

UNIVERSIDAD DE SALAMANCA

Facultad de Ciencias

GRADO EN MATEMÁTICAS

TÍTULO

MÓDULOS DE PERSISTENCIA EN EL ANÁLISIS
TOPOLÓGICO DE DATOS

TRABAJO DE FIN DE GRADO

realizado por

ALBERTO MORENO DE VEGA GARCÍA

bajo la tutela de:

DARÍO SÁNCHEZ GÓMEZ

(Departamento de Matemáticas)



VNiVERSiDAD
D SALAMANCA

CAMPUS DE EXCELENCIA INTERNACIONAL



Salamanca, 6 de julio de 2022

MÓDULOS DE PERSISTENCIA EN EL ANÁLISIS TOPOLÓGICO DE DATOS

Trabajo de fin de Grado presentado por Alberto Moreno de Vega García

AUTOR: Alberto Moreno de Vega García



TUTOR: Darío Sánchez Gómez

Salamanca, a 6 de julio de 2022.

Índice general

Introducción	1
1. De los símlices a la homología persistente	3
1.1. Complejos simpliciales geométricos	3
1.2. Complejos simpliciales abstractos	5
1.3. Complejos de cadenas y homología simplicial	7
1.4. Construcción de complejos simpliciales en el análisis topológico de datos . .	10
1.5. Homología persistente	13
2. Módulos de persistencia	15
2.1. Teoría de categorías	15
2.2. Objetos y módulos de persistencia	20
2.3. Módulos graduados	22
2.4. Módulos de persistencia como módulos sobre un anillo de polinomios . . .	25
2.5. Clasificación de módulos de persistencia por módulos de intervalo	31
2.6. Clasificación de módulos de persistencia sobre otros posets	34
2.7. Representaciones de los módulos de persistencia	37
3. Estabilidad de los módulos de persistencia	39
3.1. Intercalados generalizados	40
3.2. Distancia de cuello de botella	46
3.3. Teorema de estabilidad de las distancias	51
Conclusiones	54
Bibliografía	57

Índice de figuras

1.1. Nervio de la circunferencia	11
1.2. Filtración de un complejo de Vietoris-Rips	12
1.3. Conjunto de puntos contenidos en una circunferencia	13
2.1. Código de barras y diagrama de persistencia	38
3.1. Nubes de datos cercanas	39
3.2. Distancia cuello de botella de un código de barras	48

Introducción

En los últimos 50 años el desarrollo de la computación y el *machine learning* ha sido más que evidente. Entre otras razones debido a los avances en la microelectrónica y a la creciente disponibilidad de bases de datos con las que trabajar. En este contexto surge el análisis topológico de datos (o TDA por sus siglas en inglés) que se basa en la idea de que la estructura interna de los datos puede ser topológica y busca formas de estimarla a través de la computación. Como prueba de su utilidad el TDA se ha extendido por diversas áreas de la ciencia y las matemáticas como la ingeniería de materiales [30], la medicina [31], el análisis de series temporales [15] o las finanzas [20] y [21].

Aunque sus orígenes como idea surgen anteriormente, el concepto de persistencia, en la forma en que esta memoria lo trata, aparece formulado por primera vez en la tesis de Afra Zomorodian [36]. Este enfoque trata de medir la aparición, evolución y desaparición de las propiedades homológicas de una filtración de complejos simpliciales. Esta noción de persistencia ocupa el primer capítulo de esta memoria.

Comenzaremos con una nube de datos, presumiblemente en \mathbb{R}^n . Estos puntos serán los vértices del complejo simplicial de Vietoris-Rips que se forma al tomar bolas abiertas centradas en los puntos. El problema es que no existe una elección para el radio de estas bolas que sea más correcta que las demás, por lo que el radio ϵ se usa como parámetro indexando una filtración de complejos simpliciales. Utilizando complejos de cadenas construiremos la homología simplicial y seguidamente la homología persistente que estudia la variación de la homología simplicial de símplice a símplice.

La homología persistente es un caso particular de módulos de persistencia y en general de objetos de persistencia en una categoría. El estudio de este tipo de estructuras ocupará el segundo capítulo de la memoria, que comienza con un breve resumen de la teoría de categorías. La idea de tratar estos objetos como funtores esta inspirada en [5] y [6]. Aunque en el segundo capítulo pasa a un segundo plano para ver un enfoque más algebraico, en el capítulo 3, la teorías de categorías permite resultados muy generales. El capítulo 2 continúa con la identificación de los módulos de persistencia de tipo finito presentados y los módulos graduados sobre un anillo de polinomios para su clasificación como suma de módulos de intervalo, siguiendo [13] y [38]. Se dan también resultados de

clasificación más generales que por su extensión no serán desarrollados en su totalidad. Además de la búsqueda de la generalidad, el estudio de los módulos de persistencia sobre otros conjuntos ordenados tiene un fundamento en sus aplicaciones. A pesar de que las nubes de datos sean finitas, suelen ser consideradas como muestras estadísticas de un objeto o proceso continuo por lo que idealmente estas muestras serán una fiel aproximación del total [12]. El conjunto de los intervalos que clasifican un módulo de persistencia formará un invariante llamado código de barras y guardará información sobre el nacimiento y muerte de clases de homología en el caso de haberlo construido sobre complejos simpliciales.

Una vez conocida la estructura de los módulos de persistencia y su código de barras las técnicas de *machine learning* siguen siendo infructíferas por la falta de formas de comparar estos códigos de barras. Como solución a este inconveniente Carlsson definió la distancia de cuello de botella en [11]. Más adelante se define en [12] la distancia de *interleaving* o intercalado entre módulos de persistencia, que representa cómo de lejos están dos módulos de persistencia de ser isomorfos; y que se usa como herramienta para generalizar los resultados originales de estabilidad de [18]. El estudio de estas dos distancias y sus resultados de estabilidad ocupa el tercer capítulo de esta memoria. Se acabará conectando con la construcción original del análisis de datos mostrando cómo las variaciones en las nubes de puntos dan una cota superior de la distancia entre sus códigos de barras.

Capítulo 1

De los símplexes a la homología persistente

En este capítulo presentaremos brevemente algunas nociones básicas relacionadas con los complejos simpliciales, los cuales permiten modelizar problemas relacionados con el análisis topológico de datos. A la hora de modelizar los datos que se quieren estudiar se impondrá que se preserven ciertas propiedades topológicas relevantes para dicho estudio. La idea del análisis topológico de datos es que estas propiedades codifiquen información sobre la estructura de los datos. En general, se buscará que se preserve la homología del espacio topológico formado por la nube de puntos, pensados éstos como elementos de un espacio topológico finito. Así, la homología en grado 0 indicará el número de componentes conexas del espacio topológico y, en general, en grado n indicará el número de agujeros n -dimensionales que tiene el espacio.

1.1. Complejos simpliciales geométricos

Comenzamos viendo los complejos simpliciales desde un punto de vista geométrico, para ello partimos de $S = \{a_0, a_1, a_2, \dots, a_n\}$ un conjunto no vacío de puntos geoméricamente independientes en \mathbb{R}^N , es decir, los vectores $\{a_1 - a_0, \dots, a_n - a_0\}$ son linealmente independientes.

Definición 1.1. Llamaremos **n -símplex** generado por S , denotado por $\sigma = \langle a_0, a_1, \dots, a_n \rangle$, a la envolvente convexa de S , es decir, σ será el conjunto de los puntos de \mathbb{R}^N verificando:

$$x = \sum_{i=0}^n t_i a_i, \text{ donde } \sum_{i=0}^n t_i = 1 \text{ y } t_i \geq 0, \text{ para todo } i = 0, 1, \dots, n$$

Los puntos de S son los **vértices** de σ y n es la **dimensión** del símplex σ , denotándola por $\dim(\sigma) = n$.

Definición 1.2. Para cada subconjunto no vacío $\bar{S} = \{a_{i_1}, \dots, a_{i_k}\} \subseteq S$, el k -símplice τ que define se denomina **cara** de σ y lo denotaremos por $\tau \leq \sigma$. En particular si $\bar{S} \neq S$ la cara τ generada por \bar{S} se llama una **cara propia** y se denota por $\tau < \sigma$.

Definición 1.3. Un **complejo simplicial geométrico** K en \mathbb{R}^N es una familia finita de símplices en \mathbb{R}^N verificando las siguientes condiciones

(P1) Si $\sigma \in K$ y $\tau \leq \sigma$, entonces $\tau \in K$

(P2) Si $\sigma, \sigma' \in K$, entonces $\sigma \cap \sigma' = \emptyset$ ó $\sigma \cap \sigma' \leq \sigma$ y $\sigma \cap \sigma' \leq \sigma'$.

La **dimensión** de un complejo simplicial geométrico K es el máximo de las dimensiones de todos los símplices que forman K .

Definición 1.4. Una subfamilia de un complejo simplicial geométrico es un **subcomplejo simplicial** de K si es un complejo simplicial por sí mismo.

Un subcomplejo de particular interés es el formado por los símplices de dimensión menor o igual que s , llamado el **s-esqueleto** del complejo simplicial y que se denota por $K^{(s)}$. Los vértices del complejo simplicial K forman el conjunto $\text{Vert}(K) = K^{(0)}$.

Las nociones previas pueden extenderse para considerar un número infinito de símplices que a su vez podrían contener un número infinito de vértices. En este caso deberíamos hacer una distinción entre los complejos simpliciales finitos e infinitos. No obstante esta distinción no será reflejada en esta memoria y por complejo simplicial entenderemos siempre un complejo simplicial finito.

Definición 1.5. Sea K un complejo simplicial geométrico, llamaremos **poliedro asociado** a K a:

$$|K| := \bigcup_{\sigma \in K} \sigma.$$

Nótese que $|K|$ es un espacio topológico con la topología inducida como subespacio de \mathbb{R}^N y será por tanto Hausdorff y compacto.

Diremos que un espacio topológico X es **triangulable** si existe una pareja (K, f) formada por un complejo simplicial K y un homeomorfismo $f : |K| \rightarrow X$.

Sean K y L complejos simpliciales geométricos. Una aplicación $\phi : K^{(0)} \rightarrow L^{(0)}$ se llama **morfismo de vértices**.

Definición 1.6. Un **morfismo simplicial** $\phi : K \rightarrow L$ es un morfismo de vértices verificando que para todo conjunto $\{x_0, \dots, x_s\}$ que genera un símplice σ en K , el conjunto $\{\phi(x_0), \dots, \phi(x_s)\}$ son vértices de un símplice en L .

En particular, todo morfismo simplicial $\phi: K \rightarrow L$ induce una aplicación entre los correspondientes poliedros asociados:

$$g: |K| \rightarrow |L|$$

$$x = \sum_{i=0}^s \lambda_i x_i \mapsto \sum_{i=0}^s \lambda_i \phi(x_i)$$

la cual es continua ([29]). Además, si el morfismo de vértices es biyectivo, entonces $g: |K| \rightarrow |L|$ es un homomorfismo y decimos que K y L son isomorfos.

Se puede ver que la composición de morfismos simpliciales es un morfismo simplicial y que el morfismo 1_K originado por la identidad en K es también un morfismo simplicial.

1.2. Complejos simpliciales abstractos

Se presenta ahora un planteamiento puramente combinatorio de los complejos simpliciales, el cual es equivalente al geométrico pero que presenta ciertas ventajas con respecto a los complejos simpliciales geométricos.

Definición 1.7. Un **complejo simplicial abstracto** es una pareja $K = (V, \varphi)$ consistente en un conjunto finito no vacío V , llamado **conjunto de vértices**, junto con una familia φ de subconjuntos de V , llamada **conjunto de símlices abstractos**, verificando:

(P1) Todo $x \in V$ pertenece al menos a un símlice de φ .

(P2) Todo subconjunto τ de un símlice abstracto $\sigma \in \varphi$ también está en φ .

Si $\sigma \in \varphi$ es un símlice no vacío formado por $n+1$ vértices decimos que la dimensión de σ es n , denotándolo por $\dim(\sigma) = n$, y se dice que es un **n-símlice**. Cada uno de sus subconjuntos es una **cara** de σ . La **dimensión** de un complejo simplicial abstracto K es el máximo de las dimensiones de sus símlices abstractos.

Un **subcomplejo simplicial abstracto** Λ de K es un complejo simplicial abstracto para el cual todo símlice $\sigma \in \Lambda$ es también símlice de K .

Es común realizar un abuso de notación utilizando $\sigma \in K$ para $\sigma \in \varphi$ donde $K = (V, \varphi)$ es un complejo simplicial abstracto.

Al igual que ocurría en la versión geométrica existe una noción infinito dimensional para los complejos simpliciales abstractos, pero no responde a las necesidades de este trabajo.

Dados dos complejos simpliciales abstractos $K = (V, \varphi)$ y $L = (W, \vartheta)$, un **morfismo simplicial abstracto** $\phi: K \rightarrow L$ es una aplicación $\phi: V \rightarrow W$ tal que si $\{a_0, \dots, a_n\}$ está en φ , entonces el conjunto $\{\phi(a_0), \dots, \phi(a_n)\}$ está en ϑ . En particular dos complejos simpliciales abstractos serán **isomorfos** si la aplicación $\phi: V \rightarrow W$ es una correspondencia biyectiva.

Terminamos esta sección dando la equivalencia entre las dos nociones de complejo simplicial presentadas.

Definición 1.8. Sea K un complejo simplicial geométrico. Definimos el **esquema de vértices** de K , denotado $\mathcal{A}(K)$, como la colección de subconjuntos $\{a_0, \dots, a_n\}$ de $K^{(0)}$ para las cuales los vértices $\{a_0, \dots, a_n\}$ generan un símplice σ de K .

Por la definición de complejo simplicial geométrico $\mathcal{A}(K)$ genera un complejo simplicial abstracto junto al conjunto de vértices $K^{(0)}$. Con esto todo complejo simplicial geométrico da lugar a un complejo simplicial abstracto.

La asignación inversa también existe.

Definición 1.9. Una **realización geométrica** de un complejo simplicial $K = (V, \varphi)$ abstracto es un complejo simplicial geométrico K' en \mathbb{R}^N , para algún N , tal que K y $\mathcal{A}(K')$ son isomorfos como complejos simpliciales abstractos.

