# Notes for Week 12/30/19 - 1/3/19

Alden Green

January 10, 2020

Suppose we observe independent samples $X = \{x_1, \ldots, x_n\}$ from a distribution $P$ with density $p$ supported on $\mathcal{X} = [0,1]^d$. For a given kernel $K_r(x,y) = \frac{1}{r^d} K(\|x - y\|)$, let $G_{n,r}$ denote the random neighborhood graph formed over the samples $X$. Let $L = D - A$ be the corresponding combinatorial Laplacian. In this week's notes, we will show that as long as $r$ is chosen to be sufficiently small, the eigenvalues of $L$ we care about scale at an appropriate rate.

## 1 Random Geometric Graph

In the special case when $K(x) = 1(x \leq r)$ and $G_{n,r}$ is a random geometric graph, we can use results from optimal transport theory to will prove the following Lemma.

**Lemma 1.** *Let $\kappa = n^{2d/(4s+d)}$. Suppose $p$ is uniformly bounded away from $0$ and $\infty$ on its support,*

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

*Then there exists a constant $c$ such that for any $r \leq n^{-4/((2+d)(4s+d))}$, the eigenvalue $\lambda_\kappa$ is lower bounded,*

$$\lambda_\kappa \geq c\kappa^{2/d} n r^{2+d} \tag{1}$$

*with probability $1 - o(n^{-1})$ as $n \to \infty$.*

Lemma 1 establishes that eigenvalues of the graph Laplacian, suitably normalized, scale comparably to the eigenvalues of the continuum Laplacian operator $\mathcal{D}f := \mathrm{div}(\nabla f)$. Along with an upper bound on the graph Sobolev semi-norm $R_{s,n}(f)$ (established in the *12.20.19 Notes*), this allows us to argue that if $f$ is a smooth function with large $\mathcal{L}_2$ norm, then the projection of a function $f$ onto the eigenvectors $v_1, \ldots, v_\kappa$ is still quite large (measured in empirical norm); or, equivalently, the bias induced by projection onto these eigenvectors is small.

## 2 Higher-order kernels.

When $K$ is the uniform kernel, the aforementioned upper bound on the graph Sobolev semi-norm holds only when $s = 1$ or $s = 2$. For larger values of $s$, we require that $K$ be a higher-order kernel, that is that $K$ satisfy

$$\int K(x)\,dx = 1, \quad \int x^\ell K(x)\,dx = 0 \ \text{ for } \ell = 1, \ldots, s-1$$

It seems difficult to extend the techniques used to prove Lemma 1 to this setting. We will therefore use a different approach to prove the following Lemma.

**Lemma 2.** *Let $\kappa = n^{2d/(4s+d)}$. Suppose $p$ is uniformly bounded away from $0$ and $\infty$ on its support,*

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

*Let $K$ satisfy the following conditions:*

TODO: *Specify conditions.*

*Then there exists a constant $c$ such that the eigenvalue $\lambda_\kappa$ is lower bounded,*

$$\lambda_\kappa \geq c\kappa^{2/d} n r^{2+d} \tag{2}$$

*with probability $1 - o(n^{-1})$ as $n \to \infty$.*

# 3 Proofs

## 3.1 Proof of Lemma 1

We prove Lemma 1 by comparing $G_{n,r}$ to the tensor product of a $d$-dimensional lattice and a complete graph. The latter is a highly structured graph with known eigenvalues, which as we will see are sufficiently lower bounded for our purposes.

Let $\widetilde{r} = r/(3(\sqrt{d}+1)), M = (1/\widetilde{r})^d, N = n\widetilde{r}^d$. Assume without loss of generality that $M$ and $N$ are integers. Additionally, for $t = n^{1/d}$ and $m = M^{1/d}$ let

$$\overline{X} = \left\{\frac{1}{t}(k_1, \ldots, k_d) : k \in [t]^d\right\}, \quad \overline{Z} = \left\{\frac{1}{m}(j_1, \ldots, j_d) : j \in [m]^d\right\}.$$

For a given $\overline{z}_j \in \overline{Z}$, we write $Q(z_j) = m^{-1}[j_1 - 1, j_1] \times \cdots \times m^{-1}[j_d - 1, j_d]$ for the cube of side length $1/m$ with $z_j$ at one corner.

