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

Intro: notations

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

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

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

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

$n \to \infty$

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

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

$\varepsilon \to 0$

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

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

- Unweighted Laplace-Beltrami operator is denoted by Δ

Intro: notations

Note:

$$\frac{1}{2\rho}\operatorname{div}(\rho^2 \nabla f) = \frac{1}{2\rho}\langle 2\rho\nabla\rho, \nabla f\rangle + \frac{\rho^2}{2\rho}\Delta f$$

$$= \rho\left\langle \frac{\nabla\rho}{\rho}, \nabla f\right\rangle + \frac{\rho}{2}\Delta f$$

$$= \rho\,\langle \nabla \log\rho, \nabla f\rangle + \frac{\rho}{2}\Delta f$$

Some results: main focus of today's talk

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

Theorem (Global ε-Lipschitz (Theorem 2.1))
With probability at least $1 - C\varepsilon^{-6m}\exp(-cn\varepsilon^{m+4})$, *we have*

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

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 ε.

Q: remove ε on the right hand side due to the lack of information below ε 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)

Some results: consequences of Global ε-Lipschitz

Note: consider $\Delta_{\varepsilon,\mathcal{X}_n} f = \lambda f$ where $\|f\|_2 = 1$,

$$|f(x_i) - f(x_j)| \leqslant C(\|f\|_\infty + \lambda\|f\|_\infty)(d_{\mathcal{M}}(x_i, x_j) + \varepsilon)\,.$$

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

$$|f(x_i) - f(x_j)| \leqslant C(\lambda + 1)^{m+1}(d_{\mathcal{M}}(x_i, x_j) + \varepsilon)\,,$$

and

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

with high probability.

Denote

$$[f]_{\varepsilon,\mathcal{X}_n} \overset{\text{def}}{=} \max_{x,y\in\mathcal{X}_n} \frac{|f(x)-f(y)|}{d_{\mathcal{M}}(x,y)+\varepsilon}\,.$$

Theorem (Convergence rate of eigenvectors of graph Laplacian (Theorem 2.6))
Suppose that f is a normalized eigenvector of $\Delta_{\varepsilon,\mathcal{X}_n}$ with eigenvalue λ, i.e. $\|f\|_{L^2(\mathcal{X}_n)} = 1$. Then with probability at least $1 - C(n+\varepsilon^{-6m})\exp(-cn\varepsilon^{m+4})$, there exists a normalized eigenfunction $\tilde{f}$ of $\Delta_{\mathcal{M}}$ such that

$$\|f-\tilde{f}\|_{L^\infty(\mathcal{X}_n)} + [f-\tilde{f}]_{\varepsilon,\mathcal{X}_n} \leqslant C_\lambda \varepsilon\,.$$

Idea of Proof of convergence of eigenvectors

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

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

$$\Delta_{\varepsilon,\mathcal{X}_n} f = \lambda f$$

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

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

$$\Delta_{\mathcal{M}} \tilde{f} = \tilde{\lambda} \tilde{f}$$

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

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

and so

$$\|\Delta_{\varepsilon,\mathcal{X}_n} g\|_{L^\infty(\mathcal{X}_n)} \leqslant \lambda\|g\|_{L^\infty(\mathcal{X}_n)} + C\varepsilon(1 + \|\tilde{f}\|_{L^\infty(\mathcal{M})})\,.$$

Idea of Proof of convregence of eigenvectors (cont.)