Proposición 1.10. Para todo complejo simplicial abstracto $K = (V, \varphi)$ existe una realización geométrica K . Además, dos complejos simpliciales geométricos K y L son isomorfos si y sólo si $\mathcal{A}(K)$ y $\mathcal{A}(L)$ son isomorfos como complejos simpliciales abstractos.

Demostración. Sea $V = \{x_1, \dots, x_s\}$ el conjunto de vértices del complejo simplicial K y $\{a_i\}_{i=0}^s$ un conjunto de puntos geoméricamente independientes de \mathbb{R}^N , para $N \geq s$. Definimos la aplicación inyectiva:

$$\begin{aligned} \theta: V &\rightarrow \mathbb{R}^N \\ x_i &\mapsto \theta(x_i) := a_i \end{aligned}$$

Definimos el complejo simplicial geométrico K' como:

$$\langle \theta(x_1), \dots, \theta(x_n) \rangle \in K' \iff \{x_1, \dots, x_n\} \in \varphi.$$

Por las propiedades de K' es claro que es un complejo simplicial geométrico y que es isomorfo a $\mathcal{A}(K)$ luego es una realización geométrica. \square

Por esta equivalencia a partir de ahora nos referiremos a un complejo simplicial abstracto simplemente como complejo simplicial.

Sea un n -símlice $\sigma = \{x_0, \dots, x_n\}$, podemos definir en él una **orientación**. Para esto necesitamos dar una relación de equivalencia en el conjunto de todas las ordenaciones posibles de su conjunto de vértices. Para una permutación π en dicho conjunto diremos que dos conjuntos de vértices $\{x_0, x_1, \dots, x_n\}$ y $\{x_{\pi(0)}, x_{\pi(1)}, \dots, x_{\pi(n)}\}$ están relacionados si y sólo si $\pi: \{0, 1, 2, \dots, n\} \rightarrow \{0, 1, 2, \dots, n\}$ es una permutación par. Esta relación de equivalencia define dos clases de equivalencia en el conjunto de todas las posibles ordenaciones de los vértices llamadas **orientaciones**. Un símlice orientado $\sigma = [x_0, x_1, \dots, x_n]$ es un símlice para el cual se ha escogido una de las dos posibles orientaciones. Utilizamos $-\sigma$ para representar ese mismo símlice dotado de la orientación opuesta. Un **complejo simplicial orientado** es un complejo simplicial K para el cual se ha escogido una orientación para todos sus símlices. Claramente cuando un símlice está formado por un sólo vértice $\{x\}$ existe una única orientación posible.

1.3. Complejos de cadenas y homología simplicial

En esta sección definiremos los complejos de cadenas y veremos la noción de homología de un complejo simplicial, construida sobre un complejo de cadenas. Es común encontrar este desarrollo trabajando simplemente con grupos abelianos, no obstante mostraremos un enfoque más general trabajando sobre un anillo (conmutativo y con unidad) y formando por tanto estructuras de R -módulo en vez de grupo. Si se toma $\mathbb{Z} = R$, tenemos el desarrollo con grupos abelianos. Tanto ahora como en los siguientes capítulos de este trabajo se usarán conceptos básicos de álgebra conmutativa los cuales no serán desarrollados, como referencia puede mirarse [1].

Definición 1.11. Un **complejo de cadenas** es una pareja $\{C, \delta\}$ formada por una familia $C = \{C_n\}_{n \in \mathbb{Z}}$ de módulos sobre un anillo R y una familia de morfismos de R -módulos $\delta = \{\delta_n: C_n \rightarrow C_{n-1}\}_{n \in \mathbb{Z}}$ verificando que $Im(\delta_{n+1}) \subseteq Ker(\delta_n) \forall n \in \mathbb{Z}$.

Los elementos de C_n se denominan **n -cadenas** y llamaremos a δ_n el **operador de bordes de dimensión n -ésima**. El submódulo $Ker(\delta_n)$ se llama módulo de **n -ciclos** y lo denotaremos por $Z_n(C)$. El submódulo $Im(\delta_{n+1})$ se llama módulo de **n -bordes** y se denota $B_n(C)$. Por la propiedad del operador de bordes tenemos que, para todo $n \in \mathbb{Z}$, el R -módulo cociente:

$$H_n(C, R) = Ker(\delta_n) / Im(\delta_{n+1})$$

está bien definido y lo llamaremos **R-módulo de homología de dimensión n** de C . Para cada $n \in \mathbb{Z}$ el módulo de homología de dimensión n de C tendrá una parte libre y una parte de torsión. El rango de su parte libre se denomina **n-ésimo número de Betti** y se denota β_n .

Definición 1.12. Sean C y C' complejos de cadenas. Un **morfismo de complejos de cadenas** es una familia $f = \{f_n: C_n \rightarrow C'_n\}_{n \in \mathbb{Z}}$ de morfismos de R -módulos con la propiedad de que $\forall n \in \mathbb{Z}$ el siguiente diagrama es conmutativo:

$$\begin{array}{ccc} C_n & \xrightarrow{f_n} & C'_n \\ \downarrow \delta_n & & \downarrow \delta'_n \\ C_{n-1} & \xrightarrow{f_{n-1}} & C'_{n-1} \end{array}$$

Se puede comprobar que 1_C , la aplicación identidad de C , es un morfismo de complejos de cadenas. Además si C, C', \bar{C} son complejos de cadenas y $f: C \rightarrow C', g: C' \rightarrow \bar{C}$ son morfismos de complejos de cadenas, $(g \circ f): C \rightarrow \bar{C}$ es también un morfismo de complejos de cadenas.

Los morfismos de complejos de cadenas inducen morfismos entre los R -módulos de homología de las cadenas.

Definición 1.13. Sea $f: C \rightarrow C'$ un morfismo de complejos de cadenas. El morfismo de dimensión n inducido por f en el módulo de homología de dimensión n es:

$$\begin{aligned} H_n(f): H_n(C) &\rightarrow H_n(C') \\ [z] &\mapsto [f_n(z)] \end{aligned}$$

Este morfismo está bien definido porque $f_n(Z_n(C)) \subseteq Z_n(C')$ y $f_n(B_n(C)) \subseteq B_n(C')$ $\forall n \in \mathbb{Z}$. Además, se puede comprobar de forma inmediata que:

1. $H_n(1_C) = 1_{H_n(C)}$
2. $H_n(g \circ f) = H_n(g) \circ H_n(f)$ para cualquier morfismo de complejos de cadena f, g .

Veamos ahora cómo se puede asignar a cada complejo simplicial orientado un complejo de cadenas. Sea K un complejo simplicial orientado. Para cada $n \geq 0$ el R -módulo de n -cadenas es el módulo de combinaciones lineales formales de los elementos de dimensión n en K . Esto es, el módulo libre generado por los n -símplices de K . Se denotará por $C_n(K, R)$, aunque cuando no dé lugar a confusión la elección del anillo escribiremos simplemente $C_n(K)$. Un elemento $c \in C_n(K)$ es de la forma $c = \sum_i \lambda_i \sigma_i$, donde los σ_i 's son n -símplices de K .

Para cada n definimos la aplicación:

$$\begin{aligned} \delta_n: C_n &\rightarrow C_{n-1} \\ \sigma &\mapsto \sum_{i=0}^n (-1)^i [v_0, v_1, \dots, \hat{v}_i, \dots, v_n] \end{aligned}$$

y la extendemos por linealidad a todos los elementos de $C_n(K)$. La familia $\delta := \{\delta_n\}_{n \in \mathbb{Z}}$ verifica:

Proposición 1.14. $\delta_{n-1} \circ \delta_n = 0$.

Demostración. Por la linealidad de δ_n es suficiente probarlo sobre una n -cadena $[v_0, v_1, \dots, v_n]$.

$$\begin{aligned} \delta_{n-1} \circ \delta_n([v_0, v_1, \dots, v_n]) &= \sum_{i=0}^n (-1)^i \delta_{n-1}([v_0, v_1, \dots, \hat{v}_i, \dots, v_n]) \\ &= \sum_{i=0}^n (-1)^j \left(\sum_{j < i} (-1)^i [v_0, v_1, \dots, \hat{v}_j, \dots, \hat{v}_i, \dots, v_n] \right. \\ &\quad \left. + \sum_{j > i} (-1)^{j-1} [v_0, v_1, \dots, \hat{v}_i, \dots, \hat{v}_j, \dots, v_n] \right), \end{aligned}$$

de donde se concluye por la relación que aparece entre el signo de $(-1)^i$, $(-1)^{j-1}$ y la orientación de los símlices. \square

Con este resultado, usando la notación anterior, tenemos que para todo complejo simplicial orientado K y anillo R , el complejo de cadenas $C(K) = \{C_n(K, R), \delta_n\}$ llamado **complejo de cadenas simpliciales asociado a K** . En particular, para cada número natural tenemos el módulo cociente:

$$H_n^{simp}(C(K)) = Ker(\delta_n) / Im(\delta_{n+1})$$

llamado **R -módulo de homología simplicial n -dimensional** del complejo simplicial K , y que será denotado $H_n^{simp}(K)$.

Por otro lado, toda aplicación simplicial $\phi: K \rightarrow L$ induce un morfismo de complejos de cadenas simpliciales $\phi_{\#}: C(K) \rightarrow C(L)$ extendiendo linealmente, para cada número natural n , el morfismo: $(\phi_{\#})_n: C_n(K) \rightarrow C_n(L)$ definido por:

$$(\phi_{\#})_n[v_0, \dots, v_n] := \begin{cases} [\phi(v_0), \dots, \phi(v_n)] & \text{if } \phi(v_i) \neq \phi(v_j), \text{ para todo } i \neq j \\ 0 & \text{en los demás casos} \end{cases}$$

Se puede comprobar fácilmente que la aplicación identidad en el complejo simplicial da lugar al morfismo identidad en la cadena simplicial y que la composición de aplicaciones simpliciales se preserva componiendo los dos morfismos de cadenas simpliciales inducidos.

Juntando esto con el morfismo inducido entre los módulos de homología (Definición 1.13) tenemos que todo morfismo simplicial da lugar a un morfismo de R -módulos entre los R -módulos de homología componiendo los dos operadores:

$$H_n(\phi) := H_n(\phi_{\#}): H_n(K) \rightarrow H_n(L) \quad (1.1)$$

Observación 1.15. Este morfismo cumple de nuevo que $H_n(\phi \circ \varphi) = H_n(\phi) \circ H_n(\varphi)$ y $H_n(Id_K) = Id_{H_n(K)}$ \triangle

1.4. Construcción de complejos simpliciales en el análisis topológico de datos

La construcción de complejos simpliciales asociados a una nube de datos tiene como objetivo que la homología del espacio topológico sobre el que se encuentran los datos sea computable. Para esto será necesario, por tanto, que la construcción de los complejos simpliciales preserve la homología y además sean lo suficientemente sencillos como para poder ser computados de forma eficiente. En el análisis topológico de datos la forma de conseguir esto es construir complejos simpliciales que sean homotópicamente equivalentes al espacio X formado por el conjunto finito de puntos. Para una explicación más detallada sobre la homotopía en el TDA ver [32]. Existe variedad de complejos con estas propiedades como los complejos de banderas o los complejos alpha, sin embargo el enfoque más usual es utilizar el complejo de Čech. Aunque este es técnicamente computable, el número de símlices que genera hace su computación extremadamente costosa. Como solución se busca una aproximación al complejo de Čech que tenga ventajas computacionales con respecto a este. Para esto se utiliza el complejo de **Vitoris-Rips**. Sobre ellos se construirá la homología del complejo, que será la misma que la homología del espacio X por ser un invariante homotópico.

Sea X un espacio topológico e I un conjunto. Sea $C = \{U_i\}_{i \in I}$ un recubrimiento por abiertos. Definimos el **nervio del recubrimiento** C como:

$$N(C) := \left\{ J \subseteq I \text{ con } J \text{ un conjunto finito tal que: } \bigcap_{j \in J} U_j \neq \emptyset \right\}.$$

Se comprueba sin dificultad que el nervio $N(C)$ forma un complejo simplicial pues si $J = \{j_1, \dots, j_k\} \in N(C)$, se cumple que $\bigcap_{j \in J} U_j \neq \emptyset$ luego $\bigcap_{j \in \bar{J}} U_j \neq \emptyset$ para cualquier $1 \leq i \leq k$ y $\bar{J} = \{j_1, \dots, \hat{j}_i, \dots, j_k\}$.

Teorema 1.16 (Teorema del Nervio [32]). *Sea X un espacio topológico y sea \mathcal{C} un recubrimiento por abiertos, tal que la intersección de cualquier subconjunto de abiertos del recubrimiento es vacía o contractible. Entonces X y $N(\mathcal{C})$ son homotópicamente equivalentes.*

Sea (X, d) un espacio métrico compacto y sea \mathcal{U}_ϵ el recubrimiento del espacio topológico X dado por bolas centradas en los puntos de un subconjunto finito S y de radio ϵ .

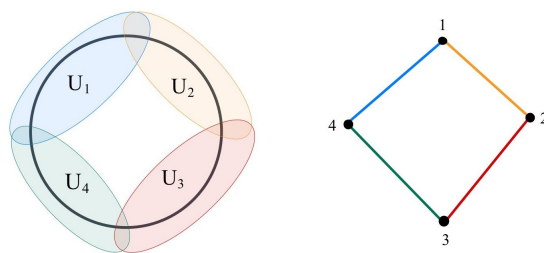


Figura 1.1: Recubrimiento de una circunferencia y el nervio del recubrimiento dado por los U_i .

Definición 1.17. El **complejo de Čech** asociado a S del espacio X de parámetro ϵ , denotado $\check{C}_\epsilon(S)$, es el nervio del recubrimiento \mathcal{U}_ϵ .

En el caso de que X sea un espacio métrico finito se toma $S = X$ y se denota $\check{C}_\epsilon(X)$. El complejo de Čech es homotópicamente equivalente al espacio X , por el **teorema del nervio**, ya que la intersección de dos bolas es o bien vacía o bien convexa luego contráctil.

