Discovering Metrics and Scale Space

Preliminary Examination

Brittany Terese Fasy

Duke University, Computer Science Department

26 July 2010

Questions

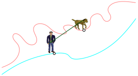
1 Tight Bound: Is the inequality

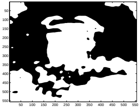
$$|\ell_1 - \ell_2| \leq \frac{4}{\pi} \cdot (\kappa_1 + \kappa_2) \cdot \mathcal{F}(\gamma_1, \gamma_2)$$

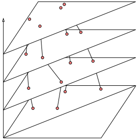
a tight bound for curves in $\mathbb{R}^n$ for $n > 3$?

2 Simultaneous Scale Space: What happens if multiple agents are diffusing at the same time?

3 Understanding V_q: Does there exist a stability result for V_q?







Part 1

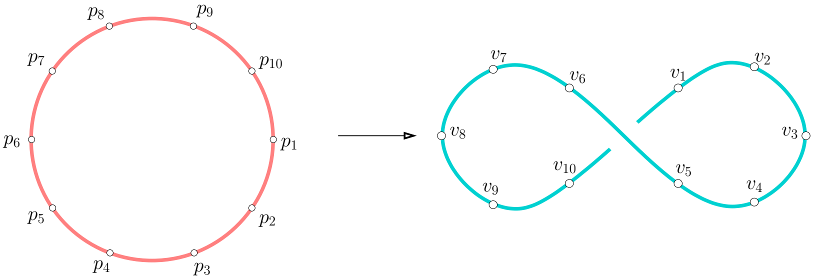
My Inequality

For curves γ_1 and γ_2,

$$|\ell_1 - \ell_2| \leq \frac{4}{\pi} \cdot (\kappa_1 + \kappa_2) \cdot \mathcal{F}(\gamma_1, \gamma_2)$$

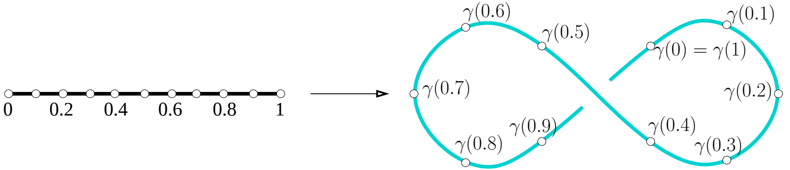
Closed Space Curves

A curve is a continuous map $\gamma_i \colon \mathbb{S}^1 \to \mathbb{R}^n$.



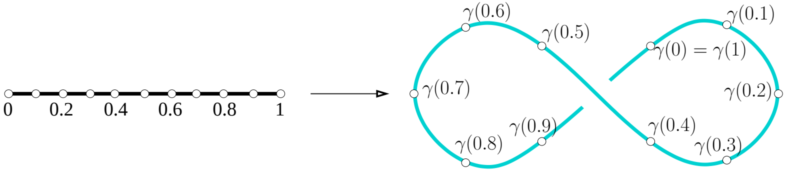
Closed Space Curves

A curve is a continuous map $\gamma_i \colon [0,1] \to \mathbb{R}^n$, such that $\gamma_i(0) = \gamma_i(1)$.

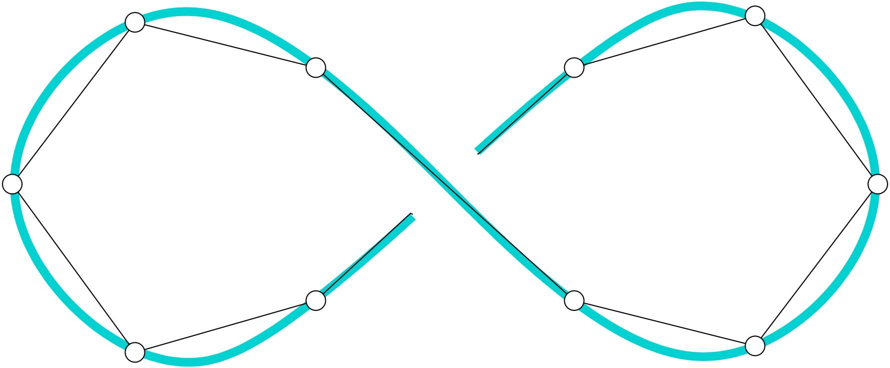


Closed Space Curves

A curve is a continuous map $\gamma_i \colon I \to \mathbb{R}^n$, such that $\gamma_i(0) = \gamma_i(1)$.

