from the previous lecture that the *local density* of the graph G is defined as

$$D = D(G) = \max_{v,r} \frac{|B(v,r)| - 1}{2r},$$

and that D is a lower bound on the optimal bandwidth of G.

21 High level idea

Our discussion will closely follow that of Feige (1998); however, our proofs and exposition often differ slightly from his. The basic idea is this: we first randomly map the points to real line $\mathbb{R}$ using some map ψ. This map ψ defines an ordering on the points (and will be $1-1$ with probability 1). We then use this ordering to get a map $f : V(G) \to [n]$ by outputting the points in the order in which their images $\psi(v)$ lie on the real line.

What do we want from the embedding? The crucial property we want is that no edge goes over large number of points. Loosely, this may translate to the fact that all points are fairly spread out in this random map. We formalize this intuition as follows.

Divide $\mathbb{R}$ into intervals of length l, for some parameter l to be specified later. Consider a set $S \subseteq V$ and look at the image of S under ψ. Call the set S *bad* if $\psi(S) \subseteq [tl, (t+1)l)$ for some t.

We need the following notion of volume of a set of points in a metric space.

Definition 21.1 *Let $S \subseteq V$; consider the complete graph G_S on S, and let the length of an edge d_e for $e = (u,v) \in S \times S$ be $d_G(u,v)$. We define*

$$\mathsf{Tvol} S = \prod_{e \in T} d_e\,,$$

where T is a minimum spanning tree on G_S. Note that this is equivalent to defining the tree volume of a T as the product of the lengths of the edges, and minimizing this over all possible subtrees.

As sketched in the previous lecture, our analysis is based on the following claims.

Claim 21.2 *The chance that a set S gets mapped inside an interval of length l under the mapping ψ is pretty small. In other words,*

$$\mathbf{Pr}[S \textit{ is bad}] \le \frac{O(\eta l)^{|S|-1}}{\mathsf{Tvol} S}$$

Claim 21.3 *Under the mapping ψ end points of an edge are not far apart.*

$$\mathbf{Pr}[\textit{an edge has length} \geq l] \leq \frac{1}{2m},$$

where $m = \#$(edges).

Claim 21.4 *For all graphs G, we have*

$$\sum_{S \subseteq V : |S| = k} \frac{1}{\mathsf{Tvol} S} \leq n \times O(D \log n)^{k-1}$$

We use these three results to prove an approximation guarantee for bandwidth minimization problem.

22 Feige's result

Theorem 22.5 *If we have a map ψ that guarantees the properties in Claim 21.2 and Claim 21.3, then the ordering given by ψ has bandwidth atmost $O(\eta l D \log^2 n)$*

Remark 22.6 *Let B^* denote the optimal bandwidth for the given graph. Then, we know that $D \leq B^*$. This means that the map ψ gives an $O(\eta l \log^2 n)$-approximation for bandwidth.*

Proof. Using Claim 21.2, we can bound the number of bad sets.

$$\mathbf{E}[\#(\text{bad sets})] \leq \sum_{S : |S| = k} \frac{O(\eta l)^{k-1}}{\mathsf{Tvol} S}$$

Using Markov's inequality, we can conclude that

$$\mathbf{Pr}[\#(\text{bad sets}) \leq 3 \cdot \text{RHS}] \geq \frac{1}{3} \tag{22.1}$$

From Claim 21.3 it follows that $\mathbf{Pr}[\text{all edges have length} \leq l] \geq \frac{1}{2}$. This, combined with Equation 22.1, implies that with probability at least $\frac{1}{6}$, we have $\#(\text{bad sets}) \leq 3n \times O(\eta l D \log n)^{k-1}$ and all edges have length atmost l.

