Topological data analysis
Lecture 1

Anton Ayzenberg

ATA Lab, FCS NRU HSE
Noeon Research

Spring 2024
Faculty of Computer Science / Yandex Data School

Topology

To start with

Topology = study of shapes.

Topological space = subset of $\mathbb{R}^d$. Which spaces are considered the same?

Topology

To start with

Topology = study of shapes.

Topological space = subset of $\mathbb{R}^d$. Which spaces are considered the same?

What you see:

What topologist sees:

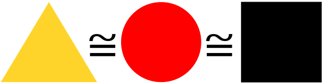
Homeomorphism

Definition

Let X, Y be topological spaces. $f\colon X \to Y$ is called a homeomorphism if

1. f is continuous;
2. f is a bijection;
3. f^{-1} is also continuous.

If there exists a homeomorphism $f\colon X \to Y$, then X, Y are called homeomorphic spaces. Notation $X \cong Y$.

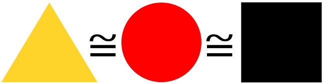


Definition

Let X, Y be topological spaces. $f\colon X \to Y$ is called a homeomorphism if

1. f is continuous;
2. f is a bijection;
3. f^{-1} is also continuous.

If there exists a homeomorphism $f\colon X \to Y$, then X, Y are called homeomorphic spaces. Notation $X \cong Y$.



Example: $(-1;1) \cong \mathbb{R}$.

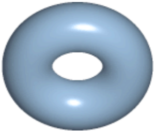
Example: Flat square is homeomorphic to flat circle.

Fact 1: $\cong$ acts like equivalence relation: if $X \cong Y$ and $Y \cong Z$, then $X \cong Z$.

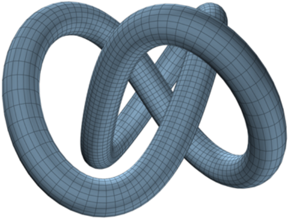
Fact 2: if $X \cong Y$ and X is compact, then Y is compact.

Coffee cup = donut

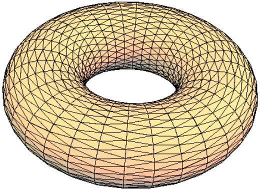
$\cong$



Also homeomorphic, but do not continuously deform to each other



$\cong$



Invariants

We want to distinguish non-homeomorphic shapes. Use **invariants** of homeomorphism.

Invariants

We want to distinguish non-homeomorphic shapes. Use **invariants** of homeomorphism.

The fundamental invariant in topology

Number of connected components. The set of connected components of X is denoted $\pi_0(X)$.

Fact: $X \cong Y$ implies X, Y have the same number of connected components.

Invariants

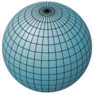
We want to distinguish non-homeomorphic shapes. Use **invariants** of homeomorphism.

The fundamental invariant in topology

Number of connected components. The set of connected components of X is denoted $\pi_0(X)$.

Fact: $X \cong Y$ implies X, Y have the same number of connected components.

But we have:

 $\ncong$  $\ncong$  $\ncong$ 

We need more invariants to distinguish them.

Ideas of invariants

- Naive invariants motivated by connectivity

Ideas of invariants

- Naive invariants motivated by connectivity
- Internal dimension

Ideas of invariants

- Naive invariants motivated by connectivity
- Internal dimension
- Number of holes, fundamental group

Ideas of invariants

- Naive invariants motivated by connectivity
- Internal dimension
- Number of holes, fundamental group
- Local properties

Ideas of invariants

- Naive invariants motivated by connectivity
- Internal dimension
- Number of holes, fundamental group
- Local properties
- Orientability

Ideas of invariants

- Naive invariants motivated by connectivity
- Internal dimension
- Number of holes, fundamental group
- Local properties
- Orientability
- Universal constructions $X \mapsto \mathcal{F}(X)$ such that $X \cong Y$ implies $\mathcal{F}(X) \cong \mathcal{F}(Y)$

Homotopy between maps

Equivalence of maps

Continuous maps $f, g\colon X \to Y$ are called homotopy equivalent, if one can be continuously deformed to another. Formally $f \sim g$ iff there exists a continuous map $F\colon X \times [0,1] \to Y$ such that $F(x,0) = f(x)$ and $F(x,1) = g(x)$.