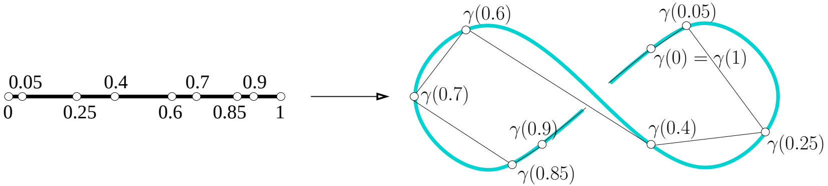


Inscribed Polygons



Closed Space Curves

A curve is a continuous map $\gamma_i \colon I \rightarrow \mathbb{R}^n$, such that $\gamma_i(0) = \gamma_i(1)$.



$$\mathrm{mesh}(P) = max_{0 \leq i < m} \left(t_{i+1} - t_i \right)$$

Arc Length

$$\ell_i = \ell(\gamma_i) = \int_0^1 ||\gamma_i'(t)||\, dt$$

$$\ell(P) = \sum_j \ell(e_j)$$

Arc Length

$$\ell_i = \ell(\gamma_i) = \int_0^1 ||\gamma_i'(t)||\, dt$$

$$\ell(P) = \sum_j \ell(e_j)$$

Lemma

If P^k is a sequence of polygons inscribed in a smooth closed curve γ such that mesh(P^k) goes to zero, then

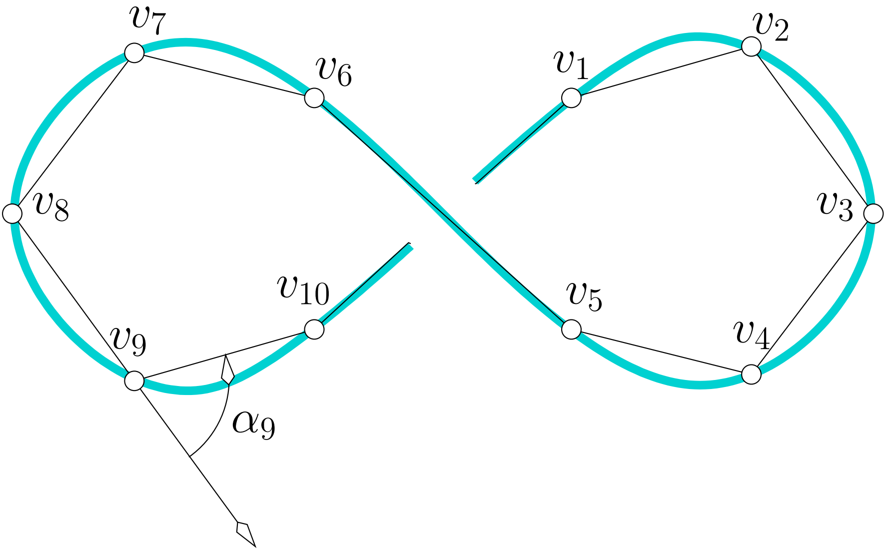
$$\ell(\gamma) = \lim_{k \to \infty} \ell(P^k).$$

Curvature

Let $x \in I$.
Let r_x denote the radius of the best approximating circle of $\gamma_i(x)$.
Then, the total curvature is:

$$\kappa_i = \kappa(\gamma_i) = \int_0^1 1/r_x \, dx.$$

Turning Angle



Curvature

Let $x \in I$.
Let r_x denote the radius of the best approximating circle of $\gamma_i(x)$.
Then, the total curvature is:

$$\kappa_i = \kappa(\gamma_i) = \int_0^1 1/r_x \, dx.$$

$$\kappa(P) = \sum_i \alpha_i.$$

Curvature

Let $x \in I$.
Let r_x denote the radius of the best approximating circle of $\gamma_i(x)$.
Then, the total curvature is:

