

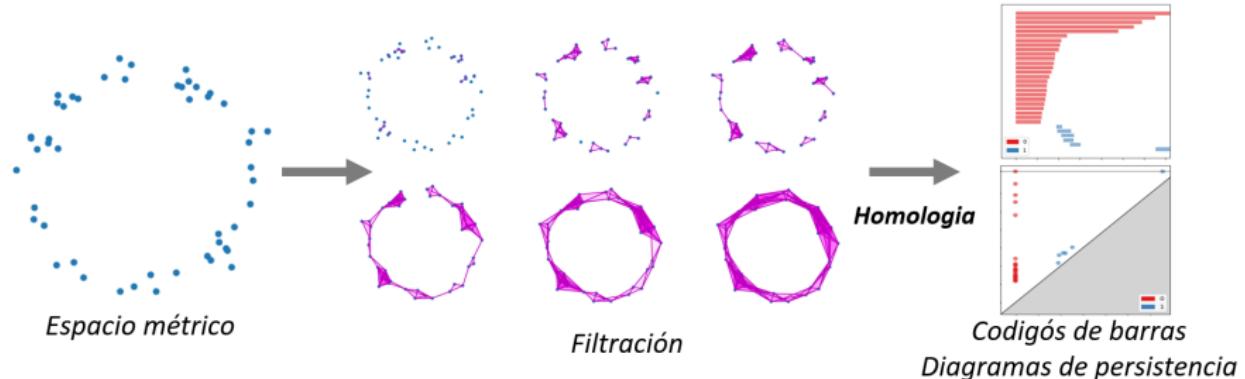
# Topología algebraica aplicada al análisis de datos: - Topological Data Analysis (TDA) -

Léo Quercioli

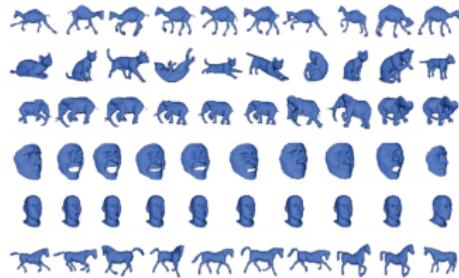
seminario relajao

15 julio 2025

# Introducción



# Introducción - ejemplos de aplicación



*Unsupervised shape  
classification with  
Bottleneck distance*

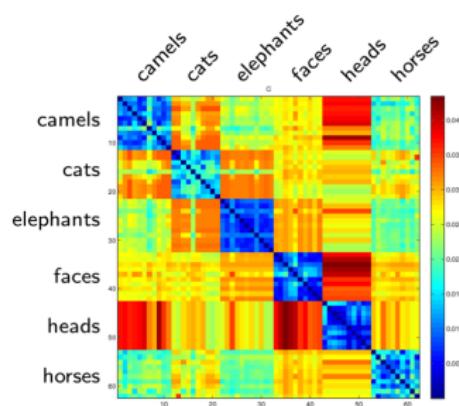


Figure: *Topological data analysis - Topological descriptors for geometric data*, S. Oudot

# Introducción - ejemplos de aplicación

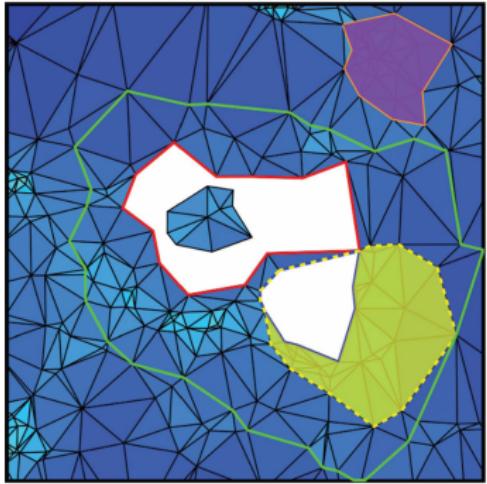


Figure: *The persistent cosmic web and its filamentary structure—I. Theory and implementation*, T. Sousbie

# Introducción - ejemplos de aplicación

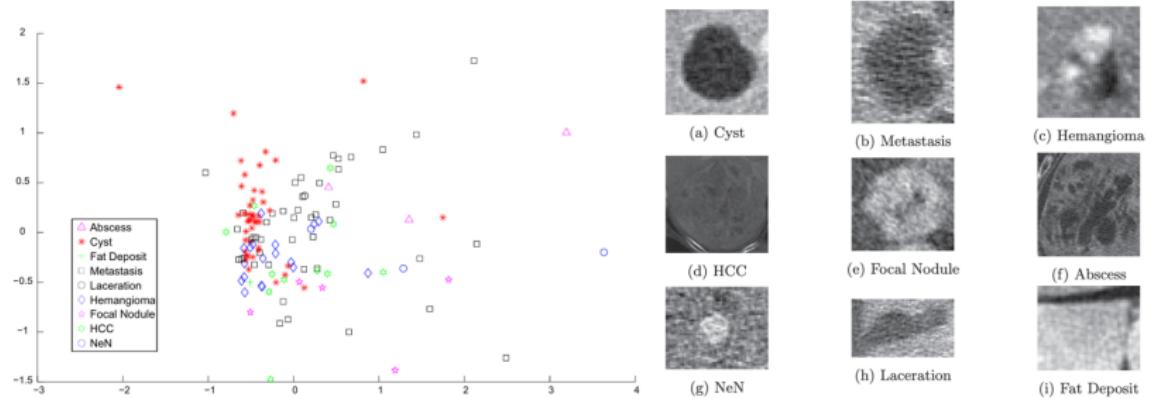


Figure: Classification of hepatic lesions using the matching metric, A.  
Adcock, D. Rubin, and G. Carlsson