Symbol $f \simeq g$.

Homotopy between maps

Equivalence of maps

Continuous maps $f, g\colon X \to Y$ are called homotopy equivalent, if one can be continuously deformed to another. Formally $f \sim g$ iff there exists a continuous map $F\colon X \times [0,1] \to Y$ such that $F(x,0) = f(x)$ and $F(x,1) = g(x)$.

Symbol $f \simeq g$.

Facts:

- Homotopy equivalence is equivalence.
- Homotopy is path in the space $Y^X = \mathrm{Maps}(X, Y)$ of continuous maps from X to Y between f and g.

Homotopy equivalent spaces

Homotopy equivalence of spaces

Two topological spaces X, Y are called homotopy equivalent, if there exist continuous maps $h\colon X \to Y$ and $k\colon Y \to X$, such that $h \circ k \simeq \mathrm{id}_Y$ and $k \circ h \simeq \mathrm{id}_X$.

Symbol $X \simeq Y$.

Homotopy equivalent spaces

Homotopy equivalence of spaces

Two topological spaces X, Y are called homotopy equivalent, if there exist continuous maps $h\colon X \to Y$ and $k\colon Y \to X$, such that $h \circ k \simeq \mathrm{id}_Y$ and $k \circ h \simeq \mathrm{id}_X$.

Symbol $X \simeq Y$. Informally: two spaces are homotopy equivalent iff one is obtained from another by thinning and thickening.

Homotopy equivalent spaces

Homotopy equivalence of spaces

Two topological spaces X, Y are called homotopy equivalent, if there exist continuous maps $h\colon X \to Y$ and $k\colon Y \to X$, such that $h \circ k \simeq \mathrm{id}_Y$ and $k \circ h \simeq \mathrm{id}_X$.

Symbol $X \simeq Y$. Informally: two spaces are homotopy equivalent iff one is obtained from another by thinning and thickening.

A space X is called contractible if $X \simeq \mathrm{pt}$ (pt is a 1-point space).

Homotopy equivalent spaces

Homotopy equivalence of spaces

Two topological spaces X, Y are called homotopy equivalent, if there exist continuous maps $h\colon X \to Y$ and $k\colon Y \to X$, such that $h \circ k \simeq \mathrm{id}_Y$ and $k \circ h \simeq \mathrm{id}_X$.

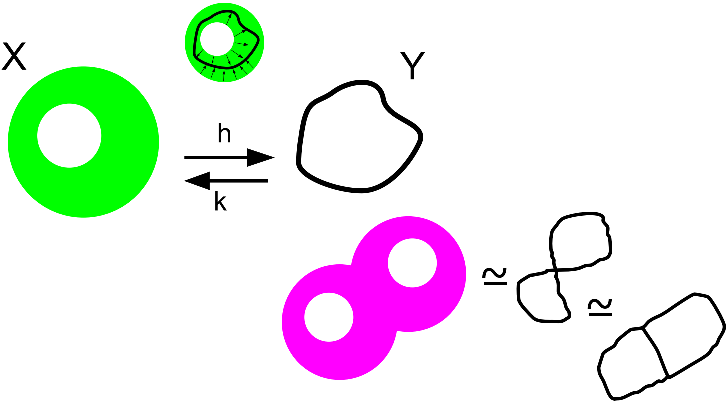
Symbol $X \simeq Y$. Informally: two spaces are homotopy equivalent iff one is obtained from another by thinning and thickening.

