

Olivier's Ricci curvature and applications on cell networks

Sunhyuk Lim

04/25/2018

1 Introduction

This article is basically summary of two papers [TSZ⁺15, PMT17]. In those papers, the authors argue that curvature and robustness of a network are positively correlated. This implies that one can measure robustness of a network by computing its curvature. Here, robustness of a network is not rigorously defined notion. Instead, roughly robustness means the capacity to remain functional in the face of random perturbations [DM05]. The second thing we have to clarify is notion of curvature for networks. Since the classical notion of curvatures for Riemannian manifolds are not applicable to networks, The authors introduce three generalized notions of Ricci curvature applicable to networks. Among them, I will especially focus on Ollivier's Ricci curvature. After developing argument to claim that robustness and curvature are positively correlated, the authors compute curvatures of cell complexes to test their hypothesis. It is already known that the normal gene interaction networks are less robust than their cancerous analogues [DM05, WBST12]. Therefore, if the authors' claim is true, then curvature of the cancerous cell networks should be greater than curvature of normal cell networks. Indeed, the result of experiment is compatible with the guess and supports the authors' hypothesis.

2 Background

Throughout this article, The term *network* always means positively edge-weighted finite undirected graphs. Now I review three important notions: curvature, entropy, and robustness.

2.1 Ricci curvature

In this subsection, let me review the notion of **Ricci curvature** for Riemannian manifolds very quickly. Suppose M is a Riemannian manifold with metric g . In general, the curvature is a way of measuring the degree to which the geometry determined by a given Riemannian metric might differ from that of ordinary Euclidean space. Especially, we are interested in Ricci curvature. Rather than defining curvature in a rigorous way, let us motivate notion of Ricci curvature in a more intuitive way.

First, consider 1-dimensional case first. In other words, consider arc length parameterized smooth plane curve $\gamma : I \rightarrow \mathbb{R}^2$. In this case, $\kappa(t) := \|\gamma''(t)\|$ is the curvature of the curve at point $\gamma(t)$. But we have more geometric interpretation of this number. Choose arbitrary three points around the point $\gamma(t)$ and draw the unique circle intersecting those three points. If we move the three points to the point $\gamma(t)$, then the circle also converges to some circle. Then, the curvature is actually the reciprocal of the radius of the converged circle. Moreover, one can give positive/negative sign to the $\kappa(t)$. If the converged circle is on the right side of the tangent vector $\gamma'(t)$, we give + sign to $\kappa(t)$, otherwise give - sign to $\kappa(t)$. Then we have signed curvature for smooth plane curves.

Second, consider 2-dimensional surface $S \subseteq \mathbb{R}^3$. Suppose $p \in S$ and $N(p)$ is a choice of unit normal vector to the tangent space $T_p S$. Let v be a unit vector in $T_p S$. The pair $\{v, N(p)\}$

determines a plane in \mathbb{R}^3 and the intersection of this plane with S gives a unit speed curve $\gamma_v(s)$ such that $\gamma'_v(0) = v$. We define $\kappa(v)$ be the signed curvature of $\gamma_v(s)$ at p . Then we are able to find two unit vectors v_1 and v_2 in $T_p S$ making $\kappa(v_1) = \kappa_1$ as maximum and $\kappa(v_2) = \kappa_2$ as minimum by Euler's theorem. The **Gaussian curvature** of S at p is defined by: $K(p) := \kappa_1 \kappa_2$.

Finally, consider general n -dimensional Riemannian manifold M . For each point $p \in M$, $T_p M$ is n -dimensional vector space. Then, for arbitrary 2-dimensional subspace σ of $T_p M$, we can define **sectional curvature** $K_p(\sigma)$ by using the exponential map and emulating the definition of the Gaussian curvature. Finally, for each unit vector $v \in T_p M$, construct orthonormal basis $\{w_1, w_2, \dots, w_n = v\}$. Then the Ricci curvature of v is defined by:

$$\text{Ric}_M(v) := \frac{1}{n-1} \sum_{i=1}^{n-1} K_p(\text{span}\{w_i, w_n\}).$$

2.2 Entropy

Consider arbitrary measure μ of Riemannian manifold M . Then, the **Boltzmann entropy** of μ is defined as follows:

$$\text{Ent}(\mu) := - \int_M \rho \log \rho \, d\text{vol}_M,$$

where vol_M is the standard Riemannian volume measure and $\rho = d\mu/d\text{vol}_M$ is the Radon-Nikodym derivative. Roughly, Boltzmann entropy measures how much "uniform" the measure is.