# Introducción

1. Homología (simplicial)
2. Homología persistente
3. Inferencia geométrica
4. Estabilidad
5. Persistencia extendida

# Homología

En una categoría abeliana  $\mathcal{C}$  un **complejo de cadenas**  $(C_\bullet, d_\bullet)$  consistente en objetos  $C_n \in \mathcal{C}$  y morfismos  $d_n : C_n \rightarrow C_{n-1}$  ( $n \in \mathbb{Z}$ ),

$$\dots \xrightarrow{d_{n+2}} C_{n+1} \xrightarrow{d_{n+1}} C_n \xrightarrow{d_n} C_{n-1} \xrightarrow{d_{n-1}} \dots$$

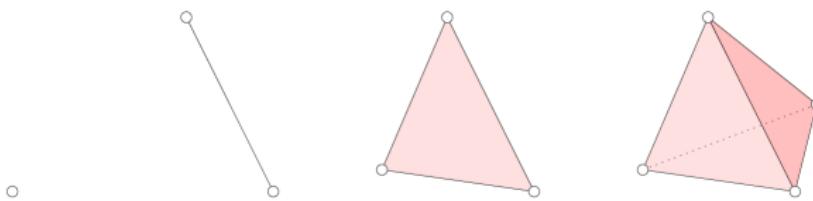
con la condición que  $\forall n \in \mathbb{Z} : d_n \circ d_{n+1} = 0$  (i.e  $\text{im } d_{n+1} \subseteq \ker d_n$ )  
Y los **grupos de homología** del complejo de cadenas  $(C_\bullet, d_\bullet)$  son para  $n \in \mathbb{Z}$ ,

$$H_n(C_\bullet) := \ker d_n / \text{im } d_{n+1}$$

# Homología (simpl.)

Un  **$k$ -símplex** de  $\mathbb{R}^n$  es la envolvente convexa de  $k + 1$  puntos independientes  $s_0, \dots, s_k \in \mathbb{R}^n$ , y se denota  $\sigma = \langle s_0, \dots, s_k \rangle$  (con orientación). Los **lados** de  $\sigma$  son las  **$(k - 1)$ -símplices**  $\tau \subseteq \sigma$ . Un **complejo simplicial** de  $\mathbb{R}^n$  es un conjunto  $K$  de simplices de  $\mathbb{R}^n$  con las condiciones que,

1.  $\forall \sigma \in K$ , los lados de  $\sigma$  están en  $K$ .
2. Si  $\sigma, \tau \in K$ ,  $\sigma \cap \tau \neq \emptyset$ , entonces  $\sigma \cap \tau \in K$



Y notamos  $K^{(i)} := \{\sigma \in K ; \dim \sigma = i\}$  el  **$i$ -ésqueleto**.

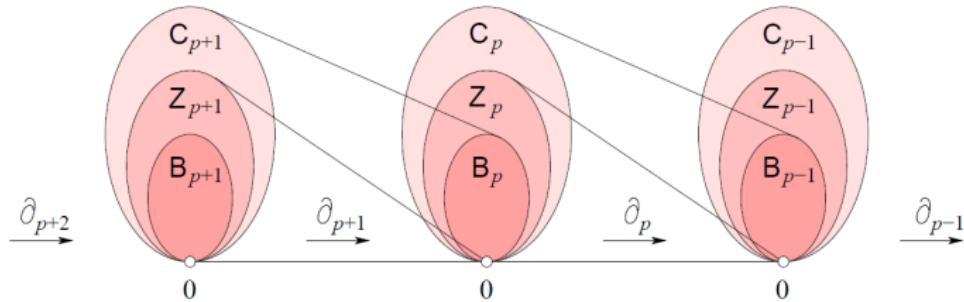
# Homología (simpl.)

Escribimos  $C_i(K)$  el grupo abeliano libre generado por  $K^{(i)}$  y el **operador de contorno**  $\partial_k : C_k(K) \rightarrow C_{k-1}(K)$  como, para  $\sigma = \langle s_0, \dots, s_k \rangle \in C_k(K)$ ,

$$\partial_k \sigma := \sum_i (-1)^i \langle s_0, \dots, \hat{s}_i, \dots, s_k \rangle$$

## Lema

$\partial^2 = 0$  i.e  $(C_\bullet(K), \partial_\bullet)$  es un complejo de cadenas.



