The coherent matching distance in 2D persistent homology

Patrizio Frosini

Department of Mathematics and ARCES, University of Bologna
patrizio.frosini@unibo.it

OSU, 5 April 2017

Outline

Mathematical setting

Extended Pareto Grid

The coherent 2-dimensional matching distance CD_U

The distance CD_U is achieved at $a = 1/2$

Mathematical setting

Extended Pareto Grid

The coherent 2-dimensional matching distance CD_U

The distance CD_U is achieved at $a = 1/2$

Mathematical setting

Let $f=(f_1,f_2), g=(g_1,g_2)$ be two continuous maps from a finitely triangulable topological space M to the real plane $\mathbb{R}^2$.

We consider the persistence diagrams $\mathrm{Dgm}(f^*_{(a,b)})$, $\mathrm{Dgm}(g^*_{(a,b)})$ associated with the admissible line $r_{(a,b)}$, where

$$f^*_{(a,b)} := \max\left\{\frac{\min\{a,1-a\}}{a}\cdot(f_1-b), \frac{\min\{a,1-a\}}{1-a}\cdot(f_2+b)\right\},$$

$$g^*_{(a,b)} := \max\left\{\frac{\min\{a,1-a\}}{a}\cdot(g_1-b), \frac{\min\{a,1-a\}}{1-a}\cdot(g_2+b)\right\}.$$

Let Λ^+ and $\mathscr{P}(\Lambda^+)$ be the set of lines with finite positive slope in $\mathbb{R}^2$ and the set $]0,1[\times\mathbb{R}$ parameterizing these lines, respectively.

Mathematical setting

Let β_f and β_g be the persistent Betti numbers functions of f and g, respectively.

We recall that the 2-dimensional matching distance $D_{match}(\beta_f, \beta_g)$ is then defined as

$$D_{match}(\beta_f, \beta_g) = \sup_{\mathscr{P}(\Lambda^+)} d_B(\mathrm{Dgm}(f^*_{(a,b)}), \mathrm{Dgm}(g^*_{(a,b)})),$$

with $d_B(\mathrm{Dgm}(f^*_{(a,b)}), \mathrm{Dgm}(g^*_{(a,b)}))$ denoting the bottleneck distance between the normalized persistence diagrams $\mathrm{Dgm}(f^*_{(a,b)})$ and $\mathrm{Dgm}(g^*_{(a,b)})$.

Mathematical setting

The following result will be of use.

Lemma

If $(a,b) \in \mathscr{P}(\Lambda^+)$ *then* $\left\| f^*_{(a,b)} - g^*_{(a,b)} \right\|_\infty \leq \|f - g\|_\infty$.

Remark

The normalization of the functions $f_{(a,b)}, g_{(a,b)}$ is crucial here. Indeed, the bottleneck distance $d_B(\mathrm{Dgm}(f^*_{(a,b)}), \mathrm{Dgm}(g^*_{(a,b)}))$ is stable against functions' perturbations when measured by the sup-norm, while this is not true for the distance $d_B(\mathrm{Dgm}(f_{(a,b)}), \mathrm{Dgm}(g_{(a,b)}))$.

The phenomenon of persistent monodromy

We recall what we said in lecture 1 about monodromy in 2D persistent homology.

We have seen that if we turn around a singular point in the parameter space $]0,1[\times\mathbb{R}$, some points in the persistence diagram $\mathrm{Dgm}(f^*_{(a,b)})$ may exchange their position. In other words, a loop around the singular point induces a permutation on the persistence diagram.

Therefore, a monodromy group is associated with the function f. In order to properly define this group, we have to give a precise definition of the path followed by a point $p \in \mathrm{Dgm}(f^*_{(a,b)})$ when (a, b) moves.

Mathematical setting

Extended Pareto Grid

The coherent 2-dimensional matching distance CD_U

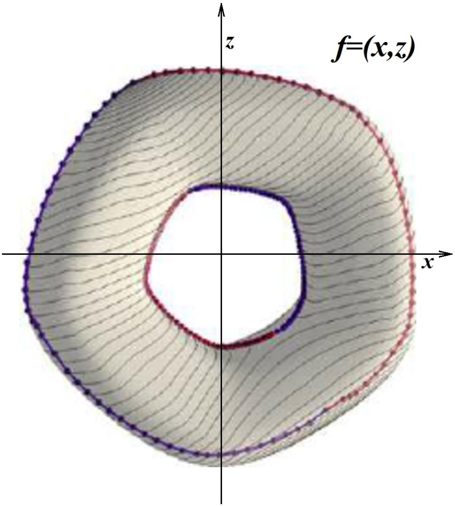
The distance CD_U is achieved at $a = 1/2$

