A theoretical framework for topological data analysis

Patrizio Frosini

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

OSU - TGDA Seminar, 4 April 2017

Outline

Our basic questions

Assumptions in our model

Mathematical setting and theoretical results

A first step towards the application of our model: GIPHOD

Some work in progress

Our basic questions

Assumptions in our model

Mathematical setting and theoretical results

A first step towards the application of our model: GIPHOD

Some work in progress

Our basic questions

We are interested in these questions:

- Is there a general metric model to compare data in TDA?
- What should be the role of the observer in such a model?
- How could we approximate the metric used in that model?

Our talk will be devoted to illustrate these questions and to propose some answers by means of a mathematical approach based on persistent homology and group invariant non-expansive operators.

Our basic questions

Assumptions in our model

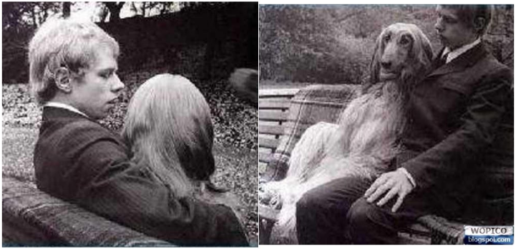
Mathematical setting and theoretical results

A first step towards the application of our model: GIPHOD

Some work in progress

Assumptions in our model

Truth often depends on the observer's perspective:



Multiple perspectives are usually unavoidable! In the past this observation was mostly confined to the philosophical debate, but nowadays it starts to be quite relevant also in several scientific applications involving Information Technology.

Assumptions in our model

We will make these assumptions:

1. No object can be studied in a direct and absolute way. Any object is only knowable through acts of measurement made by an observer.
2. Any act of measurement can be represented as a function defined on a topological space.
3. The observer usually acquires measurement data by applying operators to the functions describing these data. These operators are frequently endowed with some invariances that are relevant for the observer.
4. Only the observer is entitled to decide about data similarity.

Assumptions in our model

MEASUREMENTS

OUTLINE $\varphi : S^1 \to \mathbb{R}^2$ → OPERATORS → $T(\varphi)$

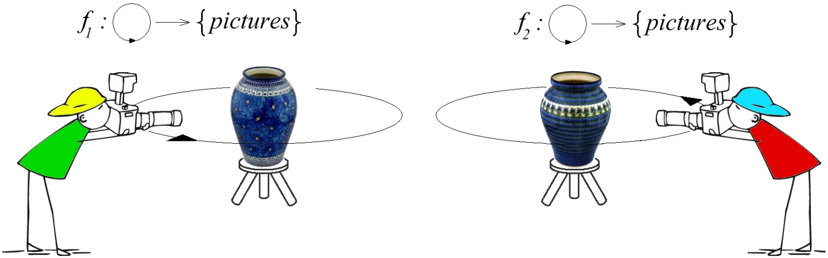
CURVATURE $\varphi : D^2 \to \mathbb{R}$ → OPERATORS → $T(\varphi)$

COLOR IMAGE $\varphi : [0,a] \times [0,b] \to \mathbb{R}^3$ → OPERATORS → $T(\varphi)$

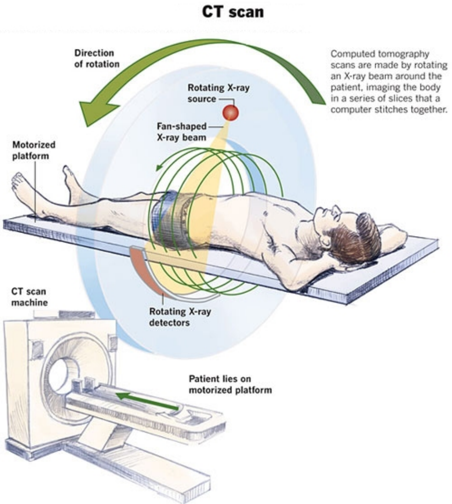
WEIGHT $\varphi : \{P\} \to \mathbb{R}$ → OPERATORS → $T(\varphi)$

$\vdots$ → OPERATORS → $\vdots$

An example of measurement



Another example of measurement



Another example of measurement

An example of operator

GRAYSCALE IMAGE

$\varphi : [0, a] \times [0, b] \to \mathbb{R}$ $\quad T(\varphi)$ = CONVOLUTION OF φ

Choice of the operators

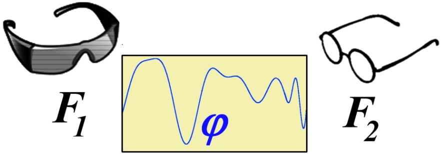
- The observer cannot usually choose the functions representing the measurement data, but he/she can often choose the operators that will be applied to those functions.
- The choice of the operators reflects the invariances that are relevant for the observer.
- In some sense we could state that the observer can be represented as a collection of (suitable) operators, endowed with the invariance he/she has chosen.