## Módulo de persistencia

Sea  $(T, \leq)$  un conjunto totalmente ordenado. Un **módulo de persistencia** sobre  $T$  es un functor  $M : (T, \leq) \rightarrow R\text{-Mod}$ . Un **morfismo** entre módulos de persistencia  $M$  y  $N$  es un conjunto de aplicaciones  $\varphi(t) : M(t) \rightarrow N(t)$  tal que el diagrama siguiente conmuta para todo  $s \leq t$ ,

$$\begin{array}{ccc} M(s) & \xrightarrow{M(s \leq t)} & M(t) \\ \varphi(s) \downarrow & & \downarrow \varphi(t) \\ N(s) & \xrightarrow{N(s \leq t)} & N(t) \end{array}$$

Un módulo de persistencia  $M$  es **reducible** si existen dos módulos de persistencia no nulos  $N, O$  tales que,

$$M \cong N \oplus O.$$

# Módulo de persistencia

Sea  $I \subseteq (T, \leq)$  un intervalo. Definimos el **intervalo de persistencia** de  $I$  como el módulo  $R_I : (T, \leq) \rightarrow R\text{-Mod}$  dado por,

- $R_I(t) = R$  si  $t \in I$  y  $0$  en el otro caso,
- $R_I(s \leq t) = id_R$  si  $s \leq t \in I$  y  $0$  en el otro caso.

## Teorema (Crawley-Boevey) [2014]

Sea  $M : (T, \leq) \rightarrow R\text{-Mod}$  un módulo de persistencia tal que  $M(t)$  es de tipo finito para todo  $t$ . Entonces existe un multiconjunto de intervalos de  $T$ ,  $\{I_j \subseteq T\}_{j \in J} =: \mathbf{Brd}(M)$ , tal que,

$$M \cong \bigoplus_{j \in J} R_{I_j}.$$

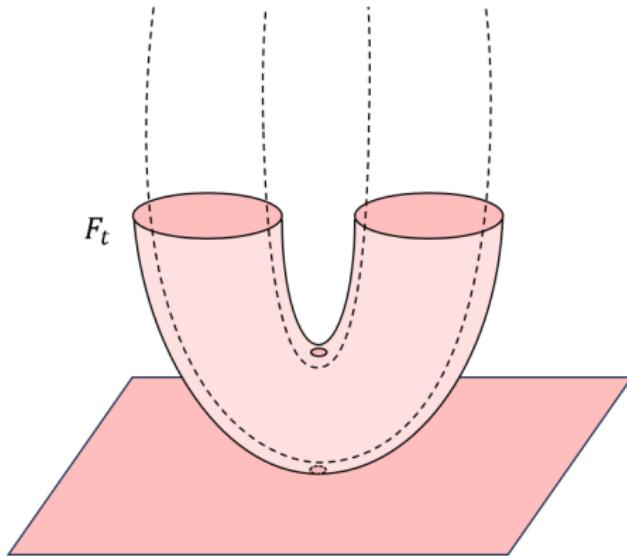
$I_j = [b_j, d_j] \in \mathbf{Brd}(M) \longrightarrow (b_j, d_j) \in \mathbf{Dgm}(M)$  (intervalos de nacimiento y muerte de ciclos)

# Homología persistente

Sea  $X$  una variedad triangulable. Sea  $f : X \rightarrow \mathbb{R}$  un mapa. Definimos los conjuntos de subnivel,  $F_t = f^{-1}(-\infty, t])$ , supongamos los  $F_t$  triangulables.

Lema (de homología persistente)

$t \mapsto H_k(F_t)$  es un módulo de persistencia, lo notamos **Dgm**( $f$ )



# Algoritmo

**INPUT:** Una filtración simplicial del complejo  $K$ ,  $F_0 \subseteq \dots \subseteq F_n$  tal que  $F_0 = \emptyset$  y  $F_n = K$ ,  $\forall t$ ,  $F_t \subseteq F_{t+1}$  y  $\forall t$ ,  $F_{t+1} \setminus F_t = \{\sigma_t\}$

**ALGO:**

1.  $A = [\partial]_{ij} = (c_1 | \dots | c_n)$  , la matriz del operador frontera en la base  $\sigma_1, \dots, \sigma_n$ .
2. Definimos,

$$\text{low}(j) := \begin{cases} 0 & \text{si } A_{ij} = 0 \forall i \\ \max\{i | A_{ij} \neq 0\} & \text{en caso contrario.} \end{cases}$$

3. Reducción de  $A$ ,

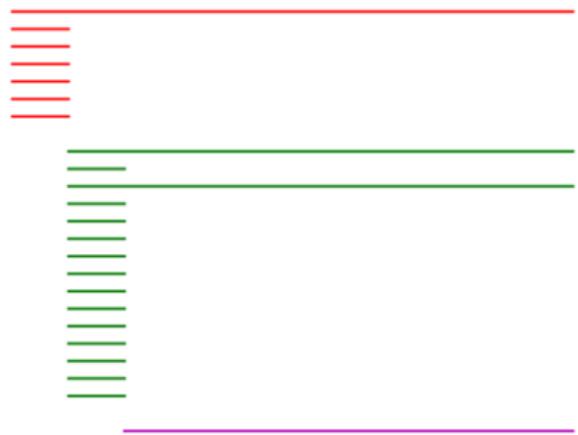
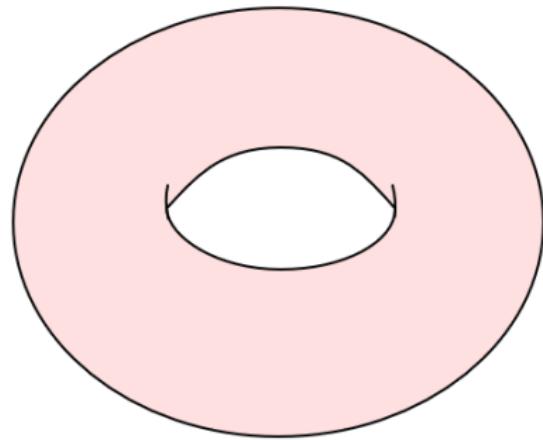
# Algoritmo

```
for  $j = 1, \dots, n$  do
    while  $\exists i < j$  tal que  $\text{low}(i) = \text{low}(j) \neq 0$  do
         $c_j \leftarrow c_j - (A_{\text{low}(i), j} / A_{\text{low}(i), i}) c_i$ 
    end while
end for
return  $A$ 
```