$$\kappa_i = \kappa(\gamma_i) = \int_0^1 1/r_x \, dx.$$

$$\kappa(P) = \sum_i \alpha_i.$$

Lemma

If P^k is a sequence of polygons inscribed in a smooth closed curve γ such that $\mathrm{mesh}(P^k)$ goes to zero, then

$$\kappa(\gamma) = \lim_{k \to \infty} \kappa(P^k).$$

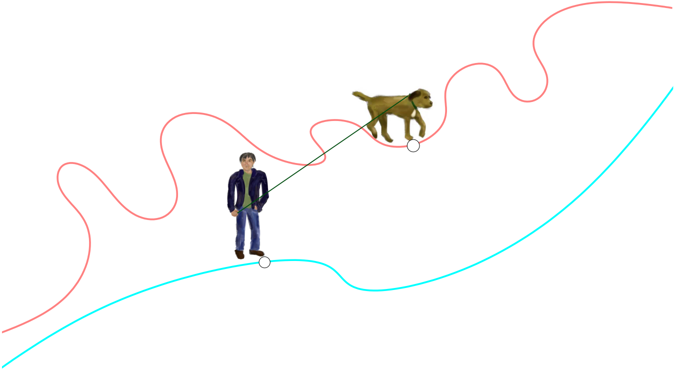
The Fréchet Distance

$$\mathcal{F}(\gamma_1, \gamma_2) = \inf_{\alpha:\, \mathbb{S}^1 \to \mathbb{S}^1} \max_{t \in \mathbb{S}^1} (\gamma(t) - \gamma(\alpha(t)))$$

Man and Dog

The Fréchet Distance

$$\mathcal{F}(\gamma_1, \gamma_2) = \inf_{\alpha\colon \mathbb{S}^1 \to \mathbb{S}^1} \max_{t \in \mathbb{S}^1} (\gamma(t) - \gamma(\alpha(t)))$$



The Fréchet Distance

$$\mathcal{F}(\gamma_1, \gamma_2) = \inf_{\alpha\colon \mathbb{S}^1 \to \mathbb{S}^1} \max_{t \in \mathbb{S}^1} (\gamma(t) - \gamma(\alpha(t)))$$

Lemma

If P^k and Q^k are sequences of polygons inscribed in smooth closed curves γ_1 and γ_2 such that mesh(P^k) and mesh(Q^k) go to zero, then

$$\mathcal{F}(\gamma_1, \gamma_2) = \lim_{k \to \infty} \mathcal{F}(P^k, Q^k).$$

My Inequality

For curves γ_1 and γ_2,

$$|\ell_1 - \ell_2| \leq \frac{4}{\pi} \cdot (\kappa_1 + \kappa_2) \cdot \mathcal{F}(\gamma_1, \gamma_2)$$

Two Related Theorems

- [Cha62, F5́0] For a closed curve contained in a disk of radius r in $\mathbb{R}^n$,

$$\ell_i \leq r \cdot \kappa_i.$$

- [CSE07] For two closed curves in $\mathbb{R}^n$,

$$|\ell_1 - \ell_2| \leq \frac{2\,\mathrm{vol}(\mathbb{S}^{n-1})}{\mathrm{vol}(\mathbb{S}^n)} \cdot (\kappa_1 + \kappa_2 - 2\pi) \cdot \mathcal{F}(\gamma_1, \gamma_2).$$

Questions

1 Tight Bound: Is the inequality

$$|\ell_1 - \ell_2| \leq \frac{4}{\pi} \cdot (\kappa_1 + \kappa_2) \cdot \mathcal{F}(\gamma_1, \gamma_2)$$

a tight bound for curves in $\mathbb{R}^n$ for $n > 3$?

2 Simultaneous Scale Space: What happens if multiple agents are diffusing at the same time?

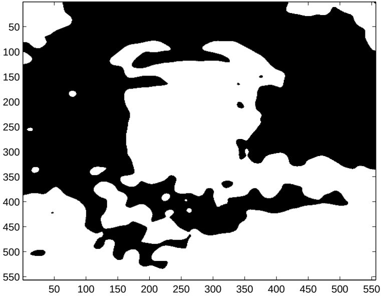
3 Understanding V_q: Does there exist a stability result for V_q?

