Lecture 7

Announcements

OH moved to Thurs 10 am

HW due Friday

project proposals next week

Today:

Pairs, Barcodes, Diagrams

Bottleneck Distance

Algebraic Features, Landscapes

Persistent Homology:

$X_0 \subseteq X_1 \subseteq X_2 \subseteq \ldots$ Filtrations

$R(X; r_0) \subseteq R(X, r_1) \subseteq \ldots$

Persistent homology:

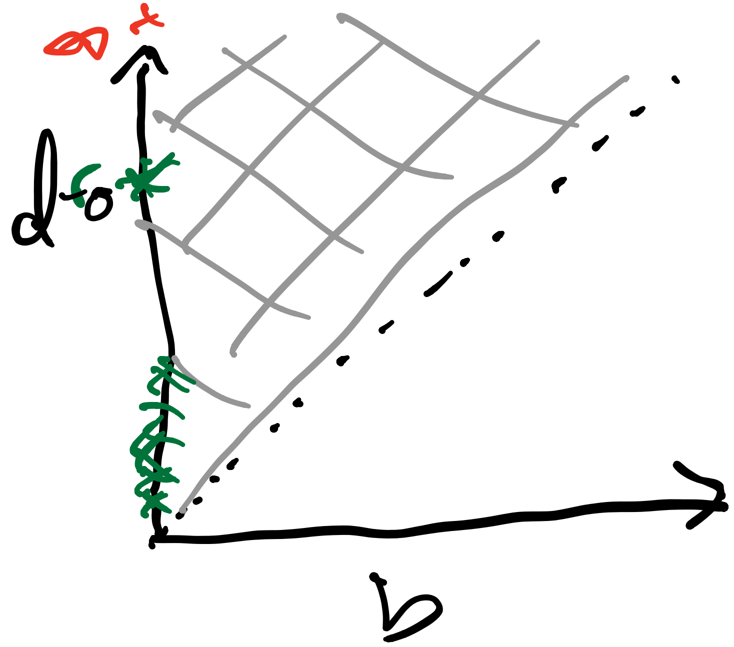
$PH_k(X): \{(b_i, d_i)\}$

$d_i = \infty$ if homology class at end

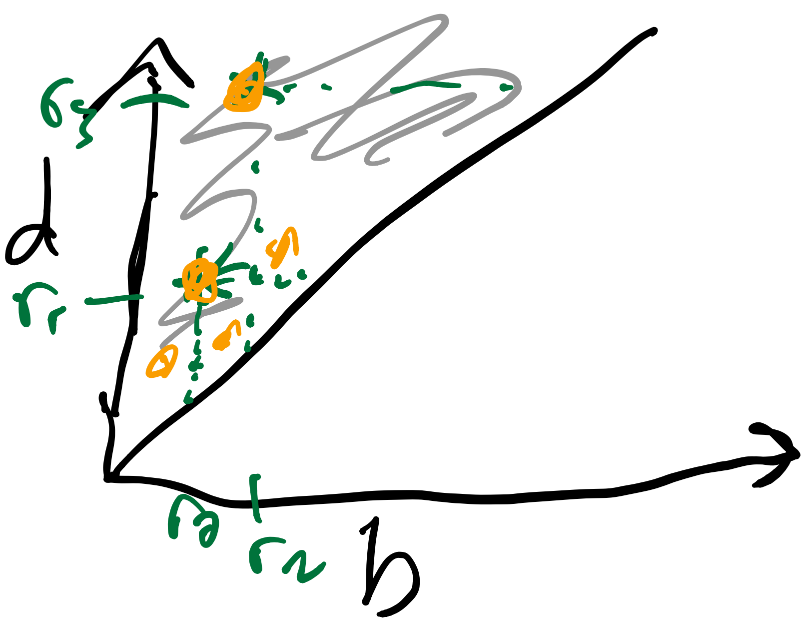
"persistence pair"