**OUTPUT:**  $\tilde{A} = (\tilde{c}_1, \dots, \tilde{c}_n)$ .  $\tilde{c}_j = 0 \Rightarrow \sigma_j$  crea un ciclo. Y  
 $\tilde{c}_j \neq 0 \Rightarrow \sigma_j$  mata un ciclo.

- (i) Cada pareja  $(\sigma_i, \tilde{c}_j)$  con  $i = \text{low}(j)$  representa un intervalo  $R_{[i,j]}$  en  $H_{\dim \sigma_i}(K)$ .
- (ii) Cada  $\tilde{c}_i = 0$  no emparejado (tal que no existe  $j$ ,  $i = \text{low}(j)$ ), representa un intervalo infinito  $R_{[i,\infty[}$  en  $H_{\dim \sigma_i}(K)$ .

# Ejemplo



## Distancia culo de botella

Sea  $M \subseteq A \times B$  (multi-conjuntos),  $M$  es un **emparejamiento parcial**, denotado  $M : A \leftrightarrow B$ , si las restricciones  $\pi_{A|_M}$  y  $\pi_{B|_M}$  son inyectivas ( $\pi_A : A \times B \rightarrow A$  y  $\pi_B : A \times B \rightarrow B$  las proyecciones canónicas), i.e para todo  $a \in A$  (resp.  $b \in B$ ), existe **a lo sumo** un  $b \in B$  (resp. un  $a \in A$ ), tales que  $(a, b) \in M$ . Le coste de  $M : A \leftrightarrow B$  se define como,

$$c(M) := \max \left\{ \sup_{(a,b) \in M} \|b - a\|_\infty, \sup_{\substack{s \in A \sqcup B \\ \text{unmatched}}} \frac{|s_y - s_x|}{2} \right\}$$

Y la **distancia culo de botella** se define como,

$$d_{\text{botella}}(A, B) := \inf_{M: A \leftrightarrow B} c(M)$$

# Estabilidad I

Para todos  $f, g : X \rightarrow \mathbb{R}$  tales que los conjuntos de subnivel  $F_t$  y  $G_t$  son triangulables y sus grupos de homología de tipo finito entonces,

$$d_{\text{botella}}(\mathbf{Dgm} f, \mathbf{Dgm} g) \leq \|f - g\|_\infty$$

# ★ Estabilidad I

Un  $\varepsilon$ -interleaving entre dos módulos de persistencia  $M, N$  es el dato de mapas lineales  $\forall t, \varphi_t : M_t \rightarrow N_{t+\varepsilon}$  y  $\psi_t : N_t \rightarrow M_{t+\varepsilon}$  tales que para todos  $s \leq t$ , los siguientes diagramas conmutan,

$$\begin{array}{ccc} M_s & \xrightarrow{M(s \leq t)} & M_t \\ \varphi_s \searrow & & \downarrow \varphi_t \\ N_{s+\varepsilon} & \xrightarrow[N(s+\varepsilon \leq t+\varepsilon)]{} & N_{t+\varepsilon} \end{array}$$

$$\begin{array}{ccc} M_{s+\varepsilon} & \xrightarrow{M(s+\varepsilon \leq t+\varepsilon)} & M_{t+\varepsilon} \\ \psi_s \nearrow & & \downarrow \psi_t \\ N_s & \xrightarrow[N(s \leq t)]{} & N_t \end{array}$$

$$\begin{array}{ccc} M_s & \xrightarrow{M(s \leq s+2\varepsilon)} & M_{s+2\varepsilon} \\ \varphi_s \searrow & & \nearrow \psi_{s+\varepsilon} \\ N_{s+\varepsilon} & & \end{array}$$

$$\begin{array}{ccc} M_{s+\varepsilon} & & \\ \psi_s \nearrow & \searrow \varphi_{s+\varepsilon} & \\ N_s & \xrightarrow[M(s \leq s+2\varepsilon)]{} & N_{s+2\varepsilon} \end{array}$$

