Optimal cuts of random geometric graphs

Mathew Penrose
(University of Bath, UK)

19th Workshop on
Stochastic Geometry, Stereology and Image Analysis
(SGSIA)
Luminy
May 2017

Can One Hear the Shape of a Drum? (Kac, 1966)

Let $D \subset \mathbb{R}^d$ be a bounded domain (open, connected) with $d = 2$.
The fundamental frequencies of a D-shaped membrane are given by:
$\lambda_0 = 0 < \lambda_1 < \lambda_2 \ldots$ the eigenvalues of the Laplacian $-\triangle$ on D:

$$\triangle u(x,y) := \frac{\partial^2 u}{\partial x^2} + \frac{\partial^2 u}{\partial y^2}, \quad u \in C^2(D), \quad u|_{\partial D} \equiv 0$$

Does $(\lambda_1, \lambda_2, \ldots)$ determine the shape of D?

Many other reasons to be interested in these eigenvalues. For example,

λ_1 is the rate of exponential decay of the survival probability for a Brownian motion in D killed when it hits ∂D.

Cheeger's inequality

Let $d \geq 2$, let $D \subset \mathbb{R}^d$ be a bounded domain. Let λ_1 be the first non-zero eigenvalue of $-\triangle$ on D. Then [Cheeger 1970]

$$\lambda_1 \geq \frac{(\mathrm{CHE}(D))^2}{4}$$

where the *Cheeger constant* of D is given by

$$\mathrm{CHE}(D) := \inf \left\{ \frac{|\partial_D A|}{|A|} : A \subset D, 0 < |A| \leq |D|/2 \right\},$$

$|A|$ denotes the volume of A,

$|\partial_D A|$ denotes the perimeter of A within D,

i.e. the surface measure of $\overline{A} \cap \overline{D \setminus A}$ (where $\overline{A}$ means closure of A).

Parini 2015: Reverse inequality $\lambda_1 < \frac{\pi^2(\mathrm{CHE}(D))^2}{4}$ for $d = 2$, D convex.

Random (weighted) geometric graphs

Let $X_1, X_2, \ldots$ be i.i.d. uniform (D).

For $n \in \mathbb{N}$ and $r > 0$, let $G(n,r)$ be the weighted graph on vertex set $V_n := \{X_1, \ldots X_n\}$ with weights

$$W_{xy} := \phi\left(\frac{|x-y|}{r}\right)$$

where $\phi(t) = \mathbf{1}_{[0,1]}(t)$, $t \geq 0$, and $|\cdot|$ is Euclidean.

i.e., connect any two points of V_n at Euclidean distance at most r_n.

[Could also consider non-uniform densities,
and other weight functions ϕ such as $\phi(t) = \exp(-t^2)$]

Cheeger constant of a graph G ($G = (V, W)$)

Also known as the *conductance* of G:

$$\mathrm{CHE}(G) = \min \left\{ \frac{\partial_G(U)}{\mathrm{vol}_G(U)} : U \subset V, 0 < \mathrm{vol}_G(U) \leq 1/2 \right\}$$

where we set

$$\partial_G(U) := \sum_{v \in U} \sum_{w \in V \setminus U} W_{vw}; \qquad \mathrm{vol}_G(U) := \frac{\#(U)}{\#(V)}$$

so $\mathrm{vol}_G(U) \in (0, 1)$. The denominator penalizes unbalanced cuts.

[Alternatively could define vol(U) by counting edges rather than vertices, and/or change the denominator to $\mathrm{vol}_G(U)\mathrm{vol}_G(V \setminus U)$.]
Uses: provides bounds on mixing times of random walk on graph, bounds on graph laplacian; reasonable criterion for optimal cut.

Machine learning

Aim: learn about D from the sample V_n. In particular:

Can we learn about $\mathrm{CHE}(D)$ **from** $\mathrm{CHE}(G(n, r_n))$**, given** $(r_n)_{n \geq 1}$**?**

[Recall $\mathrm{CHE}(D) := \inf \left\{ \frac{|\partial_D A|}{|A|} : A \subset D, 0 < |A| \leq |D|/2 \right\}$
$\mathrm{CHE}(G) = \min \left\{ \frac{\partial_G(U)}{\mathrm{vol}_G(U)} : U \subset V, 0 < \mathrm{vol}_G(U) \leq 1/2 \right\}$]

[Raised by Arias-Castro et al. 2012. Could ask similar in manifolds]

Given $U \subset V_n$, we'll use notation

$$\partial_n(U) := \partial_{G(n,r_n)}(U),$$
$$\mathrm{vol}_n(U) := \mathrm{vol}_{G(n,r_n)}(U) = \#(U)/n.$$

Assumptions on D and r_n

$D \subset \mathbb{R}^d$ open and connected.

