KTH ROYAL INSTITUTE
OF TECHNOLOGY

Lecture II. The reach of a manifold

Algebraic Geometry with a view towards applications

Sandra Di Rocco, ICTP Trieste

Plan for this course

- Lecture I: Algebraic modelling (Kinematics)
- **Lecture II: Sampling algebraic varieties: the reach.**
- Lecture III: Projective embeddings and Polar classes (classical theory)
- Lecture IV: The Euclidian Distance Degree (closest point)
- Lecture V: Bottle Neck degree from classical geometry (back to sampling)

Main goals

- Definition of reach.
- Sampling.
- Recovering the topological signature.

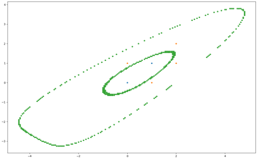
The reach of a manifold

References:

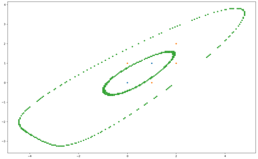
- *Estimating the reach of a Manifold.* E. Aamari, J. Kim, F. Chazal, B. Michel, A. Rinaldo, et al.. Electronic journal of Statistics, Vol. 13, **2019** 1359-1399.
- *Computing the homology of basic semialgebraic sets in weak exponential time* P. Bürgisser, F. Cucker and P. Lairez. Journal of the ACM 66(1), **2019**.
- *Learning algebraic varieties from samples*. P. Breiding, S. Kalisnik, B. Sturmfels and M. Weinstein. Revista Matemática Complutense. Vol. 31, 3, **2018**, pp 545-593.
- *Sampling Algebraic Varieties*. DR, D. Eklund, O. Gävfert. **In progress.**

(Real) Sampling

(Real) Sampling

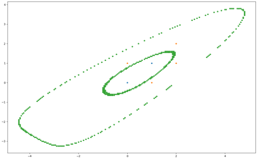


(Real) Sampling



$X \subset \mathbb{R}^N, I_X \Rightarrow$ Cloud data on $X \Rightarrow$
(CW) Complex $\Rightarrow$ Invariants of X

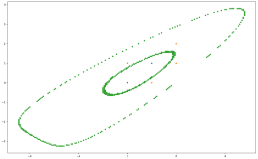
(Real) Sampling



$$X \subset \mathbb{R}^N, I_X \Rightarrow \text{ Cloud data on } X \Rightarrow$$
$$\text{(CW) Complex } \Rightarrow \text{ Invariants of } X$$

Two important steps

(Real) Sampling

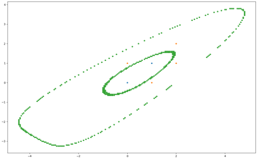


$$X \subset \mathbb{R}^N, l_X \Rightarrow \text{ Cloud data on } X \Rightarrow$$
$$\text{(CW) Complex } \Rightarrow \text{ Invariants of } X$$

Two important steps

- **density**

(Real) Sampling

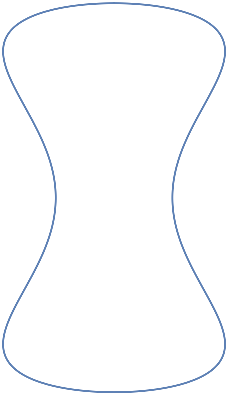


$$X \subset \mathbb{R}^N, I_X \Rightarrow \text{ Cloud data on } X \Rightarrow$$
$$\text{(CW) Complex } \Rightarrow \text{ Invariants of } X$$

Two important steps

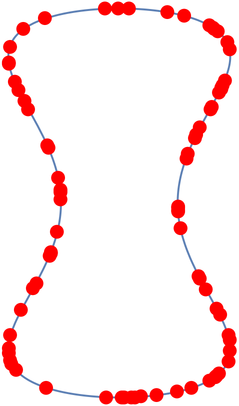
- **density**
- **complex**

Application: Variety Sampling



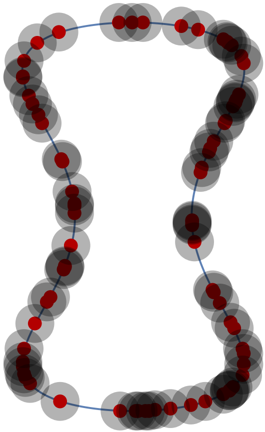
- Consider a compact submanifold $M \subseteq \mathbb{R}^n$.