Work a bit harder to get $\|g\|_{L^\infty(\mathcal{X}_n)} \leqslant C\varepsilon$ and so

$$\|\Delta_{\varepsilon,\mathcal{X}_n} g\|_{L^\infty(\mathcal{X}_n)} \leqslant C\varepsilon \,.$$

Thus, by global ε-Lipschitz, we have

$$|g(x_i) - g(x_j)| \leqslant C(\|g\|_{L^\infty(\mathcal{X}_n)} + \|\Delta_{\varepsilon,\mathcal{X}_n} g\|_{L^\infty(\mathcal{X}_n)})(d_{\mathcal{M}}(x_i, x_j) + \varepsilon) \leqslant C\varepsilon(d_{\mathcal{M}}(x_i, x_j) + \varepsilon) \,.$$

Using union bound on the complements these events, the result follows. $\square$

Idea of Proof of Global ε-Lipschitz

Discrete $\to$ Nonlocal $\to$ Local

Define the interpolation operator $\mathcal{I}_{\varepsilon,\mathcal{X}_n} : L^2(\mathcal{X}_n) \to L^2(\mathcal{M})$

$$\mathcal{I}_{\varepsilon,\mathcal{X}_n} f(x) \stackrel{\text{def}}{=} \frac{1}{d_{\varepsilon,\mathcal{X}_n}(x)} \frac{1}{n} \sum_{i=1}^{n} \frac{1}{\varepsilon^m} \eta\Big(\frac{|x - x_i|}{\varepsilon}\Big) f(x_i)$$

where the degree $d_{\varepsilon,\mathcal{X}_n}$ is defined as

$$d_{\varepsilon,\mathcal{X}_n}(x) \stackrel{\text{def}}{=} \frac{1}{n} \sum_{i=1}^{n} \frac{1}{\varepsilon^m} \eta\Big(\frac{|x - x_i|}{\varepsilon}\Big) .$$

Idea of Proof of Global ε-Lipschitz (cont.)

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

$$|\Delta_\varepsilon(\mathcal{I}_{\varepsilon,\mathcal{X}} f)(x)| \leqslant C(\|\Delta_{\varepsilon,\mathcal{X}_n} f\|_{L^\infty(\mathcal{X}_n \cap B(x,\varepsilon))} + \operatorname{osc}_{\mathcal{X}_n \cap B(x,2\varepsilon)} f)$$

for every $f \in L^2(\mathcal{X}_n)$ *and* $x \in \mathcal{M}$.

Idea of Proof of Global ε-Lipschitz (cont.)

Define "averaging" operators A_ε and $\bar{A}_\varepsilon$

More like interpolating

$$A_\varepsilon f(x) \stackrel{\text{def}}{=} \frac{1}{d_\varepsilon(x)} \int_{B_{\mathcal{M}}(x,\varepsilon)} \frac{1}{\varepsilon^m} \eta\Big(\frac{d_{\mathcal{M}}(x,y)}{\varepsilon}\Big) f(y)\rho(y) dV_{\mathcal{M}}(y)$$

$$d_\varepsilon \stackrel{\text{def}}{=} \int_{B_{\mathcal{M}}(x,\varepsilon)} \frac{1}{\varepsilon^m} \eta\Big(\frac{d_{\mathcal{M}}(x,y)}{\varepsilon}\Big) \rho(y) dV_{\mathcal{M}}(y)$$

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

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

Fact: for small enough ε, with high probability, for $C^2(\mathcal{M})$ functions

$$\mathcal{I}_{\varepsilon,\mathcal{X}_n} f(x) \stackrel{O(\varepsilon^2)}{\approx} f(x) \stackrel{O(\varepsilon^2)}{\approx} A_\varepsilon f(x) \stackrel{O(\varepsilon^2)}{\approx} \bar{A}_\varepsilon f(x)\,.$$

The last approximation only requires bounded Borel functions.

Idea of Proof of Global ε-Lipschitz (cont.)

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.,

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

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

In particular, hit this estimate on $\mathcal{I}_{\varepsilon,\mathcal{X}_n} f(x), A_\varepsilon f(x)$ and use triangle inequality.