La distancia de interleaving es

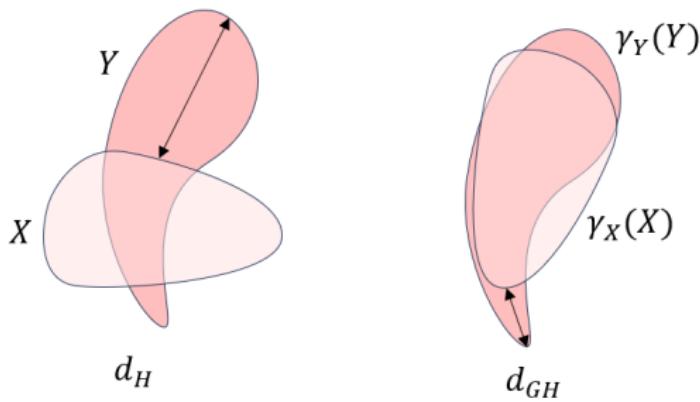
$d_{\text{inter}}(M, N) := \inf\{\varepsilon > 0 ; M, N \text{ son } \varepsilon\text{-interleaved}\}$  y tenemos la isometría  $d_{\text{botella}}(M, N) = d_{\text{inter}}(M, N)$

## Inferencia geométrica - distancia de Gromov-Hausdorff

Sea  $(X, d_X)$  y  $(Y, d_Y)$  dos espacios métricos compactos, la distancia de Gromov-Hausdorff está definida por,

$$d_{GH}(X, Y) := \inf_{\substack{(Z, d_Z) \text{ espacio métrico} \\ \gamma_X: X \rightarrow Z, \gamma_Y: Y \rightarrow Z \text{ isometrías}}} d_H(\gamma_X(X), \gamma_Y(Y))$$

Con  $d_H$  la distancia de Hausdorff.



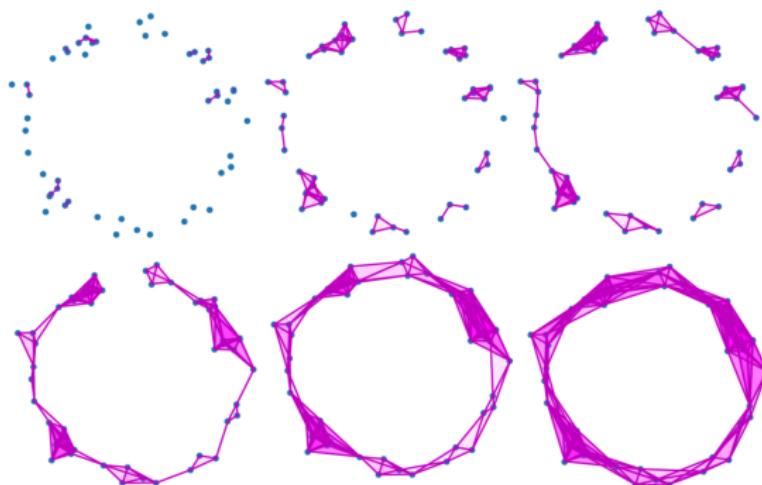
## Inferencia geométrica - filtración de Vietoris-Rips

Sea  $P$  un conjunto finito de puntos. La filtración de Vietoris-Rips está,

$$\mathcal{R}(P) := (R(P, t))_{t \in \mathbb{R}}$$

con,

$$\sigma = \{p_0, \dots, p_r\} \in R(P, t) \iff \text{diam } \sigma \leq t$$



## Estabilidad II

Teorema (Chazal, De Silva, Oudot) [2013]

Sean  $(X, d_X)$  y  $(Y, d_Y)$  espacios métricos compactos (o finitos de puntos),

$$d_{\text{botella}}(\mathbf{Dgm} \mathcal{R}(X), \mathbf{Dgm} \mathcal{R}(Y)) \leq d_{GH}(X, Y)$$

# Dromedario y persistencia ordinaria

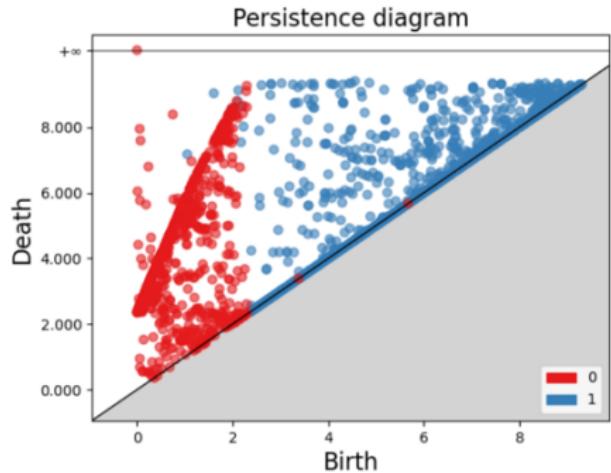
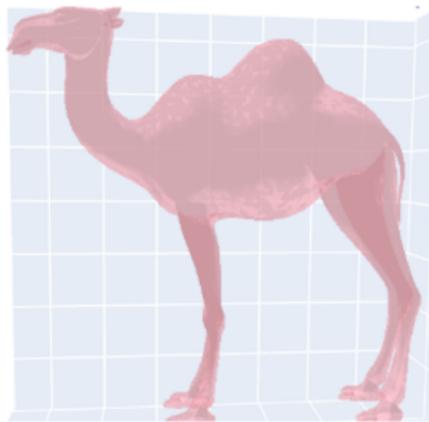
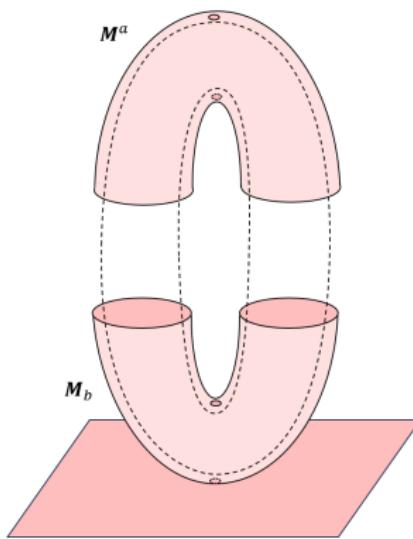


