Graph Curvature, GNNs, and TDA for Network Analysis

Amirhossein Farzam

7 novembre 2024

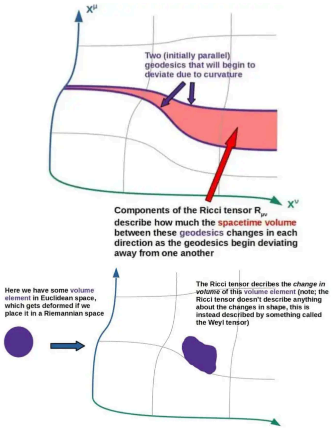
Duke University

Outline

- Ricci Curvature
- Graph Ricci Curvature
- GNN
- Attention Graphs
- TDA
- An Example : Curvature filtration for measuring graph distance

Ricci Curvature

- Sectional curvature indicates whether geodesics starting from two points remain parallel (0), converge (> 0), or diverge (< 0)
- Ricci curvature extends this to volume balls and in doing so, indicates the average sectional curvature
- Ricci curvature shows how volume balls on a manifold differ from that in the Euclidean space
- It measures **dispersion** of volume, and energy dynamics

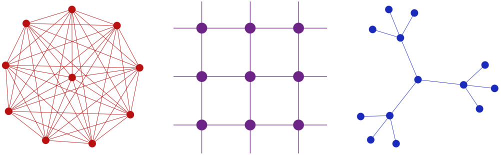


Interested in learning more ?

Carroll, S. M. (2019). Spacetime and geometry. Cambridge University Press.

Graph Ricci Curvature

- The analogue of Euclidean space on graphs is a grid/4-lattice, i.e., a 4-regular graph
- On graphs, Ricci curvature measures **dispersion** through an edge in a **"somewhat local"** neighborhood
- Various notions of graph curvature, e.g. Forman-Ricci (combinatorial) [1] and Ollivier-Ricci (optimal transport) [2]



(a) Clique (> 0) (b) Grid ($= 0$) (c) Tree (< 0)

Graph Ricci Curvature

- **Forman-Ricci Curvature**

$$\kappa_{FR}(v,u) := w_{vu}\left[\frac{w_v}{w_{vu}} + \frac{w_u}{w_{vu}} - \sum_{(v',u')\in N_v\times N_u}\left(\frac{w_v}{\sqrt{w_{vu}w_{vv'}}} + \frac{w_u}{\sqrt{w_{vu}w_{uu'}}}\right)\right],$$

  - In an unweighted graph, it simply becomes

$$\kappa_{FR}(v,u) := 4 - d_v - d_u + 3|N_v \cap N_u|$$

  - Most local (only up to triangles)
  - Known to correlate with Ollivier-Ricci curvature in empirical networks

- **Ollivier-Ricci Curvature**

$$\kappa_{OR}(v,u) := 1 - \frac{W_1(\mu_v,\mu_u)}{d_G(v,u)}$$

where, by default, μ_v is the uniform probability measure in N_v, and $d_G(.)$ is the geodesic distance on G.

  - Impacted by rectangles and pentagons too
  - Most flexible (+ best results)

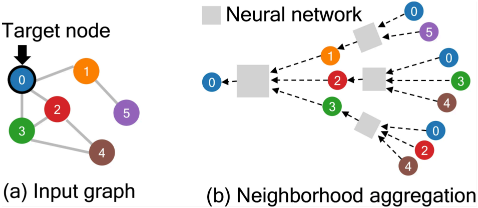
Why graph curvature ?

- Ricci curvature characterizes **local geometry** :
  - It measures **dispersion** through an edge in a **"somewhat local"** neighborhood
  - It measures *bottleneckedness* (and tends to correlate with other measures of bottleneckedness, such as edge betweenness)
- Rich **theoretical foundations**
- **Global topology** emerges from local geometry
- Ricci curvature is a proxy for **robustness :**
  - It correlates with system robustness (through entropy)
  - It also correspondence to distributional robustness distributional robustness (on a graph of dependent units)
- Ricci curvature is linked to **network dynamics and signal processing** on graphs :
  - Due to connection with the Laplace operator
  - Bounds the spectrum of the Laplacian in neighborhood graphs (hence links to many aspects of signal processing)

