Lipschitz regularity of graph Laplacians on random data clouds

(work by Jeff Calder, Nicolas Garcia Trillos and Marta Lewicka)
Presenter: Truong-Son Van

"20/2"/20/2020

CNA Working Group, Carnegie Mellon University
Overview

- Introduction
- Some results of the paper
- Regularity by stochastic coupling
As in previous talks,

- \((\mathcal{M}, d_\mathcal{M})\) is a smooth closed \(m\)-dimensional Riemannian manifold embedded in \(\mathbb{R}^d\).

- \(\eta : [0, \infty) \to [0, \infty)\) is non-decreasing, supported on \([0, 1]\) s.t.
  \[
  \int_{\mathbb{R}^m} \eta(|w|) \, dw = 1
  \]

- Denote
  \[
  \sigma_\eta \stackrel{\text{def}}{=} \int_{\mathbb{R}^m} \langle w, e_1 \rangle^2 \eta(|w|) \, dw.
  \]

For convenience let \(\sigma_\eta = 1\) in this talk.

- \(\rho : \mathcal{M} \to (0, \infty)\) is a density on \(\mathcal{M}\).
Intro: notations

- $\mathcal{X}_n = \{x_1, \ldots, x_n\}$ i.i.d samples from $\rho$

- Graph Laplacians (of length scale $\varepsilon$) $\Delta_{\varepsilon, \mathcal{X}_n} : L^2(\mathcal{X}_n) \to L^2(\mathcal{X}_n)$

$$
\Delta_{\varepsilon, \mathcal{X}_n} f(x_i) \overset{\text{def}}{=} \frac{1}{n\varepsilon^{m+2}} \sum_{j=1}^{n} \eta\left(\frac{|x_i - x_j|}{\varepsilon}\right)(f(x_i) - f(x_j))
$$

- Nonlocal Laplacians (of length scale $\varepsilon$) $\Delta_{\varepsilon} : L^2(\mathcal{M}) \to L^2(\mathcal{M})$

$$
\Delta_{\varepsilon} f(x) \overset{\text{def}}{=} \frac{1}{\varepsilon^{m+2}} \int_{\mathcal{M}} \eta\left(\frac{d_{\mathcal{M}}(x, y)}{\varepsilon}\right)(f(x) - f(y)) \, dV_{\mathcal{M}}(y)
$$

- Weighted Laplace-Beltrami operator on $\mathcal{M}$

$$
\Delta_{\mathcal{M}} f(x) = \Delta_{\mathcal{M}, \rho} f(x) \overset{\text{def}}{=} -\frac{\sigma \eta}{2\rho} \text{div}(\rho^2 \nabla f).
$$

- Unweighted Laplace-Beltrami operator is denoted by $\Delta$
Intro: notations

Note: \[
\frac{1}{2\rho} \text{div} \left( \rho^2 \nabla f \right) = \frac{1}{2\rho} \left< \rho \nabla \rho, \nabla f \right> + \frac{\rho^2}{2\rho} \Delta f
\]

\[
= \rho \left< \frac{\nabla \rho}{\rho}, \nabla f \right> + \frac{\rho}{2} \Delta f
\]

\[
= \rho \left< \nabla \log \rho, \nabla f \right> + \frac{\rho}{2} \Delta f
\]
Some results: main focus of today's talk

Things to remember: always consider \((\frac{\ln n}{n})^{1/m+4} \ll \varepsilon \ll 1\) length scale

**Theorem (Global \(\varepsilon\)-Lipschitz (Theorem 2.1))**

With probability at least \(1 - C\varepsilon^{-6m} \exp(-cn\varepsilon^{m+4})\), we have

\[
|f(x_i) - f(x_j)| \leq C(\|f\|_{L^\infty(\mathcal{X}_n)} + \|\Delta_{\varepsilon,\mathcal{X}_n} f\|_{L^\infty(\mathcal{X}_n)}) (d_{\mathcal{M}}(x_i, x_j) + \varepsilon)
\]

for all \(f \in L^2(\mathcal{X}_n)\) and all \(x_i, x_j \in \mathcal{X}_n\).