Figure: Persistencia de un dromedario (obtenido gracias a *GUDHI-library para TDA in C++ y Python*)

# Dualidad de Poincaré-Lefschetz y persistencia extendida

Construimos una homología persistente sobre la filtración creciente de los subniveles: la **persistencia ordinaria**. Vamos a construir una filtración decreciente en homología relativa de los subniveles: la **persistencia relativa**. Y una combinación de ambas: la **persistencia extendida**.



# Dualidad de Poincaré-Lefschetz

Una **variedad con borde** de dimensión  $n$  es un espacio topológico separado tal que cada punto tiene un vecindario abierto homeomorfo a un abierto de  $\mathbb{R}_+^n = \{(x_1, \dots, x_n) \in \mathbb{R}^n; x_n \geq 0\}$ . Se define el **borde** de  $X$  por,

$$\partial X = \{x \in X; H_n(X, X \setminus x) \cong 0\}$$

## Dualidad de Poincaré-Lefschetz [1926]

Sea  $X$  una variedad compacta de dimensión  $n$  con borde  $\partial X$ , entonces para todo  $k$ ,

$$H^k(X, \partial X; R) \cong H_{n-k}(X; R)$$

Se utiliza la dualidad de Poincaré sobre  $X \setminus \partial X$  que no tiene borde.

# Dualidad de Poincaré-Lefschetz y persistencia extendida

Sea  $f : M \rightarrow \mathbb{R}$  con  $M$  compacto sin borde.  $M_t = f^{-1}(-\infty, t]$  y  $M^t = f^{-1}([t, \infty[)$

Sean  $s \leq t$ , tenemos los morfismos inducidos por las inclusiones,

$$\begin{aligned} H_p(M_s) &\longrightarrow H_p(M_t), \\ H_p(M, M^t) &\longrightarrow H_p(M, M^s). \end{aligned}$$

Así,  $(H_p(M_t))_{t \in \mathbb{R}}$  y  $(H_p(M, M^t))_{t \in \mathbb{R}}$  son módulos de persistencia sobre  $\mathbb{R}$  y  $\mathbb{R}^{op}$  respectivamente. Además, definimos una tercera filtración,

$$Ext_p(t) = \begin{cases} H_p(M_t), & \text{si } t \in \mathbb{R}, \\ H_p(M) \cong H_p(M, \emptyset), & \text{si } t = \infty, \\ H_p(M, M^t), & \text{si } t \in \mathbb{R}^{op}. \end{cases}$$

sobre  $\mathbb{R}^{ext} = \mathbb{R} \coprod \{\infty\} \coprod \mathbb{R}^{op}$ , donde  $\mathbb{R} < \{\infty\} < \mathbb{R}^{op}$ . Y  $\mathbf{Ord}_p(f)$ ,  $\mathbf{Rel}_p(f)$  y  $\mathbf{Ext}_p(f)$  los diagramas correspondientes.

# Dualidad de Poincaré-Lefschetz y persistencia extendida

Tomamos los coeficientes en  $\mathbb{Z}/2\mathbb{Z}$  orientable. Tomando  $U = f^{-1}(]t + \varepsilon, +\infty[)$ , sin cruzar puntos críticos,

$$\begin{aligned} H_{n-p}(M, M^t) &\cong H_{n-p}(M \setminus U, M^t \setminus U) && \text{(lema de escisión)} \\ &\cong H_{n-p}(M_t, \partial M_t) && \text{(entre 2 puntos críticos)} \\ &\cong H^p(M_t) && \text{(Poincaré-Lefschetz)} \\ &\cong H_p(M_t) && \text{(mod 2)} \end{aligned}$$

Teorema de dualidad en persistencia extendida (Cohen-Steiner, Edelsbrunner, Haer) [2009]

Se tiene que  $\mathbf{Ord}_p(f) = \mathbf{Rel}_{n-p}^t(f)$  y  $\mathbf{Ext}_p(f) = \mathbf{Ext}_{n-p}^t(f)$

Donde  $\cdot^t$  es la reflexión del diagrama respecto a la primera diagonal.