Persistence Diagram: $\{(b_i, d_i)\}$

$\in \mathbb{R}^2$

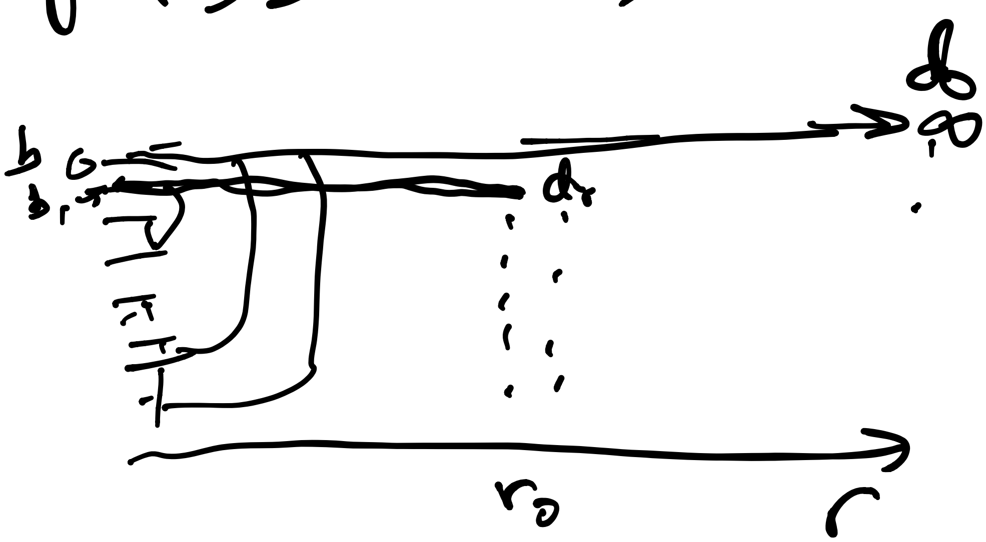


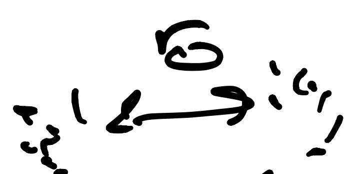
$PH_0(R(X))$



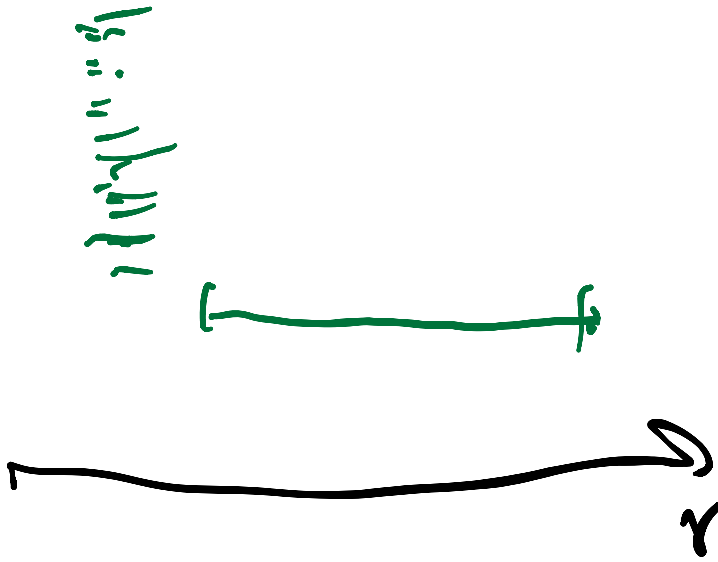
$PH_1(R(X))$

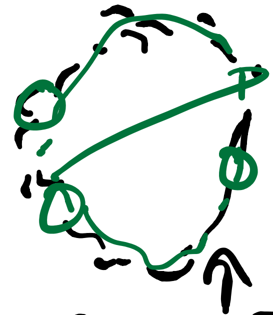
Persistence Barcode: $\{(b_i, d_i)\}$





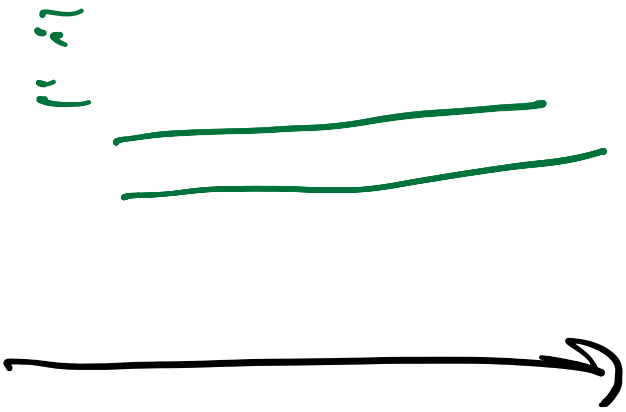
$PH_0(R(X))$

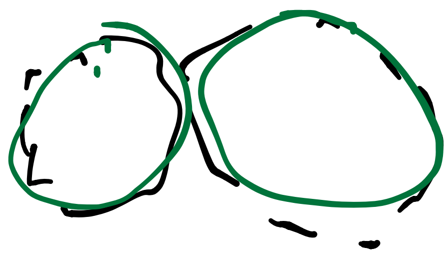




$PH_1(R(X))$

$\{(b_i, d_i)\}$





Pairs $\longleftrightarrow$ Diagrams $\longleftrightarrow$ Barcodes

Applied Math:

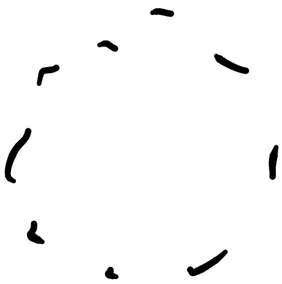
Define something computable