$$\kappa \leq \lambda_1 \leq \ldots \leq \lambda_n \leq 2 - \kappa$$

GNNs

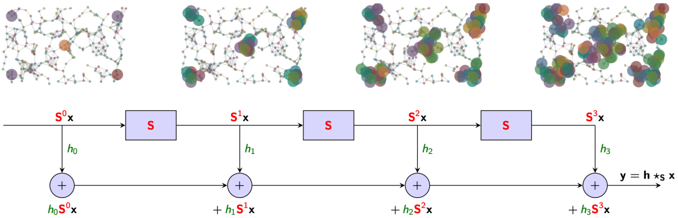
- GNNs are powerful tools for learning features of (graph) structured units
- They can propagate information across units with non-trivial dependencies, e.g., more complex than a grid (CNN) or a fully-connected graph (attention)



- Rich theory (*geometric deep learning*)
- Lots of applications
  social network analysis, molecular graphs, protein interaction networks, gene regulatory networks, autonomous driving, recommender systems, etc.

GNN

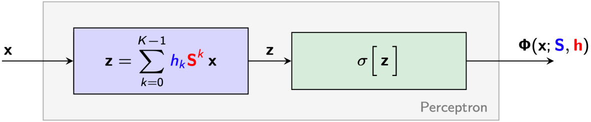
- I focus on GNNs with **graph convolution filters**
- Given $G = (V, E)$, with $|V| = N$, $|E| = M$, connectivity given in a shift operator S. Typically, $S =$
  - Adjacency matrix, A
  - Laplacian, $L = D - A$
  - Normalized laplacian, $\tilde{L} = D^{-1/2}(D - A)D^{-1/2}$
- x_i is a signal on the graph for each node $v_i \in V$.



GNN

- Graph convolutional filters (GCFs) are the building blocks of GNNs.
- GCFs process diffused signals over the graph structure.
- GCF of order K : $\sum_{k=0}^{K-1} h_k S^k$.
- A **graph perceptron** is a GCF followed by a nonlinear activation $\sigma(.)$.
- **Objective :** Learn the graph filter, $h_0, \ldots, h_{K-1}$.

$$\mathbf{h}^* = \text{argmin}_{\mathbf{h}} \sum_{(\mathbf{x},\mathbf{y}) \in \mathcal{T}} \ell\left(\mathbf{y}, \phi(\mathbf{x}; \mathbf{S}, \mathbf{h})\right)$$



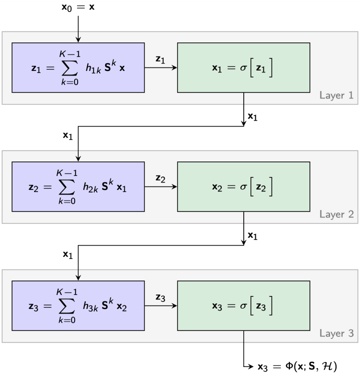
GNN

- A GNN is L layers of graph perceptron.
- The GNN is trained to learn a filter tensor $\mathcal{H} := [\mathbf{h}_1, \ldots, \mathbf{h}_L]$.

$$\mathcal{H}^* = \text{argmin}_{\mathcal{H}} \sum_{(\mathbf{x},\mathbf{y}) \in \mathcal{T}} \ell\left(\mathbf{y}, \mathbf{x}_L\right)$$

- Given a GNN, $\Phi(\mathbf{x}, S, \mathcal{H})$, with input, $\mathbf{x}_0$, compute the output $\mathbf{x}_L$ by :

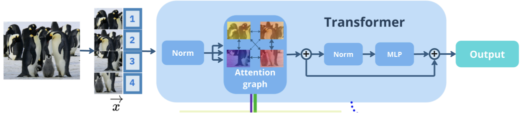
$$\mathbf{x}_l = \sigma\left(\sum_{k=0}^{K-1} \mathbf{h}_{l,k} S^k \mathbf{x}_{l-1}\right).$$



A 3-layer GNN

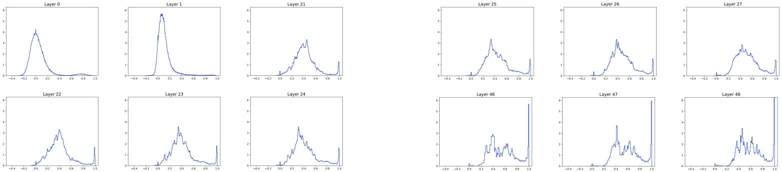
Graphs in Deep Learning : Attention

- GNNs are not the only ML tools where graph geometry topology shows up
- Transformers learn weighted graphs of interactions between tokens (often fully connected)