Some technical assumptions

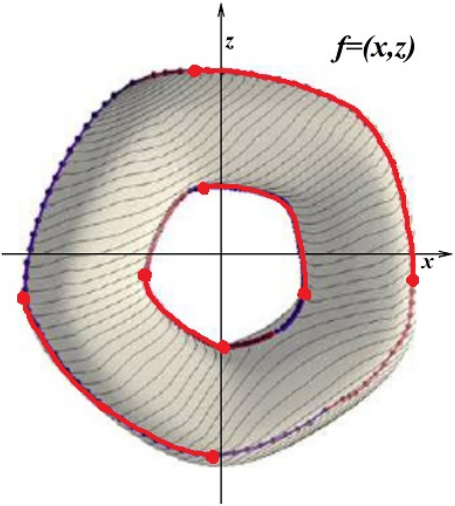
Let $f = (f_1, f_2)$ be a smooth map from a closed C^∞-manifold M of dimension $r \geq 2$ to the real plane $\mathbb{R}^2$. Choose a Riemannian metric on M so that we can define gradients for f_1 and f_2.

The Jacobi set $\mathbb{J}(f)$ is the set of all points $p \in M$ at which the gradients of f_1 and f_2 are linearly dependent, namely $\nabla f_1(p) = \lambda \nabla f_2(p)$ or $\nabla f_2(p) = \lambda \nabla f_1(p)$ for some $\lambda \in \mathbb{R}$. In particular, if $\lambda \leq 0$ the point $p \in M$ is said to be a critical Pareto point for f. The set of all critical Pareto points of f is denoted by $\mathbb{J}_P(f)$.

The Jacobi set



Critical Pareto points



Some technical assumptions

In order to proceed we will assume that our filtering functions $f : M \to \mathbb{R}^2$ are pretty regular, in the sense described in this slide.

We assume that

(i) No point $p \in M$ exists such that both $\nabla f_1(p)$ and $\nabla f_2(p)$ vanish;

(ii) $\mathbb{J}(f)$ is a smoothly embedded 1-manifold in M, consisting of finitely many circles;

(iii) $\mathbb{J}_P(f)$ is a 1-dimensional closed manifold with boundary in $\mathbb{J}(f)$.

We also consider the set $\mathbb{J}_C(f)$ of cusp points of f, that is, points of $\mathbb{J}(f)$ at which the restriction of f to $\mathbb{J}(f)$ fails to be an immersion. In other words $\mathbb{J}_C(f)$ is the subset of $\mathbb{J}(f)$ at which both ∇f_1 and ∇f_2 are orthogonal to $\mathbb{J}(f)$ (hence $\mathbb{J}_C(f) \subseteq \mathbb{J}_P(f)$).

Some technical assumptions

We also assume that (iv) the connected components of $\mathbb{J}_P(f) \setminus \mathbb{J}_C(f)$ are finite in number, each one being not a circle. The previous properties $(i), (ii), (iii), (iv)$ are generic in the set of smooth maps from M to $\mathbb{R}^2$.

Property (iv) implies that the connected components of $\mathbb{J}_P(f) \setminus \mathbb{J}_C(f)$ are open, or closed, or semi-open arcs in M. Following the notation used in previous literature, they will be referred to as critical intervals of f. If an endpoint p of a critical interval actually belongs to that critical interval, that is, p is not a cusp point, then it is a critical point for either f_1 or f_2. Along each critical interval, f_1 increases when f_2 decreases, and vice versa.

The extended Pareto grid

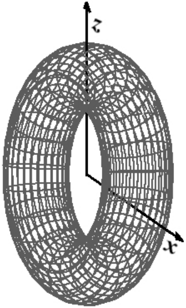
Our purpose is to establish a formal link between the position of points of $\mathrm{Dgm}(f^*_{(a,b)})$ for a function f and the intersections between the admissible line $r_{(a,b)}$ with a particular subset of the plane $\mathbb{R}^2$, called the extended Pareto grid of f, which we will define in the next slides.

The extended Pareto grid

Let us list the critical points $p_1, \ldots, p_h$ of f_1 and the critical points $q_1, \ldots, q_k$ of f_2 (our assumption (i) guarantees that $\{p_1, \ldots, p_h\} \cap \{q_1, \ldots, q_k\} = \emptyset$). Consider the following half-lines: for each critical point p_i of f_1 (resp. each critical point q_j of f_2), the half-line $\{(x, y) \in \mathbb{R}^2 | x = f_1(p_i), y \geq f_2(p_i)\}$ (resp. the half-line $\{(x, y) \in \mathbb{R}^2 | x \geq f_1(q_j), y = f_2(q_j)\}$).