Part 2

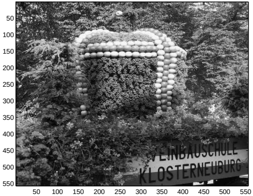
Scale Space

Scale Space

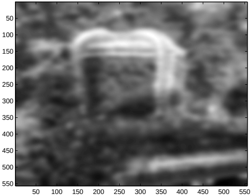

WEINBAUSCHULE
KLOSTERNEUBURG


50
100
150
200
250
300
350
400
450
500
550
50 100 150 200 250 300 350 400 450 500 550

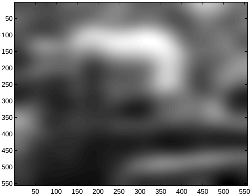
Scale Space



$H(x, 0)$

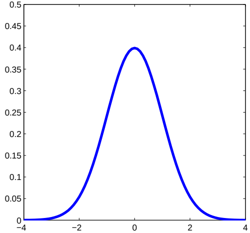


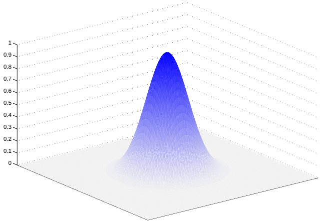
$H(x, 100)$



$H(x, 1000)$

The Gaussian





$$G_1(x,t) = \frac{1}{(\sqrt{2\pi}t)} e^{-\frac{x^2}{2t^2}} \qquad G_n(x,t) = \frac{1}{(\sqrt{2\pi}t)^n} e^{-\frac{|x|^2}{2t^2}}$$

This is the *fundamental solution* to the Heat Equation:

$$\frac{\partial u}{\partial t}(x,t) - \triangle\, u(x,t) = 0.$$

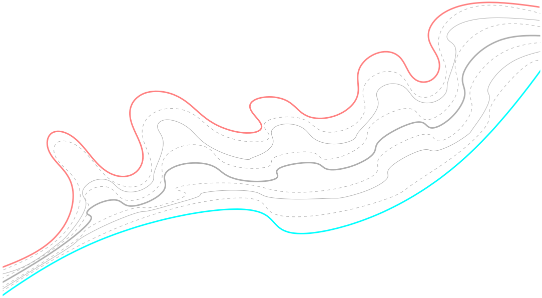
Gaussian Convolution

$$\mathrm{Blur}(y, t, h_0) = \int_{x \in \mathbb{R}^n} G_n(x - y, t) h_0(x) \, dy$$

Homotopy

$H\colon \mathbb{M} \times I \to \mathbb{R}$ is a continuous function such that

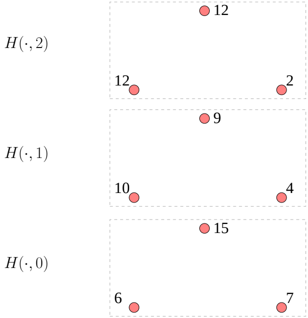
$$H(x, 0) = f(x)$$
$$H(x, 1) = g(x)$$



Discrete Homotopy

$H\colon \mathrm{vert}(K) \times \{0, 1, \ldots, \tau\} \to \mathbb{R}$ is a discrete function such that

$$H(x, 0) = f(x)$$
$$H(x, \tau) = g(x)$$

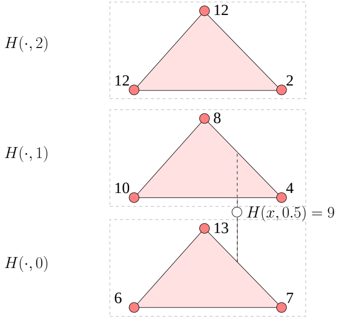


Discrete Homotopy

$H\colon K \times I_\tau \to \mathbb{R}$ is a discrete function such that

$$H(x,0) = f(x)$$
$$H(x,\tau) = g(x)$$

and the value at a general point $(x,t) \in \mathbb{M} \times I_\tau$ is determined by linear interpolation.



Heat Equation Homotopy

Let $f(x), g(x)\colon \mathbb{R}^2 \to \mathbb{R}$.
Let $h_0(x) = f(x) - g(x)$.
Then:

$$H(y,t) = h_t(y) = \mathrm{Blur}(y,t,h_0) = \int_{x \in \mathbb{R}^n} G_n(x-y,t) h_0(x)\, dy$$

is the solution to the heat equation with initial condition $H(x,0) = h_0(x)$.