Attention graph in transformers

- Attention can be thought of as network of a particles system, and can be studied via tools from network theory (e.g., clustering through a forward pass)

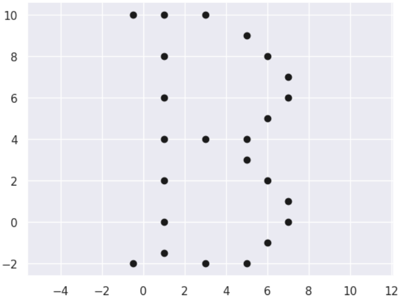


Histogram of cosign similarity of attention vectors

Topological Data Analysis (TDA)

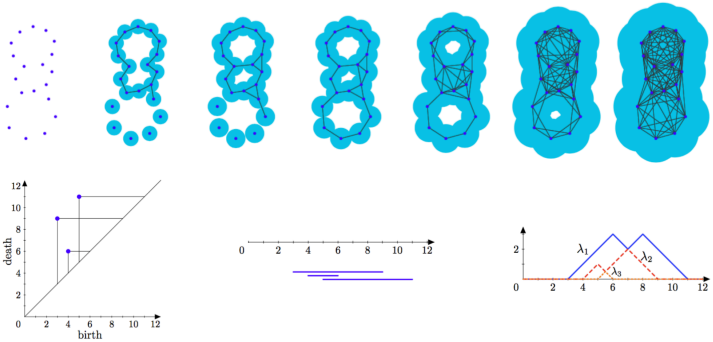
- TDA aims at characterizing the "shape of the data"
- Local geometry $\rightarrow$ global topology
- Persistent homology is a common TDA technique for characterizing the topology of point clouds

X	Y
1.0	10.0
3.0	10.0
5.0	9.0
6.0	8.0
7.0	7.0
7.0	6.0
6.0	5.0
5.0	4.0
3.0	4.0
5.0	3.0
6.0	2.0
7.0	1.0
7.0	0.0
6.0	-1.0
5.0	-2.0
3.0	-2.0
1.0	-1.5
1.0	4.0
1.0	0.0
1.0	8.0
1.0	6.0
1.0	2.0
-0.5	10.0
-0.5	-2.0



TDA : Persistent Homology

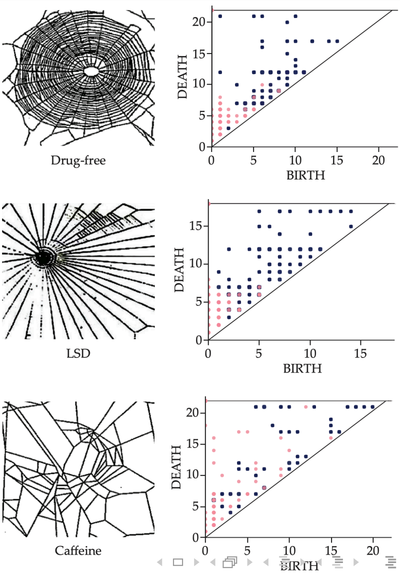
- Intuitively, persistent homology finds homology groups that best describe the shape of the data
- Well established family of **stability theorem** speak to **robustness** properties of persistence diagrams/images



(a) (b) (c)

TDA : Persistent Homology - Example

- Persistence diagram provides a signature for the topology of point cloud data
- Points farther away from the diagonal correspond to most "persistent" features



TDA : Persistent Homology on Networks

- Homology groups can be determined according to any graph filtration

Definition

A **graph filtration** is constructed from a nested family of simplicial complexes, as follows

$$\emptyset \subseteq G_0 \subseteq G_1 \ldots \subseteq G_k$$

where each $G_i = (V_i, E_i)$ satisfies $V_i = \{v \in V | f(v) \leq a_i\}$ and $E_i = \{e \in E | \max_{v \in e} f(v) \leq a_i\}$ for a scalar function $f(.)$ defined on V and a sequence of scalar values $a_1, \ldots, a_k$.