Application: Variety Sampling



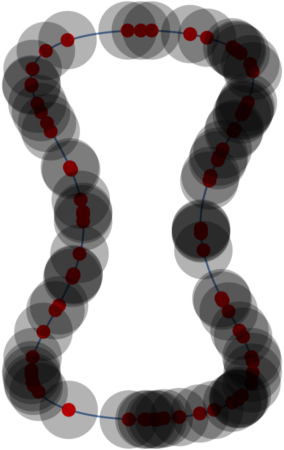
- Consider a compact submanifold $M \subseteq \mathbb{R}^n$.
- And a point cloud $E \subset M$.

Application: Variety Sampling



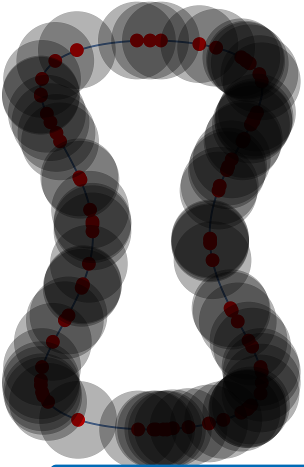
- Consider a compact submanifold $M \subseteq \mathbb{R}^n$.
- And a point cloud $E \subset M$.
- Put growing balls around the points E.

Application: Variety Sampling



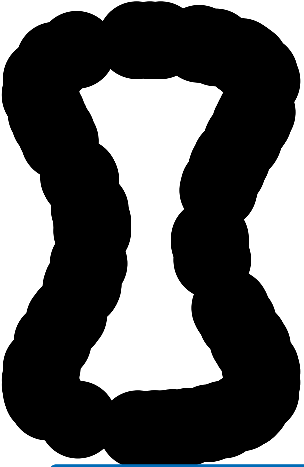
- Consider a compact submanifold $M \subseteq \mathbb{R}^n$.
- And a point cloud $E \subset M$.
- Put growing balls around the points E.

Application: Variety Sampling



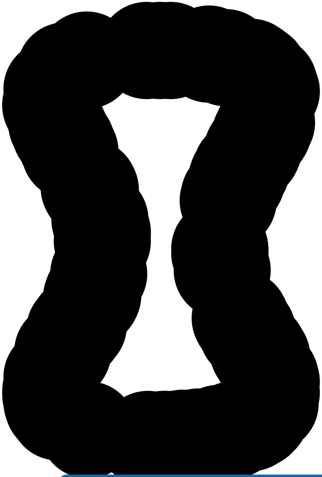
- Consider a compact submanifold $M \subseteq \mathbb{R}^n$.
- And a point cloud $E \subset M$.
- Put growing balls around the points E.
- Geometry of M captured by union of balls for ball sizes in certain interval.

Application: Variety Sampling



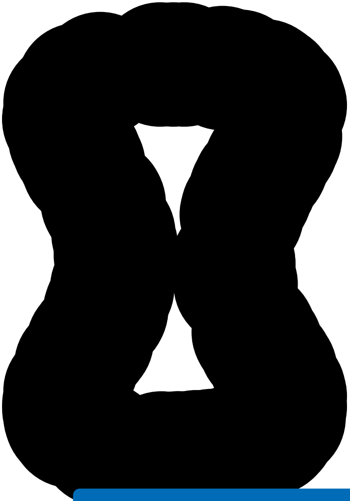
- Consider a compact submanifold $M \subseteq \mathbb{R}^n$.
- And a point cloud $E \subset M$.
- Put growing balls around the points E.
- Geometry of M captured by union of balls for ball sizes in certain interval.

