Notes for Week 8/1/19 - 8/8/19

Alden Green

August 5, 2019

1 Bound on Eigenvalue Tail Decay of Neighborhood Graph

Let $\mathcal{D} = [0,1]^d$. We consider two graphs over data on $\mathcal{D}$.

Suppose we observe the random design $X = x_1, \ldots, x_n$ independently sampled from probability measure P supported on $\mathcal{D}$. For any $r \geq 0$, let the kernel function $\eta_r : \mathcal{D} \times \mathcal{D} \to \mathbb{R}_{\geq 0}$ be given by $\eta_r(x,y) = \mathbf{1}(\|x-y\| \leq r)$. Then, let the *r-neighborhood graph* over $x_1, \ldots, x_n$ be the undirected, unweighted graph $G = (V,E)$, where $V = [n]$ and for $i,j \in [n]$, $(i,j) \in E$ if $\eta_r(x_i, x_j) = 1$.

Let B be the incidence matrix associated with G, and let $L = B^T B$ be the corresponding Laplacian. Write $L = V \Lambda V^T$ for the eigen decomposition of L, where $V = (v_1 \ldots v_n)$ is an orthonormal matrix with the eigenvectors of L as its columns, and Λ is a diagonal matrix with entries $\lambda_1 \leq \lambda_2 \leq \ldots \leq \lambda_n$. For a given $s > 0$ and $C > 0$, let $N(C) = \#\left\{k : \lambda_k^s \leq C^2\right\}$. We wish to prove the following result.

Theorem 1. *Let $r \to 0$ as $n \to \infty$ sufficiently slowly so that $r(\log n/n)^{1/d} \to \infty$. Then, for each $s > 0$ and $C \leq \sqrt{2}$,*

$$N(C) \leq n C^{d/s}. \tag{1}$$

with probability tending to one as $n \to \infty$.

2 Theory

To prove Theorem 1, we will make use of another graph on points within $\mathcal{D}$, whose spectral properties are very well-understood. Let ξ be the set of evenly spaced lattice points over $\mathcal{D}$; formally $\xi = \left\{k/n : k \in [\ell]^d\right\}$ where $\ell = n^{1/d}$, and we define $[\ell]^d = \{(\ell_1, \ldots, \ell_d) : \ell_k \in [\ell] \text{ for each } \ell_k\}$. Then, let the grid graph over ξ be given by $\widetilde{G} = (\widetilde{V}, \widetilde{E})$, where $\widetilde{V} = \xi$ and $(\xi_k, \xi_{k'})$ is in $\widetilde{E}$ if $\|\xi_k - \xi_{k'}\|_1 = 1/\ell$.

Let $\widetilde{L}$ be the Laplacian of $\widetilde{G}$, and let $\widetilde{\lambda_1} \leq \ldots \leq \widetilde{\lambda}_n$ be the ordered eigenvalues of $\widetilde{L}$. Write $\widetilde{N}(C) = \#\left\{k : \widetilde{\lambda}_k^s \leq C^2\right\}$. We have that the desired scaling rate (1) holds with respect to the grid graph $\widetilde{G}$.

Lemma 1. *For each s and every $C \leq \sqrt{2}$, we have*

$$\widetilde{N}(C) \leq 2^d \bigl(n C^{d/s} + 1\bigr) \tag{2}$$

Proof. It will be sufficient to show that

$$\lambda_k^s \leq C^2 \Rightarrow \left\lfloor k^{1/d} \right\rfloor^d \leq C^{d/s} n \tag{3}$$

To show this, observe that for any $\tau \in \mathbb{N}$ and $k = \tau^d$, we have

$$\lambda_k \geq 4 \sin^2 \left(\frac{\pi k^{1/d}}{2n^{1/d}} \right) \geq \frac{\pi^2 k^{2/d}}{4n^{2/d}} \wedge 2$$