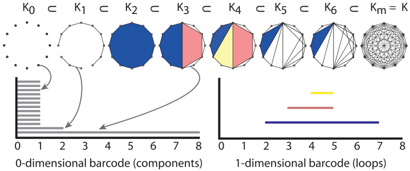


FIG. 4. Example of a weight rank clique filtration (WRCF) and the corresponding 0-dimensional and 1-dimensional barcodes. The barcode of dimension 0 indicates the connected components in every filtration step. When two components merge into one connected component, one of the bars that represent the original components dies in the barcode; the other continues to the next filtration step and now represents the newly-formed component. In filtration step 0, every node is a separate component, resulting in 12 bars in the barcode. The nodes are joined to become two components in filtration step 1, and they then become a single component in step 2. In dimension 1, we observe that as more edges are added to the filtration, the loop surrounding the blue hole born in filtration step 2 is divided first into two holes and subsequently into three holes before it is completely covered by 2-simplices and dies in filtration step 7. The colors of the bars indicate which loop they represent.

Using all of this to compare graphs

Curvature Filtrations for Graph Generative Model Evaluation

Joshua Southern*
Imperial College London
jks17@ic.ac.uk

Jeremy Wayland*
Helmholtz Munich & Technical University of Munich
jeremy.wayland@tum.de

Michael Bronstein†
University of Oxford
michael.bronstein@cs.ox.ac.uk

Bastian Rieck†
Helmholtz Munich & Technical University of Munich
bastian.rieck@tum.de

Abstract

Graph generative model evaluation necessitates understanding differences between graphs on the distributional level. This entails being able to harness salient attributes of graphs in an efficient manner. Curvature constitutes one such property that has recently proved its utility in characterising graphs. Its expressive properties, stability, and practical utility in model evaluation remain largely unexplored, however. We combine graph curvature descriptors with emerging methods from topological data analysis to obtain robust, expressive descriptors for evaluating graph generative models.

Curvature Filtration for Measuring Distance

- **Intuition :** Curvature notions, characterize the network topology through the local geometry of the edges, going beyond one-hop neighborhood (e.g. degree sequence)

Definition

A **graph filtration** is constructed from a nested family of simplicial complexes, as follows

$$\emptyset \subseteq G_0 \subseteq G_1 \ldots \subseteq G_k$$

where each $G_i = (V_i, E_i)$ satisfies $V_i = \{v \in V | f(v) \leq a_i\}$ and $E_i = \{e \in E | \max_{v \in e} f(v) \leq a_i\}$ for a scalar function $f(.)$ defined on V and a sequence of scalar values $a_1, \ldots, a_k$.

- TDA methods with a curvature filtration could improve stability, i.e. if curvature distributions are similar, so are their persistence diagrams (by a stability-type theorem)

Theorem (Southern et al, 2023, Theorem 8)

Given two graphs $F = (V_F, E_F)$ and $G = (V_G, E_G)$ with scalar-valued filtration functions f and g, and their respective persistence diagrams D_f and D_g, we have

$$d_B(D_f, D_g) \leq \max\{dis(f,g), dis(g,f)\},$$

where $dis(f,g) := \left|\max_{x \in V_F} f(x) - \min_{y \in V_G} g(y)\right|$.

$d_B(x,y) := \inf_\gamma \sup_x \|x - \gamma(x)\|_\infty$ is the Bottleneck distance.

Curvature for Distance : Stability

- The stability results are directly implied from properties of curvature.
- Derive bounds for changes in curvature upon edge addition or deletion
- **Notation** :
  - Let G be the initial graph, and G' and G'', obtained by edge addition or deletion
  - $\kappa_{.}(.,.)$, $\kappa'_{.}(.,.)$, and $\kappa''_{.}(.,.)$ refer to the cruvatures in G, G', and G'', resp.