In the first part of this talk we will mainly examine the case of operators that act on a space Φ of continuous functions and take Φ to itself. We will also assume that these operators preserve the self-homeomorphisms of X.

From comparing sets in $\mathbb{R}^n$ to comparing functions

Instead of directly focusing on the objects we are interested in, we focus on the filtering functions describing the measurements we make on them, and on the "glasses" that we use "to observe" the functions. In our approach, these "glasses" are G-operators which act on the filtering functions.
These operators represent the observer's perspective.

In some sense, the family of operators defines the observer.



Our basic questions

Assumptions in our model

Mathematical setting and theoretical results

A first step towards the application of our model: GIPHOD

Some work in progress

Natural pseudo-distance associated with a group G

First of all we need a definition allowing us to formalize the comparison of data in our model.

Definition

Let X be a compact space. Let G be a subgroup of the group $\mathrm{Homeo}(X)$ of all homeomorphisms $f : X \to X$. The pseudo-distance $d_G : C^0(X,\mathbb{R}) \times C^0(X,\mathbb{R}) \to \mathbb{R}$ defined by setting

$$d_G(\varphi,\psi) = \inf_{g\in G} \max_{x\in X} |\varphi(x) - \psi(g(x))|$$

is called the natural pseudo-distance associated with the group G.

In plain words, the definition of d_G is based on the attempt of finding the best correspondence between the functions φ, ψ by means of homeomorphisms in G.

A possible objection

A possible objection: "The use of the group of homeomorphisms makes the natural pseudo-distance d_G difficult to apply. For example, in shape comparison two objects are usually not homeomorphic, hence this pseudo-metric cannot be applied to real problems."

This objection can be faced by recalling that **the homeomorphisms do not concern the "objects" but the space where the measurements are made**. For example, if we take a grey level image, our measurement space can be modelled as the real plane and each image can be represented as a function from $\mathbb{R}^2$ to $\mathbb{R}$. Therefore, the space X is not given by the (possibly non-homeomorphic) objects displayed in the picture, but by the topological space $\mathbb{R}^2$. Analogously, each subset of the 3D space can be associated with a probability density describing the probability that each point $p \in \mathbb{R}^3$ belongs to the considered object. In this case the space X is $\mathbb{R}^3$.

G-invariant non-expansive operators

The natural pseudo-distance d_G represents our ground truth.

Unfortunately, d_G is difficult to compute. This is also a consequence of the fact that we can easily find subgroups G of $\mathrm{Homeo}(X)$ that cannot be approximated with arbitrary precision by smaller finite subgroups of G (i.e. $G =$ group of rigid motions of $X = \mathbb{R}^3$).

Nevertheless, in this talk we will show that d_G can be approximated with arbitrary precision by means of a **DUAL** approach based on persistent homology and G-invariant non-expansive operators.

This research is based on an ongoing joint research project with Grzegorz Jabłoński (Jagiellonian University - Poland and IST - Austria)

G-invariant non-expansive operators

Let us consider the following objects:

- A triangulable space X with nontrivial homology in degree k.
- A set Φ of continuous functions from X to $\mathbb{R}$, containing at least the set of the constant functions taking every finite value c with $|c| \leq \sup_{\varphi\in\Phi} \|\varphi\|_\infty$.
- A topological subgroup G of $\mathrm{Homeo}(X)$ that acts on Φ by composition on the right.
- A subset $\mathscr{F}$ of the set $\mathscr{F}^{\mathrm{all}}(\Phi, G)$ of all non-expansive G-operators from Φ to Φ.