Therefore, if $\lambda_k^s \leq C^2$ and $C \leq \sqrt{2}$, this implies

$$\frac{\pi^{2s} k^{2s/d}}{4^s n^{2s/d}} \leq C^2$$

and rearranging, we obtain

$$k \leq \frac{C^{d/s} 2^d n}{\pi^d} \leq C^{d/s} n$$

If $k^{1/d}$ is not a natural number, applying the same argument to $k' = \lfloor k^{1/d} \rfloor^d$ yields (3). $\square$

In light of Lemma 1, to prove Theorem 1 it is sufficient to show that $\widetilde{L} \preceq L$, since by the Courant-Fischer min-max theorem, the ordering $\widetilde{L} \preceq L$ implies that $\widetilde{\lambda}_k \leq \lambda_k$ for all $k \in [n]$. The next result details the conditions under which this ordering holds. This condition will be stated with respect to the min-max matching distance between ξ and X, i.e. the minimum over all bijections $T : \xi \to X$ such that

$$\min_T \max_{i \in [n]} \left| T^{-1}(x_i) - x_i \right|$$

Lemma 2. *For any radius r satisfying*

$$r \geq 2 \min_T \max_{i \in [n]} \left| T^{-1}(x_i) - x_i \right| + n^{-1/d} \tag{4}$$

we have that $\widetilde{L} \preceq L$.

Proof. Let $T_\star$ achieve the min-max matching distance, i.e

$$\max_{i \in [n]} \left| T_\star^{-1}(x_i) - x_i \right| = \min_T \max_{i \in [n]} \left| T^{-1}(x_i) - x_i \right|.$$

It will be sufficient to prove that for every pair $(T_\star^{-1}(x_i), T_\star^{-1}(x_j)) \in \widetilde{E}$, the corresponding edge (i, j) is in E. To see this, let A denote the adjacency matrix associated with the neighborhood graph G, and $\widetilde{A}$ the adjacency matrix associated with the grid $\widetilde{G}$. Precisely A is the $n \times n$ matrix with entries $A_{ij} = \eta_r(x_i, x_j)$, and $\widetilde{A}$ is the $n \times n$ matrix with entries $\widetilde{A}_{ij} = \mathbf{1}\{T_\star^{-1}(x_i), T_\star^{-1}(x_j) \in \widetilde{E}\}$. Our goal is to show that, for every $z \in \mathbb{R}^d$, we have

$$z^T L z = \frac{1}{2} \sum_{i,j=1}^n (z_i - z_j)^2 A_{ij} \leq \frac{1}{2} \sum_{i,j=1}^n (z_i - z_j)^2 \widetilde{A}_{ij} = z^T \widetilde{L} z.$$

which certainly holds if $\widetilde{A}_{ij} \leq A_{ij}$ for all $i, j \in [n]$.

Now, assume $(T_\star^{-1}(x_i), T_\star^{-1}(x_j)) \in \widetilde{E}$. This implies

$$\begin{aligned}
\|x_i - x_j\|_2 &\leq \|x_i - T_\star^{-1}(x_i)\|_2 + \|T_\star^{-1}(x_j) - T_\star^{-1}(x_j)\|_2 + \|x_j - T_\star^{-1}(x_j)\|_2 \\
&\leq 2 \max_{i \in [n]} \left| T_\star^{-1}(x_i) - x_i \right| + n^{-1/d} \\
&\leq r.
\end{aligned}$$

so $\eta_r(x_i, x_j) = 1$ and therefore $(i, j) \in E$. $\square$

Finally, the following Lemma demonstrates that, for sufficiently large r, the condition (4) will hold with high probability.

Lemma 3. *Assume P has density p which is bounded above and below uniformly over $\mathcal{D}$; that is, there exist constants $p_{\min}$ and $p_{\max}$ such that*

$$0 < p_{\min} < p(x) < p_{\max} < \infty, \quad \textit{for all } x \in \mathcal{D}.$$

Then, for any $r = r_n$ such that $r(\log(n)/n)^{1/d} \to \infty$, we have that the event

$$r < 2 \min_T \max_{i \in [n]} \left|T^{-1}(x_i) - x_i\right| + n^{-1/d}$$

occurs with probability tending to 0 as $n \to \infty$.

Together Lemmas 1, 2, and 3 imply Theorem 1.