Heat Equation Homotopy

Let $f(x), g(x)\colon \mathbb{R}^2 \to \mathbb{R}$.
Let $h_0(x) = f(x) - g(x)$.
Then:

$$H(y,t) = h_t(y) = \mathrm{Blur}(y,t,h_0) = \int_{x \in \mathbb{R}^n} G_n(x-y,t) h_0(x)\, dy$$

is the solution to the heat equation with initial condition $H(x,0) = h_0(x)$.

We define the Heat Equation Homotopy as:

$$H(x,t) := h_t(x) + g(x).$$

Questions

1. Tight Bound: Is the inequality

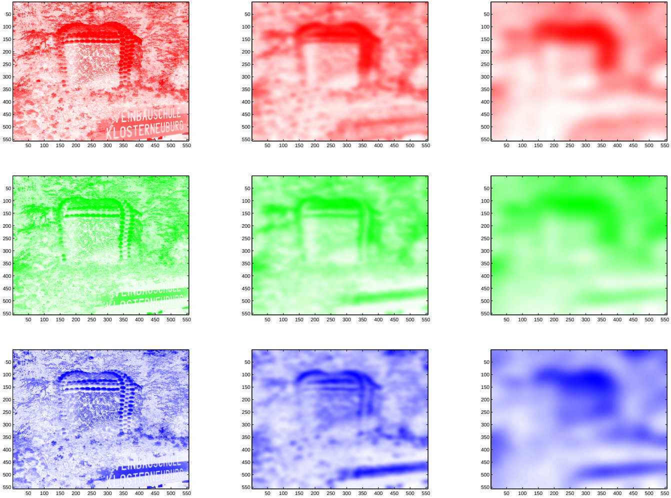
$$|\ell_1 - \ell_2| \leq \frac{4}{\pi} \cdot (\kappa_1 + \kappa_2) \cdot \mathcal{F}(\gamma_1, \gamma_2)$$

   a tight bound for curves in $\mathbb{R}^n$ for $n > 3$?
2. Simultaneous Scale Space: What happens if multiple agents are diffusing at the same time?
3. Understanding V_q: Does there exist a stability result for V_q?

Color Images

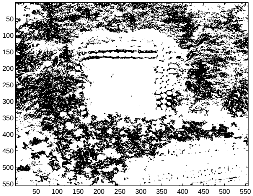
Color Images

$t = 0$ $t = 100$ $t = 1000$

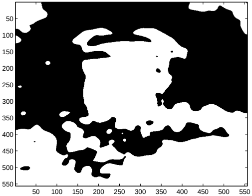


Proportion Set

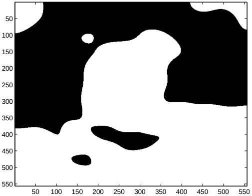
A *proportion set* for the RGB image is the set of pixels that have the same ratios of colors. For example, the boundaries in the following images depict where 4*blue = 3*green:



t = 0



t = 100



t = 1000

Questions

1. Tight Bound: Is the inequality

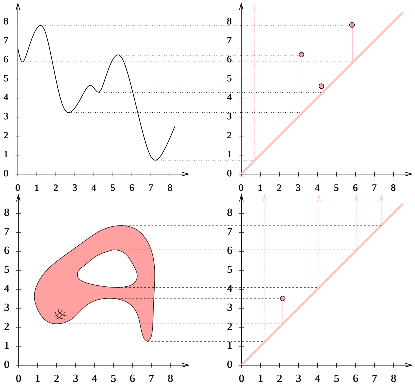
$$|\ell_1 - \ell_2| \leq \frac{4}{\pi} \cdot (\kappa_1 + \kappa_2) \cdot \mathcal{F}(\gamma_1, \gamma_2)$$

   a tight bound for curves in $\mathbb{R}^n$ for $n > 3$?
2. Simultaneous Scale Space: What happens if multiple agents are diffusing at the same time?
3. Understanding V_q: Does there exist a stability result for V_q?

