Space from data

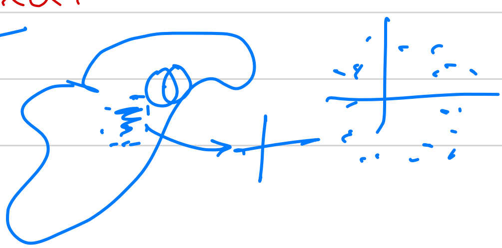
Real world data (finite set of points in $\mathbb{R}^d$ or a metric sp.)

→ as it self it is a boring topological space.

→ it is not endowed with a simplicial structure.

How to associate a Space to the data?

<u>Assumption</u> : Your data is sampled from a nice topological space, say a compact manifold.



Today's plan

- Examples of simplicial structures.
- nerve of a covering.
- Data offsets & the homology inference theorem
- The nerve complex & the nerve lemma
- Cech & VR complex.

Recall: An abstract simplicial complex on a finite set V is a collection $\mathcal{K}$ of subsets of V that have the heridetary property, i.e., if $A \in \mathcal{K}$ then $\forall\ B \subseteq A$, $B \in \mathcal{K}$.

Given an abstract simpl^l cplx here is how one constructs its geometric realization. $\|\mathcal{K}\|$

Example $\underbrace{V}_{\text{ground set}} = \{1, 2, 3\}$, $\mathcal{K} = \mathcal{P}(V)$

e_1 e_2 e_3

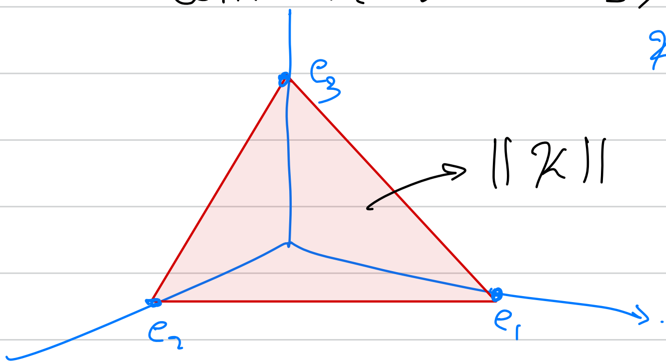
Step 1 : Consider $\mathbb{R}^{|V|}$, in our example $\mathbb{R}^3$. The standard unit vectors corr. to the elts of V.

$\|K\| = \bigcup_{B \in K} \mathrm{conv}(\{e_i \mid i \in B\})$

Step 2 : if $B \in K$ then

$\mathrm{Convex}(\{e_i \mid i \in B\}) \in \|K\|$

$K = P(V)$



$\|\emptyset$ 1, 2, 3

12 13 23

123

Two examples from graph theory

we'll consider finite simple graph.

① Independence complex : V = the set of all vertices of a graph G.

A subset $A \subset V$ is called independent if the induced subgraph $G[A]$ is discrete.

If A is an independent set then so is

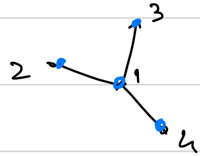
every subset of A.

$\mathcal{K}$ = Collection of all ind. subsets of G

$\therefore$ $(V, \mathcal{K})$ is an A.S.C.

$G[\{1\}] = \bullet$

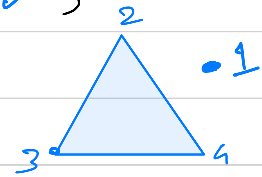
Ex.



$\mathcal{K} = \{1, 2, 3, 4, 23, 24, 34, 234\}$

$G[1,2] =$ 

$G[2,3] =$ 

$\|\mathcal{K}\| =$ 

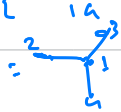
② The clique cplx : A subset $A \subset V$ is a clique if $G[A]$ is a complete graph.

($\mathcal{K}$ = coll^n of cliques)

$(V, \mathcal{K})$ is an A.S.C.

Ex.

$\mathcal{K} = \{12, 13, 14\}$

$\|\mathcal{K}\| =$ 

③ Nerve of a covering. Let X be a set and $\mathcal{U}$ be a coll^n of subsets

$$\mathcal{U} = \{U_1, \ldots, U_n\}$$

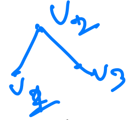
s.t. $X = U_1 \cup \cdots \cup U_n$

Let $V = \{1, \ldots, n\}$

$\mathcal{K}$ be the coll^n of subsets of $\{1, \ldots, n\}$

$X = \mathbb{R}$

$$\mathcal{U} = \{\underbrace{(-\infty, 0)}_{U_1}, \underbrace{(-1, 1)}_{U_2}, \underbrace{(0, \infty)}_{U_3}\}$$



s.t. if $J \subset \{1, \dots, n\}$ $\quad$ $U_1 \cap U_3 = \emptyset$, $\quad$ $U_1 \cap U_2 \neq \emptyset$

then $\bigcap_{j \in J} U_j$ is non empty. $\quad$ $U_2 \cap U_3 \neq \emptyset$.