Definición 1.18. Para un espacio métrico finito (X, d) el **complejo de Vietoris-Rips de parámetro $\epsilon > 0$ asociado a X** es el complejo simplicial con conjunto de vértices X y símplexes:

$$\left\{ \{x_0, x_1, \dots, x_n\} \in X \text{ tales que } d(x_i, x_j) \leq \epsilon \text{ para todo } 0 \leq i < j \leq n \right\}.$$

Se denota por $\text{VR}_\epsilon(X)$.

La siguiente definición es necesaria para la construcción de la homología persistente. Recordemos que una relación de **orden parcial** en un conjunto es una relación binaria, reflexiva, antisimétrica y transitiva. Si además toda pareja no ordenada de elementos del conjunto está en la relación, se dice que es una relación de **orden total**. Un conjunto dotado de una relación de orden parcial (respectivamente total) es un conjunto parcialmente (respectivamente totalmente) ordenado, al que llamaremos poset para abreviar.

Definición 1.19. Sea K un complejo simplicial y J un conjunto totalmente ordenado. Una **filtración** de K es una sucesión de subcomplejos simpliciales $\mathcal{F} = \{K_x\}_{x \in J}$ anidada, es decir $K_x \subseteq K_y$ si $x \leq y$, verificando que $\emptyset, K \in \mathcal{F}$. Como estamos trabajando con la versión finita de los complejos simpliciales puede describirse como una filtración finita:

$$\emptyset = K_0 \subset K_1 \subset K_2 \subset \cdots \subset K_n = K.$$

donde las inclusiones son estrictas. Un complejo simplicial K con una filtración dada se llama **complejo filtrado**.

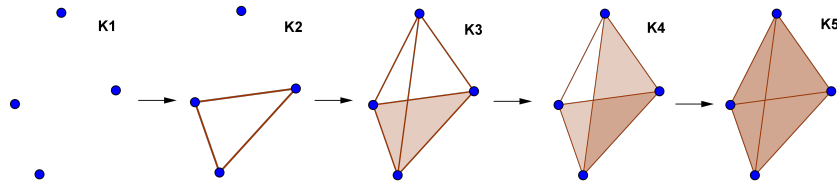


Figura 1.2: Filtración de un complejo de Vietoris-Rips sobre un espacio métrico de 4 puntos.

Nótese que tanto en el complejo de Čech como en el complejo de Vietoris-Rips cuando aumentamos el radio tenemos, para todo $\delta > 0$, las inclusiones de complejos simpliciales: $VR_\epsilon(X) \subseteq VR_{\epsilon+\delta}(X)$ y $\check{C}_\epsilon(X) \subseteq \check{C}_{\epsilon+\delta}(X)$. Luego, con los morfismos de inclusión, esto nos da filtraciones de complejos simpliciales a medida que aumentamos el radio de las bolas a los que denotaremos $VR(X) = \{VR_\epsilon(X)\}_{\epsilon \in \mathbb{R}}$ y $\mathcal{C}(X) = \{\check{C}_\epsilon(X)\}_{\epsilon \in \mathbb{R}}$. Para $\epsilon = 0$ se fija el convenio $VR_\epsilon(X) = \check{C}_\epsilon(X) = \emptyset$. Por ejemplo, en la figura 1.2 sean los vértices a_1, a_2, a_3, a_4 numerados en sentido antihorario empezando por el que se encuentra más abajo. Tenemos que

$$K2 = \{a_1, a_2, a_3, a_4, a_1a_2, a_1a_4, a_2a_4\}$$

y al aumentar el radio tenemos el nuevo complejo

$$K3 = \{a_1, a_2, a_3, a_4, a_1a_2, a_1a_3, a_1a_4, a_2a_3, a_2a_4, a_3a_4, a_1a_2a_4\}.$$

Como se decía al principio de la sección el complejo de Vietoris-Rips tiene ventajas computacionales con respecto al de Čech. El Teorema del Nervio únicamente muestra la equivalencia homotópica entre este último y el espacio topológico que triangula pero existe la siguiente inclusión entre los complejos de Čech y Vietoris-Rips que permiten aproximar uno con el otro de forma arbitrariamente precisa.

Proposición 1.20. ([8, Proposition 2.6]) Se tienen las inclusiones:

$$\check{C}_\epsilon(X) \subseteq VR_{2\epsilon}(X) \subseteq \check{C}_{2\epsilon}(X)$$

1.5. Homología persistente

Después de ver cómo obtener la homología de un complejo simplicial en la sección anterior la cuestión que debe resolverse es cuál es el radio de las bolas que debe tomarse, de manera que su elección separe las verdaderas propiedades del espacio subyacente X del ruido. Si tomamos un radio lo suficientemente pequeño ninguna bola tendrá intersección con las demás y por tanto tendremos que el complejo simplicial formado no tiene símlices más allá de los vértices, y cada punto de la nube será una componente conexa. Si por el contrario el radio tomado es más grande que un cierto valor umbral, todos los puntos estarán conectados entre sí. Un ejemplo de ello es la figura 1.3.

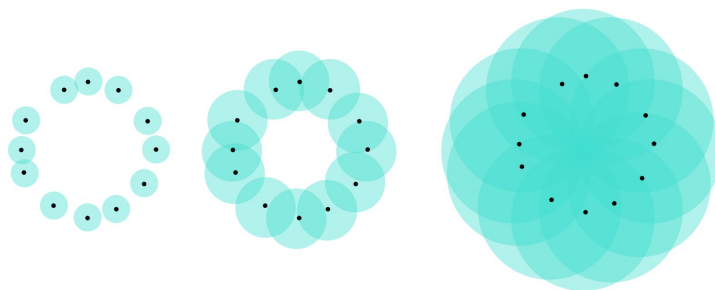


Figura 1.3: Interección de bolas abiertas para la construcción de un complejo simplicial. A simple vista puede esperarse que los puntos de la nube están contenidos en una circunferencia de \mathbb{R}^2 como de hecho se recoge al aumentar el radio ligeramente pues deja justo un agujero en medio. Cuando seguimos aumentando el radio no obstante esta información se pierde y no aparecen agujeros en la nube.

El procedimiento con el cual el análisis topológico de datos ha solucionado este inconveniente es no tomar ninguna elección concreta del radio de las bolas sino toda una familia ordenada de valores para esos radios. Así se observarán algunas características que se mantienen durante largos intervalos en esta familia de valores y otros cuya vida es muy corta. Aquellas características que persisten se considerarán que son relevantes mientras que las que no persistan en al variar el radio serán ruido y por tanto irrelevantes ([8]).

Sea una filtración finita de un complejo simplicial:

$$\emptyset = K_0 \subseteq K_1 \subseteq K_2 \subseteq \cdots \subseteq K_n = K$$

la cual es equivalente a una familia de complejos simpliciales y morfismos de complejos simpliciales dados por la inclusión natural $\iota_{i,j}: K_i \rightarrow K_j$ para todo $0 \leq i \leq j \leq n$. El ejemplo básico en el TDA será el de una filtración de una familia de complejos de Vietoris-Rips construida sobre una nube de puntos. Por el morfismo definido en 1.1 esto induce

una familia de morfismos entre los R -módulos de homología $f_p^{i,j}: H_p(K_i) \rightarrow H_p(K_j)$, con las propiedades siguientes

$$(P1) \quad f_p^{i,i} = 1_{H_p(K_i)}: H_p(K_i) \rightarrow H_p(K_i).$$

$$(P2) \quad f_p^{i,l} = f_p^{j,l} \circ f_p^{i,j}: H_p(K_i) \rightarrow H_p(K_l) \text{ for } i \leq j \leq l.$$

Tomando $f_p^i := f_p^{i,i+1}$, esta familia de morfismos está totalmente definida por:

$$0 \longrightarrow H_p(K_0) \xrightarrow{f_p^0} H_p(K_1) \xrightarrow{f_p^1} \dots \xrightarrow{f_p^{n-1}} H_p(K_n) = H_p(K)$$

Definición 1.21. Para una filtración \mathcal{F} , sean $p \geq 0$ y $0 \leq i \leq j \leq n$. El (i, j) -ésimo R -módulo de homología persistente a nivel p es el módulo:

$$H_p^{i,j}(\mathcal{F}) := \text{Im}(f_p^{i,j}).$$

El rango de este módulo se llama (i, j) -número de Betti a nivel p de \mathcal{F} .

Definición 1.22. Sea \mathcal{F} una filtración y sea $\alpha \in H_p(K_i)$ un elemento distinto del 0, decimos que α nace en K_i si $\alpha \notin H_p^{i-1,i}(\mathcal{F})$. Decimos que α muere en K_j si se cumple

$$f_p^{i,j-1}(\alpha) \notin H_p^{i-1,j-1}(\mathcal{F}) \text{ y } f_p^{i,j}(\alpha) \in H_p^{i-1,j}(\mathcal{F})$$

Es decir, el nacimiento mide la primera aparición de la clase y la muerte cuándo ha dejado de existir. Se llama **índice de persistencia** de α a $\text{pers}(\alpha) = j - i$. En el caso de que nunca muera una clase tenemos que $\text{pers}(\alpha) = \infty$. Desde un punto de vista topológico estos valores de muerte y nacimiento de una clase en el módulo de homología nos dicen cuándo aparece por primera vez y cuándo desaparece una cierta propiedad topológica que es representada por una clase en el módulo de homología.

Observación 1.23. Existen métodos para el cálculo de la homología de un complejo simplicial utilizando normalmente la forma normal de Smith. En base a ellos se ha desarrollado *software* con estos fines como *Simplicial Homology* de *GAP* y para el caso específico de su uso en el TDA paquetes de *Python*, *R* o *C++* disponibles en [4]. Dada su eficacia para el análisis de grandes bases de datos existen incluso programas comerciales como *Iris* de Ayasdi. △

Capítulo 2

Módulos de persistencia

Tras haber visto en el primer capítulo la noción de persistencia y su relevancia en el análisis topológico de datos estudiaremos ahora los objetos algebraicos que generalizan dicha noción, los módulos de persistencia. Este desarrollo se realizará utilizando el lenguaje y el formalismo de la teoría de categorías. Por esto se incluirá un breve repaso de algunas nociones fundamentales de dicha teoría, necesarias para el desarrollo de los capítulos 2 y 3 de este trabajo. Una de las contribuciones de la categorización de la persistencia es que permitirá la clasificación de los módulos de persistencia sobre el poset de los números naturales, con unas ciertas condiciones de finitud, como sumas de los llamados módulos de intervalo, gracias a que podremos encontrar una equivalencia categorial con la estructura, más familiar, de módulo graduado (Sección 2.5). También se utilizó la categorización de la homología persistente para los resultados de la Sección 2.6.

2.1. Teoría de categorías

Empezamos presentando algunas nociones básicas de la teoría de categorías. Para más detalles ver [26].

Definición 2.1. Una **categoría** \mathcal{C} es una terna formada por:

1. Una clase de **objetos**, denotada $Obj(\mathcal{C})$.
2. Una colección de **morfismos**, $Morf(\mathcal{C})$, formada por un conjunto $Hom_{\mathcal{C}}(X, Y)$ para cada par de objetos $X, Y \in Obj(\mathcal{C})$. Es común utilizar la notación $X \xrightarrow{f} Y$ cuando $f \in Hom_{\mathcal{C}}(X, Y)$.
3. Una **composición** de morfismos: para cada tres objetos $X, Y, Z \in Obj(\mathcal{C})$ existe

una aplicación:

$$\begin{aligned} \text{Hom}_{\mathcal{C}}(Y, Z) \times \text{Hom}_{\mathcal{C}}(X, Y) &\longrightarrow \text{Hom}_{\mathcal{C}}(X, Z) \\ (g, f) &\longmapsto gf \end{aligned}$$

llamada **composición**, verificando:

a) Es asociativa: $(hf)g = h(fg)$ para todos los morfismos

$$X \xrightarrow{g} Y \xrightarrow{f} Z \xrightarrow{h} T.$$

b) Para cada objeto $X \in \text{Obj}(\mathcal{C})$ existe un morfismo $Id_X: X \rightarrow X$, llamado morfismo identidad, tal que $Id_X f = f$ y $g Id_X = g$ sobre morfismos $f: X \rightarrow Y$ y $g: W \rightarrow X$.

Dos objetos en una categoría son **isomorfos** si existen elementos $f \in \text{Hom}_{\mathcal{C}}(X, Y)$ y $g \in \text{Hom}_{\mathcal{C}}(Y, X)$ tales que $gf = Id_X$ y $fg = Id_Y$.

Ejemplo 2.2. a) Todo conjunto parcialmente ordenado o **poset** (\mathcal{P}, \leq) , puede ser visto como una categoría a la que denotaremos simplemente \mathcal{P} , cuyo conjunto de objetos son los elementos en (\mathcal{P}, \leq) y cuyos morfismos son $\text{Hom}_{\mathcal{P}}(x, y) = \begin{cases} \{\cdot\} & \text{si } x \leq y \\ \emptyset & \text{si } x > y \end{cases}$.

- b) Tenemos la categoría cuyos objetos son los grupos y cuyos morfismos son los morfismos de grupos a la que se denota por **Grp**.
- c) Para un anillo R tenemos la categoría **Mod $_R$** cuyos objetos son los módulos sobre el anillo y sus morfismos son los morfismos de R -módulos. Tomando R un cuerpo \mathbb{k} tenemos la categoría de espacio vectoriales sobre \mathbb{k} , **Vect $_{\mathbb{k}}$** . Cuando no haya lugar a confusión los denotaremos simplemente por Mod y Vec .
- d) La categoría **Top** cuyos objetos son los espacios topológicos y sus morfismos son las aplicaciones continuas de espacios topológicos.
- e) Los complejos de cadenas junto a los morfismos de complejos de cadenas forman una categoría a la que denotaremos **Cad**.

Definición 2.3. Una categoría \mathcal{C}' es una **subcategoría** de una categoría \mathcal{C} si:

1. $\text{Obj}(\mathcal{C}') \subseteq \text{Obj}(\mathcal{C})$.
2. $\text{Hom}_{\mathcal{C}'}(X, Y) \subseteq \text{Hom}_{\mathcal{C}}(X, Y)$ para todo $X, Y \in \text{Obj}(\mathcal{C}')$.

3. El morfismo de composición en \mathcal{C}' es igual a la composición en \mathcal{C} y para objeto todo $X \in \text{Obj}(\mathcal{C}')$ el morfismo identidad en \mathcal{C}' es igual al morfismo identidad del objeto X en \mathcal{C} .