If an edge has length $\leq l$, then we know that it spans at most two intervals. Let B be the output bandwidth. In other words, there is an edge with B vertices between its two endpoints. Since this edge spans only two intervals, at least $\frac{B}{2}$ of them must be in one interval. And every subset of size k of these $\frac{B}{2}$ vertices is a em bad set. Therefore,

$$\#(\text{bad sets}) \geq \binom{\frac{B}{2}}{k} \geq (\frac{B}{2k})^k.$$

Using the upper bound on number of bad sets from above, we get

$$\begin{aligned} (\frac{B}{2k})^k &\leq 3n \times O(\eta l D \log n)^{k-1} \\ \Rightarrow \qquad B &\leq (3n)^{\frac{1}{k}} \times O(k\eta l D \log n) \end{aligned}$$

Setting $k = \log n$, we get

$$B \leq O(D\eta l \log^2 n)$$

■

Feige showed how to prove Claims 21.2 and 21.3 for $l = \sqrt{\log n}$ and $\eta = \log n \sqrt{\log \log n}$; we will go over most of the techniques in the following lectures, albeit with slightly worse results in some cases.

23 Volume respecting embedding

Let us now define an embedding (which we call a *"well-separated" embedding*) that is a strict generalization of low-distortion metric embeddings.

Definition 23.7 *A contracting map $\phi : V \to \mathbb{R}^d$ is called* (η, k)-well-separated *if the following condition holds.*

For each set $S \subseteq V$, s.t. $|S| \leq k$, there exists an ordering $\{s_0, s_1, \dots, s_{k-1}\}$ of S such that, for all i, if $L_i = span\{\phi(s_0), \phi(s_1), \dots, \phi(s_{i-1})\}$, then

$$dist(\phi(s_i), L_i) \geq \frac{1}{\eta}(d_G(s_i, \{s_0, \dots, s_{i-1}\}))$$

This is essentially the idea behind a *volume-respecting embedding.* Note that an (η, k)-well-separated map has the properties that:

- The images of any set S of $j \leq k$ points must span a $j-1$ dimensional subspace. Indeed, there exists a permutation such that, for every $i \leq j$, the i-th point in this permutation is not in the span of the first $i-1$ points, and hence the claim follows by a simple induction.

  Question: What is the least η such that there exists an (η, k)-well-separated map for the path graph P_n?

- Note that an (η, k)-well-separated embedding is an η-distortion embedding, since the embedding ensures that the distance between $\phi(s_0)$ and $\phi(s_1)$ is at least $d(s_0, s_1)/\eta$.

In next couple of lectures, we will show how to construct (η, k) well-separated embeddings. For now, we will assume that we have such a map ϕ.

Define the map $\psi : V \to \mathbb{R}$ as

$$\psi = (\text{"projection of } \mathbb{R}^d \text{ onto random line"}) \circ \phi,$$

where the "projection" involves picking a random d-dimensional normal with the identity covariance matrix, and will be formally described below in (23.2).

Remark 23.8 *Technically speaking, we are abusing the term "projection" here, since the length of the "projection" of a unit vector could be as large as d; furthermore, it will be concentrated around* 1*, and not around* $1/\sqrt{d}$*, as would be the case if we chose a random vector* $\hat{r} \in S^{d-1}$ *and projected onto it.*

We will use the following simple fact about the probability of a random variable distributed according to Gaussian distribution taking value in a particular interval.

Fact 23.9 *Consider* $X \sim N(0, \hat{\sigma}^2)$ *with* $\hat{\sigma} \geq \sigma$*. Let* I *be an interval of length* l *in real line. Then,*

$$\mathbf{Pr}[X \in I] \leq \frac{l}{\sqrt{2\pi}\sigma}$$

Proof. Let f denote the probability distribution function for X. Clearly, $f(0) \geq f(x)$ for all $x \in R$. Hence,

$$\mathbf{Pr}[X \in I] = \int_I f(x)dx \leq f(0) \cdot l = \frac{l}{\sqrt{2\pi}\hat{\sigma}} \leq \frac{l}{\sqrt{2\pi}\sigma}$$

■

Now we are all set to prove the Claim 21.2

Proof of Claim 21.2: Fix a set $S = \{s_0, s_1, \ldots, s_{k-1}\}$. Let $\phi(s_i)$ be referred to as v_i. Let $L_i = \text{span}\{v_0, \ldots, v_{i-1}\}$ be the affine span of the first i vectors.