The operator space $\mathscr{F}^{\mathrm{all}}(\Phi, G)$

In plain words, $F \in \mathscr{F}^{\mathrm{all}}(\Phi, G)$ means that

1. $F : \Phi \to \Phi$
2. $F(\varphi \circ g) = F(\varphi) \circ g$. ($F$ is a G-operator)
3. $\|F(\varphi_1) - F(\varphi_2)\|_\infty \leq \|\varphi_1 - \varphi_2\|_\infty$. ($F$ is non-expansive)

The operator F is not required to be linear.

Some simple examples of F, taking Φ equal to the set of all continuous functions $\varphi : \mathbf{S}^1 \to \mathbb{R}$ and G equal to the group of all rotations of $\mathbf{S}^1$:

- $F(\varphi) :=$ the constant function $\psi : \mathbf{S}^1 \to \mathbb{R}$ taking the value $\max \varphi$;
- $F(\varphi)$ defined by setting $F(\varphi)(x) := \max\left\{\varphi\left(x - \frac{\pi}{8}\right), \varphi\left(x + \frac{\pi}{8}\right)\right\}$;
- $F(\varphi)$ defined by setting $F(\varphi)(x) := \frac{1}{2}\left(\varphi\left(x - \frac{\pi}{8}\right) + \varphi\left(x + \frac{\pi}{8}\right)\right)$.

The pseudo-metric $D^{\mathscr{F}}_{\mathrm{match}}$

For every $\varphi_1, \varphi_2 \in \Phi$ we set

$$D^{\mathscr{F}}_{\mathrm{match}}(\varphi_1, \varphi_2) := \sup_{F \in \mathscr{F}} d_{match}(\rho_k(F(\varphi_1)), \rho_k(F(\varphi_2)))$$

where $\rho_k(\psi)$ denotes the persistent Betti number function (i.e. the rank invariant) of ψ in degree k, while d_{match} denotes the usual bottleneck distance that is used to compare the persistence diagrams associated with $\rho_k(F(\varphi_1))$ and $\rho_k(F(\varphi_2))$.

Proposition

$D^{\mathscr{F}}_{match}$ is a G-invariant and stable pseudo-metric on Φ.

The G-invariance of $D^{\mathscr{F}}_{match}$ means that for every $\varphi_1, \varphi_2 \in \Phi$ and every $g \in G$ the equality $D^{\mathscr{F}}_{\mathrm{match}}(\varphi_1, \varphi_2 \circ g) = D^{\mathscr{F}}_{\mathrm{match}}(\varphi_1, \varphi_2)$ holds.

An equivalence result

We observe that the pseudo-distance $D^{\mathscr{F}}_{\text{match}}$ and the natural pseudo-distance d_G are defined in quite different ways.

In particular, the definition of $D^{\mathscr{F}}_{\text{match}}$ is based on persistent homology, while the natural pseudo-distance d_G is based on the group of homeomorphisms G.

In spite of this, the following statement holds:

Theorem

If $\mathscr{F} = \mathscr{F}^{\text{all}}(\Phi, G)$, then the pseudo-distance $D^{\mathscr{F}}_{\text{match}}$ coincides with the natural pseudo-distance d_G on Φ.

Our main idea

The previous theorem suggests to study $D_{\mathrm{match}}^{\mathscr{F}}$ instead of d_G.

To this end, let us choose a finite subset $\mathscr{F}^*$ of $\mathscr{F}$, and consider the pseudo-metric

$$D_{\mathrm{match}}^{\mathscr{F}^*}(\varphi_1, \varphi_2) := \max_{F \in \mathscr{F}^*} d_{match}(\rho_k(F(\varphi_1)), \rho_k(F(\varphi_2)))$$

for every $\varphi_1, \varphi_2 \in \Phi$.

Obviously, $D_{\mathrm{match}}^{\mathscr{F}^*} \leq D_{\mathrm{match}}^{\mathscr{F}}$.

Furthermore, if $\mathscr{F}^*$ is dense enough in $\mathscr{F}$, then the new pseudo-distance $D_{\mathrm{match}}^{\mathscr{F}^*}$ is close to $D_{\mathrm{match}}^{\mathscr{F}}$.

In order to make this point clear, we need the next theoretical result.