Ejemplo 2.4. En la categoría de **Grp**, los grupos abelianos forman una subcategoría **Ab**.

Definición 2.5. Dadas dos categorías \mathcal{C} y \mathcal{D} , un **functor covariante** \mathcal{F} entre ellas es una manera de asignar a cada $X \in \text{Obj}(\mathcal{C})$ un objeto $F(X) \in \text{Obj}(\mathcal{D})$ y a cada morfismo $f \in \text{Hom}_{\mathcal{C}}(X, Y)$, un morfismo $F(f) \in \text{Hom}_{\mathcal{D}}(F(X), F(Y))$, de manera que $F(\text{Id}_X) = \text{Id}_{F(X)}$ y $F(gf) = F(g)F(f)$.

Llamamos la **composición** de dos funtores $F : \mathcal{C} \rightarrow \mathcal{D}$ y $G : \mathcal{D} \rightarrow \mathcal{E}$ al functor $G \circ F : \mathcal{C} \rightarrow \mathcal{E}$ definido por $(G \circ F)(X) = G(F(X))$ y $(G \circ F)(f) = G(F(f))$. Cuando no haya lugar a confusión se podrá denotar la composición de funtores simplemente por GF . Para cualquier categoría \mathcal{C} existe el functor $\text{Id} : \mathcal{C} \rightarrow \mathcal{C}$ definido por $\text{Id}(X) = X$ sobre los objetos y $\text{Id}(f) = f$ sobre los morfismos.

Definición 2.6. Sean $\mathcal{F}, \mathcal{G} : \mathcal{C} \rightarrow \mathcal{D}$ funtores covariantes. Una **transformación natural** o **morfismo de funtores** entre ellos es una familia de morfismos para cada objeto $\{\eta_X : F(X) \rightarrow G(X)\}_{X \in \text{Obj}(\mathcal{C})}$ tal que para toda pareja X, Y de objetos y todo morfismo $f \in \text{Hom}_{\mathcal{C}}(X, Y)$ el diagrama:

$$\begin{array}{ccc} F(X) & \xrightarrow{\eta_X} & G(X) \\ F(f) \downarrow & & \downarrow G(f) \\ F(Y) & \xrightarrow{\eta_Y} & G(Y) \end{array}$$

son conmutativos. La denotaremos por $\eta : \mathcal{F} \rightarrow \mathcal{G}$ o simplemente $\eta : \mathcal{F} \rightarrow \mathcal{G}$.

Para todo functor $F : \mathcal{C} \rightarrow \mathcal{D}$ tenemos el morfismo de funtores

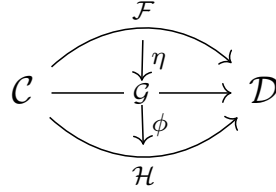
$$\text{Id}_F = \{\text{Id}_{F(X)} : F(X) \rightarrow F(X)\}_{X \in \text{Obj}(\mathcal{C})}.$$

Dos funtores $F, G : \mathcal{C} \rightarrow \mathcal{D}$ se dicen **isomorfos** si existen transformaciones naturales $\eta : \mathcal{F} \rightarrow \mathcal{G}$ y $\rho : \mathcal{G} \rightarrow \mathcal{F}$ tales que $\eta \circ \rho = \text{Id}_{\mathcal{G}}$ y $\rho \circ \eta = \text{Id}_{\mathcal{F}}$.

Decimos que para dos categorías, \mathcal{C} y \mathcal{D} , un functor $\mathcal{F} : \mathcal{C} \rightarrow \mathcal{D}$ es una **equivalencia de categorías** si existe otro functor \mathcal{G} verificando que $\mathcal{F} \circ \mathcal{G} \simeq \text{Id}$ y $\mathcal{G} \circ \mathcal{F} \simeq \text{Id}$. Al functor \mathcal{G} se le llama **quasi-inversa** de \mathcal{F} .

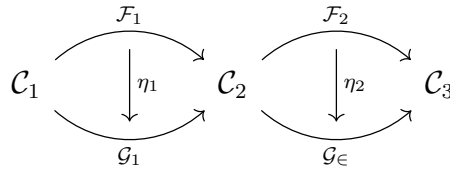
A diferencia de los funtores, existen dos posibles formas de componer morfismos de funtores:

Definición 2.7. Sean $\mathcal{F}, \mathcal{G}, \mathcal{H} : \mathcal{C} \rightarrow \mathcal{D}$ funtores y sean $\eta : \mathcal{F} \rightarrow \mathcal{G}$, $\phi : \mathcal{G} \rightarrow \mathcal{H}$ morfismos de funtores. Definimos la **composición vertical** de η y ϕ , denotada $\phi \circ \eta : \mathcal{F} \rightarrow \mathcal{H}$ como el morfismo de funtores dada por la familia $(\phi \circ \eta)_X := \phi_X \circ \eta_X$.



Definición 2.8. Sean los funtores $\mathcal{F}_1, \mathcal{G}_1 : \mathcal{C}_1 \rightarrow \mathcal{C}_2$ y $\mathcal{F}_2, \mathcal{G}_2 : \mathcal{C}_2 \rightarrow \mathcal{C}_3$. Supongamos que tenemos morfismos de funtores $\eta_i : \mathcal{F}_i \rightarrow \mathcal{G}_i$ para $i = 1, 2$. La **composición horizontal** de η_1 y η_2 , denotada $\eta_2 * \eta_1$, es un morfismo de funtores $\eta_2 * \eta_1 : \mathcal{F}_2 \mathcal{F}_1 \rightarrow \mathcal{G}_2 \mathcal{G}_1$, definido por cualquiera de las siguientes dos fórmulas equivalentes:

$$(\eta_2 * \eta_1)_X := (\eta_2)_{\mathcal{G}_1(X)} \circ \mathcal{F}_2((\eta_1)_X) = \mathcal{G}_2((\eta_1)_X) \circ (\eta_2)_{\mathcal{F}_1(X)}$$



Se dan ahora las nociones necesarias para definir una categoría abeliana. Diremos que un objeto, \emptyset , en una categoría \mathcal{C} es **inicial** si para todo objeto $X \in \text{Obj}(\mathcal{C})$ existe un único elemento en $\text{Hom}_{\mathcal{C}}(\emptyset, X)$. Diremos que un objeto, $*$, es **terminal** si para todo objeto $X \in \text{Obj}(\mathcal{C})$, existe un único elemento en $\text{Hom}_{\mathcal{C}}(X, *)$. Si un objeto es terminal e inicial a la vez se dice que es **cero**. Es fácil ver que en caso de existir cualquiera de estos objetos son únicos salvo isomorfismo. Cuando exista un objeto cero podemos definir para cada par de elementos X, Y en $\text{Obj}(\mathcal{C})$ un morfismo 0 como la única composición $X \rightarrow 0 \rightarrow Y$ verificando $f0 = 0$ y $0f = 0$ para cualquier otro morfismo f .

Un **monomorfismo** es un morfismo $f : X \rightarrow Y$ verificando que si dos morfismos $g_1, g_2 : W \rightarrow X$ cumplen $fg = fh$ entonces $g = h$. Análogamente será un **epimorfismo** si verifica que para dos morfismos $h_1, h_2 : Y \rightarrow Z$ con $h_1f = h_2f$ entonces $h_1 = h_2$.

Sea una categoría \mathcal{C} con un objeto cero y sea $f : X \rightarrow Y$ un morfismo de \mathcal{C} . El **núcleo** de f es un morfismo $j : \text{Ker} f \rightarrow X$ tal que $fj = 0$. El **conúcleo** es un morfismo $q : Y \rightarrow \text{coker} f$ tal que si $h : Y \rightarrow Z$ entonces $hf = 0$.

Definición 2.9. Sean $X, Y \in \text{Obj}(C)$. El **producto** de X e Y , si existe, es un elemento de $\text{Obj}(C)$ denotado $X \times Y$ con morfismos $p_X : X \times Y \rightarrow X$ y $p_Y : X \times Y \rightarrow Y$ verificando la siguiente propiedad universal: para todo $W \in \text{Obj}(C)$ y morfismos $f_X : W \rightarrow X$ y $f_Y : W \rightarrow Y$ existe un único morfismo $f : W \rightarrow X \times Y$ tal que $f_X = p_X f$ y $f_Y = p_Y f$. Si existe, el producto es único salvo isomorfismo.

Definición 2.10. Sean $X, Y \in \text{Obj}(C)$. El **coproducto** de dos objetos X, Y , si existe, es un objeto denotado $X \oplus Y$ junto con morfismos $j_X : X \rightarrow X \oplus Y$ y $j_Y : Y \rightarrow X \oplus Y$ verificando la siguiente propiedad universal: para todo objeto U con un par de morfismos $g_X : X \rightarrow U$ y $g_Y : Y \rightarrow U$ existe un único morfismo $g : X \oplus Y \rightarrow U$ tal que $g_X = g j_X$ y $g_Y = g j_Y$. Si existe, el coproducto es único salvo isomorfismo.

Definición 2.11. Una categoría es **abeliana** si tiene un objeto cero y todos los productos directos y sumas directas, todo morfismo tiene un núcleo y un conúcleo y todo monomorfismo es un núcleo y todo epimorfismo es un conúcleo.

Cabe indicar que en las categorías abelianas el morfismo natural $X \oplus Y \rightarrow X \times Y$ es un isomorfismo, para cualesquiera objetos X, Y .

Ejemplo 2.12. Un ejemplo de categoría abeliana es la categoría, $\mathbf{Vec}_{\text{fin}}$, de los espacios vectoriales de dimensión finita sobre un cuerpo k , siendo los morfismos las aplicaciones lineales.

Definición 2.13. Decimos que una categoría \mathcal{C} es de **Krull-Schmidt** si verifica las siguientes condiciones:

1. \mathcal{C} es abeliana.
2. Para todo $X \in \text{Obj}(\mathcal{C})$ existe una familia esencialmente única de objetos indecomponibles, $\mathbf{W}(X)$, tales que:

$$X \cong \bigoplus_{I \in \mathbf{W}(X)} I \tag{2.1}$$

es decir, si existen dos familias con la misma propiedad, $\mathbf{W}(X)$ y $\mathbf{W}'(X)$, entonces existe una biyección $\tau : \mathbf{W}(X) \rightarrow \mathbf{W}'(X)$.

3. Para cualquier objeto y descomposición como arriba, $\prod_{I \in \mathbf{W}(X)} I$ existe en \mathcal{C} y el morfismo:

$$\bigoplus_{I \in \mathbf{W}(X)} I \rightarrow \prod_{I \in \mathbf{W}(X)} I \tag{2.2}$$

es un isomorfismo.

2.2. Objetos y módulos de persistencia

Se definen ahora los objetos persistentes y los módulos de persistencia, los cuales son una generalización de la homología persistente de la Sección 1.5.

Definición 2.14. Sea \mathcal{C} una categoría y (\mathcal{P}, \leq) un conjunto parcialmente ordenado. Un **objeto \mathcal{P} -persistente en \mathcal{C}** es un functor $\Psi: (\mathcal{P}, \leq) \rightarrow \mathcal{C}$. Esto es, una familia de objetos en \mathcal{C} $\{c_x\}_{x \in \mathcal{P}}$ y un morfismo $\Psi_{x \leq y}: c_x \rightarrow c_y$ cuando $x \leq y$, verificando además que $\Psi_{x \leq x} = Id$ y $\Psi_{y \leq z} \circ \Psi_{x \leq y} = \Psi_{x \leq z}$ para todo $x \leq y \leq z$. Estos morfismos se llaman **morfismos de transición**. Esta colección de objetos y morfismos será denotada $\{c_x, \Psi_{x \leq y}\}$.

Observación 2.15. Los objetos persistentes también pueden definirse en base a una topología de Alexandrov como hacen en [2] pero no abarcaremos ese resultado por motivos de extensión y resultados necesarios. \triangle

Tomando como morfismos los morfismos de funtores (con su composición vertical) entre dos objetos \mathcal{P} -persistentes en \mathcal{C} , Ψ y Φ , los objetos \mathcal{P} -persistentes forman una categoría denotada $\mathcal{C}^{\mathcal{P}}$. Nótese que en este caso un morfismo de funtores de $\Psi = \{c_x, \Psi_{x \leq y}\}$ a $\Phi = \{d_x, \Phi_{x \leq y}\}$ será una familia de morfismos $f = \{f_x: c_x \rightarrow d_x\}_{x \in \mathcal{P}}$ tales que para todo $x \leq y$ el siguiente diagrama es conmutativo:

$$\begin{array}{ccc} c_x & \xrightarrow{\Psi_{x \leq y}} & c_y \\ f_x \downarrow & & \downarrow f_y \\ d_x & \xrightarrow{\Phi_{x \leq y}} & d_y \end{array}$$

Observación 2.16. Utilizaremos el nombre objeto persistente cuando se trate de un objeto de $\mathcal{C}^{\mathcal{P}}$ para una categoría y poset cualquiera. Llamaremos módulo de persistencia a todo objeto de $Mod_R^{\mathcal{P}}$, para un poset cualquiera. Para diferenciarlos en las siguientes secciones, se usará la notación caligráfica $\mathcal{M} = \{M_x, M_{x \leq y}\}$ para módulos de persistencia y M para los módulos de persistencia de los módulos graduados. \triangle

Si la categoría \mathcal{C} tiene un objeto cero, un objeto persistente de especial relevancia es el llamado **cero** o **nulo** denotado 0 que es aquel que verifica: $0_x = 0 \in \mathcal{C}$ para todo $x \in \mathcal{P}$.

Definición 2.17. Una aplicación entre posets $\sigma: \mathcal{P} \rightarrow \mathcal{Q}$ es **monótona** si preserva las relaciones de orden, es decir, si $\sigma(x) \leq \sigma(y) \Leftrightarrow x \leq y$. Toda aplicación monótona se puede ver como un functor entre los posets, pensados como categorías, al que denotaremos T_σ .