A space X is called contractible if $X \simeq \mathrm{pt}$ (pt is a 1-point space).

Facts:

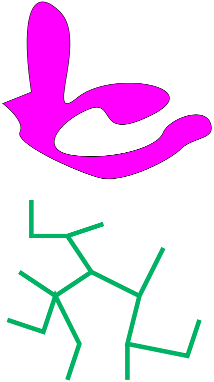
- Homotopy equivalence is an equivalence.
- Convex sets of $\mathbb{R}^n$ are contractible.
- A graph (its picture) is contractible iff it is a tree.

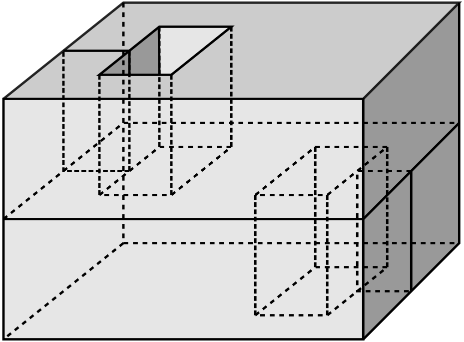
Homotopy equivalence: some pictures



We allow to make objects thin and thick.

Contractible spaces





Bing's house is an example of a contractible space which cannot be contracted to a point in a tree-like manner.

Homotopy invariants

Invariants of homotopy equivalence:

- Number of connected components

Homotopy invariants

Invariants of homotopy equivalence:

- Number of connected components
- Fundamental group $\pi_1(X)$

Homotopy invariants

Invariants of homotopy equivalence:

- Number of connected components
- Fundamental group $\pi_1(X)$
- Homology vector spaces and Betti numbers (to be discussed)

Homotopy invariants

Invariants of homotopy equivalence:

- Number of connected components
- Fundamental group $\pi_1(X)$
- Homology vector spaces and Betti numbers (to be discussed)

Not invariants:

- Dimension

Homotopy invariants

Invariants of homotopy equivalence:

- Number of connected components
- Fundamental group $\pi_1(X)$
- Homology vector spaces and Betti numbers (to be discussed)

Not invariants:

- Dimension
- Local things

Homotopy invariants

Invariants of homotopy equivalence:

- Number of connected components
- Fundamental group $\pi_1(X)$
- Homology vector spaces and Betti numbers (to be discussed)

Not invariants:

- Dimension
- Local things
- Orientability

Constructivity

How can we explain shapes to computer?

Constructivity

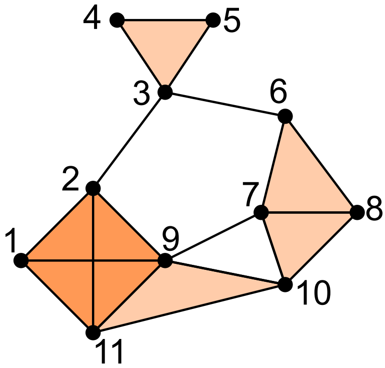
How can we explain shapes to computer?

Basically:

1. Formulas (and their geometrical interpretations)
2. Discrete data structures (and their topological interpretations)

such as graphs, simplicial complexes, partially ordered sets, etc.

Simplicial complex



(-1)-мерный: $\varnothing$

0-мерные:
{1},{2},{3},{4},{5},{6},{7},{8},{9},{10},{11}

1-мерные:
{1,2},{1,9},{2,9},{1,11},{2,11},{9,11},{2,3},
{3,4},{2,5},{4,5},{3,6},{6,7},{6,8},{7,8},
{7,10},{8,10},{7,9},{9,10},{10,11}

2-мерные:
{1,2,9},{1,2,11},{1,9,11},{2,9,11},
{3,4,5},{6,7,8},{7,8,10},{9,10,11}

3-мерные:
{1,2,9,11}

Simplicial complex

Definition

Simplicial complex on a finite vertex set V is a collection $K \subset 2^V$ satisfying the properties:

1. if $I \in K$ and $J \subset I$, then $J \in K$;
2. $\varnothing \in K$.