Compactness of $\mathscr{F}^{\mathrm{all}}(\Phi, G)$

The following result holds:

Theorem

If Φ is a compact metric space with respect to the sup-norm, then $\mathscr{F}^{\mathrm{all}}(\Phi, G)$ is a compact metric space with respect to the distance d defined by setting

$$d(F_1, F_2) := \max_{\varphi \in \Phi} \|F_1(\varphi) - F_2(\varphi)\|_\infty$$

for every $F_1, F_2 \in \mathscr{F}$.

Approximation of $\mathscr{F}^{\mathrm{all}}(\Phi, G)$

This statement follows:

Corollary

Assume that the metric space Φ is compact with respect to the sup-norm. Let $\mathscr{F}$ be a subset of $\mathscr{F}^{\mathrm{all}}(\Phi, G)$. For every $\varepsilon > 0$, a finite subset $\mathscr{F}^*$ of $\mathscr{F}$ exists, such that

$$\left| D_{match}^{\mathscr{F}^*}(\varphi_1, \varphi_2) - D_{match}^{\mathscr{F}}(\varphi_1, \varphi_2) \right| \leq \varepsilon$$

for every $\varphi_1, \varphi_2 \in \Phi$.

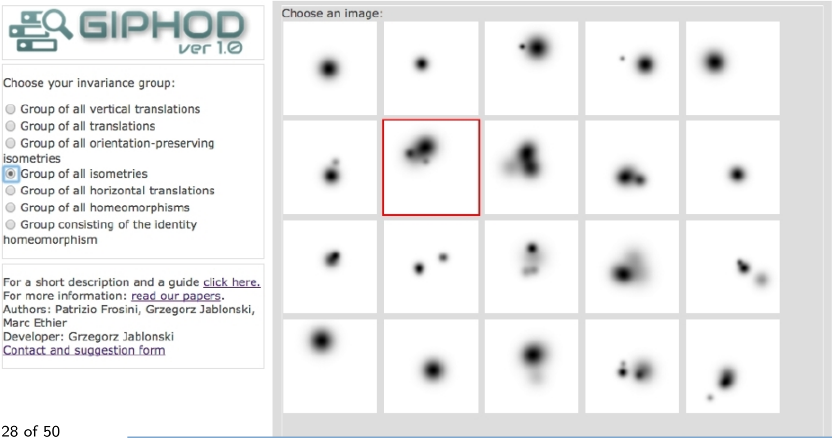
This corollary implies that the pseudo-distance $D_{match}^{\mathscr{F}}$ can be approximated computationally, at least in the compact case.

Two references

- Patrizio Frosini, Grzegorz Jabłoński, *Combining persistent homology and invariance groups for shape comparison*, Discrete & Computational Geometry, vol. 55 (2016), n. 2, pages 373-409.
- Patrizio Frosini, *Towards an observer-oriented theory of shape comparison*, Proceedings of the 8th Eurographics Workshop on 3D Object Retrieval, Lisbon, Portugal, May 7-8, 2016, A. Ferreira, A. Giachetti, and D. Giorgi (Editors), 5-8.

GIPHOD

Joint project with Grzegorz Jabłoński (Jagiellonian University and IST Austria) and Marc Ethier (Université de Saint-Boniface - Canada)



GIPHOD (Group Invariant Persistent Homology On-line Demonstrator

GIPHOD is an on-line demonstrator, allowing the user to choose an image and an invariance group. GIPHOD searches for the most similar images in the dataset, with respect to the chosen invariance group.

Purpose: to show the use of our theoretical approach for image comparison.

Dataset: 10.000 quite simple grey-level synthetic images obtained by adding randomly chosen bell-shaped functions. The images are coded as functions from $\mathbb{R}^2$ to $[0,1]$.

GIPHOD can be tested at `http://giphod.ii.uj.edu.pl`.
All suggestions are greatly appreciated and welcomed (please send them to grzegorz.jablonski@uj.edu.pl)

GIPHOD (Group Invariant Persistent Homology On-line Demonstrator)

We will now show some results obtained by GIPHOD when the invariance group G is the group of isometries:
Some data about the pseudo-metric $D^{\mathscr{F}}_{\mathrm{match}}$ in this case:

- The images are coded as functions from $\mathbb{R}^2 \to [0,1]$;
- Mean distance between images: 0.2984;
- Standard deviation of distance between images: 0.1377;
- Number of GINOs that have been used: 5.