Además, para toda aplicación monótona σ , T_σ induce un functor que permite cambiar

la forma en que indexamos los objetos persistentes:

$$(T_\sigma)^* : \mathcal{C}^{\mathcal{Q}} \longrightarrow \mathcal{C}^{\mathcal{P}}$$

$$\Phi \mapsto (T_\sigma)^*(\Phi) := \Phi \circ T_\sigma$$

En la anterior sección ya hemos visto cómo construir complejos simpliciales \mathbb{R} -persistentes o complejos de cadenas \mathbb{N} -persistentes. En este trabajo nos centraremos en el estudio de los R -módulos persistentes, es decir en $Mod_R^{\mathcal{P}}$ para un poset \mathcal{P} . Los dos ejemplos de relevancia para el análisis topológico de datos son los siguientes:

Ejemplo 2.18. 1. Sea K un complejo simplicial con una filtración

$$F : \emptyset = K_0 \subset K_1 \subset \dots \subset K_n = K$$

Definimos $[n] := \{0, 1, 2, \dots, n\}$. Tomando $C(i) := C(K_i)$ y $C(i \leq j)$ como el morfismo de cadenas dado por la inclusión natural entre los complejos simpliciales tenemos que $C \in Cad^{[n]}$. Para cualquier número natural $p \in \mathbb{N}$, y un anillo R , podemos tomar para cada $i \in [n]$ el R -módulo $H_p(i) := H_p^{simp}(C(K_i))$ y para $i \leq j$ el morfismo inducido $H_p(i \leq j) := H_p(C(K_i)) \rightarrow H_p(C(K_j))$. La functorialidad se sigue de la observación 1.15, luego $H_p \in Mod_R^{[n]}$.

2. Sea $f : X \rightarrow \mathbb{R}$ una función, donde X es un espacio topológico (no necesariamente continua), para cada $\alpha \in \mathbb{R}$ definimos **el subconjunto de nivel de f de parámetro α** como

$$f^{-1}(-\infty, \alpha] := \{x \in X \text{ tal que } f(x) \leq \alpha\}$$

Utilizando la topología inducida en el subespacio, $f^{-1}(-\infty, \alpha]$ es un espacio topológico. Si $\beta \geq \alpha$ se cumple además que $f^{-1}(-\infty, \alpha] \subseteq f^{-1}(-\infty, \beta]$ y además la inclusión es una aplicación continua pues si A es un abierto en $f^{-1}(-\infty, \beta]$ debe ser intersección con un abierto de X , digamos U , y se cumple que la antiimagen de A por la inclusión es $U \cap f^{-1}(-\infty, \alpha]$. Definimos $F = \{F_x, F(x \leq y)\} \in Top^{\mathbb{R}}$ como: para cada $x \in \mathbb{R}$ $F(x) := f^{-1}((-\infty, x])$, y para $x \leq y$, $F(x \leq y)$ es la inclusión natural $f^{-1}((-\infty, \alpha]) \rightarrow f^{-1}((-\infty, \beta])$. Si \mathbb{k} un cuerpo, $p \in \mathbb{N}$ y H_p el functor de p -ésima homología singular con coeficientes en \mathbb{k} , entonces componiendo $H_k \circ F$ tenemos un objeto de $Vec^{\mathbb{R}}$.

Recordamos las siguientes definiciones de la teoría de módulos:

Definición 2.19. Decimos que un módulo M sobre un anillo R es **finito generado** si existe una familia de elementos $\{m_i\}_{i=1}^k \in M$, llamado **conjunto de generadores de**

M , tal que el morfismo de módulos:

$$f_{\{m_i\}_{i=1}^k} : \bigoplus_{i=1}^k R \rightarrow M$$

$$\{r_i\}_{i=1}^k \mapsto \sum_{i=1}^k r_i \cdot m_i$$

es epiyectivo. Decimos que un módulo M es **finito presentado** o de presentación finita si es finito generado y el núcleo del morfismo $f_{\{m_i\}_{i=1}^k}$ es también un módulo finito generado. Esto es equivalente a decir que M es finito generado si existe una sucesión exacta:

$$R^m \rightarrow R^n \rightarrow M \rightarrow 0. \quad (2.3)$$

Damos ahora algunas condiciones de finitud para los módulos de persistencia sobre un poset cualquiera \mathcal{P} .

Definición 2.20. Sea R un anillo y $\mathcal{M} \in \text{Mod}_R^{\mathcal{P}}$. Decimos que \mathcal{M} es **puntualmente finito generado** si $\forall x \in \mathcal{P}$ el R -módulo M_x es finito generado. Si todo M_x es también finito presentado diremos que \mathcal{M} es **puntualmente finito presentado**.

Definición 2.21. Sea \mathcal{M} un objeto de $\text{Mod}_R^{\mathcal{P}}$. Decimos que \mathcal{M} es de **tipo finito** si existe un índice $x_0 \in \mathcal{P}$ tal que $M_{x \leq y} : M_x \rightarrow M_y$ es un isomorfismo de módulos, para todo $y \geq x \geq x_0$. Un módulo de persistencia será de tipo finito generado/presentado si es de tipo finito y puntualmente finito generado/presentado.

En el caso de nubes de datos finitas, estas condiciones serán satisfechas y tendremos por tanto módulos de persistencia de tipo finito. Que cada componente es finito generada se deduce de su construcción por ser símlices finitos. Además, una vez hayamos alcanzado el valor umbral a partir del cual todos los puntos están conectados, al seguir aumentando el radio de las bolas que los forman no cambiarán los símlices que aparecen y por tanto tampoco su homología, dado que ya se han formado todos los símlices posibles.

2.3. Módulos graduados

Esa sección es un breve resumen de algunas nociones de la teoría de módulos graduados, necesaria para la clasificación de los módulos de persistencia.

Definición 2.22. Se dice que un anillo R es \mathbb{N} -graduado o simplemente **graduado** si R puede ser escrito como $R = \bigoplus_{i \in \mathbb{N}} R_i$, donde para cada índice $i \in \mathbb{N}$ R_i es un grupo abeliano, y si $r_i \in R_i$ y $r_j \in R_j$ se cumple que $r_i \cdot r_j \in R_{i+j}$. Cada uno de los R_i se llama componente homogénea en grado i y sus elementos son los elementos homogéneos de grado i .

También es común encontrar como conjunto de índices los números enteros y puede hacerse en general con un monoide cualquiera.

Ejemplo 2.23. Para cualquier anillo R , el anillo de polinomios $R[x]$ es un anillo graduado siendo $R[x]_i$ el R -módulo generado por x^i . Aunque puedan existir otras posibles graduaciones para $R[x]$ nosotros únicamente consideraremos la aquí expuesta.

Definición 2.24. Un **ideal graduado** es un ideal generado por elementos homogéneos.

Definición 2.25. Para un anillo graduado R , un R -**módulo graduado** M es un módulo para el cual existe una descomposición $M = \bigoplus_{i \in \mathbb{N}} M_i$, tal que cada M_i es un grupo abeliano, y $r_i \cdot m_j \in M_{i+j}$ para todo $r_i \in R_i$ y $m_j \in M_j$. Cada M_i se llama componente homogénea de grado i y, por definición, todo elemento $m \in M$ puede expresarse como:

$$m = m_{i_1} + \cdots + m_{i_l}$$

donde $m_{i_j} \in M_{i_j}$.

Un ejemplo de módulo graduado es simplemente considerar el propio anillo graduado, R , como un R -módulo.

Definición 2.26. Para un anillo graduado R y dos módulos graduados sobre R , M y M' , un **morfismo graduado de grado j** $f: M \rightarrow M'$ es un morfismo de R -módulos con la propiedad de que $f(M_i) \subset M'_{i+j}$ para todo $i \in \mathbb{N}$.

Observación 2.27. En esta memoria trabajaremos únicamente con morfismos graduados de grado 0, a los que se llamarán simplemente morfismos graduados. \triangle

Es fácil comprobar que para un módulo graduado M , la identidad Id_M es un morfismo graduado, y la composición de dos morfismos graduados $M \xrightarrow{f} M' \xrightarrow{g} M''$ también es un morfismo graduado. Por tanto, para un anillo graduado R la colección de todos los R -módulos graduado junto con los morfismos graduados forma una categoría, que denotaremos por $R_{Grad}Mod$.

Para los módulos graduados tenemos las mismas condiciones de finitud: finito generado, finito presentado o libre (en este caso las bases deben ser homogéneas, esto es, sus elementos son todos homogéneos). Recordamos que un módulo es **Noetheriano** si todos sus submódulos son finito generados, o equivalentemente si satisface las condiciones de cadena ascendente. Un anillo R es un anillo **Noetheriano** si como R -módulo es Noetheriano. Un ejemplo de anillo Noetheriano son los cuerpos, pues solo tienen dos ideales.

Observación 2.28. Dado que un ideal graduado está generado por elementos homogéneos, si tenemos un ideal homogéneo en $\mathbb{k}[x]$ éste debe estar generado por x^n para algún n , luego todos sus ideales homogéneos son de la forma (x^n) . \triangle

Proposición 2.29. *Si R es un anillo Noetheriano todo R -módulo graduado finito generado es también finito presentado.*

Demostración. Si M es finito generado existirá un morfismo $f_G: R^n \rightarrow M$ epiyectivo. Entonces $\ker(f_G)$ es finito generado por ser un submódulo de R^n , que es Noetheriano. \square

La siguiente proposición nos dice que, sobre un dominio de ideales principales, todo submódulo de un módulo libre es libre, y además da una relación entre sus sistemas de generadores, necesaria para demostrar el teorema de estructura que veremos posteriormente.

Proposición 2.30. *([17, Theorem 4, Chapter 12]) Sea D un dominio de ideales principales y M un D -módulo libre de rango n . Si N es un submódulo de M entonces N es libre de rango $m \leq n$. Además, existe una base $\{\theta_1, \dots, \theta_n\}$ de M tal que $\{\alpha_1\theta_1, \dots, \alpha_m\theta_m\}$ es base del submódulo N , para elementos no nulos $\alpha_i \in D$ verificando $\alpha_i | \alpha_{i+1}$ para $1 \leq i \leq m$.*

El motivo por el cual son importantes los módulos graduados es que para ellos conocemos un teorema de la estructura que nos permite dar una clasificación de los mismos. El siguiente teorema es cierto para cualquier dominio de ideales principales, pero aquí nos reduciremos al caso de $\mathbb{k}[x]$, siendo \mathbb{k} es un cuerpo.

Teorema 2.31 (Teorema de la estructura para módulos graduados sobre un anillo de polinomios). *Sea $\mathbb{k}[x]$ el anillo de polinomios sobre un cuerpo. Para cada módulo graduado finito generado M sobre $\mathbb{k}[x]$ existen a_i, b_i y $q_i \in \mathbb{Z}$ tales que M descompone como:*

$$M \cong \left(\bigoplus_{i=1}^{n-m} x^{a_i} \cdot \mathbb{k}[x] \right) \oplus \left(\bigoplus_{i=j}^m x^{b_i} \cdot \mathbb{k}[x] / (x^{q_j}) \right). \quad (2.4)$$

Esta descomposición es única salvo isomorfismo.

Demostración. Sea $\{m_i\}_{i=1}^n$ un sistema de generadores mínimo y $\mathbb{k}[x]^n$ el módulo libre de base estándar $\{e_i\}_{i=1}^n$. Tenemos entonces el morfismo epiyectivo de módulos $\phi: \mathbb{k}[x]^n \rightarrow M$ definido por $\phi(e_i) = m_i$. Por el primer teorema de isomorfía:

$$\mathbb{k}[x]^n / \ker f(\phi) \cong M. \quad (2.5)$$

Usando la Proposición 2.30, como $\mathbb{k}[x]^n$ es libre sobre un dominio de ideales principales y $\ker(\phi)$ es un submódulo suyo, $\ker(\phi)$ es libre de rango $m \leq n$ y existe una base $\{\theta_i\}_{i=1}^n$ de $\mathbb{k}[x]^n$ tal que para ciertos elementos $\alpha_i \in \mathbb{k}[x]$ el conjunto $\{\alpha_1\theta_1, \dots, \alpha_m\theta_m\}$ forma una base de $\ker(\phi)$. Además, para $1 \leq j \leq n$ los ideales (θ_j) de $\mathbb{k}[x]$ deben ser de la forma (x^{q_j}) (ver Observación 2.28).

Definimos ahora el morfismo epiyectivo:

$$\varphi : \bigoplus_{j=1}^n \langle \theta_j \rangle \rightarrow x^{\deg(\theta_1)} \cdot \mathbb{k}[x] / (x^{q_1}) \oplus \cdots \oplus x^{\deg(\theta_m)} \cdot \mathbb{k}[x] / (x^{q_m}) \oplus \left(\bigoplus_{i=1}^{n-m} x^{\deg(\theta_{m+i})} \cdot \mathbb{k}[x] \right) \quad (2.6)$$

por

$$\varphi(\lambda_1 \theta_1, \dots, \lambda_n \theta_n) := ([\lambda_1], \dots, [\lambda_m], \lambda_{m+1}, \dots, \lambda_n)$$

donde $[\lambda_i]$ es $\lambda_i \bmod(x^{q_i})$ para $1 \leq i \leq m$. El núcleo de φ es $\bigoplus_{j=1}^m \langle x^{q_j} \theta_j \rangle$ que es isomorfo al núcleo de ϕ y claramente $\bigoplus_{j=1}^n \langle \theta_j \rangle \cong \mathbb{k}[x]^n$. Luego tenemos que:

$$M \cong \mathbb{k}[x]^n / \ker f(\phi) \cong \mathbb{k}[x]^n / \ker f(\varphi) \cong \text{Im}(\varphi) \quad (2.7)$$

con lo que acabamos la parte de la existencia.

Para ver la unicidad puede consultarse [17, Theorem 9, Chapter 12].