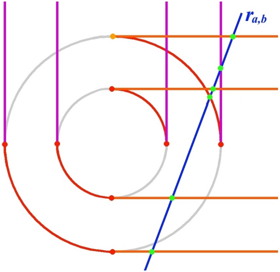
The extended Pareto grid $\boldsymbol{\Gamma(f)}$ will be the union of $f(\mathbb{J}_P(f))$ with these half-lines. The closures of the images of critical intervals of f will be called proper contours of f, while the half-lines will be known as improper contours of f. We observe that every contour is a closed set. For each point $p \in \mathbb{R}^2$, we say that the number of (proper or improper) contours containing p is the multiplicity of p with respect to the function f. (This definition should not be confused with the definition of multiplicity for points in persistence diagrams.)

The extended Pareto grid



The torus endowed with the filtering function $f(p) := (x(p), z(p))$.

The extended Pareto grid



The extended Pareto grid for the torus endowed with the filtering function $f(p) := (x(p), z(p))$. The images of the critical intervals are in red, the vertical half-lines with abscissa equal to a critical value of f_1 are in purple, and the horizontal half-lines with ordinate equal to a critical value of f_2 are in orange. A blue admissible line $r_{(a,b)}$ is also represented.

Assumptions about the extended Pareto grid

We recall that, by definition, a pair $(a, b) \in]0, 1[\times \mathbb{R}$ is singular for f if and only if the persistence diagram $\mathrm{Dgm}(f^*_{(a,b)})$ contains at least one point not belonging to Δ with multiplicity strictly greater than 1. A pair (a, b) that is not singular is called regular.

Definition

We say that the function $f : M \to \mathbb{R}^2$ is normal if

1. The number of proper and improper contours in $\Gamma(f)$ is finite;
2. The number of multiple points of $\Gamma(f)$ is finite;
3. Each multiple point of $\Gamma(f)$ is double;
4. No line $r_{(a,b)}$ contains more than two multiple points of $\Gamma(f)$;

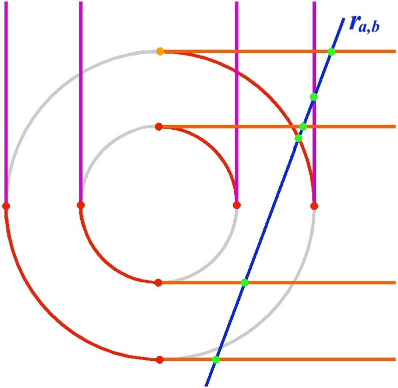
Assumptions about the extended Pareto grid

Definition

5. Let i_*^k be the map $H_k(M_{u-\varepsilon,v-\varepsilon}) \to H_k(M_{u+\varepsilon,v+\varepsilon})$ induced by the inclusion $M_{u-\varepsilon,v-\varepsilon} \hookrightarrow M_{u+\varepsilon,v+\varepsilon}$. Every contour γ of $\Gamma(f)$ is associated with a pair $(d(\gamma), s(\gamma)) \in \mathbb{Z} \times \{-1,1\}$ such that at each internal point (u,v) of γ the following properties hold for every small enough $\varepsilon > 0$:
   - If $k \neq d(\gamma)$, i_*^k is an isomorphism;
   - If $k = d(\gamma)$ and $s(\gamma) = 1$, i_*^k is injective and $\mathrm{rank}(H_k(M_{u+\varepsilon,v+\varepsilon}) = \mathrm{rank}(H_k(M_{u-\varepsilon,v-\varepsilon}) + 1$;
   - If $k = d(\gamma)$ and $s(\gamma) = -1$, i_*^k is surjective and $\mathrm{rank}(H_k(M_{u+\varepsilon,v+\varepsilon}) = \mathrm{rank}(H_k(M_{u-\varepsilon,v-\varepsilon}) - 1$.

Assumptions about the extended Pareto grid

In plain words, property 5 guarantees that the passage across a contour γ just creates (if $s(\gamma) = 1$) or destroy (if $s(\gamma) = -1$) one homological class in degree $d(\gamma)$.



The Position Theorem

With the concept of extended Pareto grid at hand, we can state and prove the following result, which gives a necessary condition for P to be a point of $\mathrm{Dgm}(f^*_{(a,b)})$.