Application: Variety Sampling



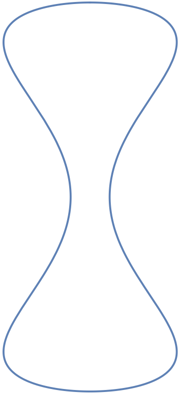
- Consider a compact submanifold $M \subseteq \mathbb{R}^n$.
- And a point cloud $E \subset M$.
- Put growing balls around the points E.
- Geometry of M captured by union of balls for ball sizes in certain interval.

Application: Variety Sampling

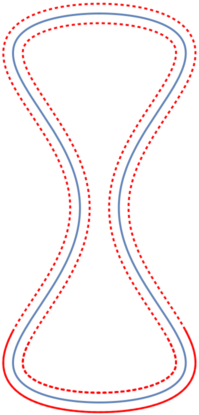


- Consider a compact submanifold $M \subseteq \mathbb{R}^n$.
- And a point cloud $E \subset M$.
- Put growing balls around the points E.
- Geometry of M captured by union of balls for ball sizes in certain interval.

Application: Variety Sampling

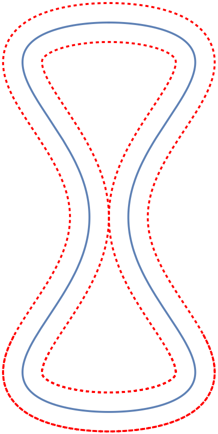


Application: Variety Sampling



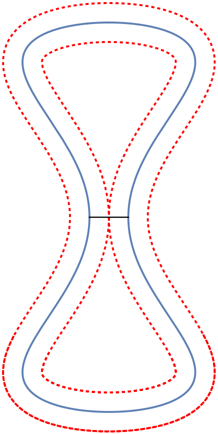
- Consider a growing tubular neighborhood of M.

Application: Variety Sampling



- Consider a growing tubular neighborhood of M.
- At some point it becomes singular ($M \neq \{p\}$).

Application: Variety Sampling



- Consider a growing tubular neighborhood of M.
- At some point it becomes singular ($M \neq \{p\}$).
- Half the black distance is called the *reach* of M.
- And the line is a *bottleneck*.

Set up

Consider $\mathbb{R}^n$, endowed with the euclidean inner product $< x, y >= \sum x_i y_i$.
Let $X \subset \mathbb{R}^N$ be a smooth compact algebraic variety defined by an ideal $I = (f_1, ..., f_k) \subset \mathbb{R}[x_1, ..., x_n]$.
Let $p \in X$ and consider the Jacobian matrix

$$J(f_1, ..., f_k) = (\frac{\partial f_i}{\partial x_j})_{i,j}$$

evaluated at p.

- Since X is smooth, the rows of the Jacobian matrix span an $(n-d)$-dimensional linear subspace of $\mathbb{R}^n$, where d is the local dimension of X at p.
- This subspace translated to the point p is called the **normal space** of X at p, and is denoted $N_p(X)$ (fibers of the *Normal Bundle*).

For $p \in \mathbb{R}^n$, Consider the **distance function**

$$d_X(p) = min_{x \in X}||x - p||$$

and

$$\pi_X(p) = \{x \in X \mid ||x - p|| = d_X(p)\}$$

For $p \in \mathbb{R}^n$, Consider the **distance function**

$$d_X(p) = min_{x \in X}||x - p||$$

and

$$\pi_X(p) = \{x \in X \,|\, ||x - p|| = d_X(p)\}$$

For $r \geqslant 0$ let X_r the *tubular neighborhood* of X of radius r

$$X_r = \{p \in \mathbb{R}^n \,|\, d_X(p) < r\}$$

For $p \in \mathbb{R}^n$, Consider the **distance function**

$$d_X(p) = min_{x \in X} ||x - p||$$

and

$$\pi_X(p) = \{x \in X \mid ||x - p|| = d_X(p)\}$$

For $r \geqslant 0$ let X_r the *tubular neighborhood* of X of radius r

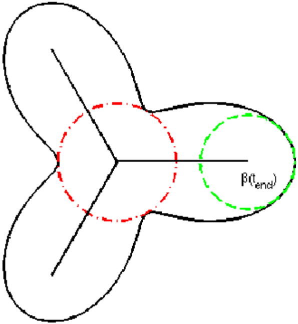
$$X_r = \{p \in \mathbb{R}^n \mid d_X(p) < r\}$$