□

Observación 2.32. Nuestro trabajo se centrará en los módulos graduados finito generados/presentados. En el caso de un módulo finito generado graduado, siempre podremos encontrar un conjunto de elementos homogéneos que lo generan, al que llamaremos **sistema de generadores homogéneo**. Gracias a esto podremos definir una graduación en $\mathbb{k}[x]^n$. En efecto, si M es un módulo graduado finito generado, $G = \{g_1, \dots, g_n\}$ un sistema de generadores homogéneo y $\{e_1, \dots, e_n\}$ la base estándar de $\mathbb{k}[x]^n$ tal que $f_G : \mathbb{k}[x]^n \rightarrow M$ manda e_i a g_i , definimos $\deg(e_i) = \deg(g_i)$ siendo \deg el grado correspondiente en cada módulo. Además $\deg(s_i) = i$ para todo elemento de $\mathbb{k}[x]_i$ (usando la graduación usual). Todo elemento de M descompone como una suma de elementos homogéneos de acuerdo a la graduación de M , con ello podremos decir que un elemento de $\mathbb{k}[x]^n$, digamos $e = \sum_{i=1}^n s_i e_i$, será homogéneo si $f_G(e)$ es homogéneo, y así todo elemento descompone como suma de elementos homogéneos. Si además M es finito presentado los generadores de $\ker(f_G)$ también pueden ser tomados formando un sistema de generadores homogéneo. \triangle

2.4. Módulos de persistencia como módulos sobre un anillo de polinomios

En esta sección se muestra que, bajo ciertas condiciones de finitud, las categorías $\text{Mod}_R^{\mathbb{N}}$ y $R_{\text{Grad}}\text{Mod}$ son equivalentes. Los primeros resultados de clasificación de módulos \mathbb{N} -persistentes en la literatura se basan en esta equivalencia de categorías, dado que en

$R_{Grad}Mod$ tenemos el teorema de la estructura 2.31 para la calificación de módulos graduados. Esta equivalencia suele darse con el siguiente teorema:

Teorema 2.33 (Teorema de representación Zomorodian-Carlsson). *Sea R un anillo unitario y conmutativo. La categoría de los módulos \mathbb{N} -persistentes de tipo finito es equivalente a la categoría de $R[x]$ -módulos graduados finito generados*

Este teorema se encuentra de distintas formas en la literatura relacionada con el análisis topológico de datos (ver, por ejemplo [6], [38] y [7]). Sin embargo, no aparece una demostración del mismo y suele referenciarse a la teoría de Artin-Rees (en particular [19]). Según los autores la demostración sería una consecuencia inmediata. En [37] aparece un esquema de la demostración pero contiene un error, como se muestra en el contraejemplo de [13]. La razón por la que este planteamiento funcionaría en el análisis topológico de datos es que suele considerarse que el anillo base es un cuerpo \mathbb{k} , y por tanto es Noetheriano, pudiendo aplicarse la teoría de Artin-Rees. Sin embargo, si quitamos esa suposición no es cierto en general el teorema de representación.

En este trabajo se ha tratado de seguir un planteamiento más general, desarrollando en esta sección los pasos del artículo de Corbet y Kerber [13], donde las condiciones para el teorema de representación son bastante generales. La mayor dificultad de esta sección es asegurar que las condiciones de finitud en cada categoría se satisfacen.

Comenzamos dando una equivalencia de categorías entre los módulos de persistencia y los módulos graduados:

Proposición 2.34. *Existen funtores:*

$$\alpha: Mod_R^{\mathbb{N}} \rightarrow R[x]_{Grad}Mod \quad y \quad \beta: R[x]_{Grad}Mod \rightarrow Mod_R^{\mathbb{N}}$$

de manera que $Mod_R^{\mathbb{N}}$ y $R[x]_{Grad}Mod$ son categorías equivalentes.

Demostración. Primero damos su definición sobre objetos.

Sea $\mathcal{M} = \{M_i, M_{i \leq j}\}$ un módulo de persistencia sobre un anillo R , definimos:

$$\alpha(\mathcal{M}) := \bigoplus_{i \in \mathbb{N}} M_i \tag{2.8}$$

que es un $R[x]$ -módulo graduado definiendo

$$x \cdot (m_i)_{i \in \mathbb{N}} := (n_{i+1})_{i \in \mathbb{N}}$$

donde $n_{i+1} = M_{i \leq i+1}(m_i)$, y después extendiendo linealmente.

Para la quasi-inversa, dado un $R[x]$ -módulo graduado $\bigoplus_{i \in \mathbb{N}} M_i$ definimos

$$\beta \left(\bigoplus_{i \in \mathbb{N}} M_i \right) := (M_i, M_{i \leq j})_{i \in \mathbb{N}} \quad (2.9)$$

donde los morfismos de transición $M_{i \leq j}$ vienen dados por la multiplicación por x^{j-i} .

Ahora tenemos que dar su definición sobre los morfismos de cada categoría, de manera que tengan propiedades functoriales.

Para un morfismo $\eta : \mathcal{M} \rightarrow \mathcal{N}$ de R -módulos de persistencia se define

$$\begin{aligned} \alpha(\eta) : \bigoplus_{i \in \mathbb{N}} M_i &\rightarrow \bigoplus_{i \in \mathbb{N}} N_i \\ (m_i) &\mapsto (\eta_i(m_i)) \end{aligned}$$

donde $\alpha(\mathcal{M}) = \bigoplus_{i \in \mathbb{N}} M_i$ y $\alpha(\mathcal{N}) = \bigoplus_{i \in \mathbb{N}} N_i$. Tenemos que comprobar que $\alpha(\eta)$ está bien definido, como morfismo graduado, y las propiedades functoriales de α . En primer lugar, $\alpha(\eta)$ es un morfismo de grupos porque η_i es un morfismo de grupos, para cada índice i . Para todo $i \in \mathbb{N}$ se cumple que $\alpha(\eta)(M_i) \subset N_i$ pues si $m_i \in M_i$ se tiene que $\alpha(\eta)(m_i) = \eta_i(m_i) \in N_i$. Por último, conmuta con el producto por polinomios pues al ser η morfismo se tiene

$$\eta(x \cdot m_i) = x \cdot \eta(m_i)$$

luego

$$\alpha(\eta)(x \cdot (m_i)_{i \in \mathbb{N}}) = x \alpha(\eta)((m_i)_{i \in \mathbb{N}})$$

Además, la identidad va a ir a la identidad, ya que $Id_{\mathcal{M}} = (Id_i)_{i \in \mathbb{N}}$, y para dos morfismos de funtores:

$$(\alpha(\eta' \circ \eta))(m_i)_{i \in \mathbb{N}} = (\eta'_i(\eta_i(m_i)))_{i \in \mathbb{N}} = (\alpha(\eta')(\eta_i(m_i)))_{i \in \mathbb{N}} = (\alpha(\eta') \circ \alpha(\eta))(m_i)_{i \in \mathbb{N}}.$$

Por otro lado, si tenemos un morfismo graduado ω en $R[x]_{Grad}Mod$, restringiéndolo a cada componente homogénea nos da un morfismo de R -módulos $\omega_i : M_i \rightarrow N_i \forall i \in \mathbb{N}$ que es compatible con la multiplicación por x , es decir $x \cdot \omega_i(m_i) = \omega_{i+1}(x \cdot m_i)$. Definimos entonces el morfismo de funtores $\beta(\omega) := (\omega_0, \omega_1, \dots)$. Falta probar las propiedades functoriales de β . De nuevo la functorialidad sobre la identidad se comprueba de forma inmediata ya que si M es un módulo graduado los morfismos inducidos por Id_M en las componentes homogéneas son $(Id_M)_i = Id_{M_i}$. Finalmente, para cualesquiera dos morfismos de $R[x]$ -módulos graduados ω y ω' se tiene— que

$$\beta(\omega' \circ \omega) = ((\omega' \circ \omega)_i)_{i \in \mathbb{N}} = (\omega'_i \circ \omega_i)_{i \in \mathbb{N}} = (\beta(\omega') \circ \beta(\omega)).$$

donde la segunda igualdad se debe a la definición de los ω_i y ω'_i . Por construcción es inmediato que $\alpha \circ \beta$ y $\beta \circ \alpha$ son isomorfos a la identidad en ambos casos, luego las categorías son isomorfas. \square

El siguiente lema prueba que el functor α mantiene las condiciones de finitud. En la demostración se incluyen solo los detalles más importantes por motivos de extensión.

Lema 2.35. *Si un módulo de persistencia \mathcal{M} es de tipo finito presentado, entonces $\alpha(\mathcal{M})$ es un módulo graduado finito presentado.*

Demostración. Se ve sin dificultad que $\alpha(\mathcal{M})$ es finito generado por el conjunto $G = \cup_{i=1}^{n_0} G_i$ donde G_i es un generador de M_i y n_0 es el índice a partir del cual todos son isomorfos (tipo finito). La dificultad está en demostrar que también es finito presentado. Como cada M_i era finito presentado tenemos morfismo epiyectivos $f_{G_i} : R^{n_i} \rightarrow M_i$. Definimos $n = \sum_{i=1}^{n_0} n_i$ y tenemos que ver que el núcleo de $F_G : R[x]^n \rightarrow \alpha(\mathcal{M})$ es finito generado. Para cada $1 \leq i \leq n_0$ definimos Ξ_i como el conjunto de generadores de $\ker(f_{G_i})$, que es finito por ser cada M_i finito presentado. Para cada índice i y $\xi_i \in \Xi_i$ se cumple que $(0, \dots, 0, \xi_i, 0, \dots, 0) \in R[x]^n$ está en $\ker(F_G)$. Además, para cada pareja (i, j) , con $0 \leq i < j \leq n_0$, y cada elemento $g_i \in G_i$ tal que $M_{i \leq j}(g_i) \neq 0$ tenemos que:

$$M_{i \leq j}(g_i) = \sum_{k=0}^{l_j} \lambda_k g_j^{(k)}$$

donde $\lambda_k \in R$ y $G_j = \{g_j^{(1)}, \dots, g_j^{(n_j)}\}$. El elemento $x^{j-i} \cdot e_i - \sum_{k=0}^{l_j} \lambda_k e_j^{(k)}$ está en $\ker(F_G)$, por la definición de α , siendo e_i y $e_j^{(k)}$ los generadores de $R[x]^n$ que por F_G van a parar a g_i y $g_j^{(k)}$ respectivamente. Denotemos por $\Xi_{(i,j)}$ al conjunto finito obtenido eligiendo un elemento como los anteriores para cada g_i con $M_{i \leq j}(g_i) \neq 0$. En ese caso el conjunto de generadores de $\ker(F_G)$ es:

$$\Xi := \bigcup_{i=0}^{n_0} \Xi_i \cup \bigcup_{0 \leq i < j \leq n_0} \Xi_{(i,j)}. \quad (2.10)$$

Para demostrarlo consideramos un elemento del núcleo $e = \sum_j \lambda_j e_j$ con $\lambda_j \in R[x]$ y e_j generador de $R[x]^n$. Por la graduación que hemos dado en $R[x]^n$ podemos asumir que es homogéneo pues en un elemento $e' = e'_1 + \dots + e'_l$ del núcleo cada e'_i va a parar a una única componente homogénea M_i y deben ser todas 0. La demostración se termina viendo los siguientes casos: primero, suponemos que $\deg(e) \leq n_0$ y todos los λ_j son de grado 0; después, que $\deg(e) \leq n_0$ y algunos λ_j son de grado no nulo; finalmente, suponemos $\deg(e) > n_0$. \square

Sea ahora un $R[x]$ -módulo graduado finito presentado $M := \bigoplus_{i \in \mathbb{N}} M_i$, con generadores $\mathcal{G} := \{g_1, \dots, g_n\}$ y el morfismo de módulos $F_{\mathcal{G}} : R[x]^n \rightarrow M$ y donde $\Xi = \{\xi_1, \dots, \xi_m\}$ genera $\ker(F_{\mathcal{G}})$. Podemos asumir por la Observación 2.32 que todos los elementos de \mathcal{G} y Ξ son homogéneos y, adicionalmente, que sus elementos están ordenados de menor a mayor grado. Los dos siguientes lemas muestran que β también mantiene las condiciones de finitud que tengan los módulos graduados al pensarlos como módulos de persistencia.

Lema 2.36. *Cada una de las componentes homogéneas M_i de M es finito presentada como R -módulo.*

Demostración. Claramente tiene estructura de módulo puesto que estamos multiplicando por elementos del anillo que tienen grado 0 y además M_i era un grupo abeliano por ser una componenete homogénea.

Denotemos por n_i al número de elementos de G de grado hasta i . Sea $e_j^{(i)}$ el j -ésimo generador de $\bigoplus_{i=1}^{n_i} R$. La aplicación:

$$\begin{aligned} \theta_i : \bigoplus_{i=1}^{n_i} R &\rightarrow M_i \\ e_j^{(i)} &\mapsto t^{i-\deg(g_j)} g_j \end{aligned}$$

es epimorfismo por ser los g_j generadores de M luego cada componente es finito generada.

Tenemos que ver la generación finita del núcleo de los θ_i . Probamos ahora que el núcleo de cada uno de los morfismos de la parte anterior es finito generado. Sean $\{e_1, \dots, e_n\}$ los generadores de $\bigoplus_{i=1}^n R[x]$ que van a $\{g_1, \dots, g_n\}$ por $F_{\mathcal{G}}$. Denotemos por m_i al número de elementos en Ξ de grado hasta i . Para cada $1 \leq j \leq m_i$ y ξ_j tomamos el elemento $x^{i-\deg(\xi_j)}$ que puede escribirse como:

$$x^{i-\deg(\xi_j)} \xi_j = \sum_{k=1}^{n_i} r_k x^{i-\deg(g_k)} e_k \quad (2.11)$$

con los $r_k \in R$. Definimos:

$$\xi_j^{(i)} := \sum_{k=1}^{n_i} r_k e_k^{(i)} \quad (2.12)$$

y $\Xi_i := \{\xi_j^{(i)} \mid 1 \leq j \leq m_i\}$. Entonces Ξ_i genera $\ker(\theta_i)$. Es claro que $\theta_i(\xi_j^{(i)}) = F_{\mathcal{G}}(\xi_j) = 0$. Ahora para un elemento de $\eta \in \ker \theta_i$ se cumple que η es una combinación lineal de elementos $e_j^{(i)}$ con $1 \leq j \leq n_i$ y toma coeficientes en R . Sustituyendo $e_j^{(i)}$ por $x^{i-\deg(g_j)} e_j$, obtenemos un elemento η' homogéneo en grado i de $\bigoplus_{i=1}^n R[x]$ al que llamaremos η' . Este

elemento se puede escribir como combinación lineal de elementos de Ξ :

$$\eta' = \sum_{k=1}^{m_i} r'_k x^{i-\deg(\xi_k)} \xi_k \quad (2.13)$$

con $r'_k \in R$. Para $j \in \{1, \dots, n_i\}$ sea $c_j \in R$ el coeficiente de $e_j^{(i)}$ en η . Sea c'_j el coeficiente de $e_j^{(i)}$ en la suma $\sum_{k=1}^{m_i} r'_k \xi_k^{(i)}$, escribiendo cada $\xi_k^{(i)}$ como una combinación lineal. Por construcción, c_j es el coeficiente de $x^{i-\deg(g_j)} e_j$ en η' , y c'_j es el coeficiente de $x^{i-\deg(g_j)} \cdot e_j$ en la suma $\sum_{k=1}^{m_i} r'_k x^{i-\deg(\xi_k)} \cdot e_j$. Esta suma es igual a η' y se sigue que $c_j = c'_j$ luego se cumple:

$$x = \sum_{k=1}^{m_i} r'_k \xi_k^{(i)} \quad (2.14)$$

y hemos obtenido un sistema de generadores para el núcleo de θ_i . \square

Lema 2.37. ([13, Lemma 8]) *Si M es un $R[x]$ -módulo graduado finito presentado, entonces $\beta(M)$ es un módulo de persistencia de tipo finito presentado.*

Tenemos ahora ya todos los resultados necesarios para probar el teorema general de representación de módulos de persistencia de tipo finito presentados.

