Scaling limits of graph Laplacians

Ernesto Garcia

LAAS - CNRS

01/02/2024

Plan

Definition/Notation

A **weighted graph** *will be* $G = (V, w)$ *where:*

- $V = \{x_1, \ldots, x_n\}$ *is the set of nodes and*
- $w : V \times V \to [0, +\infty)$ *is some positive* **symmetric** *function and* W *is the (symmetric) matrix* $(w(x, y))_{x,y \in V} \in \mathbb{R}^{n \times n}$.
- *We take the set of edges* E *to be the set of pairs of nodes* (x_i, x_j), $i \neq j$ *such that* $w(x_i, x_j) > 0$.
- *Most of the time will assume that the graph in question is connected.*

Definition/Notation

The **(out-)degree** *of a vertex* $x \in V$ *is*

$$d(x) = \sum_{y \in V} w(x, y),$$

the **degree matrix** *is* $D = diag(d(x_1), \dots, d(x_n)) \in \mathbb{R}^{n \times n}$*. Also, for any set of nodes* $A \subset V$*, its* **volume** *is the sum of the degrees of nodes in* A*:*

$$vol(A) = \sum_{x \in A} d(x).$$

Graph cutting problem

Problem

Provide a partition of V into k (predefined or not) subsets using W as a similarity matrix (large $w(x, y)$ implies x, y "similar" or "close" to each other).

Simple random walk on a weighted graph

Definition/Notation

Let P denote the row stochastic matrix $D^{-1}W$. Its entries are given by $p(x,y) = \frac{w(x,y)}{d(x)},\ x, y \in V$.

Proposition

In the symmetric case the matrix P has stationary measure $\pi_P = \frac{1}{vol(V)}(d(x_1), \ldots, d(x_n))$*, that is:*

$$\pi_P \cdot P = \pi_P.$$

Moreover, they generate a reversible Markov chain, i.e:

$$\pi_P(x)p(x,y) = \pi_P(y)p(y,x),\ x, y \in V.$$

Definition

The inner product on $L^2(V)$ induced by π_P is

$$\langle f, g \rangle_{\pi_P} = \sum_{x \in V} \pi_P(x) f(x) g(x)$$

for $f, g : V \to \mathbb{R}$.

Corollary

Reversibility is equivalent to P being self-adjoint, as an operator on $L^2(V)$, for $\langle , \rangle_{\pi_P}$:

$$\langle Pf, g \rangle_{\pi_P} = \langle f, Pg \rangle_{\pi_P}, \ \forall f, g : V \to \mathbb{R}.$$

Hence, if the weights are symmetric, P is diagonalizable over $\mathbb{R}$.

Graph Laplacians

Definition

The graph Laplacians we are interested in are the following $n \times n$ matrices:

- **unnormalized graph Laplacian**:

$$L = D - W;$$

- **random walk graph Laplacian**:

$$L_{rw} = D^{-1}L = I - D^{-1}W = I - P.$$

Main properties

Proposition

The following properties hold (under symmetry assumption):

- *For any function* $f : V \to \mathbb{R}$ *we have:*

$$\langle f, Lf \rangle = \sum_{x,y \in V} w(x,y)(f(y) - f(x))^2.$$

- L *is a positive semi-definite, symmetric matrix.*
- 0 *is an eigenvalue of* L *and its multiplicity is equal to the number of connected components of the graph.*

Proposition

Since P *is diagonalizable so is* L_{rw} *and its spectrum lies in* $[0, 2)$. *The eigendecomposition of* L_{rw} *is central in some balanced graph-cut problems ([vL07]).*

Random geometric graphs

Let D be some precompact domain in $\mathbb{R}^d$ and $p : D \to \mathbb{R}$ some probability density function on D. Consider n independent samples $V_n = \{x_1, \ldots, x_n\}$ with distribution p. For each n we may build weighted graphs in several ways. Some examples are:

- **gaussian weights**.
- $\varepsilon-$**graph**, where we take $\varepsilon_n \to 0$ and connect nodes x and y if $||x - y|| < \varepsilon_n$ with weight one and zero otherwise.
- $knn-$**graph**: for each n we take $k_n \in \mathbb{N}$ and put a unit-weight edge from each vertex x to each of its k_n nearest neighbors: for each vertex x let

$$\varepsilon_n(x) = \inf\{r > 0 : \ |V_n \cap B(x, r)| \geq k_n\}$$

  and connect x to y with weight one if $||x - y|| \leq \varepsilon_n(x)$ and zero otherwise.