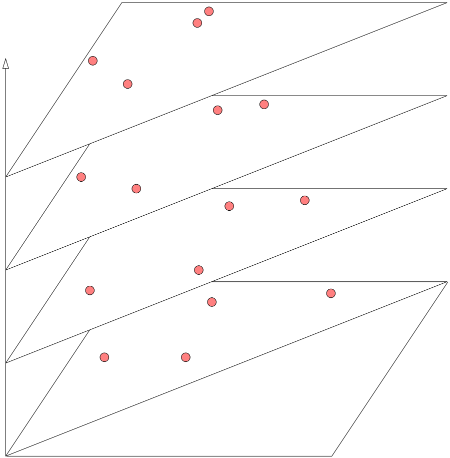
Part 3

Persistence Diagrams

A set of points in $\bar{\mathbb{R}}^2$ that describe the changing homology of the sublevel sets of a function

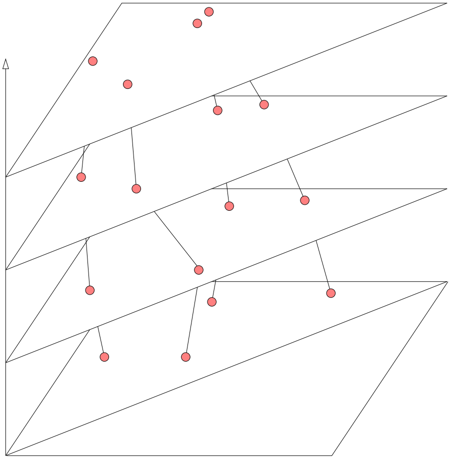


Stacking the Persistence Diagrams



We stack the diagrams so that $\mathrm{Dgm}_p(h_t)$ is drawn at height $z = t$.

Stacking the Persistence Diagrams



Then, we match the diagrams using a linear time algorithm [CSEM].

Vineyards

- The path of an off-diagonal point is called a *vine*. A vine is represented by a function $s\colon I_\tau \to \mathbb{R}^3$.
- The collection of vines is called a *vineyard*.
- Matching of $\mathrm{Dgm}_p(f)$ and $\mathrm{Dgm}_p(g)$ is obtained by looking at the endpoints of the vines.

Another Representation of a Vineyard

[Movie]

Total Movement in a Vineyard

For a vine s, we can compute the weighted distance traveled by a point in the persistence diagrams. Then, we sum this distance over all vines in the vineyard V.

$$D_s = \int_0^1 \omega_{s(t)} \cdot \frac{\partial s(t)}{\partial t}\, dt$$

$$V_q(H) = \left(\sum_{s \in V} D_s^q \right)^{1/q}$$

Total Movement in a Vineyard

For a vine s, we can compute the weighted distance traveled by a point in the persistence diagrams. Then, we sum this distance over all vines in the vineyard V.

$$D_s = \sum_{i \in \{1,2,\ldots,\tau\}} \omega_s(i) \cdot ||s(t_i) - s(t_{i-1})||_\infty$$

$$V_q(H) = \left(\sum_{s \in V} D_s^q \right)^{1/q}$$

Questions

1 Tight Bound: Is the inequality

$$|\ell_1 - \ell_2| \leq \frac{4}{\pi} \cdot (\kappa_1 + \kappa_2) \cdot \mathcal{F}(\gamma_1, \gamma_2)$$

a tight bound for curves in $\mathbb{R}^n$ for $n > 3$?

2 Simultaneous Scale Space: What happens if multiple agents are diffusing at the same time?

3 Understanding V_q: Does there exist a stability result for V_q?

Related Metrics

Let $A = \mathrm{Dgm}(f)$ and $B = \mathrm{Dgm}(g)$.
We find a bijection between A and B by minimizing some quantity, such as:

- Bottleneck Distance
- Wasserstein Distance

Bottleneck Matching

The bottleneck cost of a matching is the maximum L_∞ distance between matched points:

$$W_\infty(P) = \max_{(a,b)\in P} ||a-b||_\infty.$$

We seek to minimize the bottleneck distance over all perfect matchings:

$$W_\infty(A,B) = \min_P \{W_\infty(P)\}.$$

Wasserstein Matching

The Wasserstein cost is measures the cumulative distance as follows:

$$W_q(P) = \left(\sum_{(a,b) \in P} ||a - b||_\infty^q \right)^{1/q}.$$

We seek to minimize the Wasserstein distance over all perfect matchings:

$$W_q(A, B) = \min_P \{ W_q(P) \}.$$

Related Stability Results

We say that the matching of persistence diagrams is *stable* if the cost of the matching is bounded by some reasonable function of $||f - g||_\infty$.