Also assume $|D| = 1$, and that D has a *Lipschitz boundary* ∂D

[this holds e.g. if ∂D is smooth or D is a cube]

Also, assume that $r_n \ll 1$ and (unless stated otherwise) that

$$nr_n^d \gg \log n,$$

where $a_n \ll b_n$ or $b_n \gg a_n$ means $(a_n/b_n) \to 0$ as $n \to \infty$.

Note: $\exists c > 0$: if $nr_n^d \leq c \log n$ then G is not connected so $\mathrm{CHE}(G) = 0$. Need at least $nr_n^d \geq c \log n$ to have any chance of learning anything from $\mathrm{CHE}(G(n, r_n))$. But want r_n small for computational reasons.

Asymptotic upper bound for $\mathrm{CHE}(G)$ in general D

Choose $A \subset D$ to minimize $|\partial_D A|/|A|$ subject to $0 < |A| \leq \frac{1}{2}$.
Let $U_n = V_n \cap A$. By the SLLN, $\mathrm{vol}_n(U_n) \to |A|$. Also,

$$\mathbb{E}[\partial_n(U_n)] = n^2 \int_A \int_{D\setminus A} \mathbf{1}_{[0,r_n]}(|y-x|)dydx$$

$$\sim |\partial_D A| \sigma n^2 r_n^{d+1},$$

with $\sigma := (1/2)\int_{\mathbb{R}^d} x_1 \mathbf{1}_{[0,1]}(|x|)dx$. ['Surface tension' of $\phi = \mathbf{1}_{[0,1]}$], at least if $\partial_D A$ is smooth. So assuming $n^{-2}r_n^{-d-1}\partial_n(U_n)$ is concentrated, and using the Strong Law of Large Numbers for $\mathrm{vol}_n(U_n)$, this gives

$$\limsup n^{-2}r_n^{-d-1}\mathrm{CHE}(G(n,r_n)) \leq \limsup n^{-2}r_n^{-d-1}\left(\frac{\partial_n(U_n)}{\mathrm{vol}_n(U_n)}\right)$$

$$= \frac{\sigma|\partial_D A|}{|A|} = \sigma \mathrm{CHE}(D)$$

THEOREM (García Trillos et al. '16; Müller and P.)

[Recall $\mathrm{CHE}(D) := \inf\left\{\frac{|\partial_D A|}{|A|} : A \subset D, 0 < |A| \leq |D|/2\right\}$
$\mathrm{CHE}(G) = \min\left\{\frac{\partial_G(U)}{\mathrm{vol}_G(U)} : U \subset V(G), 0 < \mathrm{vol}_G(U) \leq 1/2\right\}$]

Under our conditions ($|D| = 1$, ∂D Lipschitz, $r_n \to 0$, $nr_n^d \gg \log n$), a.s.:

- $n^{-2} r_n^{-d-1} \mathrm{CHE}(G(n, r_n)) \to \sigma \mathrm{CHE}(D)$. [already shown $\leq$]
- If $A \subset D$ is the (essentially) unique Cheeger minimizer, i.e. $|A| < 1/2$ and $\frac{|\partial_D A|}{|A|} < \frac{\partial_D A'}{|A'|}$ for all $A' \subset D$ with $|A' \triangle A| \neq 0$, then

$$n^{-1} \sum_{x \in A_n} \delta_x \to \mathrm{Leb}_d|_A \quad \text{weakly.}$$

- If A is not unique, we still have convergence on a subsequence.

G. Trillos et al. needed the additional conition $nr_n^2 \gg (\log n)^{3/2}$ if $d = 2$.

Differences between the proofs

- García Trillos *et al.*, in their proof, divide D into n cubes of side $n^{-1/d}$.
- They use minimax grid matching results (Leighton and Shor '89, Shor and Yukich '91) to associate each point of V_n with a nearby cube.
- Hence convert discrete set $U_n \subset V_n$ into a union of cubes.
- Grid matchings need $nr_n^2 \geq C(\log n)^{3/2}$ if $d = 2$.
- In the Müller and P. proof, instead divide D into larger cubes (boxes) of side $\gamma_n r_n$, with $\gamma_n \to 0$ but $n(\gamma_n r_n)^d \gg \log n$.
- Convert U_n into set of boxes, namely boxes containing 'mostly' points of U_n.

References

- Arias-Castro, E., Pelletier, B. and Pudlo, P. (2012) The normalized graph cut and Cheeger constant: from discrete to continuous. *Adv. Appl. Probab.* **44**, 907-937.
- García Trillos, N. and Slepčev, D. (2016) Continuum limit of total variation on point clouds. *Arch. Ration. Mech. Anal.* **220**, 193-241.
- García Trillos, N., Slepčev, D., von Brecht, J., Laurent, T. and Bresson, X. (2016) Consistency of Cheeger and ratio cuts. *Journal of Machine Learning Research*.
- Müller, T. and Penrose, M.D. (2017+) Optimal Cheeger cuts and bisections of random geometric graphs. In preparation.