Note: There is also an interior Lipschitz regularity with length scale bigger than \(\varepsilon\).

Q: remove \(\varepsilon\) on the right hand side due to the lack of information below \(\varepsilon\) length scale?
Related works

- On discrete regularity: Kuo and Trudinger, “Schauder Estimates for Fully Nonlinear Elliptic Difference Operators” (2002) – only on $Z^d$ (difficulties: definition of derivative, manifold)

- On related application by CMU local: Pegden and Smart, “Convergence of the Abelian sandpile” (2013), “Stability of patterns in the Abelian sandpile” (2020)
Note: consider $\Delta_{\varepsilon, \mathcal{X}} f = \lambda f$ where $\|f\|_2 = 1$,

$$|f(x_i) - f(x_j)| \leq C(\|f\|_\infty + \lambda \|f\|_\infty)(d_M(x_i, x_j) + \varepsilon).$$

When $\varepsilon \leq c/(\lambda + 1)$, work a bit harder using concentration of measure type inequalities to get

$$|f(x_i) - f(x_j)| \leq C(\lambda + 1)^{m+1}(d_M(x_i, x_j) + \varepsilon),$$

and

$$\|f\|_\infty \leq C(\lambda + 1)^{m+1}\|f\|_1$$

with high probability.
Some results: consequences of global $\varepsilon$-Lipschitz

Denote

$$[f]_{\varepsilon, X_n} \overset{\text{def}}{=} \max_{x, y \in X_n} |f(x) - f(y)|.$$  

**Theorem (Convergence rate of eigenvectors of graph Laplacian (Theorem 2.6))**

Suppose that $f$ is a normalized eigenvector of $\Delta_{\varepsilon, X_n}$ with eigenvalue $\lambda$, i.e.

$$\|f\|_{L^2(X_n)} = 1.$$  

Then with probability at least $1 - C(n + \varepsilon^{-6m})\exp(-cn\varepsilon^{m+4})$, there exists a normalized eigenfunction $\tilde{f}$ of $\Delta_M$ such that

$$\|f - \tilde{f}\|_{L^\infty(X_n)} + [f - \tilde{f}]_{\varepsilon, X_n} \leq C\lambda\varepsilon.$$  

Idea of Proof of convergence of eigenvectors

By an earlier work of Calder and Trillos, with probability $1 - C n \exp(-c n \varepsilon^{m+4})$, there exists a normalized eigenfunction $\tilde{f}$ of $\Delta_M$ with eigenvalue $\tilde{\lambda}$ s.t.

$$|\lambda - \tilde{\lambda}| + \|f - \tilde{f}\|_{L^2(\mathcal{X}_n)} \leq C \varepsilon.$$ 

By the same work, with probability $1 - 2n \exp(-c n \varepsilon^{m+4})$,

$$\|\Delta_M \tilde{f} - \Delta_{\varepsilon, \mathcal{X}_n} \tilde{f}\|_{L^\infty(\mathcal{X}_n)} \leq C \varepsilon.$$ 

Let $g \overset{\text{def}}{=} f - \tilde{f}$. Then

$$\Delta_{\varepsilon, \mathcal{X}_n} g = (\Delta_{\varepsilon, \mathcal{X}_n} f - \Delta_M \tilde{f}) + (\Delta_M \tilde{f} - \Delta_{\varepsilon, \mathcal{X}_n} \tilde{f}) = \lambda (f - \tilde{f}) + (\lambda - \tilde{\lambda}) \tilde{f} + O(\varepsilon)$$

and so

$$\|\Delta_{\varepsilon, \mathcal{X}_n} g\|_{L^\infty(\mathcal{X}_n)} \leq \lambda \|g\|_{L^\infty(\mathcal{X}_n)} + C \varepsilon (1 + \|\tilde{f}\|_{L^\infty(M)})$$.
Work a bit harder to get $\|g\|_{L^\infty(X_n)} \leq C\varepsilon$ and so