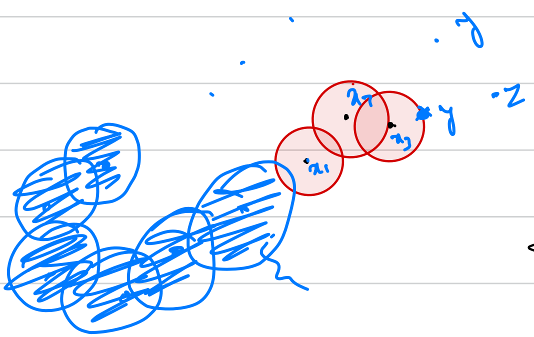
$(V, \mathcal{K})$ is called the nerve of this covering. it is an ASC.

Let X be a pt. cloud in $\mathbb{R}^d$, $|X| = n$.

The r-offset of X, $r \in [0, \infty)$

$$X^r := \bigcup_{x \in X} B(x; r) \subseteq \mathbb{R}^d$$

closed ball of center x & radius r.



For $y \in \mathbb{R}^d$

$$d(y, X) := \inf_{x \in X} \| x - y \|$$

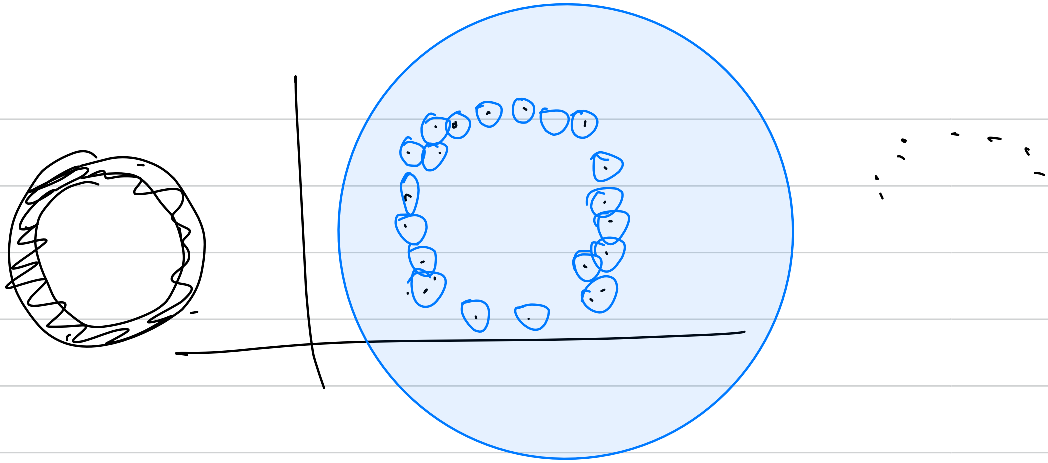
$$X^r = d^{-1}([0, r])$$

$$\boxed{\underset{(-,X)}{d} : \mathbb{R}^d \to \mathbb{R}_{\geq 0}}$$

Homology inference theorem : adjectives

The Betti numbers of nice Riemanian mflds can be recovered with high probability from r-offsets of a sample on (or close to) the mfld.

w.r.t. the Gromov-Hausdorff dist.



<u>Nerve of an open Covering</u> Let X be a top. space and <u>$\mathcal{U}$ be an open</u> cover (i.e., $\mathcal{U}$ is a collection of open sets in X s.t. their union covers X).

$$\text{nerve}(\mathcal{U}) = \mathcal{N}(\mathcal{U}) := \{\underline{U_i \in \mathcal{U} \mid \cap U_i \neq \emptyset}\}$$

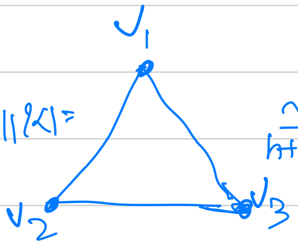
The vertices or the 0-simplices of nerve $(\mathcal{U})$ corr. to open sets.

1 - simplices $\longleftrightarrow$ pairs of open sets which intersect non trivially

2 - simplices $\longleftrightarrow$ triples with nonempty intersection

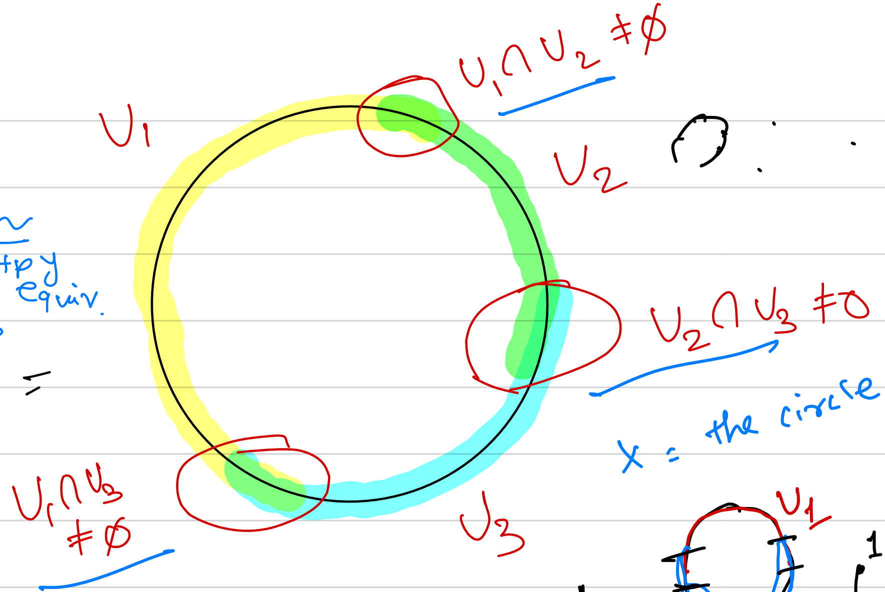
& so on.