We will use the spherical symmetry of choosing a random line heavily. Using suitable a rotation and translation, we can assume that $v_0 = \vec{0}$, and that $L_i = \text{span}\{\vec{e_1}, \ldots, \vec{e_i}\}$ (i.e., the span of the vectors $\{v_0, \ldots, v_{i-1}\}$ is the subspace spanned by the first $i-1$ basis vectors.

This affine transform allows us to re-interpret the well-separatedness property of the map ϕ very conveniently: let $v_i = (v_{i1}, \ldots, v_{ii}, 0, \ldots, 0)$ and let $d_G(s_i, \{s_0, \ldots, s_{i-1}\}) = q_i$. Then

$$\text{well-separatedness } \Rightarrow v_{ii} \geq \frac{q_i}{\eta}$$

Let us now define the map ψ formally as follows.

$$\psi(x) = \langle \phi(x), \vec{r} \rangle, \tag{23.2}$$

where $\vec{r} = (r_1, \ldots, r_d)$ and each $r_i \sim N(0,1)$. Hence,

$$\psi(s_i) = \langle \vec{v_i}, \vec{r} \rangle. \tag{23.3}$$

Finally, we can now bound the probability that all of S is mapped into the interval $I = [0, l)$ by the map ψ. (Note that, since we performed the translations, this is the only interval in which $\phi(s_0)$ can lie.)

$$\begin{aligned}\mathbf{Pr}[\psi(S) \subseteq I] = & \mathbf{Pr}[\psi(s_0) \in I] \times \\ & \mathbf{Pr}[\psi(s_1) \in I | \psi(s_0) \in I] \times \\ & \dots \\ & \mathbf{Pr}\Big[\psi(s_k) \in I | \wedge_{j=0}^{k-1} \psi(s_j) \in I\Big] \end{aligned} \tag{23.4}$$

In order to compute $\mathbf{Pr}[\psi(s_j) \in I]$ for $j < i$, note that it suffices to look at the values of $\{r_1, r_2, \dots, r_{i-1}\}$. Indeed, the vectors v_j for $j < i$ are non-zero only in their first $i - 1$ coordinates. Hence, we can assume that we have not yet "looked at" the values r_j for $j \geq i$, and will prove that there is "enough randomness" in r_i to ensure that $\psi(s_i)$ lies in I with small probability.

Formally, let us consider the expression

$$\mathbf{Pr}[\psi(s_i) \in I | \wedge_{j<i} \psi(s_j) \in I \quad \wedge \quad (r_1 = \widehat{r_1} \wedge \dots \wedge r_{i-1} = \widehat{r_{i-1}})].$$

Recall that $\psi(s_i) = \sum_{j=1}^{i} v_{ij} r_j$, and let $Z = \sum_{j<i} v_{ij} \widehat{r_j}$. Note that for $\psi(s_i)$ to fall in the interval I (conditioned on all the previous things), it must be the case that $v_{ii} r_i \in [-Z, l - Z)$. Since the value r_i is independent of all the conditioning, $v_{ii} r_i \sim N(0, v_{ii}^2)$ with $v_{ii} \geq q_i/\eta$. Now, by Fact 23.9, the chance of this is at most $\eta\, l/\sqrt{2\pi} q_i$, and hence

$$\mathbf{Pr}[\psi(s_i) \in I | \wedge_{j<i} \psi(s_j) \in I \quad \wedge \quad (r_1 = \widehat{r_1} \wedge \dots \wedge r_{i-1} = \widehat{r_{i-1}})] \leq \frac{\eta\, l}{\sqrt{2\pi} q_i}.$$