Teorema 2.38. *Sea R un anillo unitario. Las subcategoría de $R[x]_{\text{GradMod}}$ formada por los módulos finito presentados y la subcategoría de $\text{Mod}_R^{\mathbb{N}}$ formada por los módulos de persistencia de tipo finito presentados son equivalentes a través de las restricciones de los funtores α y β a estas subcategorías.*

Demostración. Si \mathcal{M} es un módulo de persistencia de tipo finito presentado, entonces $\alpha(\mathcal{M})$ es un módulo graduado de presentación finita (por el Lema 2.35). Si tenemos un módulo graduado M' , por los Lemas 2.36 y 2.37 se sigue que $\beta(M')$ es un módulo de persistencia de tipo finito. Como β era la quasi-inversa de α esto se mantiene para sus restricciones. \square

Como puede observarse de la demostración de los Lemas 2.35 y 2.36, sin necesidad de restricciones sobre el anillo, se cumple que si $\mathcal{M} \in \text{Mod}_R^{\mathbb{N}}$ es de tipo finito generado $\alpha(\mathcal{M})$ es finito generado y si $M \in R[x]_{\text{GradMod}}$ es finito generado, $\beta(M)$ es puntualmente finito generada. La parte que falla en el teorema de representación original de Zomorodian-Carlsson es que no se garantiza que $\beta(M)$ sea de tipo finito, para lo cual necesitamos la presentación finita.

Obtenemos el siguiente corolario:

Corolario 2.39. *Sea \mathbb{k} un cuerpo. Las subcategorías $\mathbb{k}[x]_{\text{GradMod}}$ finitos generados y $\text{Mod}_{\mathbb{k}}^{\mathbb{N}}$ de tipo finito generado son equivalentes.*

2.5. Clasificación de módulos de persistencia por módulos de intervalo

Al igual que con cualquier estructura algebraica, en el caso de los módulos de persistencia existen unos objetos básicos que clasifican y construyen cualquiera de ellos. En el caso de los módulos de persistencia son los llamados módulos de intervalo, así que comenzamos esta sección con la definición de intervalo.

Definición 2.40. Un **intervalo** en un conjunto totalmente ordenado (\mathcal{P}, \leq) es un subconjunto no vacío $I \subseteq \mathcal{P}$ verificando que si $x \leq t \leq y$ con $x, y \in I$, entonces $t \in I$.

Para un intervalo I y un cuerpo \mathbb{k} definimos el **módulo de intervalo**, denotado $\chi^I = \{\chi_x^I, \chi_{x \leq y}^I\}$ como el módulo de persistencia:

$$\chi_x^I = \begin{cases} \mathbb{k} & \text{si } x \in I \\ 0 & \text{en otro caso} \end{cases}$$

con los morfismos de transición:

$$\chi_{x \leq y}^I = \begin{cases} 1 & \text{si } x, y \in I \\ 0 & \text{en los demás casos} \end{cases}$$

Cuando se trate de un módulo de persistencia sobre \mathbb{Z} o \mathbb{N} será frecuente trabajar con intervalos finitos de la forma $[m, n)$ e infinitos $[m, \infty)$. La explicación de esto es que estos intervalos en realidad representan la primera vez que aparece una clase en el módulo de homología; por tanto debe estar cerrado, y la primera vez en la que muere esa clase, y por tanto estará abierto pues deja de aparecer. En el caso infinito es una clase que nunca muere.

Recordamos que dados $f: M \rightarrow M'$ y $g: N \rightarrow N'$ la suma directa

$$f \oplus g: M \oplus N \rightarrow M' \oplus N'$$

está definida por :

$$(f \oplus g)(m, n) := (f(m), g(n)).$$

Definición 2.41. Sea \mathcal{P} un poset y sean \mathcal{M} y \mathcal{N} dos elementos de $Mod_R^{\mathcal{P}}$. Se define la **suma directa** de \mathcal{M} y \mathcal{N} como el elemento de $Mod_R^{\mathcal{P}}$ siguiente:

$$\mathcal{M} \oplus \mathcal{N} := ((\mathcal{M} \oplus \mathcal{N})_x, (\mathcal{M} \oplus \mathcal{N})_{x \leq y})$$

donde $(\mathcal{M} \oplus \mathcal{N})_x := M_x \oplus N_x$ y $(\mathcal{M} \oplus \mathcal{N})_{x \leq y} := M_{x \leq y} \oplus N_{x \leq y}$.

Esto se puede generalizar a sumas arbitrarias de módulos de persistencia.

Definición 2.42. Decimos que un módulo de persistencia es **descomponible** si existen módulos de persistencia \mathcal{M}_1 y \mathcal{M}_2 , distintos del módulo 0, tales que $\mathcal{M} \cong \mathcal{M}_1 \oplus \mathcal{M}_2$. Si no existen tales módulos se dice que \mathcal{M} es **indescomponible**.

En lo que sigue veremos que los módulos de intervalo son objetos indescomponibles en la categoría de módulos de persistencia y cómo se interpretan en términos de módulos graduados.

Dado un módulo de persistencia \mathcal{M} , el conjunto de todos sus endomorfismos (esto es, los morfismos de funtores $\tau : \mathcal{M} \rightarrow \mathcal{M}$) es un anillo con la suma y la composición, lo denotamos por $End(\mathcal{M})$. En el caso particular de los módulos de intervalo se tiene:

Lema 2.43. $End(\chi^I) = \mathbb{k}$, para todo \mathbb{k} -módulo de intervalo χ^I .

Demostración. Sea τ un endomorfismo. Para todo índice x , τ_x es una multiplicación por un elemento de \mathbb{k} por ser un endomorfismo de \mathbb{k} . Por ser τ un morfismo de funtores, debe cumplirse para todos $x \leq y \in \mathcal{P}$ que $1 \circ \tau_x = \tau_y \circ 1$. Por tanto tiene que ser el mismo escalar para todos los índices, luego $End(\chi^I) = \mathbb{k}$. \square

Lema 2.44. Los módulos de intervalo son indescomponibles.

Demostración. Supongamos que existe una descomposición de un módulo de intervalo $\chi^I = \mathcal{M}_1 \oplus \mathcal{M}_2$. Tomamos la proyección natural en uno de ellos:

$$\begin{aligned} \pi_{\mathcal{M}_1} : \mathcal{M}_1 \oplus \mathcal{M}_2 &\rightarrow \mathcal{M}_1 \oplus \mathcal{M}_2 \\ (m_1 \oplus m_2) &\rightarrow (m_1 \oplus 0) \end{aligned}$$

que tiene que ser idempotente, es decir $\pi_{\mathcal{M}_1} \circ \pi_{\mathcal{M}_1} = \pi_{\mathcal{M}_1}$. Pero como $\pi_{\mathcal{M}_1} \in End(\chi^I)$, el endomorfismo $\pi_{\mathcal{M}_1}$ debe ser un elemento idempotente de \mathbb{k} , esto es 0 o 1. En el primer caso $\chi^I \cong \mathcal{M}_2$ y en el segundo $\chi^I \cong \mathcal{M}_1$. \square

Para dos módulos graduados M, N su suma directa es el módulo graduado $M \oplus N$ cuyas componentes homogéneas son $(M \oplus N)_n = M_n \oplus N_n$.

Lema 2.45. Los funtores α, β definidos en 2.8 y 2.9 son aditivos, es decir preservan las sumas directas finitas.

Demostración. Es suficiente mostrarlo para dos módulos de persistencia $\mathcal{M} = \{M_i, M_{i \leq j}\}$ y $\mathcal{N} = \{N_i, N_{i \leq j}\}$. Definimos una aplicación $\alpha(\mathcal{M} \oplus \mathcal{N}) \rightarrow \alpha(\mathcal{M}) \oplus \alpha(\mathcal{N})$ por:

$$\left(\sum_{i=0}^{\infty} (n_i, n'_i) \right) \rightarrow \left(\sum_{i=0}^{\infty} n_i, \sum_{i=0}^{\infty} n'_i \right).$$

Esta aplicación es biyectiva, conserva la graduación y es un morfismo de grupos. Además, es lineal con el producto por polinomios pues se tiene que:

$$x \cdot \left(\sum_{i=0}^{\infty} (n_i, n'_i) \right) = \sum_{i=0}^{\infty} (M_{i \leq i+1} \oplus N_{i \leq i+1})(n_i, n'_i) = \left(\sum_{i=0}^{\infty} xn_i, \sum_{i=0}^{\infty} xn'_i \right) = x \cdot \left(\sum_{i=0}^{\infty} n_i, \sum_{i=0}^{\infty} n'_i \right)$$

Entonces, el functor α preserva las sumas directas y por tanto el functor β también lo hace por ser su quasi-inversa. \square

Este lema nos da la relación entre los objetos indescomponibles de cada categoría:

Lema 2.46. *Existen los siguientes isomorfismos de $\mathbb{k}[x]$ -módulos graduados:*

$$\begin{aligned} \alpha(\chi^{[i,j]}) &\cong x^i \cdot \mathbb{k}[x] / \langle x^{j-i} \rangle \\ \alpha(\chi^{[i,+\infty)}) &\cong x^i \cdot \mathbb{k}[x]. \end{aligned}$$

Demostración. Consideremos el caso de los intervalos finitos. Por definición tenemos

$$\alpha(\chi^{[i,j]}) = \bigoplus_{n=0}^{\infty} (\chi^{[i,j]})_n$$

luego su componente homogénea en grado n valdrá \mathbb{k} si $i \leq n < j$ y 0 en cualquier otro caso. Por otra parte, los elementos de $x^i \cdot \mathbb{k}[x] / \langle x^{j-i} \rangle$ son de la forma

$$x^j \cdot \sum_{l=0}^{j-i-1} a_l \cdot x^l = \sum_{l=0}^{j-i-1} a_l \cdot x^{l+i} \quad (2.15)$$

es decir, considerando la graduación usual en los polinomios, tendremos que

$$\left(x^i \cdot \mathbb{k}[x] / \langle x^{j-i} \rangle \right)_n = \begin{cases} 0 & \text{si } n < i \\ \mathbb{k} & \text{si } i \leq n < j \\ 0 & \text{si } j \leq n \end{cases}$$

Para los intervalos de la forma $[i, +\infty)$ tenemos, por un razonamiento similar, que son isomorfos a $x^m \cdot \mathbb{k}[x]$. \square

Teorema 2.47. *Los módulos de \mathbb{N} -persistencia sobre un cuerpo \mathbb{k} de tipo finito generado descomponen como suma directa de módulos de intervalo de modo único salvo el orden.*

Demostración. Sea \mathcal{M} un \mathbb{k} -módulo de \mathbb{N} -persistencia de tipo finito generado. Por el teorema de representación tenemos que $\alpha(\mathcal{M})$ es finito generado luego, por el teorema de la estructura de módulos graduados, tiene una descomposición única de la forma:

$$\alpha(\mathcal{M}) \cong \left(\bigoplus_{i=1}^{n-m} x^{a_i} \cdot \mathbb{k}[x] \right) \oplus \left(\bigoplus_{i=j}^m x^{b_i} \cdot \mathbb{k}[x] / \langle x^{q_j} \rangle \right).$$

Utilizando los Lemas 2.45 y 2.46:

$$\begin{aligned} \alpha(\mathcal{M}) &\cong \left(\bigoplus_{i=1}^{n-m} \alpha(\chi^{[a_i, +\infty)}) \right) \oplus \left(\bigoplus_{i=1}^m \alpha(\chi^{[b_j, b_j+q_j)}) \right) \\ &\cong \alpha \left(\left(\bigoplus_{i=1}^{n-m} \chi^{[a_i, +\infty)} \right) \oplus \left(\bigoplus_{i=1}^m \chi^{[b_j, b_j+q_j)} \right) \right). \end{aligned}$$

Utilizando finalmente β en ambos lados de la igualdad se tiene:

$$\mathcal{M} \cong \left(\bigoplus_{i=1}^{n-m} \chi^{[a_i, +\infty)} \right) \oplus \left(\bigoplus_{i=1}^m \chi^{[b_j, b_j+q_j)} \right). \quad (2.16)$$

Si hubiera dos descomposiciones de \mathcal{M} como suma de módulos de intervalo D_1 y D_2 entonces $\alpha(\mathcal{M}) = \alpha(D_1) = \alpha(D_2)$ y usando la unicidad del teorema de la estructura y β tenemos $D_1 = D_2$. \square

Por este resultado tenemos que dos módulos de persistencia son isomorfos si y sólo si tienen los mismo factores invariantes, es decir, descomponen como la misma suma directa de módulos de intervalo. Con el anterior teorema tenemos que la subcategoría de $Vec^{\mathbb{N}}$ formada por los objetos de tipo finito generado, junto con sus morfismos de funtores, son una categoría de Krull-Schmidt. El tercer axioma no hace falta verificarlo al tratarse únicamente de descomposiciones finitas, para las cuales la suma y el producto directo son isomorfos por ser una categoría abeliana. La razón de que sea una categoría abeliana es que $Vec_{\mathbb{k}}$ lo es y podemos hacer la construcción punto a punto conservando las propiedades.