$\Delta_X = \{p \in \mathbb{R}^n : \pi_X(p) > 1\}$, the **medial axis**

$$M_X = \overline{\Delta_X}$$

$\Delta_X = \{p \in \mathbb{R}^n : \pi_X(p) > 1\}$, the **medial axis**

$$M_X = \overline{\Delta_X}$$



Picture: *International Conference on Cyberworlds (CW'07).* Henning Naß, Franz-Erich Wolter and Hannes Thielhelm.DOI:10.1109/CW.2007.55

The Reach

The Reach

Define **the reach of** X as:

$$\tau_X = inf_{x \in X, y \in \Delta_X} \{d(x, y)\}$$

The Reach

Define **the reach of** X as:

$$\tau_X = inf_{x \in X, y \in \Delta_X} \{d(x, y)\}$$

Notice that for each point $x \in X_r$ where $r < \tau$ $|\pi_X(p)| = 1$. This point is called the **closest point** of p on X.

The Reach

Define **the reach of** X as:

$$\tau_X = inf_{x \in X, y \in \Delta_X}\{d(x, y)\}$$

Notice that for each point $x \in X_r$ where $r < \tau$ $|\pi_X(p)| = 1$.
This point is called the **closest point** of p on X.
Facts:

- X is compact, thus $\tau_X > 0$.
- dim $X > 0$ (not convex) $\tau_X < \infty$.
- τ_X is a combination of local and global estimates:

- *Locally*: ρ_X is the **minimal radius of curvature** on X (radius of curvature at $x \in X$ is the reciprocal of the maximal curvature of a geodesic passing through x.)

- *Locally*: ρ_X is the **minimal radius of curvature** on X (radius of curvature at $x \in X$ is the reciprocal of the maximal curvature of a geodesic passing through x.)
- *Globally*: **bottlenecks**:

$$\eta_X = \frac{1}{2} \inf\{||x - y|| : (x, y) \in X \times X,\ x \neq y,\ y \in N_x X,\ x \in N_y X\}.$$

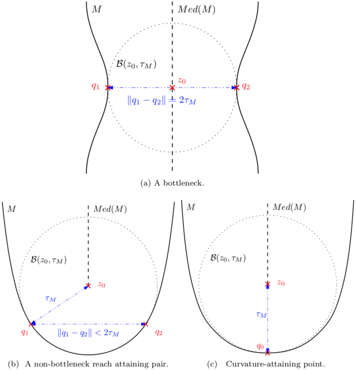
- *Locally*: ρ_X is the **minimal radius of curvature** on X (radius of curvature at $x \in X$ is the reciprocal of the maximal curvature of a geodesic passing through x.)
- *Globally*: **bottlenecks**:

$$\eta_X = \frac{1}{2} \inf\{||x-y|| : (x, y) \in X \times X,\ x \neq y,\ y \in N_x X,\ x \in N_y X\}.$$

$$\tau_X = min\{\rho_X, \eta_X\}$$

From:Estimating the reach of a Manifold

From:Estimating the reach of a Manifold



Picture: *Estimating the reach of a Manifold.* E. Aamari, J. Kim, F. Chazal, B. Michel, A. Rinaldo, et al.

Sample

Sample

$X \subset \mathbb{R}^N$ smooth, compact variety.

Definition

Let $\varepsilon > 0$. A finite set of points $E \subset X$ is called an ε-*sample* of X if for every $x \in X$ there is a point $e \in E$ such that $llx - ell < \varepsilon$.