Finally, since this holds for all values of $\widehat{r_j}$ for $j < i$, we can remove the conditioning and claim that

$$\mathbf{Pr}[\psi(s_i) \in I | \wedge_{j<i} \psi(s_j) \in I] \leq \frac{\eta\, l}{\sqrt{2\pi} q_i}.$$

Substituting this into equation (23.4), we get

$$\begin{aligned}\mathbf{Pr}[\psi(S) \subseteq I] &\leq \prod_{i=1}^{k-1} \frac{\eta l}{\sqrt{2\pi} q_i} \\ &\leq \frac{O(\eta\, l)^{k-1}}{\mathsf{Tvol} S}\end{aligned}$$

■

24 Relating tree volumes and the local density

Let us now begin the proof of Claim 21.4.

Claim 21.4. *For all graphs G, we have*

$$\sum_{S \subseteq V : |S| = k} \frac{1}{\mathsf{Tvol} S} \leq n \times O(D \log n)^{k-1}$$

The proof requires the following property, which we will prove in the next class.

Proposition 24.10 *For all $S \subseteq V$, s.t. $|S| = k$, we have*

$$\frac{2^{k-1}}{\mathsf{Tvol} S} \leq \sum_{\pi:[k]\to[k]} \frac{1}{d(v_{\pi(1)}, v_{\pi(2)}) \cdot d(v_{\pi(2)}, v_{\pi(3)}) \cdot \ldots \cdot d(v_{\pi(k-1)}, v_{\pi(k)})}$$

Proof of Claim 21.4: Using Proposition 24.10,

$$\sum_{S:|S|=k} \frac{1}{\mathsf{Tvol} S} \leq \sum_{S:|S|=k} \sum_{\pi:[k]\to[k]} \frac{1}{d(v_{\pi(1)}, v_{\pi(2)}) \cdot d(v_{\pi(2)}, v_{\pi(3)}) \cdot \ldots \cdot d(v_{\pi(k-1)}, v_{\pi(k)})},$$

However, the RHS is at most a sum over sequences from V of length k; i.e.,

$$\sum_{(u_1,u_2,\ldots,u_k)\in V^k} \frac{1}{d(u_1, u_2) \cdot d(u_2, u_3) \cdot \ldots \cdot d(u_{k-1}, u_k)} \tag{24.5}$$

Now, let us fix a tuple $(a_1, \ldots, a_{k-1})$, with $a_i \in \{1, 2, \ldots, \log n\}$, and fix a vertex $u_1 \in V$. Now look at all the sequences of k vertices $(u_1, u_2, \ldots, u_k)$ where the first one is u_1, and $d(u_i, u_{i+1}) \in [2^{a_i}, 2 \cdot 2^{a_i})$ for all i.

We claim that the sum over this subset of sequences is $O(D)^k$. Indeed, recall that D was defined as $\max_{v,r} \frac{|B(v,r)|}{2r}$; i.e., any ball around v of radius r has atmost $D \cdot (2r)$ vertices in it. Therefore, once we choose u_i, there are at most $O(D \cdot 2^{a_i})$ choices for u_{i+1}. Hence, the total number of such sequences is at most $O(D)^k \cdot 2^{\sum a_i}$. But the contribution of any such sequence to the sum is $\frac{1}{\prod_i d(u_i,u_{i+1})}$, which is at most $1/2^{\sum a_i}$. Thus the sum over such sequences is just $O(D)^k$.

Finally, since there are n choices for u_0, and $(\log n)^{k-1}$ choices for vector of a_i's, we get that (24.5) is at most $n \times O(D \log n)^{k-1}$, which proves Claim 21.4. ∎

References

[Fei00] Uriel Feige. Approximating the bandwidth via volume respecting embeddings. *Journal of Computer and System Sciences*, 60(3):510–539, 2000. Also in *Proc. 30th STOC*, 1998, pp. 90–99.