We recall that

$$f^*_{(a,b)} := \max\left\{ \frac{\min\{a, 1-a\}}{a} \cdot (f_1 - b), \frac{\min\{a, 1-a\}}{1-a} \cdot (f_2 + b) \right\}.$$

Theorem (Position Theorem)

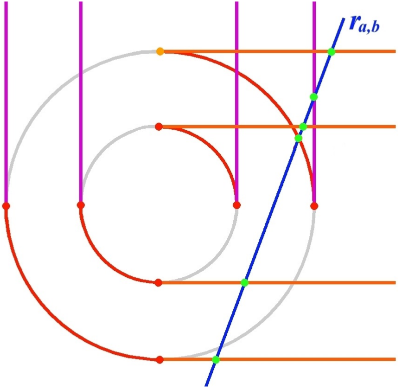
Let $(a,b) \in \mathscr{P}(\Lambda^+)$*,* $P \in Dgm(f^*_{(a,b)}) \setminus \Delta$*. Then, for each finite coordinate* c *of* P *a point* $(x,y) \in r_{(a,b)} \cap \Gamma(f)$ *exists, such that* $c = \frac{\min\{a,1-a\}}{a} \cdot (x-b) = \frac{\min\{a,1-a\}}{1-a} \cdot (y+b)$*.*

Using the extended Pareto grid

The Position Theorem suggests a way to find the possible positions for points of $\mathrm{Dgm}(f^*_{(a,b)})$. It consists in drawing the extended Pareto grid $\Gamma(f)$ and considering its intersections $(x_1, y_1), \ldots, (x_l, y_l)$ with the admissible line $r_{(a,b)}$. For each proper point of $\mathrm{Dgm}(f^*_{(a,b)})$, both its coordinates belong to the set

$$\left\{ \frac{\min\{a, 1-a\}}{a} \cdot (x_i - b) = \frac{\min\{a, 1-a\}}{1-a} \cdot (y_i + b) \right\}_{1 \le i \le l} \cup \{\infty\}.$$

Using the extended Pareto grid

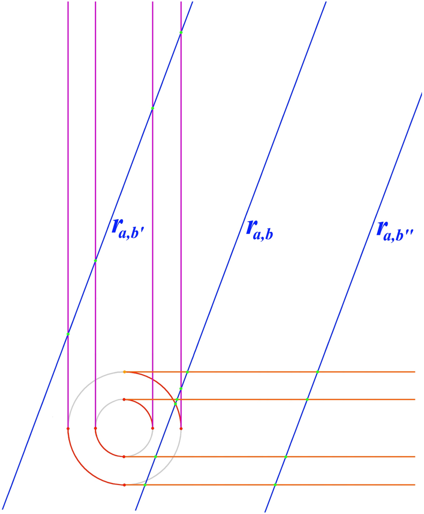


Each coordinate of a point in $\mathrm{Dgm}(f^*_{(a,b)})$ equals $\frac{\min\{a,1-a\}}{a} \cdot (x - b)$, where (x, y) is a green point.

Using the extended Pareto grid

Note that when $b < 0$ and $|b|$ is sufficiently large, the admissible line $r_{(a,b)}$ may intersect $\Gamma(f)$ only at the vertical half-lines. In this case, $f^*_{(a,b)} := \frac{\min\{a,1-a\}}{a} \cdot (f_1 - b)$, and the values $x_1, \ldots, x_l$ are the critical values of f_1. Similarly, when $b > 0$ and $|b|$ is large enough, $r_{(a,b)}$ intersects $\Gamma(f)$ only at the horizontal half-lines. Then $f^*_{(a,b)} := \frac{\min\{a,1-a\}}{1-a} \cdot (f_2 + b)$, and the values $y_1, \ldots, y_l$ are the critical values of f_2. (See next slide)

Using the extended Pareto grid



Using the extended Pareto grid

The Position Theorem allows us to deduce where singular pairs can be in $\mathscr{P}(\Lambda^+)$.

Proposition

If $(a,b) \in \mathscr{P}(\Lambda^+)$ is a singular pair for f, then $r_{(a,b)}$ contains two double points of $\Gamma(f)$.

Corollary

The set of singular pairs in $\mathscr{P}(\Lambda^+)$ for f is finite.

Singular pairs

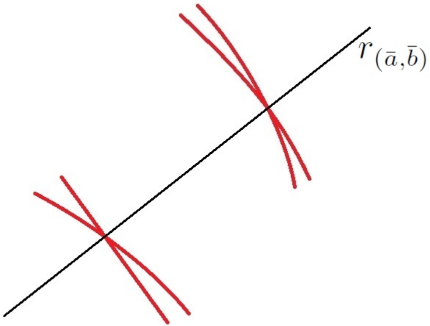


Figure: A line $r_{(\bar{a},\bar{b})}$ associated with a singular pair $(\bar{a},\bar{b}) \in \mathscr{P}(\Lambda^+)$. Parts of four proper contours are displayed in red.

Creation and distruction of points in $\mathrm{Dgm}(f^*_{(a,b)})$ when (a,b) varies in $\mathscr{P}(\Lambda^+)$

The Position Theorem allows us to deduce at which points of Δ points of $\mathrm{Dgm}(f^*_{(a,b)})$ can be created or destroyed.