Sample

$X \subset \mathbb{R}^N$ smooth, compact variety.

Definition

Let $\varepsilon > 0$. A finite set of points $E \subset X$ is called an ε-*sample* of X if for every $x \in X$ there is a point $e \in E$ such that $llx - ell < \varepsilon$.

Definition

- Let $q \in \mathbb{R}^N$ and let $c(q, \varepsilon) \subset \mathbb{R}^N$ denote the closed *ball* centered at q and of radius ε.
- Consider

$$C(\varepsilon, E) = \bigcup_{q \in E} c(q, \varepsilon).$$

Sample gives homology

Sample gives homology

Theorem

Let $\varepsilon > 0$ and consider E an $\frac{\varepsilon}{2}$ sample of X. If $\varepsilon < \frac{1}{2}\tau$ for all $p \in E$ then $X \hookrightarrow C(E, \varepsilon)$ is a homotopy equivalence.

Sample gives homology

Theorem

Let $\varepsilon > 0$ and consider E an $\frac{\varepsilon}{2}$ sample of X. If $\varepsilon < \frac{1}{2}\tau$ for all $p \in E$ then $X \hookrightarrow C(E, \varepsilon)$ is a homotopy equivalence.

proof:

- $\pi_X : C(E, \varepsilon) \to X$ is continuous.
- $C(E, \varepsilon) \times [0, 1] \to C(E, \varepsilon)$ defines as $(x, t) \mapsto t\pi_X(x) + (1 - t)x$ is a deformation retract. To prove: $t\pi_X(x) + (1 - t)x \in C(E, \varepsilon)$.

One way to sample via Numerical AG

One way to sample via Numerical AG

Let $X \in \mathbb{R}^N$ be a variety of dimension d. Let $T_d \subset \{1, 2, \ldots, N\}$ be the set of unordered d-tuples. Let $e_1, \ldots, e_N$ be the standard basis. For $t = (t_1, \ldots, t_d) \in T_d$ we let $V_t = Span(e_{t_i}, \ldots, e_{t_d})$. For $\delta > 0$ consider the grid

$$G_t(\delta) = \delta\mathbb{Z} \cap V_t$$

and the projection $\pi_t : \mathbb{R}^N \to V_t$. Then

$$E_\delta = \bigcup_{t \in T_d, g \in G_t} X \cap \pi_t^{-1}(g)$$

is a finite sample (up to random perturbation).

One way to sample via Numerical AG

One way to sample via Numerical AG

Theorem

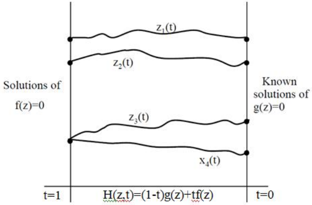
If $0 < \delta < \frac{\varepsilon}{\sqrt{N}}$ *and* $E_\delta \neq \varnothing$*, then* E_δ *is* ε *sample of* X.

One way to sample via Numerical AG

Theorem

If $0 < \delta < \frac{\varepsilon}{\sqrt{N}}$ *and* $E_\delta \neq \varnothing$, *then* E_δ *is* ε *sample of* X.

Use **numerical methods** to construct sample:

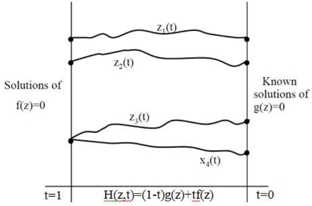


One way to sample via Numerical AG

Theorem

If $0 < \delta < \frac{\varepsilon}{\sqrt{N}}$ *and* $E_\delta \neq \emptyset$*, then* E_δ *is* ε *sample of* X.

Use **numerical methods** to construct sample:



- Bertini: Bates-Hauenstein-Sommese-Wampler
- HomotopyContinuation.jl: Paul Breiding and **Sascha Timme**

Simplicial Complex

A *simplicial complex* on X is a collection K of subsets of X such that if $\sigma \in K$ and $\tau \subseteq \sigma$, then $\tau \in K$.