$\|K\| =$



$\simeq$ htpy equiv.

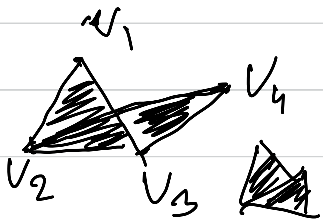
$X =$

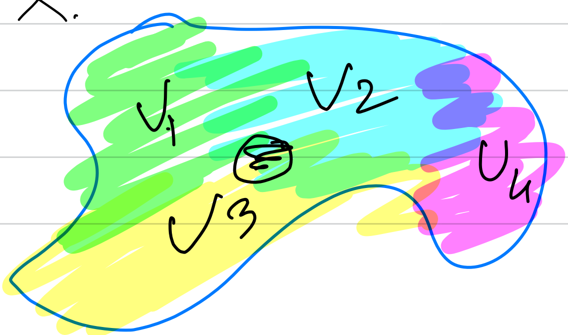


nerve$(\mathcal{U})$, $\left(V = \{1, 2, 3\} \right.$

$\left. K = \{1, 2, 3, 12, 23, 13\} \right)$

The nerve theorem: Let $\mathcal{U}$ be an open or closed cover of X s.t. each $U \in \mathcal{U}$ is Contractible and so are all the non empty intersections. Then the nerve$(\mathcal{U})$ has the same htpy type as that of X.





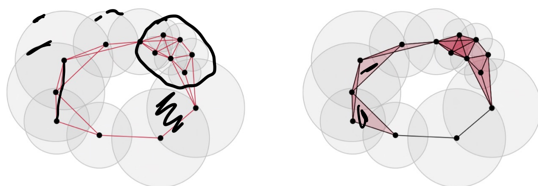
check

The Čech complex

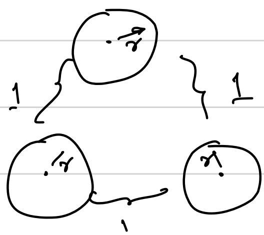
Let X be a PCD

γ is a real number

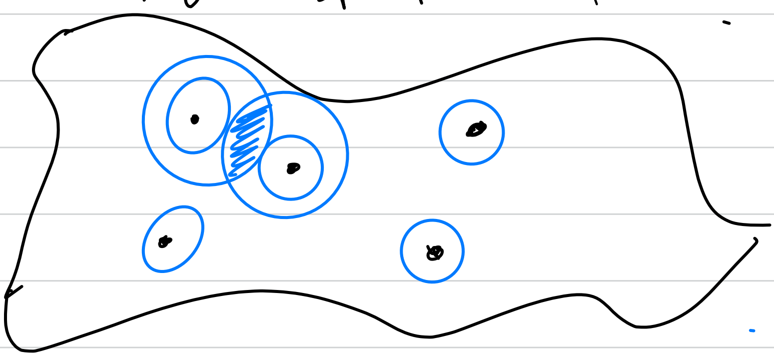
$$Cech_\gamma(X) := \left\{ \sigma \subset X \;\middle|\; \bigcap_{x \in \sigma} B(x, \gamma) \neq \emptyset \right\}$$



The 1-skeleton of the nerve   The 2-skeleton of the nerve



Metric space



running parameter γ

$\gamma = 1$

1 2 3 4 5

Vietoris - Rips Complex

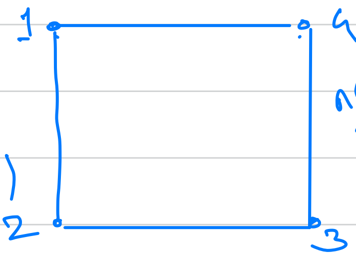
A finite simple graph can be considered as a metric space.

The vertices are elts of the metric space.

"every edge has length 1"

distance betⁿ two pts is the graph distance.

$X = (G, d)$



network

$d(1,2) = 1$

$d(1,4) = 1$

$d(1,3) = 2$

$d(2,4) = 2$

$d(3,4) = 1$

$d(2,3) = 1$

$r = 0$

$VR_0(X) = \{1, 2, 3, 4\}$

$VR_1(X) = \{1, 2, 3, 4, 12, 23, 34, 14\}$

$VR_2(X) = VR_1(X) \cup \{123, 134, 234, 123, 13, 24, 1234\}$