GIPHOD (Group Invariant Persistent Homology On-line Demonstrator)

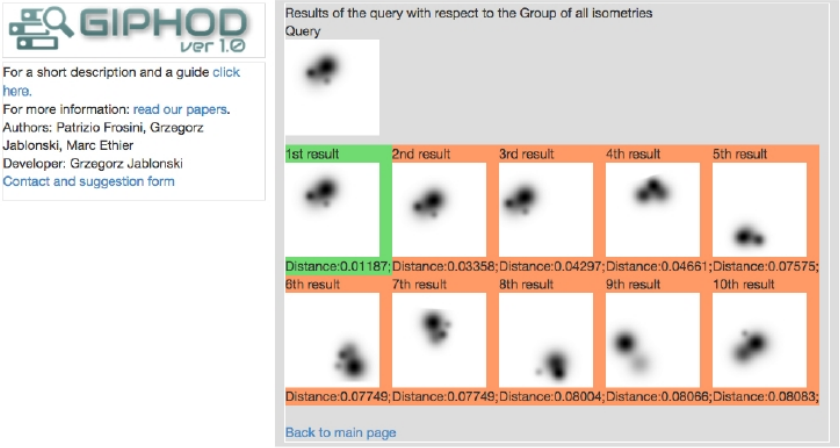
List of GINOs that have been used in the following image retrievals, where the invariance group G is the group of isometries:

- $F(\varphi) = \varphi$.
- $F(\varphi) :=$ constant function taking each point to the value $\int_{\mathbb{R}^2} \varphi(\boldsymbol{x})\ d\boldsymbol{x}$.
- $F(\varphi)$ defined by setting

$$F(\varphi)(\boldsymbol{x}) := \int_{\mathbb{R}^2} \varphi(\boldsymbol{x} - \boldsymbol{y}) \cdot \beta\left(\|\boldsymbol{y}\|_2\right)\ d\boldsymbol{y}$$

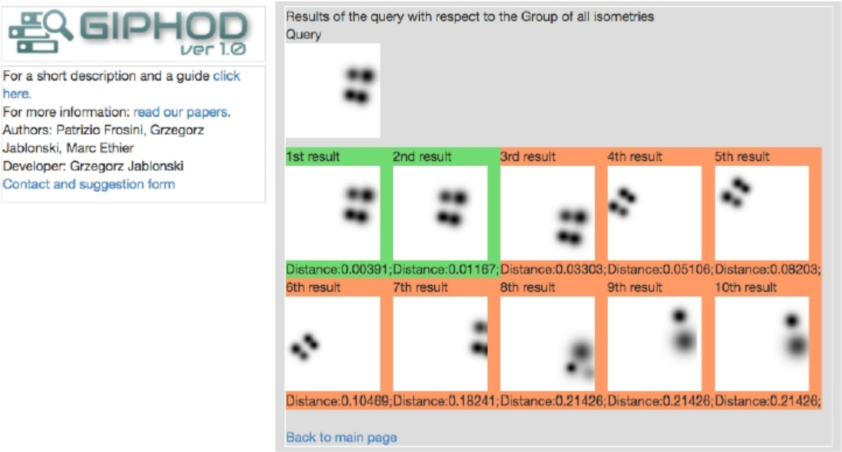
  where $\beta : \mathbb{R} \to \mathbb{R}$ is an integrable function with $\int_{\mathbb{R}^2} \left|\beta\left(\|\boldsymbol{y}\|_2\right)\right|\ d\boldsymbol{y} \leq 1$. Three GINOs of this kind have been used.