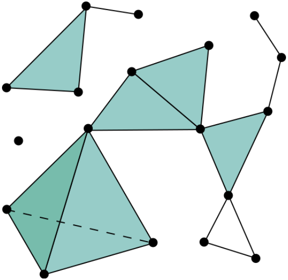


Figure: Geometric realization of a simplicial complex. Photo credit: Wikipedia

From Data to Simplicial Complexes

Consider a finite metric space (X, d).
The Vietoris-Rips complex on X at scale ε, $\mathrm{VR(X)}_\varepsilon$ consists of :

- singletons $\{x\}$, for all $x \in X$.
- sets $\{x_0, \ldots, x_n\} \subseteq X$, such that $d(x_i, x_j) \leqslant \varepsilon$ for all $0 \leqslant i, j \leqslant n$.

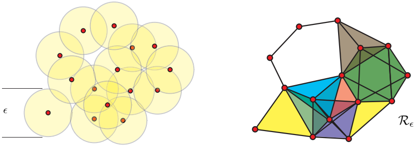
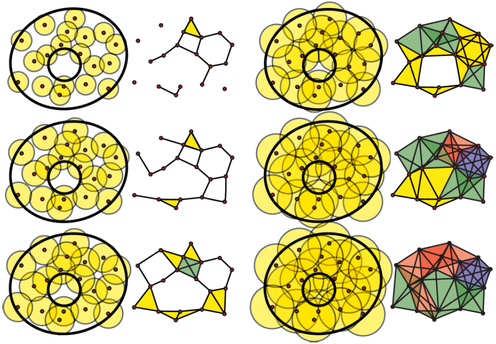


Figure: $\mathrm{VR(X)}_\varepsilon$, Photo credit: R. Ghrist, 2008, Barcodes: The Persistent Topology of Data.

From Data to Simplicial Complexes

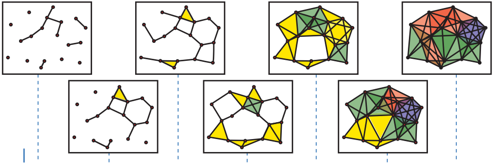
If $\varepsilon \leqslant \varepsilon'$, VR(X)$_\varepsilon$ is a sub-complex of VR(X)$_{\varepsilon'}$.
Therefore by considering increasing values of ε, we obtain a filtration of simplicial complexes.



Persistent Homology

Given a filtration of simplicial complexes:

$$X_0 \overset{i_0}{\hookrightarrow} X_1 \overset{i_1}{\hookrightarrow} \dots \overset{i_{k-1}}{\hookrightarrow} X_k \hookrightarrow \dots$$



fixed $n \in \mathbb{N}$ and a field K, we compute the n-th homology, with coefficients in K, of the filtration to obtain a parametrized vector space.

$$H_n(X_0) \overset{H_n(i_0)}{\rightarrow} H_n(X_1) \overset{H_n(i_1)}{\rightarrow} \dots \overset{H_n(i_{k-1})}{\rightarrow} H_n(X_k) \rightarrow \dots$$

Barcode

A parametrized vector space is completely described by a multi-set of half open intervals, commonly visualized as a barcode.

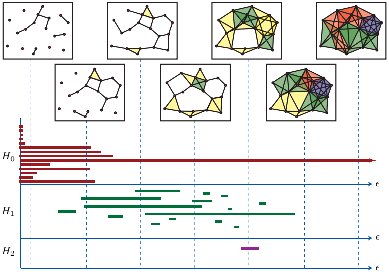


Figure: A filtration of simplicial complexes and associated barcode, Photo credit: R. Ghrist, 2008, Barcodes: The Persistent Topology of Data.

Summary

Summary

- Sampling provides data clouds to analyse
- The sampling recovers the underline topology if it is sufficiently fine
  - Reach
  - Complex
- Persistence may be applied in algebraic settings.