$$\|\Delta_{\varepsilon,X_n}g\|_{L^\infty(X_n)} \leq C\varepsilon.$$ 

Thus, by global $\varepsilon$-Lipschitz, we have

$$|g(x_i) - g(x_j)| \leq C(\|g\|_{L^\infty(X_n)} + \|\Delta_{\varepsilon,X_n}g\|_{L^\infty(X_n)})(d_M(x_i, x_j) + \varepsilon) \leq C\varepsilon(d_M(x_i, x_j) + \varepsilon).$$

Using union bound on the complements these events, the result follows.
Idea of Proof of Global $\varepsilon$-Lipschitz

Discrete $\rightarrow$ Nonlocal $\rightarrow$ Local

Define the interpolation operator $I_{\varepsilon, x_n} : L^2(\mathcal{X}_n) \rightarrow L^2(\mathcal{M})$

$$I_{\varepsilon, x_n} f(x) \overset{\text{def}}{=} \frac{1}{d_{\varepsilon, x_n}(x)} \frac{1}{n} \sum_{i=1}^{n} \frac{1}{\varepsilon^m} \eta \left( \frac{|x - x_i|}{\varepsilon} \right) f(x_i)$$

where the degree $d_{\varepsilon, x_n}$ is defined as

$$d_{\varepsilon, x_n}(x) \overset{\text{def}}{=} \frac{1}{n} \sum_{i=1}^{n} \frac{1}{\varepsilon^m} \eta \left( \frac{|x - x_i|}{\varepsilon} \right).$$
Lemma (Discrete to nonlocal)

With probability at least $1 - C\varepsilon^{-6m} \exp(-cn\varepsilon^{m+4})$, we have

$$|\Delta_\varepsilon(I_\varepsilon,x f)(x)| \leq C(\|\Delta_\varepsilon,\delta f\|_{L^\infty(x_n \cap B(x,\varepsilon))} + \text{osc}_{x_n \cap B(x,2\varepsilon)} f)$$

for every $f \in L^2(x_n)$ and $x \in \mathcal{M}$. 
Idea of Proof of Global $\varepsilon$-Lipschitz (cont.)

Define “averaging” operators $A_\varepsilon$ and $\bar{A}_\varepsilon$

$$A_\varepsilon f(x) \overset{\text{def}}{=} \frac{1}{d_\varepsilon(x)} \int_{B_M(x, \varepsilon)} \frac{1}{\varepsilon^m} \eta \left( \frac{d_M(x, y)}{\varepsilon} \right) f(y) \rho(y) dV_M(y)$$

$$d_\varepsilon \overset{\text{def}}{=} \int_{B_M(x, \varepsilon)} \frac{1}{\varepsilon^m} \eta \left( \frac{d_M(x, y)}{\varepsilon} \right) \rho(y) dV_M(y)$$

For $B_x(0, t) \subseteq T_x M$,

$$\bar{A}_\varepsilon f(x) \overset{\text{def}}{=} \int_{B_x(0,1)} \eta(|\omega|)(1 + \varepsilon \langle w, \nabla \log \rho(0) \rangle) f(\varepsilon w) \, dw$$

Fact: for small enough $\varepsilon$, with high probability, for $C^2(M)$ functions

$$I_{\varepsilon, x, \eta} f(x) \overset{O(\varepsilon^2)}{\approx} f(x) \overset{O(\varepsilon^2)}{\approx} A_\varepsilon f(x) \overset{O(\varepsilon^2)}{\approx} \bar{A}_\varepsilon f(x).$$

The last approximation only requires bounded Borel functions.
By the discrete-to-nonlocal lemma and the above discussion, we will be in business if we have a similar estimate for the non-local Laplacian on $\mathcal{M}$, i.e.,

$$|f(x) - f(y)| \leq C(\|f\|_{L^\infty(\mathcal{M})} + \|\Delta_\varepsilon f\|_{L^\infty(\mathcal{M})})(d_{\mathcal{M}}(x, y) + \varepsilon)$$

for all bounded Borel function $f$ on $\mathcal{M}$.