Scaling limit in the classical $\varepsilon-$graph

Proposition

Let $x \in D \subset \mathbb{R}^d$, $f : D \to \mathbb{R}$ be twice continuously differentiable. For each n, the $n-$th $\varepsilon-$graph Laplacian of f at x is given by

$$L_n f(x) = \frac{1}{d_n(x)} \sum_{y \in V_n} \mathbb{1}(||y - x|| < \epsilon_n)(f(y) - f(x)).$$

Assume the following:

1. *p is continuously differentiable and strictly positive in D,*
2. $\varepsilon_n = (n^{-1} \log n)^{\frac{1}{d+\alpha}}$ *for some* $\alpha > 0$.

Then with probability one:

$$\lim_n \frac{c_d}{\varepsilon_n^2} L_n f(x) = \Delta f(x) + \langle \nabla \log p(x), \nabla f(x) \rangle$$

Sketch of proof for $\varepsilon-$graphs

Proof.

1. **Step 1:** Compute expectation and Taylor expand the integrands:

$$\mathbb{E}L_n f(x) = \frac{\int_{B(x,\varepsilon_n)}(f(y) - f(x))p(y)dy}{\int_{B(x,\varepsilon_n)} p(y)dy}$$
$$= \frac{\varepsilon_n^2}{c_d}\left(\frac{p(x)\Delta f(x) + \langle \nabla \log p(x), \nabla f(x)\rangle + o(\varepsilon_n^2)}{p(x) + O(\varepsilon_n^2)}\right)$$

2. **Step 2:** Use some concentration inequality (such as Hoeffding or Bernstein) to obtain exponential upper bounds on

$$\mathbb{P}\left(\frac{c_d}{\varepsilon_n^2}\left|L_n f(x) - \mathbb{E}L_n f(x)\right| > \delta\right).$$

□

Remark

The above proof makes use of the fact that ε_n *is not a function of the sample* $\{x_1, \dots, x_n\}$.

Generalized graph Laplacian and its scaling limit

Definition ([SJK22])

Let $G = (V, W)$ be a weighted graph, with W not necesarilly symmetric. We also consider some finite measure $v : V \to \mathbb{R}$. The associated **generalized random walk graph Laplacian** (GGL) is the operator L^v on $L^2(V)$ given by:

$$L^v f(x) = \frac{1}{\tilde{d}(x)} \sum_{y \in V} \left(v(x)p(x,y) + v(y)p(y,x)\right)\left(f(y) - f(x)\right),$$

where as before

$$p(x,y) = \frac{w(x,y)}{d(x)}, \; p(y,x) = \frac{w(y,x)}{d(y)};$$

and

$$\tilde{d}(x) = \sum_{y \in V} v(x)p(x,y) + v(y)p(y,x) = v(x) + (vP)(x).$$

An intuition on GGL

Observation

We may take $\tilde{p}(x, y)$ *to be* $\frac{v(x)p(x,y)+v(y)p(y,x)}{\tilde{d}(x)}$ *and write* $L^v f$ *more compactly as*

$$L^v f(x) = \sum_y \tilde{p}(x, y)(f(y) - f(x)).$$

$\tilde{p}(\cdot, \cdot)$ *can be thought of as the transition kernel of the following Markov dynamics: At node* x *toss a fair coin. Now, to go to the node* y*:*

- *if heads pick the out-going edge* $x \to y$ *with probability* $\frac{v(x)p(x,y)}{\tilde{d}(x)}$.
- *If tails pick the in-going edge* $y \to x$ *with probability* $\frac{v(y)p(y,x)}{\tilde{d}(x)}$.