Proposition

Let $(a(t),b(t))$ be a continuous curve in $\mathscr{P}(\Lambda^+)$ such that the distance between $Dgm(f^*_{(a,b)})\setminus\Delta$ and $(c,c)\in\Delta$ tends to 0 for $t\to\bar{t}$. Then two contours γ_1,γ_2 of f exist, such that γ_1,γ_2 have a common extremum $E=(\bar{x},\bar{y})$ and $c=\frac{\min\{a,1-a\}}{a}\cdot(\bar{x}-b)=\frac{\min\{a,1-a\}}{1-a}\cdot(\bar{y}+b)$.

In plain words, the previous result shows that points of $\mathrm{Dgm}(f^*_{(a,b)})$ can be created or destroyed just when the line $r_{(a,b)}$ goes across a common extremum of two contours.

Distruction of a point in $\mathrm{Dgm}(f^*_{(a,b)})$

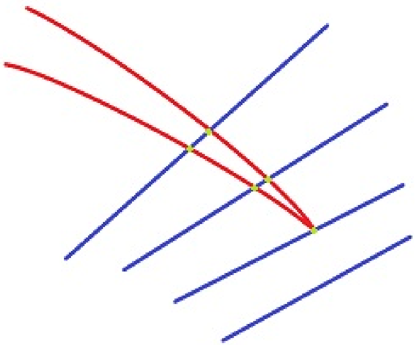


Figure: A point of $\mathrm{Dgm}(f^*_{(a,b)})$ reaches the diagonal Δ and disappears.

Ghost points

Definition

If two contours γ_1, γ_2 like the ones cited in the previous proposition are given, and the line $r_{(a,b)}$ does not meet γ_1, γ_2 at two of their internal points, then we set $D_{(a,b)}(E) := (c, c) \in \Delta$ with $c = \frac{\min\{a, 1-a\}}{a} \cdot (\bar{x} - b)$ and call $D_{(a,b)}(E)$ a ghost point of E at (a, b). The set of all ghost points at (a, b) varying the contours γ_1, γ_2 is denoted by the symbol $\Delta_{(a,b)}(f)$.

The concept of ghost point allows us to follow points in the persistence diagrams $\mathrm{Dgm}(f^*_{(a,b)})$ while (a, b) varies in the parameter space $\mathscr{P}(\Lambda^+)$, even after these points have reached the diagonal Δ.

Mathematical setting

Extended Pareto Grid

The coherent 2-dimensional matching distance CD_U

The distance CD_U is achieved at $a = 1/2$

The coherent 2-dimensional matching distance

Our next step is the definition of the coherent 2-dimensional matching distance.
The existence of monodromy implies that each loop in the set $\mathrm{Reg}(f)$ of all regular pairs for f induces a permutation on $\mathrm{Dgm}(f^*_{(a,b)})$. In other words, it is not possible to establish which point in $\mathrm{Dgm}(f^*_{(a,b)})$ corresponds to which point in $\mathrm{Dgm}(f^*_{(a',b')})$ for $(a,b) \neq (a',b')$, since the answer depends on the path that is considered from (a,b) to (a',b') in the set $\mathrm{Reg}(f)$. As a consequence, different paths going from (a,b) to (a',b') might produce different results while "transporting" a matching $\sigma_{(a,b)} : \mathrm{Dgm}(f^*_{(a,b)}) \to \mathrm{Dgm}(g^*_{(a,b)})$ to another point $(a',b') \in \mathscr{P}(\Lambda^+)$.
Despite this problem, it is possible to define a notion of coherent 2-dimensional matching distance.

Transporting a matching along a path

First, we need to specify the concept of transporting a point $X \in \mathrm{Dgm}(f^*_{(a(0),b(0))})$ along a path $(a(t),b(t))$ in $\mathrm{Reg}(f)$.

Definition (Induced path)

A continuous path $P:[0,1]\to\mathbb{R}^2$ is said to be induced by the path $\pi:[0,1]\to\mathrm{Reg}(f)$ if for every $t\in[0,1]$ it holds that $P(t)\in(\mathrm{Dgm}(f^*_{\pi(t)})\setminus\Delta)\cup\Delta_{\pi(t)}(f)$.

Proposition

Let $\pi=(a,b):[0,1]\to Reg(f)$ be a continuous path. For every point $X\in(Dgm(f^*_{\pi(0)})\setminus\Delta)\cup\Delta_{\pi(0)}(f)$, a unique path $P:[0,1]\to\mathbb{R}^2$ induced by π exists, such that $P(0)=X$.