In particular, hit this estimate on $I_{\varepsilon, \chi_n} f(x), A_\varepsilon f(x)$ and use triangle inequality.
Nonlocal global $\varepsilon$-Lipschitz by stochastic coupling

- Observe that $Bf(x) \overset{\text{def}}{=} \frac{1}{\rho} \Delta_M f(x) = \Delta f(x) + 2 \langle \nabla f, \nabla \log \rho \rangle_x$.

- This suggest that one should look at an Itô process with drift $\nabla \log \rho$

$$dX_t = \nabla \log \rho(X_t) \, dt + dB_t.$$  

(SDE)

Intuition: for small $t \approx \varepsilon^2$, $|X_t - x| \approx O(\varepsilon)$ most of the time. So

$$A_\varepsilon f(x) - f(x) \approx \tilde{A}_\varepsilon f(x) - f(x)$$

Taylor expand $f(\varepsilon w)$
term in $\tilde{A}_\varepsilon f$

$$\approx \varepsilon^2 \left( \langle \nabla f(x), \nabla \log \rho(x) \rangle + \frac{1}{2} \Delta f(x) \right) = \varepsilon^2 2Bf(x)$$

(KEY)

$$\approx E_x \int_0^t \langle f'(X_s), \nabla \log \rho(X_s) \rangle ds + \frac{1}{2} E_x \int_0^t \Delta f(X_s) \, ds$$

$$= E_x (f(X_t) - f(x))$$
Given two processes $X_t$, $Y_t$ such that $X_0 = x$, $Y_0 = y$ and a stopping time

$$\tau \overset{\text{def}}{=} \inf\{t > 0 : |X_t - Y_t| > r \text{ or } |X_t - Y_t| < \varepsilon\}.$$

Suppose that $f(x) - f(y) > 0$ and $f(X_t) - f(Y_t)$ is a submartingale, then

$$f(x) - f(y) \leq E(f(X_\tau) - f(Y_\tau))$$

$$= E(f(X_\tau) - f(Y_\tau); |X_\tau - Y_\tau| > r) + E(f(X_\tau) - f(Y_\tau); |X_\tau - Y_\tau| < \varepsilon)$$

$$\leq 2\|f\|_\infty P(|X_\tau - Y_\tau| > r) + \sup\{|f(a) - f(b)| : a, b \in B(x, r), |a - b| < \varepsilon\}$$

$$= 2\|f\|_\infty P(|X_\tau - Y_\tau| > r) + \Theta(r, \varepsilon).$$

Game: find $X_t$ and $Y_t$ so that the above inequality is nice.
Nonlocal global $\varepsilon$-Lipschitz by stochastic coupling

Suppose $p$ = uniform density, $M = IR^m$.

Let $X_t$ be B.M.

Reflect $X_t$ along the $x$-axis to find $Y_t$.

Then $P(\|X_t - Y_t\| > r) \approx \frac{|x - y|}{r}$.

Think: 1-D BM hopping left and right.

$P(B_t = -a) = \frac{b}{a+b}$.

$\Theta(r, \varepsilon) \approx O(\varepsilon)$ by local estimates.
Nonlocal global $\varepsilon$-Lipschitz by stochastic coupling

Unfortunately, this is very hard. But

$$M_t = f(X_t) - f(Y_t) + 2\langle f(X_t) - f(Y_t) \rangle_t$$

is always a submartingale!

By (KEY), it is wise to cook up a process that follows the dynamic of (SDE) such that

$$\langle f(X_t) - f(Y_t) \rangle_t \approx \| A_\varepsilon f(x) - f(x) \|_{L^\infty(B(x,r))}.$$
- Let $X_t$ be a discrete approximation of size $\varepsilon$ of the stochastic flow (SDE).

- Construct $Y_t$ by reflect $X_t$ along the flow. (Draw picture)