GIPHOD: Examples for the group of isometries



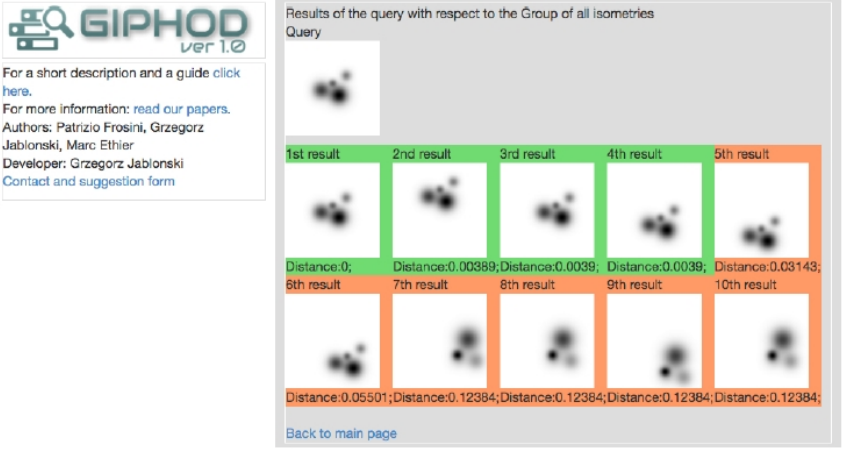
Mean distance: 0.2984; Standard deviation of distance: 0.1377; Number of GINOs: 5.

GIPHOD: Examples for the group of isometries



Mean distance: 0.2984; Standard deviation of distance: 0.1377; Number of GINOs: 5.

GIPHOD: Examples for the group of isometries



Mean distance: 0.2984; Standard deviation of distance: 0.1377; Number of GINOs: 5.

Our basic questions

Assumptions in our model

Mathematical setting and theoretical results

A first step towards the application of our model: GIPHOD

Some work in progress

Some work in progress

Some work is in progress, concerning these three lines of research:

- Change of the topologies used on X and G.
- Extension of our approach to operators taking the space Φ (where a group G acts) to a different space of functions Ψ (where another group H acts).
- Study of the algebra of GINOs both in the case $(\Phi, G) = (\Psi, H)$ and in the case $(\Phi, G) \neq (\Psi, H)$.

(Joint work with Nicola Quercioli)

Some work in progress: New topologies on X and G

Let X be a set. Let Φ be a non-empty subset of the set of all bounded functions from X to $\mathbb{R}$, endowed with the norm $\|\cdot\|_\infty$. We assume that Φ is compact and contains at least the constant functions taking every value c with $|c| \leq \sup_{\varphi\in\Phi} \|\varphi\|_\infty$. We also consider a group $G \subseteq \mathrm{Homeo}(X)$, acting on Φ by composition on the right.

We endow X with the initial topology, i.e. the coarsest topology on X such that every function in Φ is continuous. In other words, on X we consider this pseudo-metric: $d_X(x_1, x_2) := \sup_{\varphi\in\Phi} |\varphi(x_1) - \varphi(x_2)|$.

We endow the group G with the pseudo-metric $D_G(g_1, g_2) := \sup_{\varphi\in\Phi} \|\varphi \circ g_1 - \varphi \circ g_2\|_\infty$. G is a topological group that acts continuously on Φ by composition on the right.

We will also assume that X and G are compact, and say that (Φ, G) is a perception pair.

Some work in progress: Changing (Φ, G) into (Ψ, H)

We wish to extend our theoretical approach to the case of different perception pairs. In order to do that, we consider (Φ, G) as a **category** whose objects are the elements of the compact space Φ and whose arrows are the elements of the topological group G.

Each functor $F : (\Phi, G) \to (\Psi, H)$ is called a Group Invariant Non-expansive Operator (GINO) if:

- F is group invariant: $F(\varphi \circ g) = F(\varphi) \circ F(g)$ for every $\varphi \in \Phi, g \in G$;
- F is non-expansive on Φ: $\|F(\varphi_1) - F(\varphi_2)\|_\infty \leq \|\varphi_1 - \varphi_2\|_\infty$ for every $\varphi_1, \varphi_2 \in \Phi$;
- F is non-expansive on G: $D_H(F(g_1), F(g_2)) \leq D_G(g_1, g_2)$ for every $g_1, g_2 \in G$.

Some work in progress: The metric space of GINOs from (Φ, G) to (Ψ, H)

We can endow the set of all GINOs from (Φ, G) to (Ψ, H) with this metric: $d_{\mathscr{F}}(F_1, F_2) :=$
$\max\left\{\sup_{\varphi\in\Phi} \|F_1(\varphi) - F_2(\varphi)\|_\infty, \sup_{g\in G} D_G\left(F_1(g), F_2(g)\right)\right\}$.

Theorem

The metric space of GINOs from (Φ, G) to (Ψ, H) is compact.

Corollary

The metric space of GINOs from (Φ, G) to (Ψ, H) can be ε-approximated by a finite subset.