Transport of points and matchings

With reference to the previous Proposition, we say that π transports X to $X' = P(1)$ with respect to f and write $T_{\pi}^{f}(X) = X'$. Now, we need to define the concept of transporting a matching along a path $\pi : [0,1] \to \mathrm{Reg}(f) \cap \mathrm{Reg}(g)$ with $\pi(0) = (a,b)$. Let $\sigma_{(a,b)}$ be a matching between $\mathrm{Dgm}(f^*_{(a,b)})$ and $\mathrm{Dgm}(g^*_{(a,b)})$, with (a,b) an element of $\mathrm{Reg}(f) \cap \mathrm{Reg}(g)$. We can naturally associate to $\sigma_{(a,b)}$ a matching $\sigma_{\pi(1)} : \mathrm{Dgm}(f^*_{\pi(1)}) \to \mathrm{Dgm}(g^*_{\pi(1)})$. Suppose that $\sigma_{(a,b)}(X) = Y$. We set $\sigma_{\pi(1)}(X') = Y'$ if and only if π transports X to X' with respect to f and Y to Y' with respect to g. We also say that π transports $\sigma_{(a,b)}$ to $\sigma_{\pi(1)}$ along π with respect to the pair (f,g). The transported matching will be denoted by the symbol $T_{\pi}^{(f,g)}(\sigma_{(a,b)})$.

Transport of points

The next property trivially follows from the definition of transport.

Proposition

Let π_1, π_2 be two continuous paths in $Reg(f)$, with $\pi_1(1) = \pi_2(0)$. Let $\pi_1 * \pi_2$ be their composition, i.e. the loop $\pi_1 * \pi_2 : [0,1] \to Reg(f)$ defined by setting $\pi_1 * \pi_2(t) := \pi_1(2t)$ for $0 \leq t \leq 1/2$ and $\pi_1 * \pi_2(t) := \pi_2(2t-1)$ for $1/2 \leq t \leq 1$. Then $T^f_{\pi_2} \circ T^f_{\pi_1} = T^f_{\pi_1 * \pi_2}$.

Continuity of the transport w.r.t. the path

By using the 1-dimensional Stability Theorem and the Position Theorem, we can prove that the transport along a path in $\mathrm{Reg}(f)$ is continuous with respect to the path, as stated by the following proposition.

Proposition

Let $X \in Dgm(f^*_{\pi(0)})$. The function $T^f_\pi(X)$ is continuous in the variable π, when π varies in the set $S^f_{(\bar{a},\bar{b})}$ of the paths in $Reg(f)$ starting from a fixed point $(\bar{a},\bar{b})$ and $S^f_{(\bar{a},\bar{b})}$ is endowed with the uniform convergence metric.

Each loop in $\mathrm{Reg}(f)$ induces a permutation on $\mathrm{Dgm}(f^*_{(\bar{a},\bar{b})})$

From the previous proposition the next result immediately follows.

Proposition

If two paths π_1, π_2 in $Reg(f)$ are homotopic to each other relatively to their common extrema, then $T^f_{\pi_1} \equiv T^f_{\pi_2}$.

Corollary

The map T^f taking each equivalence class $[\pi]$ to the permutation T^f_π is a well-defined homomorphism from the fundamental group of $Reg(f)$ at $(\bar{a}, \bar{b})$ to the group of permutations of $Dgm(f^*_{(\bar{a},\bar{b})})$.

Turning twice around a singular point produces the identical permutation on persistence diagrams

The following interesting property holds.

Proposition

Let $\pi : [0,1] \to Reg(f)$ *be a loop turning once around exactly one singular pair. Then* T^f_π *is either a transposition or the identity.*

Turning twice around a singular point produces the identical permutation on persistence diagrams

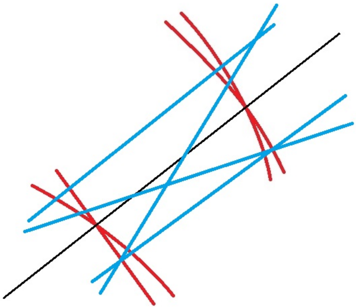


Figure: A loop around a singular pair in $\mathscr{P}(\Lambda^+)$. Parts of four proper contours are displayed in red. The lines $r_{(a,b)}$ are in blue.

Computing permutations on persistence diagrams

Proposition

For each singular pair (a_i, b_i) let us choose a loop π_i starting at a regular pair $(\bar{a}, \bar{b})$ and turning once only around (a_i, b_i). The image of T^f is generated by the permutations $T^f_{\pi_i}$.

Remark

The previous proposition implicitly gives a simple method to compute the image of T^f. We know that if G is a subgroup of the symmetric group S_n and G is generated by m transpositions, then $|G| \leq (m+1)!$. It follows that the cardinality of the image of T^f is bounded by the factorial of the number of singular pairs in $\mathscr{P}(\Lambda^+)$ plus one.