Consider the graph $H = (\overline{X}, E_H)$, where $(\overline{x}_k, \overline{x}_\ell) \in E_H$ if

$$\text{there exists } \overline{z}_i, \overline{z}_j \in \overline{Z} \text{ such that } \overline{x}_k \in Q(\overline{z}_i),\ \overline{x}_\ell \in Q(\overline{z}_j), \text{ and } \|i - j\|_1 \leq 1.$$

On the one hand $H \cong \overline{G}_d^M \otimes K_N$ where $\overline{G}_d^M$ is the $d$-dimensional lattice on $M$ nodes, and $K_N$ is the complete graph on $N$ nodes. On the other hand, we now show that with high probability $G_{n,r} \succeq H$. If $(\overline{x}_k, \overline{x}_\ell) \in E_H$, then there exist $\overline{z}_i, \overline{z}_j$ such that

$$\|\overline{x}_k - \overline{x}_\ell\|_2 \leq m^{-1} + \|\overline{x}_k - \overline{z}_i\|_2 + \|\overline{x}_\ell - \overline{z}_j\|_2 \leq \widetilde{r}(1 + \sqrt{d}) = r/3.$$

By Theorem 1.1 of Garcia Trillos and Slepcev, with probability at least $1 - n^{-1}$ there exists a bijection $\pi : \overline{X} \to X$ such that

$$\max_{k \in [t]^d} |\overline{x}_k - \pi(\overline{x}_k)| \leq c\left(\frac{\log n}{n}\right)^{1/d} \tag{3}$$

Assuming (3) holds, if $(\overline{x}_k, \overline{x}_\ell) \in E_H$, then for sufficiently large $n$

$$\|\pi(\overline{x}_k) - \pi(\overline{x}_\ell)\|_2 \leq 2c\left(\frac{\log n}{n}\right)^{1/d} + \frac{r}{3} \leq r,$$

implying that $(\pi(\overline{x}_k), \pi(\overline{x}_\ell)) \in E$. Therefore, $G_{n,r} \succeq \overline{G}_d^M \otimes K_N$ whenever (3) holds.

The eigenvalues of lattices and complete graphs are known to satisfy, respectively

$$\lambda_k(\overline{G}_d^M) \geq \frac{k^{2/d}}{M^{2/d}} \text{ for } k = 0, \ldots, M-1, \quad \text{and } \lambda_j(K_N) \geq N\mathbf{1}\{j > 0\} \text{ for } j = 0, \ldots, N-1.$$

and by standard facts regarding the eigenvalues of tensor product graphs (reference), we have that the spectrum $\Lambda(H)$ satisfies

$$\Lambda(H) = \Big\{ N\lambda_k(\overline{G}_d^M) + M\lambda_j(K_N) : \text{for } k = 0, \dots, M-1 \text{ and } j = 0, \dots, N-1 \Big\}$$

For all $j = 1, \dots, N-1$, we have that $M\lambda_j(K_N) = MN = n$. Therefore,

$$\begin{aligned}
\lambda_\kappa(H) &\geq \{n \wedge N\lambda_\kappa(\overline{G}_d^M)\} \\
&\geq \{n \wedge n\widetilde{r}^d \frac{\kappa^{2/d}}{M^{2/d}}\} \\
&\geq \{n \wedge (3\sqrt{d}+3)^{-(2+d)} n r^{d+2} \kappa^{2/d}\} \\
&\geq (3\sqrt{d}+3)^{-(2+d)} n r^{d+2} \kappa^{2/d},
\end{aligned}$$

where the last inequality can be verified by a quick calculation in light of $\kappa = n^{2d/(4s+d)}$ and $r \leq n^{-4/((2+d)(4s+d))}$. Since we've already shown $\lambda_\kappa(G_{n,r}) \geq \lambda_\kappa(H)$ with probability $1 - o(n^{-1})$, this completes the proof of Lemma 1.

## 3.2 Proof of Lemma 1

Let $D_{\min} = \min_{i \in [n]} D_{ii}$ be the minimum degree of any sample $x \in X$ in the graph $G_{n,r}$. We have that

$$\lambda_\kappa(L) \geq D_{\min} - \lambda_{n-\kappa}(A)$$