- [CSEH] The Bottleneck Distance is stable for monotone Functions $f, g \colon \mathbb{M} \to \mathbb{R}$.

$$W_\infty(A, B) \leq ||f - g||_\infty$$

- [CSEHM10] The Wasserstein Distance is stable for tame Lipschitz Functions with bounded degree k total persistence.

$$W_q(A, B) \leq C^{1/q} ||f - g||_\infty^{1-k/q}$$

Related Stability Results

We say that the matching of persistence diagrams is *stable* if the cost of the matching is bounded by some reasonable function of $||f - g||_\infty$.

- [CSEH] The Bottleneck Distance is stable for monotone Functions $f, g \colon \mathbb{M} \to \mathbb{R}$.

$$W_\infty(A, B) \leq ||f - g||_\infty$$

- [CSEHM10] The Wasserstein Distance is stable for tame Lipschitz Functions with bounded degree k total persistence.

$$W_q(A, B) \leq C^{1/q} ||f - g||_\infty^{1-k/q}$$

- Is the Vineyard Metric stable too?

$$V_q(f, g) \leq ???$$

Preliminary Findings

Let $A = \mathrm{Dgm}(f)$ and $B = \mathrm{Dgm}(g)$.

$$W_1(A, B) \leq V_1(f, g)$$

$$W_\infty(A, B) \leq V_\infty(f, g)$$

Questions

1 Tight Bound: Is the inequality

$$|\ell_1 - \ell_2| \leq \frac{4}{\pi} \cdot (\kappa_1 + \kappa_2) \cdot \mathcal{F}(\gamma_1, \gamma_2)$$

a tight bound for curves in $\mathbb{R}^n$ for $n > 3$?

2 Simultaneous Scale Space: What happens if multiple agents are diffusing at the same time?

3 Understanding V_q: Does there exist a stability result for V_q?

Questions

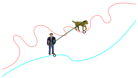
1 Tight Bound: Is the inequality

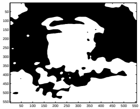
$$|\ell_1 - \ell_2| \leq \frac{4}{\pi} \cdot (\kappa_1 + \kappa_2) \cdot \mathcal{F}(\gamma_1, \gamma_2)$$

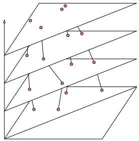
a tight bound for curves in $\mathbb{R}^n$ for $n > 3$?

2 Simultaneous Scale Space: What happens if multiple agents are diffusing at the same time?

3 Understanding V_q: Does there exist a stability result for V_q?







Thank You

- My adviser, Herbert Edelsbrunner
- My committee: Hubert Brey, John Harer, and Carlo Tomasi
- Those who read my prelim document and provided comments, including Michael Kerber and Amit Patel.
- Michelle Phillips (for making the graphics of dog and person)
- Everyone here!

Questions?

References

- G. Donald Chakerian, *An Inequality for Closed Space Curves*, Pacific J. Math. **12** (1962), no. 1, 53–57.
- David Cohen-Steiner and Herbert Edelsbrunner, *Inequalities for the Curvature of Curves and Surfaces*, Found. Comput. Math. **7** (2007), no. 4, 391–404.
- David Cohen-Steiner, Herbert Edelsbrunner, and John Harer, *Stability of persistence diagrams*, 2005, pp. 263–271.
- David Cohen-Steiner, Herbert Edelsbrunner, John Harer, and Yuriy Mileyko, *Lipschitz Functions have L_p-Stable Persistence*, Found. Comput. Math. **10** (2010), no. 2, 127–139.
- David Cohen-Steiner, Herbert Edelsbrunner, and Dmitriy Morozov, *Vines and Vineyards by Updating Persistence in Linear Time*, 2006, pp. 119–126.
- Istvan Fáry, *Sur Certaines Inégalités Géométriques*, Acta Sci. Math. (Szeged) **12** (1950), no. aa, 117–124.