Elements $I \in K$ are called simplices. If $|I| = k$, we say that I is a $(k-1)$-dimensional simplex.

1. Vertices $\{i\}$ — simplices of dim 0;
2. Edges $\{i, j\}$ — simplices of dim 1;
3. Triangles $\{i, j, k\}$ — simplices of dim 2;
4. etc.

$\dim K$ is the maximal dimension of simplices of K.

Geometrical realizations

- Graph is (a) a discrete object, (b) a picture.
- Some graphs cannot be drawn in $\mathbb{R}^2$ without self-intersections.
- But all graphs can be drawn in $\mathbb{R}^3$

Geometrical realizations

- Graph is (a) a discrete object, (b) a picture.
- Some graphs cannot be drawn in $\mathbb{R}^2$ without self-intersections.
- But all graphs can be drawn in $\mathbb{R}^3$

Similarly:

- Simplicial complex K is a discrete object.

Geometrical realizations

- Graph is (a) a discrete object, (b) a picture.
- Some graphs cannot be drawn in $\mathbb{R}^2$ without self-intersections.
- But all graphs can be drawn in $\mathbb{R}^3$

Similarly:

- Simplicial complex K is a discrete object.
- In order to understand it as a continuous topological space, some picture in $\mathbb{R}^d$ should be drawn. It is called **the geometrical realization** of K and denoted $|K|$.

Geometrical realizations

- Graph is (a) a discrete object, (b) a picture.
- Some graphs cannot be drawn in $\mathbb{R}^2$ without self-intersections.
- But all graphs can be drawn in $\mathbb{R}^3$

Similarly:

- Simplicial complex K is a discrete object.
- In order to understand it as a continuous topological space, some picture in $\mathbb{R}^d$ should be drawn. It is called **the geometrical realization** of K and denoted $|K|$.
- It is easy to draw simplicial complex in the space of dimension $d = |V|$.

Fact: simplicial complex of dim k can be drawn in $\mathbb{R}^{2k+1}$ without self-intersections.

Geometrical realizations

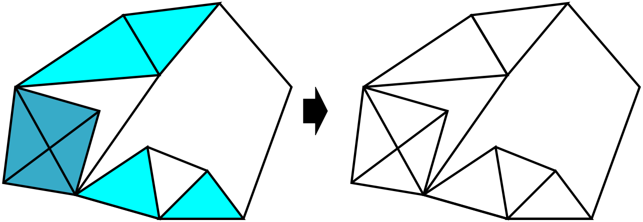
Simplicial complex is a discrete structure and can be encoded in computer. But in general:

- There is no algorithm to check $|K| \cong |L|$ given K and L.
- There is no algorithm to check $|K| \simeq |L|$ given K and L.
- There is no even an algorithm to check $|K| \simeq \mathrm{pt}$ given K!

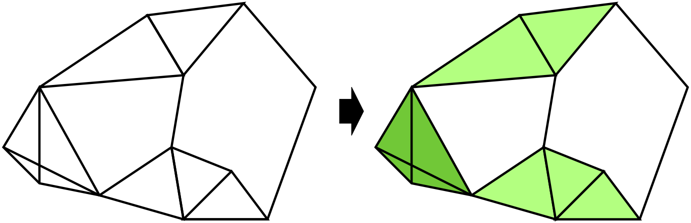
1-dimensional simplicial complexes (aka graphs) are simpler:

- Homeomorphism of two graphs can be checked algorithmically.
- Homotopy equivalence of two graphs can be checked algorithmically.

1-skeleton



Clique complex



*Clique = subgraph isomorphic to full graph.

Clique complex

Clique complex makes it possible to transform a graph into a high-dimensional structure.

Since graphs are everywhere, this observation opens a way to use topological invariants everywhere.