Coherent bouquets of matchings

The idea of coherent matchings consists in requiring that the matchings at close points in $\mathscr{P}(\Lambda^+)$ are close to each other. We have already seen that monodromy prevents us from transporting single matchings in a coherent way. Fortunately, this can be done for bouquets of matchings. We start by fixing a connected open subset U of $\mathscr{P}(\Lambda^+)$ and defining $\boldsymbol{\Phi_{U,c}}$ as the set of all continuous functions $f : M \to \mathbb{R}^2$ such that $\mathrm{Reg}(f) \supseteq U$ and the minimal distance between two points of $\mathrm{Dgm}(f^*_{(a,b)}) \setminus \Delta$ is strictly greater than $2c > 0$ for every $(a, b) \in U$. Let us assume that $f, g \in \Phi_{U,c}$.

Remark

The definition of the set $\Delta_{(a,b)}$ implies that if $f \in \Phi_{U,c}$ then the minimal distance between two points of $(\mathrm{Dgm}(f^*_{(a,b)}) \setminus \Delta) \cup \Delta_{(a,b)}$ is strictly positive for every $(a, b) \in U$.

Bouquets of matchings

Definition

Let $\sigma_{(\bar{a},\bar{b})}$ be a matching between $\mathrm{Dgm}(f^*_{(\bar{a},\bar{b})})$ and $\mathrm{Dgm}(g^*_{(\bar{a},\bar{b})})$, with $(\bar{a},\bar{b}) \in U$. The bouquet $Bq_U(\sigma_{(\bar{a},\bar{b})})$ of $\sigma_{(\bar{a},\bar{b})}$ is the set of all matchings we obtain by transporting $\sigma_{(\bar{a},\bar{b})}$ along any loops in U based at $(\bar{a},\bar{b})$. In symbols, if we denote the set of all continuous paths $\pi : [0,1] \to U$ with $\pi(0) = \pi(1) = (\bar{a},\bar{b})$ by $L^U_{(\bar{a},\bar{b})}$, we define

$$Bq_U(\sigma_{(\bar{a},\bar{b})}) := \left\{ T^{(f,g)}_{\pi}(\sigma_{(\bar{a},\bar{b})}) \middle| \pi \in L^U_{(\bar{a},\bar{b})} \right\}.$$

The independence property

If $Bq_U(\sigma_{(\bar{a},\bar{b})})$ is a bouquet of matchings at $(\bar{a},\bar{b}) \in U$, then for every $(a,b) \in U$ we can take a continuous path γ from $(\bar{a},\bar{b})$ to (a,b) in U and define the set

$$T^{(f,g)}_{(\bar{a},\bar{b})\mapsto(a,b)}\left(Bq_U(\sigma_{(\bar{a},\bar{b})})\right) := \left\{ T^{(f,g)}_{\gamma}(\sigma) \mid \sigma \in Bq_U(\sigma_{(\bar{a},\bar{b})}) \right\}.$$

It is easy to prove the following property.

Independence property: The set $T^{(f,g)}_{(\bar{a},\bar{b})\mapsto(a,b)}\left(Bq_U(\sigma_{(\bar{a},\bar{b})})\right)$ is a bouquet of matchings at (a,b). It does not depend on the chosen path γ in U from $(\bar{a},\bar{b})$ to (a,b), but only on its endpoints.

Coherent families of bouquets of matchings

Definition

A family of bouquets of matchings $\left\{Bq_U(\sigma_{(a,b)})\right\}_{(a,b)\in U}$ that are obtained by transporting a bouquet of matchings $Bq_U(\sigma_{(\bar{a},\bar{b})})$ all over the set U is said to be coherent in U for the pair (f,g). The set of all family of bouquets of matchings that are coherent in U for the pair (f,g) will be denoted by the symbol $Coh_U(f,g)$.

Composition of coherent families of bouquets of matchings

It is easy to show that the following definition is well-posed.

Definition

Let $f, g, h \in \Phi_{U,c}$. If $Bq_U(\sigma_{(a,b)})$ and $Bq_U(\tau_{(a,b)})$ are two bouquets of matchings in U for (f,g) and (g,h) (respectively), we can define their composition $Bq_U(\tau_{(a,b)}) \circ Bq_U(\sigma_{(a,b)})$ as the bouquet at (a,b) for (f,h) given by the set $\{\tau \circ \sigma : \sigma \in Bq_U(\sigma_{(a,b)}), \tau \in Bq_U(\tau_{(a,b)})\}$. If two coherent families of bouquets in U of matchings $E = \left\{Bq_U(\sigma_{(a,b)})\right\}_{(a,b)}$ and $F = \left\{Bq_U(\tau_{(a,b)})\right\}_{(a,b)}$ for (f,g) and (g,h) (respectively) are given, we can define the coherent family $F \circ E$ by taking at each point $(a,b) \in U$ the composition of the bouquets for E and F at that point.