Theorem (Southern et al, 2023, Theorems 2 and 3)

$$\kappa_{FR}(v,u) - 1 \leq \kappa'_{FR}(v,u) \leq \kappa_{FR}(v,u) + 2$$
$$\kappa_{FR}(v,u) - 2 \leq \kappa''_{FR}(v,u) \leq \kappa_{FR}(v,u) + 1$$

Curvature for Distance : Stability

- **More notation** :
  - $\tilde{.}$ marks the modified variables in G' or G''
  - $W_1(.,.)$ is the 1-Wasserstein distance

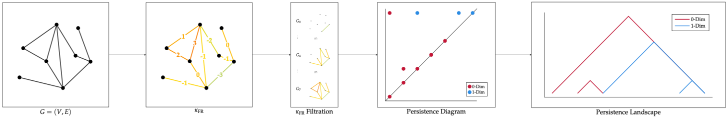
Theorem (Southern et al, 2023, Theorems 4 and 5)

$$1 - \frac{1}{d_{\tilde{G}}(v,u)} \left[2\tilde{W}_{max} + \tilde{W}_1(\mu_v, \mu_u) \right] \leq \tilde{\kappa}_{OR} \leq \frac{\tilde{J}(v) + \tilde{J}(u)}{d_{\tilde{G}}(v,u)},$$

where $\tilde{J}(v) := \tilde{W}_1(\delta_v, \mu_v)$ *is the jump probability at* v*, with* δ_v *being the Dirac measure, and* $\tilde{W}_{max} := \max_{x \in V} \tilde{W}_1(\mu_x, \tilde{\mu}_x)$.

Curvature for Distance : Experimental Results

- Curvature filtration pipeline :



- Stability experiment :
  - Increase edge removal rate from 0.0 to 0.95 and compute the pairwise distance between consecutive modified graphs
  - Pearson correlations show higher stability of curvature-based distance

Measure	Adding Edges	Removing Edges
Laplacian	0.46	0.42
Clustering Coefficient	0.48	0.50
Degrees	0.76	1.00
κ_{FR}	0.42	0.43
κ_{OR}	0.90	0.91
κ_{R}	0.42	0.44
Filtration (κ_{FR})	0.57	1.00
Filtration (κ_{OR})	1.00	0.97
Filtration (κ_{R})	0.73	0.95

- Remarkable stability of the Ollivier-Ricci curvature and significant improvement by the TDA-based method

References

Bauer, F., Hua, B., Jost, J., Liu, S., Wang, G. (2017). The geometric meaning of curvature : Local and nonlocal aspects of ricci curvature. *Modern Approaches to Discrete Curvature*, 1-62.

Pouryahya, M., Mathews, J., Tannenbaum, A. (2017). Comparing three notions of discrete Ricci curvature on biological networks. *arXiv* preprint arXiv :1712.02943.

Farzam, A., Tannenbaum, A., Sapiro, G. (2024, July). From Geometry to Causality-Ricci Curvature and the Reliability of Causal Inference on Networks. In Forty-first *International Conference on Machine Learning*.

Bauer, F., Jost, J., Liu, S. (2011). Ollivier-Ricci curvature and the spectrum of the normalized graph Laplace operator. *arXiv* preprint arXiv :1105.3803.

Samal, A., Sreejith, R. P., Gu, J., Liu, S., Saucan, E., Jost, J. (2018). Comparative analysis of two discretizations of Ricci curvature for complex networks. *Scientific reports*, 8(1), 8650.

Alejandro Riebeiro, Graph Neural Networks, Graph Neural Networks, University of Pennsylvania, 2022.

Geshkovski, B., Letrouit, C., Polyanskiy, Y., Rigollet, P. (2023). A mathematical perspective on transformers. *arXiv* preprint arXiv :2312.10794.

Otter, N., Porter, M. A., Tillmann, U., Grindrod, P., Harrington, H. A. (2017). A roadmap for the computation of persistent homology. *EPJ Data Science*, 6, 1-38.

References

Stolz, B. J., Harrington, H. A., Porter, M. A. (2017). Persistent homology of time-dependent functional networks constructed from coupled time series. *Chaos : An Interdisciplinary Journal of Nonlinear Science*, 27(4).

Fugacci, U., Scaramuccia, S., Iuricich, F., De Floriani, L. (2016, October). Persistent Homology : a Step-by-step Introduction for Newcomers. In *STAG* (pp. 1-10).

Feng, M., Porter, M. A. (2020). Spatial applications of topological data analysis : Cities, snowflakes, random structures, and spiders spinning under the influence. *Physical Review Research*, 2(3), 033426.

Southern, J., Wayland, J., Bronstein, M., Rieck, B. (2024). Curvature filtrations for graph generative model evaluation. *Advances in Neural Information Processing Systems*, 36.

Thank You !