Perturbation Theory

How robust is our procedure?

PH_1

PH_1



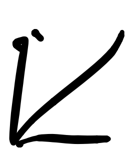
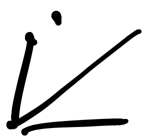
X_0

X_1

$X_0, X_1 \underset{\text{fd}}{\simeq} S^1$

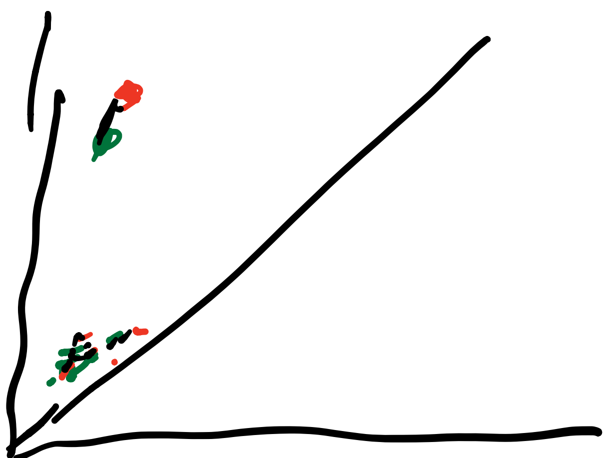
Bottleneck distance on Persistence D.

$d_B(D_1, D_2)$

$d_B($  ,  $)$

$d_B : \inf_{\gamma} \sup_{x} \underbrace{\|x - \gamma(x)\|_\infty}_{\text{distortion}}$

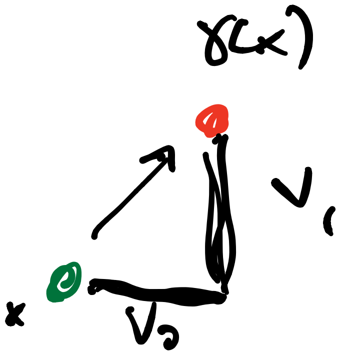
γ matching between pairs



d

b

$x\ (b_0, d_0)$	(b_0, d_0)
(b_1, d_1)	(b_1, d_1)
(b_2, d_2)	(b_2, d_2)
$\vdots$	$\vdots$