Stability of the transport of points

The next result will be of use.

Lemma

Let $f, g \in \Phi_{U,c}$ *with* $\|f-g\|_\infty < c$. *If* π *is a continuous path in* U, *and* $X \in Dgm(f^*_{\pi(0)})$, $Y \in Dgm(g^*_{\pi(0)})$ *are two points whose distance is less than* $\|f-g\|_\infty$, *then* $\|T^g_\pi(Y) - T^f_\pi(X)\| \leq \|f-g\|_\infty$.

The definition of the coherent matching distance

Definition

The cost of a bouquet $Bq_U(\sigma_{(a,b)})$ of matchings at $(a,b) \in U$ is the value $\mathrm{cost}\left(Bq_U(\sigma_{(a,b)})\right) := \max_{\sigma \in Bq_U(\sigma_{(a,b)})} \mathrm{cost}(\sigma)$.

Definition

Let E be a coherent family $\left\{Bq_U(\sigma_{(a,b)})\right\}_{(a,b)\in U}$ of bouquets in U of matchings for (f,g). We set $\mathrm{cost}(E) := \sup_{(a,b)\in U} \mathrm{cost}\left(Bq_U(\sigma_{(a,b)})\right)$.

We observe that the set $Coh_U(f,g)$ can be constructed by taking each possible bouquet in U of matchings at an arbitrarily fixed point $(\bar{a},\bar{b}) \in U$ and extending these bouquets to coherent families of bouquets of matchings.

The coherent 2-dimensional matching distance

Definition

The coherent 2-dimensional matching distance between β_f and β_g is defined as

$$CD_U(\beta_f, \beta_g) = \inf_{E \in Coh_U(f,g)} \mathrm{cost}(E).$$

Proposition

$CD_U(\beta_f, \beta_g)$ is a pseudo-distance.

Stability of the coherent 2-dimensional matching distance

The next result shows that the coherent 2-dimensional matching distance is stable, in a suitable sense.

Theorem

If $f, g \in \Phi_{U,c}$ *and* $\|f - g\|_\infty < c$ *then* $CD_U(\beta_f, \beta_g) \leq \|f - g\|_\infty$.

Mathematical setting

Extended Pareto Grid

The coherent 2-dimensional matching distance CD_U

The distance CD_U is achieved at $a = 1/2$

The distance CD_U is achieved at $a = 1/2$

Definition

Let $(\bar{a}, \bar{b})$ be a point of the line $l_{\frac{1}{2}}$ of equation $a = 1/2$ in $\mathscr{P}(\Lambda^+)$, and assume that $(\bar{a}, \bar{b}) \in U$ is a regular pair. Let $\sigma : Dgm(f^*_{(\bar{a},\bar{b})}) \to Dgm(g^*_{(\bar{a},\bar{b})})$ be a matching. The function γ_σ that associates each $(a, b) \in U \cap l_{\frac{1}{2}}$ to the bouquet of matchings obtained by transporting σ to (a, b) in U will be called the coherent family of bouquets of matchings of σ on $l_{\frac{1}{2}}$ (note that transporting σ may require to move out of $l_{\frac{1}{2}}$). We define *cost* γ_σ as the maximum cost of the bouquets of matchings in the image of γ_σ.

Definition

Let $\mathscr{C}_{\frac{1}{2}}$ be the set of coherent families of bouquets of matchings on $l_{\frac{1}{2}}$.

The distance CD_U is achieved at $a = 1/2$

From the independence property the next result easily follows.

Proposition

The set $\mathscr{C}_{\frac{1}{2}}$ does not depend on the basepoint $(\bar{a}, \bar{b}) \in I_{\frac{1}{2}}$.

Definition

We set $CD_{\frac{1}{2}}(\beta_f, \beta_g) := \inf_{\gamma \in \mathscr{C}_{\frac{1}{2}}} cost\ \gamma$.

Theorem

$CD_U \equiv CD_{\frac{1}{2}}$.

Conclusions

In this lecture we have presented a new approach to metric comparison in 2D persistent homology, introducing the concept of coherent matching distance and studying some of its properties. In order to do that, we have also introduced the concept of extended Pareto grid and shown its use to manage the phenomenon of monodromy. Finally, we have proved a theorem that makes clear the importance of filtrations associated with lines of slope 1 in 2D persistent homology.

Further research

In our opinion, many problems should deserve further research. First of all, it would be interesting to extend the presented concepts to filtering functions taking values in $\mathbb{R}^m$ with $m > 2$. Secondly, the genericity of our assumptions concerning the extended Pareto grid should be possibly proved. Finally, methods for the efficient computation of the coherent matching distance should be developed.

Thanks for your attention!