Topological Time Series Analysis - Jose A. Perea

https://www.ams.org/journals/notices/201905/rnoti-p686.pdf

Defining and measuring the shape of a topological space (e.g., an attractor) from a finite set X (the data). This is the type of problem driving advances in Topological Data Analysis (where tools like persistent homology are relevant). The shape of a space refers to those properties that are invariant under continuous deformations (e.g., is it connected? Are there holes?) Said properties can be formalized and quantified using homology.

Betti numbers

> gave examples of Betti numbers of sphere and a torus

- Betty number capture the homology of a space giving succinct shape descriptors for its topology
- Persistent homology describes the "multi-scale evolution of homological features underlying the data"

Persistemt homology (barcodes)

- Looking at the Rips complex at varying scales of alpha and the intervals that comprise the barcode, we can infer the invariants of the data set. The barcodes can also be used to quantify/identify properties of dynamical systems
  - If the barcode (bcd-1) for an attractor is one then we know that it follows a periodic motion (
  - If the barcode (bcd-2) for an attractor follows that of a torus (b0 = b2 =1, b1 = 2) then we know it follows **quasiperiodicity**
- In practice, it is exceedingly rare to have an explicit mathematical description of a dynamical system of interest.

Dynamical system

- Intuitively, a dynamical system consists of two pieces of data: a set of states M (e.g., all possible atmospheric conditions at a given location on earth) along with rules symbol-phi describing how each state in M changes over time
  - Statistical inference (you see the data and wants to see which model generated the data)
- Assuming a dataset is a dynamical system (we have data and rules that govern the data) we are implying that there is derivatives that governs the data (?)

Attractor

- Intuitively it is a subset of a dynamical system M that "attract the evolution of states in close

proximity."
    - Compact, invariant, has an open basin of attraction (pg 688 — I dont get 2 and 3)
- <u>The shape of attractors carries a great deal of information about the global structure of a dynamical system</u>
    - Attractor with the shape of a circle give rise to periodic processes
    - Non-integral Hausdorff dimension is evidence of chaotic behavior
            - Whoa this is so cool

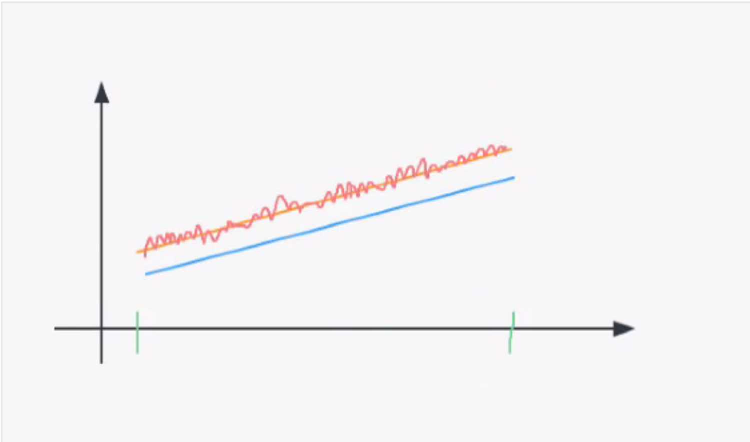
Challenges with barcodes

- 01 Parameter selection appropriate values for the parameters need to be determined. How to determine appropriate parameters? How do people select parameters << — ??
- 02 How to validate the results? (And further interpret the wrestles)
    - In tasks where the barcodes are used as features, one must quantify the likelihood that positive results are due to random fluctuation in the data. In order to interpret the results, one needs a theoretical understanding of barcodes as a function of all the parameters and the time series involved. The challenge here is that our understanding of homotopy type/homology of the Rips complex for arbitrary alpha is limited.

Lorenz Butterfly Effect (Lorenz attractor looks like a butterfly)

Edward Norton Lorenz - set the foundations for what we know today as chaos theory. Understanding that slight alteration in initial condition can have a diverging consequence of reconstructing a data (?) how can one reconstruct a dataset (?) using topological invariants ??

- <u>Weather prediction</u> (blue and orange curves from figure 1 (687 image) is an example of time series from the Lorenz system. => a single time series may appear to be a complete oversimplification of the underlying dynamics.
- The system can be extremely sensitive to initial conditions in that any errors are compounded exponentially with time (which is now know today as the butterfly effect)
- Dynamical systems are mathematical abstraction soo time-dependent physical processes.
- **Whitney topology**, in which, roughly speaking, two functions are close if and only of the functions and all their derivatives up to degree k (closeness of the curves and not the first/second/kth derivative) are close on compact subsets of M (when k=0, red one is closer to the orange one , but when k=1 is considering the first derivative of the given function and the experimental model (the one we made up), then blue is closer to the orange one; the 1st derivative of red fluctuates too much unlike the ones of orange)



Conclusion: Taken's theory motivates an existence of the sliding window point cloud that provides a topological copy that can help reconstruct attractors.

- The underlying shape can be then quantified with persistent homology.
- The barcodes can be used as features in inference, classification, and learning tasks.
-

Do research plot things with python

Assignment 1: For all values of t = 0, h, 2h, 3h, ..., 50h, plot all the points (t,sin(t))

Signal; sin(t) for t in [0, h, 2h, 3h, 4h, ...]

- For different values of h (on the t-axis) plot (t, sin(t)) as dots
    - Values of h : 0.01, 0.1, 1, 10, 0.47, pi, pi+0.1, 22/7
- Plot with lines between them and without
- Time series

Assignment 2: plot (sin(kh), sin((k+1)h)) for k = 0, 1, ..., 50

- For different values of h (on the t-axis) plot (t, sin(t)) as dots
    - Values of h : 0.01, 0.1, 1, 10, 0.47, pi, pi+0.1, 22/7
- Plot with lines between them and without
- Spatial series (where we eliminated time) and we have points in 2d