Nonlocal global ε-Lipschitz by stochastic coupling

- Observe that $Bf(x) \overset{\text{def}}{=} \frac{1}{\rho}\Delta_{\mathcal{M}} f(x) = \Delta f(x) + 2\langle \nabla f, \nabla \log \rho \rangle_x$.
- This suggest that one should look at an Itô process with drift $\nabla \log \rho$

$$dX_t = \nabla \log \rho(X_t)\, dt + dB_t\,. \tag{SDE}$$

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

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

$$\begin{aligned} A_\varepsilon f(x) - f(x) &\approx \bar{A}_\varepsilon f(x) - f(x) \\ &\approx \varepsilon^2 \Big(\langle \nabla f(x), \nabla \log \rho(x) \rangle + \frac{1}{2}\Delta f(x) \Big) = \varepsilon^2 2Bf(x) \\ &\approx \boldsymbol{E}^x \int_0^t \langle f'(X_s), \nabla \log \rho(X_s) \rangle ds + \frac{1}{2}\boldsymbol{E}^x \int_0^t \Delta f(X_s)\, ds \\ &= \boldsymbol{E}^x (f(X_t) - f(x)) \end{aligned} \tag{KEY}$$

Nonlocal global ε-Lipschitz by stochastic coupling

Given two processes X_t, Y_t such that $X_0 = x$, $Y_0 = y$ and a stopping time

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

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

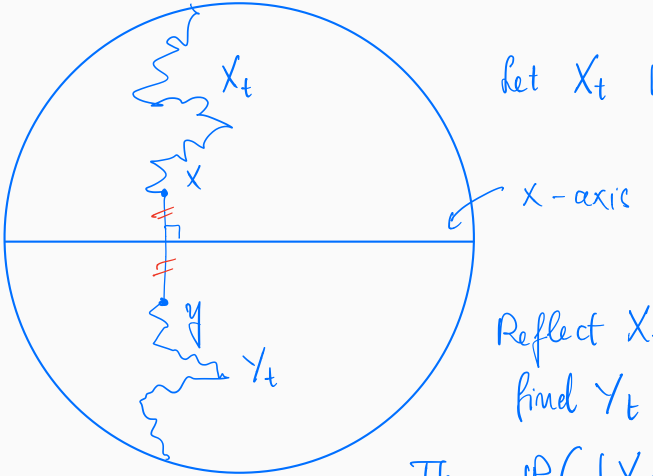
$$\begin{aligned}
f(x) - f(y) &\leqslant \boldsymbol{E}(f(X_\tau) - f(Y_\tau)) \\
&= \boldsymbol{E}(f(X_\tau) - f(Y_\tau); |X_\tau - Y_\tau| > r) + \boldsymbol{E}(f(X_\tau) - f(Y_\tau); |X_\tau - Y_\tau| < \varepsilon) \\
&\leqslant 2\|f\|_\infty \boldsymbol{P}(|X_\tau - Y_\tau| > r) + \sup\{|f(a) - f(b)| : a, b \in B(x, r), |a - b| < \varepsilon\} \\
&= 2\|f\|_\infty \underbrace{\boldsymbol{P}(|X_\tau - Y_\tau| > r)}_{|x-y|/r} + \underbrace{\Theta(r, \varepsilon)}_{O(\varepsilon)}\,.
\end{aligned}$$

want:

Game: find X_t and Y_t so that the above inequality is nice.

Nonlocal global ε-Lipschitz by stochastic coupling

* Suppose ρ = uniform density, $M = \mathbb{R}^m$.



Let X_t be B.M.

Reflect X_t along the x-axis to find Y_t

Then $\mathbb{P}(|X_\tau - Y_\tau| > r) \approx \frac{|x-y|}{r}$

$\Theta(r, \varepsilon) \approx O(\varepsilon)$

By local estimates.

Think: 1-D BM hopping left and right

$\mathbb{P}(B_\tau = -a) = \frac{b}{a+b}$

Nonlocal global ε-Lipschitz by stochastic coupling

in general

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

$$\langle f(X_t) - f(Y_t)\rangle_t \approx \|A_\varepsilon f(x) - f(x)\|_{L^\infty(B(x,r))}\,.$$

Nonlocal global ε-Lipschitz by stochastic coupling

In general,

- Let X_t be a discrete approximation of size ε of the stochastic flow (SDE) .

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