2.3 Robustness

As I already mentioned in the introduction, **robustness** of a network is the capability to remain functional in the face of random perturbations [DM05, PMT17]. Since it is not mathematically defined notion, robustness is measured empirically. For example, [HMJ⁺00] shows yeast cells can maintain their function after gene deletions by doing experimental perturbation studies in yeast cells. Another example is that Computational analysis of network observables under node deletion [AJB00].

3 Positive correlation between Robustness and Curvature

The theoretical main contribution of [TSZ⁺15, PMT17] is the claim that robustness and curvature of a network are positively correlated. The authors support their claim by using following argument. First, they argue that curvature and entropy are positively correlated. Next, they argue that entropy and robustness are positively correlated. Then, finally one can say robustness and curvature are positively correlated.

First, consider relationship between curvature and entropy. It is well known that a metric space (X, d_X) is a compact length space if and only if $P_2(X) := (P(X), W_2)$ is a compact length space [LV09, Stu06]. Hence, one can consider geodesics in $P_2(M)$ when M is a compact Riemannian manifold. Moreover, Lott, Sturm, and Villani discovered following connection between Ricci curvature and entropy [LV09, Stu06]: $\text{Ric}_M(v) \geq k \|v\|^2$ for any $v \in \text{TM}$ if and only if

$$\text{Ent}(\mu_t) \geq (1-t) \text{Ent}(\mu_0) + t \text{Ent}(\mu_1) + k \frac{t(1-t)}{2} W_2(\mu_0, \mu_1)^2,$$

where $(\mu_t)_{0 \leq t \leq 1}$ is the 2-Wasserstein geodesics between μ_0 and μ_1 in $P_2(M)$. This inequality indicates the positive correlation between entropy and curvature.

$$\Delta \text{Ent} \times \Delta \text{Ric} \geq 0.$$

Second, in [DGO04, DM05], the authors claimed that entropy and robustness of a network are positively correlated by invoking theory of large deviations and suggesting some computational results. Hence we also get:

$$\Delta\text{Ent} \times \Delta R \geq 0.$$

Therefore, one can conclude that there is a positive correlation between robustness and curvature.

$$\Delta R \times \Delta\text{Ric} \geq 0.$$

4 Three generalized notions of the Ricci curvature

However, still we don't have any notion of curvature for networks. Hence, the author introduce three generalized notion of curvature. Of course, all of them are applicable to networks.

- **Ollivier-Ricci curvature** are defined for metric spaces with Markov chain structure, or metric measure spaces. For networks, we will get curvature value $k_{OR}(x, y) \in \mathbb{R}$ for each edge xy .
- **Bakry-Émery Ricci curvature** are defined for graphs. We will get curvature value $k_{BER}(x) \in \mathbb{R}$ for each vertex x .
- **Forman-Ricci curvature** are defined for CW-complexes. For networks, we will get curvature value $k_{FR}(e) \in \mathbb{R}$ for each edge e .

In this article I only focus on Olliver's Ricci curvature because it is most understandable and also involves Wasserstein distance in its definition. To motivate Olliver's definition, we have to know following theorem first.

Theorem 1 [vRS05] *For any compact connected Riemannian manifold M and $k \in \mathbb{R}$, the following properties are equivalent:*

1. $\text{Ric}_M(v) \geq k\|v\|^2$ for any $v \in \text{TM}$
2. *The normalized Riemannian uniform distribution on balls*

$$m_{x,r}(A) := \text{vol}_M(A \cap B(x,r)) / \text{vol}_M(B(x,r))$$

satisfies the asymptotic estimate

$$W_1(m_{r,x}, m_{y,r}) \leq \left(1 - \frac{k}{2(n+2)}r^2 + o(r^2)\right) \cdot d_M(x, y)$$

where d_M is the geodesic distance and the error term is uniform with respect to $x, y \in M$.

In particular, if $k > 0$, small balls are closer in transportation distance than their centers are. Following this line of ideas, we can define Ollivier-Ricci curvature.

Definition 4.1 (Ollivier-Ricci curvature) [Oll09] *Let (X, d_X) be a metric space with a Markov chain m_X . Let $x, y \in X$ be two distinct points. The **coarse Ricci curvature** of (X, d_X, m_X) along (xy) is:*

$$k_{OR}(x, y) := 1 - \frac{W_1(m_X(x, \cdot), m_X(y, \cdot))}{d_X(x, y)}.$$

Observe that if the Wasserstein distance between $m_X(x, \cdot)$ and $m_X(y, \cdot)$ are smaller than $d_X(x, y)$, then we have positive Ollivier-Ricci curvature, just like the case of classical Ricci curvature.

Now, how can we apply this definition to the networks? As I already mentioned, a network $G = (V, E)$ is a positively edge-weighted finite undirected graph. For each $x, y \in V$, define $m_X(x, y) := \frac{w_{xy}}{\sum_{z \in V} w_{xz}}$ where w_{xy} is the weight on edge (xy) . Then G has a Markov chain structure. Also, the metric on G is usual graph metric, which is the number of edges in the shortest path.