Some work in progress: Extending the definition of $D^{\mathscr{F}}_{\mathrm{match}}$ to the case $(\Phi, G) \neq (\Psi, H)$

The previous corollary opens the way to the computational approximation of the following pseudo-metric.
Let us consider a set $\mathscr{F}$ of GINOs from (Φ, G) to (Ψ, H).

Fore every $\varphi_1, \varphi_2 \in \Phi$ we set

$$D^{\mathscr{F}}_{\mathrm{match}}(\varphi_1, \varphi_2) := \sup_{F \in \mathscr{F}} d_{match}(\rho_k(F(\varphi_1)), \rho_k(F(\varphi_2)))$$

for every $\varphi_1, \varphi_2 \in \Phi$, where $\rho_k(\psi)$ denotes the persistent Betti numbers function (i.e. the rank invariant) of $\psi \in \Psi$ in degree k, while d_{match} denotes the usual bottleneck distance that is used to compare the persistence diagrams associated with $\rho_k(F(\varphi_1))$ and $\rho_k(F(\varphi_2))$.

Some work in progress: Extending the definition of $D^{\mathscr{F}}_{\text{match}}$ to the case $(\Phi, G) \neq (\Psi, H)$

Proposition

$D^{\mathscr{F}}_{\text{match}}$ is a G-invariant and stable pseudo-metric on Φ.

The G-invariance of $D^{\mathscr{F}}_{\text{match}}$ means that for every $\varphi_1, \varphi_2 \in \Phi$ and every $g \in G$ the equality $D^{\mathscr{F}}_{\text{match}}(\varphi_1, \varphi_2 \circ g) = D^{\mathscr{F}}_{\text{match}}(\varphi_1, \varphi_2)$ holds.

The stability of $D^{\mathscr{F}}_{\text{match}}$ means that $D^{\mathscr{F}}_{\text{match}}$ is upper-bounded by the natural pseudo-distance d_G:
$D^{\mathscr{F}}_{\text{match}}(\varphi_1, \varphi_2) \leq d_G(\varphi_1, \varphi_2) \leq \|\varphi_1 - \varphi_2\|_\infty$.

The algebra of GINOs

Our approach to G-invariant TDA is based on the availability of GINOs.

How could we build new GINOs from other GINOs?

A simple method consists in using the properties of functors and producing new GINOs by composition of other GINOs: If F_1 is a GINO from (Φ, G) to (Ψ, H) and F_2 is a GINO from (Ψ, H) to (χ, K), then $F_2 \circ F_1$ is a GINO from (Φ, G) to (χ, K).

Building GINOs via 1-Lipschitzian functions

We can also produce new GINOs by means of a 1-Lipschitzian function applied to other GINOs:

Proposition

Assume that two perception categories (Φ, G), (Ψ, H) *are given. Let* $\mathscr{L}$ *be a* 1*-Lipschitzian map from* $\mathbb{R}^n$ *to* $\mathbb{R}$, *where* $\mathbb{R}^n$ *is endowed with the norm* $\|(x_1, \ldots, x_n)\|_\infty := \max_{1 \le i \le n} |x_i|$. *Assume also that* $F_1, \ldots, F_n$ *are GINOs from* (Φ, G) *to* (Ψ, H) *that coincide on the homeomorphisms in* G. *Let us define* $\mathscr{L}^*(F_1, \ldots, F_n) : \Phi \to \Psi$ *by setting* $\mathscr{L}^*(F_1, \ldots, F_n)(\varphi)(x) := \mathscr{L}(F_1(\varphi)(x), \ldots, F_n(\varphi)(x))$. *We also set* $\mathscr{L}^*(F_1, \ldots, F_n)(g) = F_1(g) = \ldots = F_n(g)$ *for every* $g \in G$. *Then* $\mathscr{L}^*(F_1, \ldots, F_n)$ *is a GINO from* (Φ, G) *to* (Ψ, H).

From this proposition the following three results follow.

Building new GINOs via translations, weighted averages and the maximum operator

Proposition (Translation)

Let F be a GINO from (Φ, G) to (Ψ, H). Then the operator F_b that is defined as $F_b(\varphi) = \varphi - b$ on Φ and as $F_b(g) = F(g)$ on G is a GINO from (Φ, G) to (Ψ, H) for every $b \in \mathbb{R}$.

