Applied Algebraic Topology

Mathematics Winter School

February 25, 2021

Gustavo Granja
Department of Mathematics
IST, University of Lisbon

PLAN

1. What is topology?
2. Measuring shape with algebra: **homology**
3. Topological data analysis

1

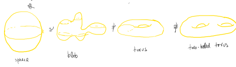
Topology is the part of mathematics that studies **shape** or **coarse** geometric properties of **spaces**

Space = subset of euclidean n-dimensional space $\mathbb{R}^n$

Examples:

Some subsets of $\mathbb{R}^2$:

Some subsets of $\mathbb{R}^3$:



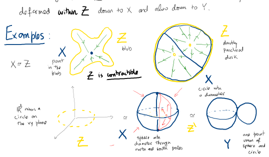
Definition: Two spaces X and Y are **homeomorphic** if there is a function $f: X \to Y$ (called a homeomorphism) such that
i) f is continuous
ii) f is bijective (i.e. one to one and onto)
iii) $f^{-1}: Y \to X$ is also continuous

Matters: Number of separate pieces, number of "holes", dimension

Doesn't matter: Position, length/area/volume, angles

Another notion of shape: More useful and even coarser notion of shape: Say two spaces X, Y are **homotopy equivalent** if they can both be embedded in a larger space Z such that Z can be continuously deformed **within Z** down to X and also down to Y.

Examples:

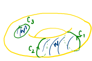


2

The question of how to quantify shape was addressed by **Henri Poincaré** in a famous series of papers which he started publishing in 1895 on **Analysis Situs**

One of his fundamental ideas was to use

Homology



$\vec{F}$ vector field on the torus with $\text{curl}\,\vec{F} = 0$

Stokes' Theorem $\Rightarrow$

$$\oint_{C_1} \vec{F}\cdot d\vec{r} = \oint_{C_2} \vec{F}\cdot d\vec{r} \qquad (\partial W = C_1 - C_2)$$

boundary; opposite orientation

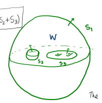
$$\oint_{C_3} \vec{F}\cdot d\vec{r} = 0 \qquad (\partial W' = C_3)$$

We say that the **closed curves** C_1 and C_2 are **homologous** and that C_3 is **homologous to zero**

means $\partial C_1 = \partial C_2 = \phi$

$\boxed{\partial W = S_1 - (S_2 + S_3)}$

opposite orientation



If $\text{div}\,\vec{F} = 0$ then

Divergence Theorem $\Rightarrow$

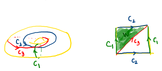
$$\iint_{S_1} \vec{F}\cdot\vec{n} = \iint_{S_2} \vec{F}\cdot\vec{n} + \iint_{S_3} \vec{F}\cdot\vec{n}$$

The **closed surface** S_1 is homologous to the union of S_2 and S_3

Poincaré: given a space X, count the number of **independent, closed** k-dimensional "surfaces" in X **up to homology**.

$\rightsquigarrow b_k(X)$ number which should depend only on the shape of X. This number is called the k-th **Betti number** of X.

Example:



$\partial W = C_3 - (C_1 + C_2)$

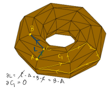
C_1 and C_2 are independent but C_3 is homologous to $C_1 + C_2$

$b_1(\text{torus}) = 2$ (corresponding to C_1 and C_2)

How to make this precise? Discretize by decomposing a space X into the simplest possible shapes:

dimension	0	1	2	3	...	n
shape	point	line segment	triangle	tetrahedron	...	**n-simplex**

$\{(x_1, \dots, x_n) \in \mathbb{R}^n : x_i \geq 0,\ x_1 + \dots + x_n \leq 1\}$



← discretized or **triangulated** torus

Poincaré: Consider only k-dimensional surfaces that can be expressed as combinations of the k-simplices in your discretized space

There, it is easy to define boundary and "closed":

Definition: Consider the chain of vector spaces

$$\cdots \to C_k(X) \xrightarrow{\partial_k} C_{k-1}(X) \xrightarrow{\partial_{k-1}} \cdots \to C_2(X) \xrightarrow{\partial_2} C_1(X) \xrightarrow{\partial_1} C_0(X)$$

vector space spanned by the k-simplices $\cong \mathbb{R}^N$ with $N = \#$ of k-simplices (with chosen orientations)

boundary naturally defined on the basis elements and extended linearly

$\boxed{\partial_{k-1} \circ \partial_k = 0}$

closed k-surfaces = Kernel of $\partial: C_k(X) \to C_{k-1}(X)$

$b_k(X)$ = dimension of the vector space obtained by identifying elements in $\text{Ker}\,\partial_k$ which differ by the image of ∂_{k+1}

Example: X

$\partial m = C - A$

$C_1(X) = \mathbb{R}^3$ basis corresponding to m, k, l

$$\downarrow \partial = \begin{bmatrix} -1 & 0 & -1 \\ 0 & -1 & 1 \\ 1 & 1 & 0 \end{bmatrix}$$

$C_0(X) = \mathbb{R}^3$ basis corresponding to A, B, C

Kernel $\partial = \mathbb{R}\cdot(1, -1, 1)$ ← $m - k + l$

opposite orientation

(algebraic manifestation of curve going around the triangle)

no 2d simplices $\Rightarrow$ no identifications $b_1(X) = 1$

In $C_0(X)$ we have to identify the points which differ by boundaries. They all do (e.g. $B - A = \partial l$) so $b_0(X) = 1$

Exercise: In general $b_0(X)$ = number of separate pieces (connected components) of X

General idea:

$b_0(X) = \#$ of components of X

$b_1(X) = \#$ of independent, closed, oriented curves in X up to the boundary of 2d pieces of X

$b_2(X) = \#$ of independent closed oriented surfaces in X up to boundary of 3d pieces of X

⋮

→ It turns out that **homotopy equivalent spaces have the same b_k's**

3 Topological data analysis

What is the shape of this point cloud?

Given a set S of points in $\mathbb{R}^N$ can associate to it a list of triangulated spaces X_r, one for each $r > 0$ (the **Vietoris-Rips** complexes of S):

the k-simplices correspond to the sets of $k+1$ points $x_0, x_1, \dots, x_k \in S$ such that $d(x_i, x_j) < r$

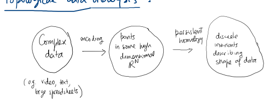
Example: S

r	r = 0.01	r = 1	r = 1.5	r = 2	r = 10
X_r					5-simplex
$b_0(X_r)$	5	3	1	1	1
$b_1(X_r)$	0	0	1	1	0

X_r always "starts" as a discrete set of points
"ends" as a (giant) simplex

Persistent homology: "In between" the betti numbers $b_i(X_r)$ reflect the perceived shape

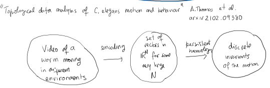
Topological data analysis:



Since the encoding is always somewhat arbitrary it is good that the result is not very dependent on details!

Random example from a paper appearing last week:

"Topological data analysis of C. elegans motion and behavior" A. Thomas et al. arxiv 2102.09380



Authors find that they can determine properties of the movement (e.g. backward/forward) and environment (e.g. acidity) from topological signatures.

To learn more:

- analysis-situs.math.cnrs.fr (a fantastic site about Poincaré's work by a collective of french mathematicians)
- "Topology and data" Gunnar Carlsson, Bull. Amer. Math. Soc. 46 (2009), 255-308
- " video of talk at Microsoft Research (youtube)