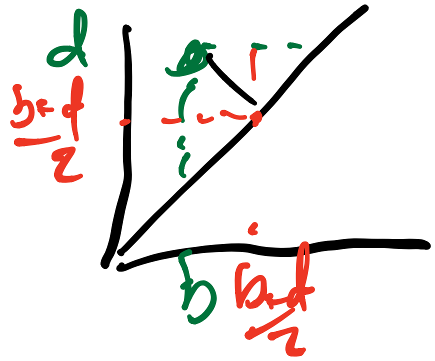


$\|v\|_\infty = \max_i |v_i|$

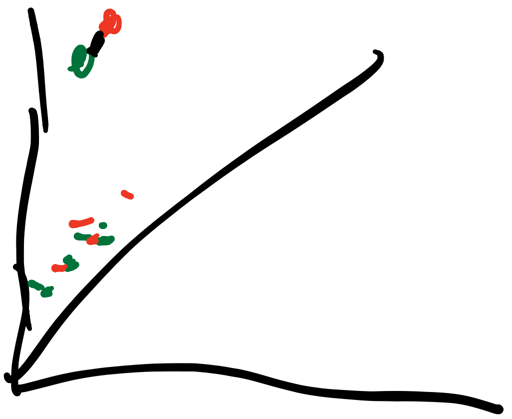
$\max\{|b - \gamma b|, |d - \gamma d|\}$

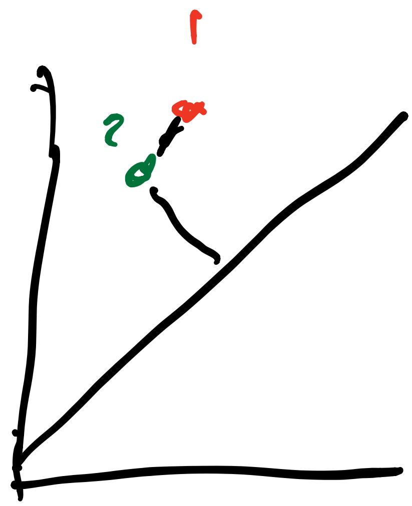
what if the pairs is not the same?

- add all points on the diagonal



$d_B(K', K) = \sup_{(b_i, d_i)} \frac{d_i - b_i}{2}$





$\{(b, d), (b, d)\}$

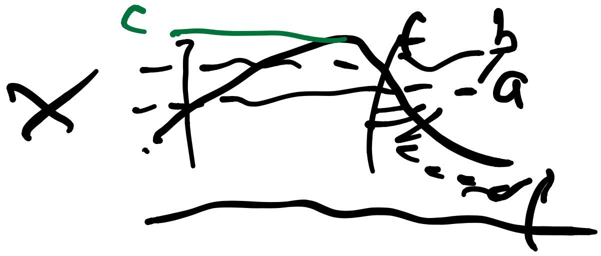
$$d_B = \inf_\gamma \sup_x \|x - \gamma(x)\|$$

$x \in \{(b_i, d_i)\} \cup \Delta$

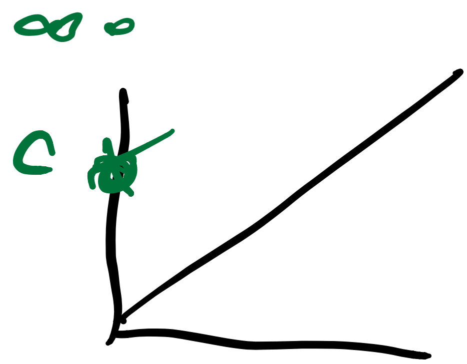
notion of distance

perturbation theory?