# Vaca y persistencia extendida

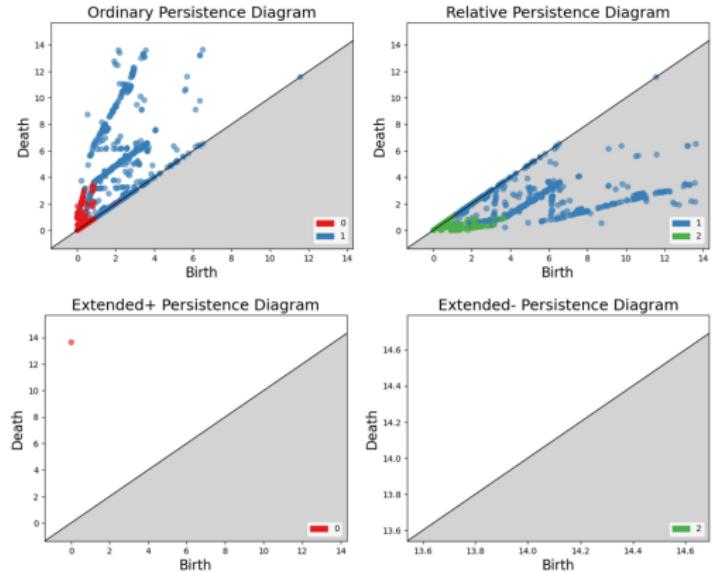
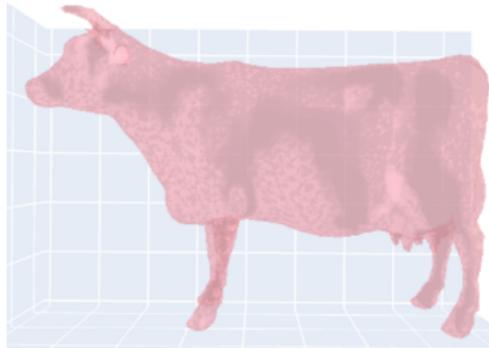


Figure: Observamos reflexiones entre  $\text{Ord}_p$  y  $\text{Rel}_{n-p}$ .

## ★ Diagramas de persistencia extendida de grafos de Reeb

Sean  $X$  un espacio topológico y  $f : X \rightarrow \mathbb{R}$ . Para todos  $x, y \in X$ ,

$x \sim y \iff f(x) = f(y)$  y  $x, y$  pertenecen a la misma CC. de  $f^{-1}(f(x))$   
y el **grafo de Reeb** is  $R_f(X) := X / \sim$ .

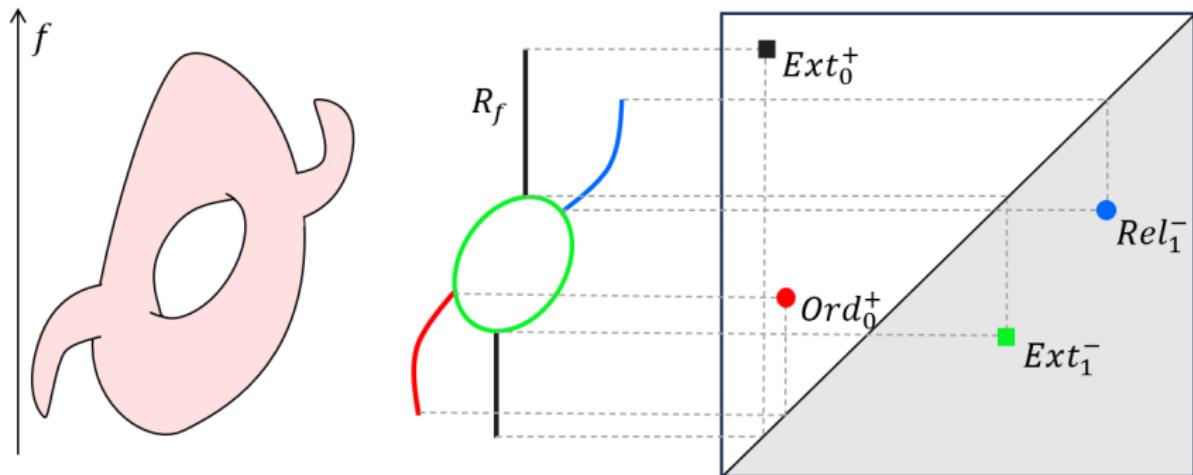


Figure: **Ord**: aparece/muere en subniveles, **Rel**: aparece/muere en superniveles y **Ext**: aparece en subniveles, muere en superniveles.  
(Carrière, Oudot) [2015]

# Referencias

1. S. Oudot, *Persistence Theory: From Quiver Representations to Data Analysis*, AMS (2015).
2. H. Edelsbrunner, J. Harer, *Computational Topology*, AMS Press. (2009)
3. D. Cohen-Steiner, H. Edelsbrunner, J. Harer, *Extending Persistence Using Poincaré and Lefschetz Duality*, Found. Comput. Math. (2009)
5. A. Hatcher, *Algebraic topology*, Cambridge Univ. Press (2001)
6. M. Carrière, S. Oudot, *Structure and Stability of the 1-Dimensional Mapper* (2015).