Observation

Note that $\tilde{p}(x, y) = \tilde{p}(y, x)$ *and hence* L^v *is self-adjoint for* $\langle , \rangle_{v+vP}$.

Setting for the main theorem

Remark

Let:

1. *p be a positive differentiable probability density on the domain* $D \subset \mathbb{R}^d$,
2. *for each* $n \in \mathbb{N}$*:*
   - $V_n = \{x_1, \ldots, x_n\}$ *n independent samples from p,*
   - *a finite measure* v_n *on* V_n,
   - *a positive integer* k_n*: the number of neighbors to connect with at time n.*

The knn−GGL associated to (V_n, v_n, k_n) *is given by*

$$L_n f(x) = \sum_{y \in V_n} \tilde{p}(x, y)(f(y) - f(x))$$

where: $\tilde{p}(x, y) = \frac{v_n(x)\mathbb{1}(||x-y|| \leq \varepsilon_n(x)) + v_n(y)\mathbb{1}(||x-y|| \leq \varepsilon_n(y))}{k_n \tilde{d}_n(x)}$.

Pointwise convergence of $knn-$GGL

Theorem (Puricelli,Jonckheere,G.)

Take $f \in C^2(D)$ and $x \in D$. If:

- $k_n \in \mathbb{N}$ *is of order* $\left(\frac{n}{\log(n)}\right)^{\alpha}$ *for some* $\alpha \in \left(\frac{2}{d+2}, 1\right)$,
- $c_n = c_d \left(\frac{n}{k_n}\right)^{2/d}$ *with* c_d *some positive constant depending only on the dimension* d,
- $\left|\frac{\nu_n(x)}{k_n} - \nu(x)\right| = o\left(\left(\frac{n}{k_n}\right)^{1/d}\right)$ *for some positive* $\nu \in C^1(D)$;

then $\frac{c_n}{k_n} L_n f(x)$ *converges almost surely to:*

$$\mathcal{L}f(x) = \frac{1}{p(x)^{2/d}} \left(\Delta f(x) + \langle \nabla f(x), \nabla \log(\nu p^{1/d})(x) \rangle\right).$$

Main difficulties

Computing the expectation is not straightforward due to correlations involved in the ε_n's. To overcome this we use the following estimates:

Proposition ([CGT22])

Let $\overline{\varepsilon_n}(x) = \left(\frac{k_n}{c_d p(x) n}\right)^{1/d}$ *There are constants* $C, c > 0$ *such that*

$$\mathbb{P}\left(|\varepsilon_n(x)^d - \overline{\varepsilon_n}(x)^d| \geq C\left(\frac{k_n}{n}\right)^{2/d}\overline{\varepsilon_n}(x)^d\right) \leq n\exp\left(-c\left(\frac{k_n}{n}\right)^{4/d}k_n\right).$$

Using this we were able to prove that the limit is the same if we change $\varepsilon_n(y)$ for $\overline{\varepsilon_n}(y)$ for every node y and adapt the steps of the proof of the classic setting.

To do list

- Prove convergence of eigenvalues and eigenvectors.
- Compare spectral clustering from the continuous operator of the GGL.
- Examples and simulations with the limit operators.

Bibliography

Jeff Calder and Nicolás García Trillos.
Improved spectral convergence rates for graph Laplacians on ε-graphs and k-NN graphs.
Appl. Comput. Harmon. Anal., 60:123–175, 2022.

Harry Sevi, Matthieu Jonckheere, and Argyris Kalogeratos.
Generalized spectral clustering for directed and undirected graphs, 2022.

Ulrike von Luxburg.
A tutorial on spectral clustering.
Stat. Comput., 17(4):395–416, 2007.

Suggestions/ideas are most wecome. **Thanks for your attention!**