Sub-levelset filtrations



$f: X \to \mathbb{R}$



$f^{-1}((-\infty, a]) : X_a$

$X_a \subseteq X_b$ if $a \leq b$

Def: Let X be a top space, $f: X \to \mathbb{R}$

A <u>homological critical value</u> of f is a real number $\underline{a}$ for which $\exists$ an integer k, for which $\forall \varepsilon > 0$ the map induced by inclusion

$$H_k(f^{-1}(-\infty, a-\varepsilon]) \longrightarrow H_k(f^{-1}(-\infty, a+\varepsilon]))$$

is <u>not</u> an isomorphism

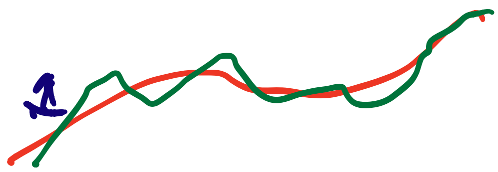
Def: A function $f: X \to \mathbb{R}$ is <u>tame</u> if it has a finite number of hom. crit. values, and and the homology groups $H_k(f^{-1}((-\infty, a]))$ are finite dimensional $\forall k \in \mathbb{Z}$, $\forall a \in \mathbb{R}$

Cohen-Steiner, Edelsbrunner, Harer 2007

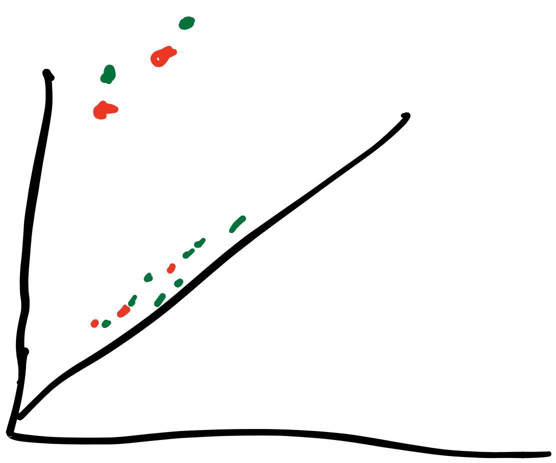
Let X be a triangulable space w/ cts. tame functions $f, g: X \to \mathbb{R}$

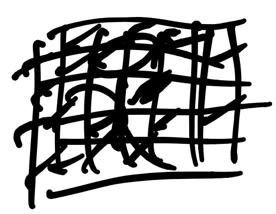
Then

$$d_B(D_k(f), D_k(g)) \le \|f-g\|_\infty$$



$$\|f-g\|_\infty := \sup_{x\in X} |f(x)-g(x)|$$





Gromov-Hausdorff Stable Signatures for Shapes using Persistence

Chazal et al 2009

$X, Y \subseteq \mathbb{R}^2$
Z

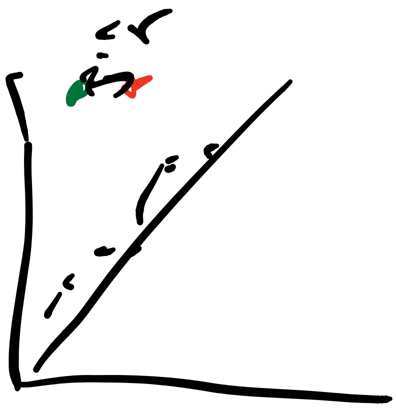
def: Hausdorff distance

Let $X, Y \subseteq Z$ (some ambient metric space)
so we can compute $d_Z(x,y)$ $\forall x \in X, y \in Y$

$$d_H = \max\left\{ \sup_x \inf_y \|x-y\|_\infty , \sup_y \inf_x \|x-y\|_\infty \right\}$$

$$\max\{r, s\} = r$$

thm: $d_B(D_k(R(X)), D_k(R(Y))) \leq d_H(X,Y)$



---

def: The Gromov-Hausdorff distance btw compact metric spaces (X, d_X), (Y, d_Y):
$d_{GH}((X, d_X), (Y, d_Y))$:

$d_{GH}((X,d_X),(Y,d_Y))$

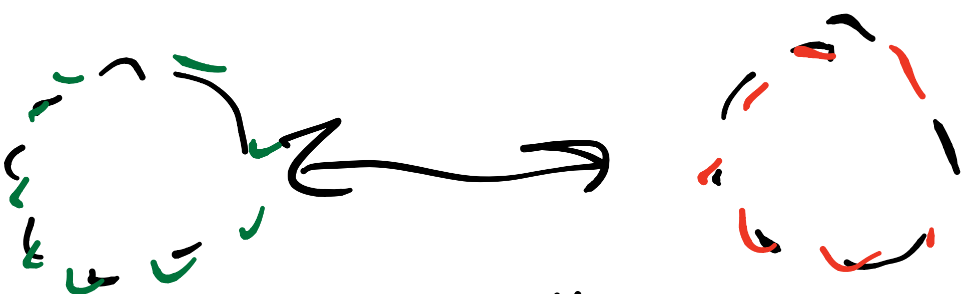
$$\inf_{Z,\gamma_X,\gamma_Y} d_H^Z(\gamma_X(X),\gamma_Y(Y)))$$

Z some metric space

γ_X: isometric embedding $X \to Z$

γ_Y " " $Y \to Z$

$$d_{GH}((X,d_1),(X,d_2))$$



d_{GH} small

Thm: $d_B(D_K(R(X,d_X)), D_K(R(Y,d_Y)))$
$\le d_{GH}((X,d_X),(Y,d_Y))$

Comment: d_B is a version of d_H for persistence diagrams.

Exercise: prove d_B is a metric

Challenge of using PDs:
are multisets of points in $\mathbb{R}^2$

are matrices of p...

stats/ML assumes data in $\mathbb{R}^n$

Idea 1: Adcock, Carlsson, Carlsson 2014

Algebraic fns of barcodes

$$\{(b_i, d_i)\} \to \sum_i (d_i - b_i)^p (d_i + b_i)^q$$

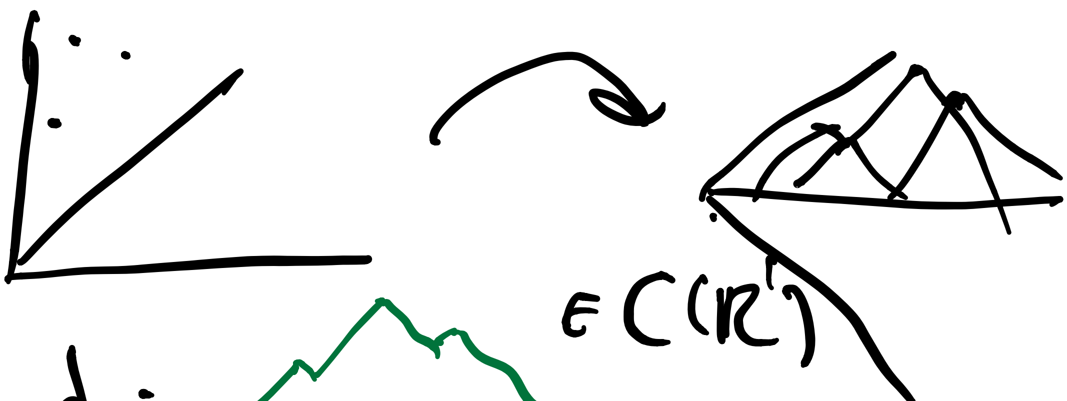
p, q can be anything

choose a couple values of p, q

$\to$ PDs $\to \mathbb{R}^n \to$ ML pipeline

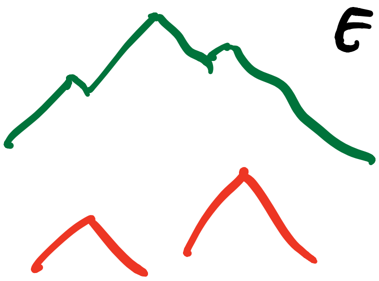
How to take avgs of PDs?

Persistence Landscapes: Bubenik 2016



$\in C(\mathbb{R}^1)$

λ_1 :



λ_2 :

discretize to form vector in $\mathbb{R}^n$