**GUDHI**, para hacer TDA en C++ y Python <https://gudhi.inria.fr>

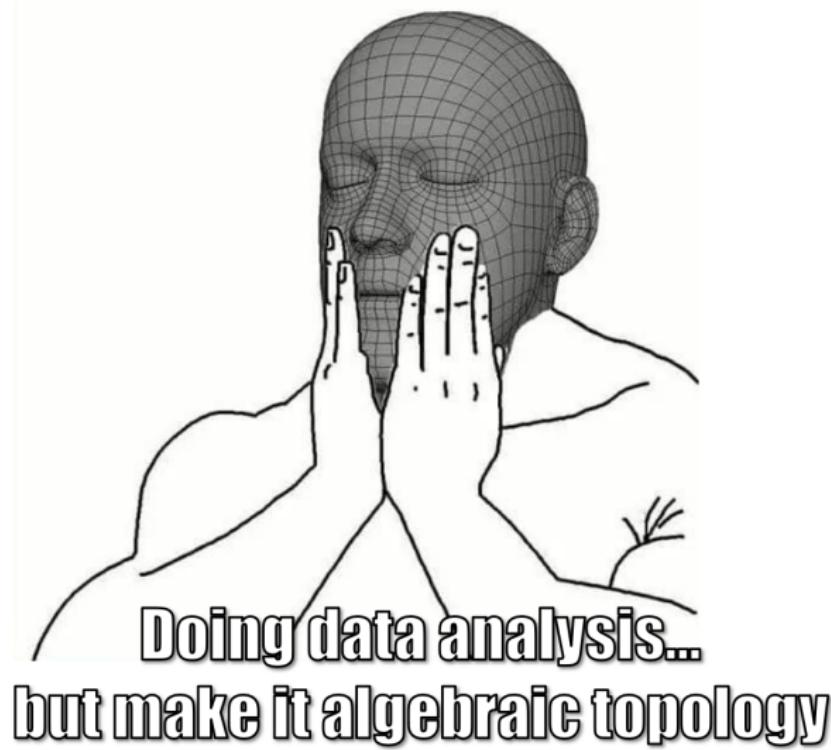
¿Preguntas?



**Trastorno por  
Déficit de  
Atención**

**Topological  
Data Analysis**

¿Otras preguntas?



## ★ Modulos de persistencia y haces

Sea  $(T, \leq)$  un conjunto parcialmente ordenado (poset).

Teorema (Kashiwara-Schapira) [2018]

Existe una equivalencia de categorías

$$\text{Pers}(T, k) \simeq_{\text{cat}} \text{Sch}_X(k)$$

donde  $X = (T^{\text{op}}, \text{Alexandrov})$ .

$$\star \partial X = \{x \in X ; H_n(X, X \setminus x) \cong 0\}$$

Sea  $X$  una variedad de dimensión  $n$  con borde  $\partial X$  y  $x \in X \setminus \partial X$  tal que  $x \in U \cong V$  un abierto de  $\mathbb{R}^n$ . Por el **lema de escisión** tenemos  $\forall i$ ,

$$H_i(X, X \setminus x) \cong H_i(\mathbb{R}^n, \mathbb{R}^n \setminus 0),$$

Y por la **sucesión exacta larga de homología relativa** tenemos,

$$\longrightarrow \tilde{H}_i(\mathbb{R}^n) \longrightarrow \tilde{H}_i(\mathbb{R}^n, \mathbb{R}^n \setminus 0) \xrightarrow{\delta} \tilde{H}_{i-1}(\mathbb{R}^n \setminus 0) \longrightarrow \tilde{H}_{i-1}(\mathbb{R}^n) \longrightarrow$$

Y como  $\mathbb{R}^n$  es contráctil,

$$\tilde{H}_i(X, X \setminus x) \cong \tilde{H}_{i-1}(\mathbb{R}^n \setminus 0) \cong \tilde{H}_{i-1}(\mathbb{S}^{n-1}) \cong \mathbb{Z} \text{ si } i = n \text{ y } 0 \text{ si no.}$$

Y con  $x \in \partial X$  de misma manera tenemos,

$$\tilde{H}_i(X, X \setminus x) \cong \tilde{H}_i(\mathbb{R}_+^n, \mathbb{R}_+^n \setminus 0) \cong \tilde{H}_{i-1}(\mathbb{R}_+^n \setminus 0) \cong 0$$

$$\star H^i(X, \mathbb{Z}_2) \cong H_i(X, \mathbb{Z}_2)$$

Por el teorema de los coeficientes universales para la cohomología tenemos que,

$$\begin{aligned} H^i(X, \mathbb{Z}_2) &\cong \text{Hom}_{\mathbb{Z}_2}(H_i(X, \mathbb{Z}_2), \mathbb{Z}_2) \oplus \text{Ext}_{\mathbb{Z}_2}^1(H_{i-1}(X, \mathbb{Z}_2), \mathbb{Z}_2) \\ &\cong \text{Hom}_{\mathbb{Z}_2}(\mathbb{Z}_2^{\beta_i(X, \mathbb{Z}_2)}, \mathbb{Z}_2) \quad (\text{$H_{i-1}$ libre}) \\ &\cong \text{Hom}_{\mathbb{Z}_2}(\mathbb{Z}_2, \mathbb{Z}_2)^{\beta_i(X, \mathbb{Z}_2)} \\ &\cong \mathbb{Z}_2^{\beta_i(X, \mathbb{Z}_2)} \cong H_i(X, \mathbb{Z}_2) \end{aligned}$$