Proposition (Weighted average)

Assume that $F_1, \ldots, F_n$ are GINOs from (Φ, G) to (Ψ, H), that they coincide on the homeomorphisms in G, and that $(a_1, \ldots, a_n) \in \mathbb{R}^n$ with $\sum_{i=1}^n |a_i| \leq 1$. Then the operator F that is defined as $F(\varphi) = \sum_{i=1}^n a_i F_i(\varphi)$ on Φ and as $F(g) = F_1(g) = \ldots = F_n(g)$ on G is a GINO from (Φ, G) to (Ψ, H).

Building new GINOs via translations, weighted averages and the maximum operator

Proposition (Maximum)

Assume $F_1, \ldots, F_n$ are GINOs from (Φ, G) to (Ψ, H), and that they coincide on the homeomorphisms in G. Then the operator F that is defined as $F(\varphi) = \max_i F_i(\varphi)$ on Φ and as $F(g) = F_1(g) = \ldots = F_n(g)$ on G is a GINO from (Φ, G) to (Ψ, H).

An interesting GINO in kD persistent homology

Previous propositions imply the following statement.

Proposition

Assume $F_1, \dots, F_n$ are GINOs from (Φ, G) to (Ψ, H), and that they coincide on the homeomorphisms in G. Assume also that $(a_1, \dots, a_n)$, $(b_1, \dots, b_n) \in \mathbb{R}^n$, with $a_1, \dots, a_n > 0$, $\sum_{i=1}^n a_i = 1$ and $\sum_{i=1}^n b_i = 0$. Then the operator F that is defined as

$$F(\varphi) = \max\left\{ \frac{\min_j a_j}{a_1} \cdot (F_1(\varphi) - b_1), \dots, \frac{\min_j a_j}{a_n} \cdot (F_n(\varphi) - b_n) \right\}$$

on Φ and as $F(g) = F_1(g) = \ldots = F_n(g)$ on G is a GINO from (Φ, G) to (Ψ, H).

An interesting GINO in kD persistent homology

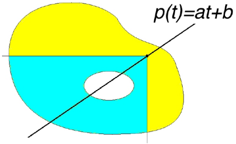
The previous proposition shows an interesting link between the theory of GINOs and the reduction of multidimensional persistent Betti numbers to 1D persistent Betti numbers.

We recall that this reduction is done by means of the operator that takes each $\mathbb{R}^n$-valued function $\varphi = (\varphi_1, \ldots, \varphi_n)$ to the $\mathbb{R}$-valued functions

$$\varphi_{(a,b)} := \max\left\{ \frac{\min_j a_j}{a_1} \cdot (\varphi_1 - b_1), \ldots, \frac{\min_j a_j}{a_n} \cdot (\varphi_n - b_n) \right\}$$

with $a_1, \ldots, a_n > 0$, $\sum_{i=1}^n a_i = 1$ and $\sum_{i=1}^n b_i = 0$. The key point is that each sublevel set of φ can be represented as a sublevel set of the filtering function $\varphi_{(a,b)}$, so that our operator can be used to reduce a multidimensional filtration to a collection of 1D-filtrations.

An interesting GINO in kD persistent homology



The collection of the 1D-filtrations associated with the lines $p(t) = (a, 1-a)t + (b, -b)$ such that $a, b \in \mathbb{R}$ with $0 < a < 1$ is equivalent to the 2D-filtration associated with the filtering function $p \mapsto (x(p), y(p))$.

[A. Cerri, B. Di Fabio, M. Ferri, P. Frosini, C. Landi, Betti numbers in multidimensional persistent homology are stable functions, Mathematical Methods in the Applied Sciences, vol. 36 (2013), 1543-1557.]

Conclusions

In this talk we have supported these statements:

- Data comparison is based on acts of measurement made by an observer. The acts of measurement can be represented as a function defined on a topological space X. The observer can be seen as a collection of G-invariant operators, applied to the functions describing the data.
- These functions can be compared by means of the natural pseudo-distance associated with any subgroup G of $\mathrm{Homeo}(X)$.
- Persistent homology can be used to approximate the natural pseudo-metric d_G. This can be done by means of a method that is based on non-expansive G-operators. This method is stable with respect to noise.


Thanks for your attention!