5 Curvatures of cell complexes

Now we are ready to test the authors' claim. It is already well known that the normal gene interaction networks are less robust than their cancerous analogues [DM05, WBST12]. I think this is intuitively natural result. Since cancerous tissues are already "broken", it will still function as cancer even after some deletions or changes of some of critical genes. To check the validity of the hypothesis: curvature and robustness are positively correlated, the authors computed three curvatures for cell complexes. But before discussing the result of experiments, let's clarify what cell complexes are first.

We will consider seven kinds of cancer types. Breast, Head/Neck, Kidney, Liver, Lung, Prostate and Thyroid cancers. For each cancer type, the authors used normal tissue and cancerous tissue data from 3000 samples. Then, we will have networks, depending on types and normal/cancerous. The vertices of the networks consist of 500 cancer related genes. The edges are weighted by correlation values of gene-to-gene expressions. More precisely, let's denote the correlation value between gene A and B by $Corr(A, B)$. This is some real number between 1 and -1 . To make positive weight, we use $\frac{1+Corr(A, B)}{2}$ as the weight between gene A and gene B. Here, expression value of a gene is biological notion and it measures activity of genes.

The authors computed all three kinds of curvatures for those cell networks. This is the result.

1. All three generalized curvatures have higher values in the seven cancer networks to the normal ones. Hence, the result is consistent with the authors' hypothesis.
2. There are a few common top ranked genes in breast cancer. One can guess that those genes are most contributing for the "robustness" of the the cell complexes. Especially, there are three genes, SDHB, EPS15, and ERG found among the top ranked genes with respect to all three FR, BER and OR curvatures.
3. A number of genes have known clinical implications with regards to breast cancer. For example, EPS15 plays a crucial role in the degradation of growth factor receptions. It is reported that over-expression of EPS15 is significantly associated with a favorable clinical outcome. Also, SDHB gene is another known tumor suppressor. Both genes are top ranked genes with respect to all three kinds of curvatures. So, maybe we can use curvatures to find target genes for clinical applications.
4. However, there are some important cancer-related gene mutations known to play a significant role in breast cancer such as BRCA1 and BRCA2 which are not ranked among the top ranked genes.

References

- [AJB00] Réka Albert, Hawoong Jeong, and Albert-László Barabási. Error and attack tolerance of complex networks. *nature*, 406(6794):378, 2000.

- [DGO04] Lloyd Demetrius, Volker Matthias Gundlach, and Gunter Ochs. Complexity and demographic stability in population models. *Theoretical population biology*, 65(3):211–225, 2004.
- [DM05] Lloyd Demetrius and Thomas Manke. Robustness and network evolution—an entropic principle. *Physica A: Statistical Mechanics and its Applications*, 346(3-4):682–696, 2005.
- [HMJ⁺00] Timothy R Hughes, Matthew J Marton, Allan R Jones, Christopher J Roberts, Roland Stoughton, Christopher D Armour, Holly A Bennett, Ernest Coffey, Hongyue Dai, Yudong D He, et al. Functional discovery via a compendium of expression profiles. *Cell*, 102(1):109–126, 2000.
- [LV09] John Lott and Cédric Villani. Ricci curvature for metric-measure spaces via optimal transport. *Annals of Mathematics*, pages 903–991, 2009.
- [Oll09] Yann Ollivier. Ricci curvature of markov chains on metric spaces. *Journal of Functional Analysis*, 256(3):810–864, 2009.
- [PMT17] Maryam Pouryahya, James Mathews, and Allen Tannenbaum. Comparing three notions of discrete ricci curvature on biological networks. *arXiv preprint arXiv:1712.02943*, 2017.
- [Stu06] Karl-Theodor Sturm. On the geometry of metric measure spaces. *Acta mathematica*, 196(1):65–131, 2006.
- [TSZ⁺15] Allen Tannenbaum, Chris Sander, Liangjia Zhu, Romeil Sandhu, Ivan Kolesov, Eduard Reznik, Yasin Senbabaoglu, and Tryphon Georgiou. Graph curvature and the robustness of cancer networks. *arXiv preprint arXiv:1502.04512*, 2015.
- [vRS05] Max-K von Renesse and Karl-Theodor Sturm. Transport inequalities, gradient estimates, entropy and ricci curvature. *Communications on pure and applied mathematics*, 58(7):923–940, 2005.
- [WBST12] James West, Ginestra Bianconi, Simone Severini, and Andrew E Teschendorff. Differential network entropy reveals cancer system hallmarks. *Scientific reports*, 2:802, 2012.