2.6. Clasificación de módulos de persistencia sobre otros posets

La clasificación de módulos de persistencia como suma directa de módulos de intervalo puede extenderse a otros conjuntos totalmente ordenados bajo ciertas condiciones de finitud. Sin embargo, en general no se conocen resultados de estabilidad, en el sentido que mostraremos en el próximo capítulo. Una de las líneas actuales de investigación en el análisis topológico de datos es estudiar, a la vista de estos últimos resultados de clasificación, qué resultados de estabilidad pueden ser generalizados.

Sea \mathbb{k} un cuerpo y \mathcal{M} un \mathbb{k} -módulo de \mathbb{N} -persistencia o \mathbb{Z} -persistencia puntualmente finito generado, aunque no necesariamente de tipo finito. Usando el functor α de 2.8 podemos pensarlo como un módulo graduado, donde cada una de sus componentes es un

espacio vectorial finito generado. La parte de la existencia del siguiente teorema se debe a Webb [35] y la unicidad puede encontrarse en la tesis de Lesnick [24, Lemma 2.2.3.]

Teorema 2.48. *Sea M un \mathbb{Z} -módulo graduado cuyas componentes homogéneas son finito generadas. Entonces existe la descomposición única*

$$M \cong \left(\bigoplus_{j \in J} x^{a_j} \cdot \mathbb{k}[x] \right) \oplus \left(\bigoplus_{i \in I} x^{b_i} \cdot \mathbb{k}[x] / (x^{q_j}) \right).$$

donde I y J son subconjuntos (posiblemente infinitos) de \mathbb{Z} .

Con planteamientos análogos a los de la Sección 2.5 podemos adaptar el Teorema 2.47 y dar una clasificación de los módulos de persistencia sobre \mathbb{Z} .

El siguiente planteamiento da un marco general para la clasificación de un modulo de persistencia indexado por los números reales con un número finito de morfismos de transición que no sean isomorfismos. Esto permite, por ejemplo, dar un marco teórico para la clasificación de los módulos de persistencia asociados a dos nubes de datos.

Definición 2.49. Sea \mathcal{M} un elemento de $Vec^{\mathbb{R}}$ y sea $I \subseteq \mathbb{R}$ un intervalo. Decimos que \mathcal{M} es constante en I si para todo $x \leq y \in I$ el morfismo de transición $M_{x \leq y}$ es un isomorfismo. Diremos que x es un **valor regular** de \mathcal{M} si existe un entorno I_x de x , tal que \mathcal{M} es constante en I_x . Si no existe tal entorno se dice que x es un **valor crítico**. Diremos que un módulo de \mathbb{R} -persistencia, \mathcal{M} , es **moderado** si sólo existe un número finito de valores críticos para \mathcal{M} .

Observación 2.50. Un módulo de persistencia moderado es de tipo finito pues a partir de un índice todos los módulos son isomorfos. \triangle

Lema 2.51. *Si \mathcal{M} no contiene valores críticos en I , \mathcal{M} es constante en I .*

Demostración. Sean $x \leq y \in I$. Para todo $z \in [x, y]$ existe un intervalo I_z tal que $z \in I_z$ y \mathcal{M} es constante en I_z . Como $[x, y]$ es compacto, existe un número finito de valores z_1, \dots, z_k cuyos intervalos I_{z_i} recubren $[x, y]$. Sean $\{d_j\}_{j=1}^{m+1}$ con $x = d_1$ y $y = d_{m+1}$ cumpliendo que $d_j, d_{j+1} \in I_{z_l}$ para algún l . Entonces, $M_{d_j \leq d_{j+1}}$ es un isomorfismo, para todo j , y por tanto también lo es $M_{x \leq y}$. Luego es constante en I . \square

Con esto podemos clasificar los módulos de \mathbb{R} -persistencia moderados como suma de módulos de intervalo de la siguiente forma. Sea \mathcal{M} un módulo de \mathbb{R} -persistencia moderado y sea $a_1 < a_2 < \dots < a_n$ sus valores críticos. Sean $\{b_i\}$ de modo que

$b_0 < a_1 < b_1 < \dots < a_n < b_n$. Definimos la aplicación $i : [2n] \rightarrow \mathbb{R}$ por:

$$i(k) = \begin{cases} b_{k/2} & \text{si } k \text{ es par} \\ a_{(k+1)/2} & \text{si } k \text{ es impar} \end{cases}$$

Tenemos a su vez la aplicación $r : \mathbb{R} \rightarrow [2n]$ definido por:

$$r(x) = \begin{cases} 2k - 1 & \text{si } x = a_k \text{ para } k \in 1, \dots, n \\ 2k & \text{si } a_k < x < a_{k+1} \end{cases}$$

Componiendo las dos tenemos $ir : \mathbb{R} \rightarrow \mathbb{R}$:

$$ir(x) = \begin{cases} a_k & \text{si } x = a_k \text{ para } k \in 1, \dots, n \\ b_k & \text{si } a_k < x < a_{k+1} \text{ para } k \in 1, \dots, n \end{cases}$$

Con la aplicación i , pensada como un functor entre los conjuntos parcialmente ordenados, podemos clasificar usando el teorema de clasificación (Teorema 2.47) el módulo de persistencia $i^*\mathcal{M}$, luego

$$\mathcal{M} \cong (T_i T_r)^* \mathcal{M} = T_r^*(T_i^* \mathcal{M}) \cong \left(\bigoplus_{i=1}^{n_1} T_r^* \chi^{[w_i, +\infty)} \right) \oplus \left(\bigoplus_{i=1}^{n_2} T_r^* \chi^{[z_j, z_j + q_j]} \right)$$

siendo

$$T_r^* \chi^{[k,l)} = \begin{cases} \chi^{[a_{\frac{k+1}{2}}, a_{\frac{l+1}{2}}]} & \text{si } k, l \text{ son impares} \\ \chi^{[a_{\frac{k+1}{2}}, a_{\frac{l}{2}}]} & \text{si } k \text{ impar } l \text{ par} \\ \chi^{[a_{\frac{k}{2}}, a_{\frac{l+1}{2}}]} & \text{si } k \text{ par } l \text{ impar} \\ \chi^{[a_{\frac{k}{2}}, a_{\frac{l}{2}}]} & \text{si } k, l \text{ pares,} \end{cases} \quad \text{y } T_r^* \chi^{[k, +\infty)} = \begin{cases} \chi^{[a_{\frac{k+1}{2}}, +\infty)} & \text{si } k \text{ es impar} \\ \chi^{[a_{\frac{k}{2}}, +\infty)} & \text{si } k \text{ es par.} \end{cases}$$

En 2012 Crawley-Boowey demostró desde un punto de vista puramente categorial que la clasificación podía extenderse a casos más generales:

Teorema 2.52. [14, Teorema 1.1] *Todo \mathbb{k} -módulo de persistencia puntualmente finito generado sobre un conjunto totalmente ordenado descompone como suma directa de módulos de intervalo.*

Proposición 2.53 ([2, Proposition 2.1.6]). *Sea $T \subseteq \mathbb{R}$ y $\{M_j\}_{j \in J}$ una familia de módulos de persistencia T -indexados de forma que $\sum_{j \in J} M_j$ sea puntualmente finito generado.*

Entonces:

$$\bigoplus_{j \in J} M_j \cong \prod_{j \in J} M_j.$$

Con esto se tiene entonces el corolario siguiente:

Corolario 2.54. *Para $T \subseteq \mathbb{R}$, la subcategoría de Vec^T formada por elementos puntualmente finito generados es una categoría de Krull-Schmidt.*

2.7. Representaciones de los módulos de persistencia

En las secciones anteriores hemos visto la clasificación de algunos módulos de persistencia. En esta sección daremos las dos formas más comunes de representar módulos de persistencia utilizando la teoría de clasificación anterior: los códigos de barras y diagramas de persistencia. Al igual que la clasificación antes mostrada, estas representaciones no tienen porque existir para un módulo de persistencia en general.

Estas representaciones serán de utilidad para la computación y también para los teoremas de estabilidad que veremos en el siguiente capítulo. Además permiten (en el caso finito) entender de forma gráfica propiedades de los módulos de persistencia que representan. En el caso de un módulo de persistencia construido sobre homología simplicial como en el Capítulo 1, la aparición y desaparición de propiedades homológicas de la nube de datos.

Definición 2.55. Un **multiconjunto** es una pareja (S, n) , donde S es un conjunto en el que se permite la repetición de sus elementos, llamado conjunto **subyacente** de elementos, y n es una función $n : S \rightarrow \mathbb{Z}_+ \cup \{\infty\}$ que mide la multiplicidad de los elementos de S . Normalmente nos referiremos a un multiconjunto (S, n) simplemente por S .

Definición 2.56. Sea \mathcal{P} un conjunto totalmente ordenado. Un **código de barras** es un multiconjunto $\{I_j\}_{j \in J}$, donde los I_j son intervalos de \mathcal{P} .

Definición 2.57. Para un módulo de persistencia $\mathcal{M} \in \text{Mod}^{\mathbb{P}}$, su **códigos de barras asociado**, cuando exista, es el único código de barras $B_{\mathcal{M}}$ tal que:

$$\mathcal{M} \cong \bigoplus_{I \in \mathbf{B}_{\mathcal{M}}} \chi^I$$

Por la sección anterior tenemos que para cualquier subconjunto $T \subseteq \mathbb{R}$, los elementos puntualmente finito generados de Vec^T tienen un código de barras asociado. En la siguiente definición se utiliza el convenio usual de la recta real extendida $-\infty < a < \infty$ para todo $a \in \mathbb{R}$.

Definición 2.58. Un **diagrama de persistencia** es un multiconjunto de parejas (a, b) de elementos de $\mathbb{R} \cup \{\pm\infty\}$, con $a \leq b$. Para un módulo de persistencia \mathcal{M} su **diagrama de persistencia asociado**, cuando exista, es el único diagrama de persistencia formado por los extremos de los intervalos de su códigos de barras. Su representación gráfica serán los puntos en el plano de coordenadas (a, b) , que estarán por encima de la diagonal o contenidos en ella.

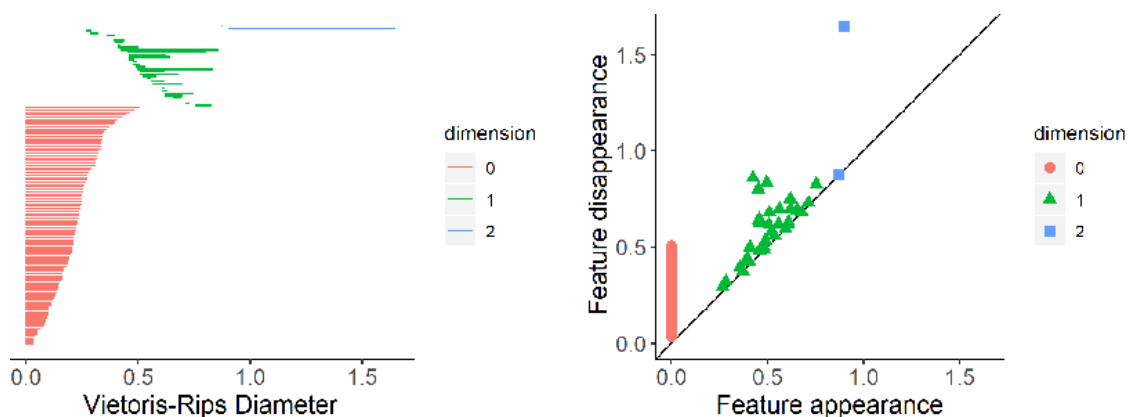


Figura 2.1: A la izquierda código de barras asociado al módulo de persistencia dado por una filtración de complejos simpliciales de Vietoris-Rips. A la derecha su diagrama de persistencia equivalente. Esta imagen es de [34]

Se ve que ambas representaciones existen o no existen de forma simultánea.

La ventaja que tiene la representación por código de barras es que permite dar información sobre distintos módulos de persistencia (la homología para distintos grado p) de forma simultánea. Para realizar esto en el diagrama de persistencia se usan distintos símbolos para los distintos módulos, pudiendo quedar solapados al representarlos en el plano. Además, al representar el diagrama de persistencia, se pierde información de la multiplicidad en el multiconjunto de cada uno de los intervalos. Esto se debe a que los mismos puntos ocupan las mismas coordenadas en el plano. Para solucionar este inconveniente puede acompañarse cada punto de un valor numérico que indica su multiplicidad.

En su aplicación al análisis topológico de datos estas representaciones permiten estudiar las propiedades de una nube de datos. En el código de barras se considerará que los intervalos más cortos son ruido y aquellos más largos son las verdaderas propiedades de la nube de datos. En el diagrama de persistencia, los puntos cercanos a la diagonal se consideran ruido mientras que aquellos lejos de ella indican propiedades de la nube de datos.

Observación 2.59. No es el caso general que, aunque sí existan formalmente el código de barras o el diagrama de persistencia, se puedan mostrar gráficamente, pues estos pueden tener infinitos elementos. \triangle

Capítulo 3

Estabilidad de los módulos de persistencia

Los conjuntos de datos en muestras del mundo real suelen ser más ruidosos de lo que se desearía, por tanto, el uso de las técnicas de la homología persistente para analizarlos requiere que éstas sigan funcionando bien aún cuando se expongan a muestras ruidosas. Se necesita entonces que pequeñas variaciones en los datos supongan también pequeñas variaciones en los invariantes y las representaciones del módulo de persistencia construido a partir de ellos. En este capítulo expondremos algunos resultados sobre la estabilidad, en el sentido que precisaremos más adelante, de las estructuras estudiadas en las anteriores secciones. Para ello necesitamos introducir una noción de distancia en la categoría de módulos de persistencia. Esto se hará desde el enfoque de la teoría de categorías y, en la medida de lo posible, daremos las definiciones y resultados más generales conocidos. Un ejemplo de esto es la noción de intercalado generalizado que presentaremos a continuación siguiendo [25]. Después, nos centraremos en el caso de los módulos de \mathbb{R} -persistencia sobre el que demostrará un teorema de isometría.

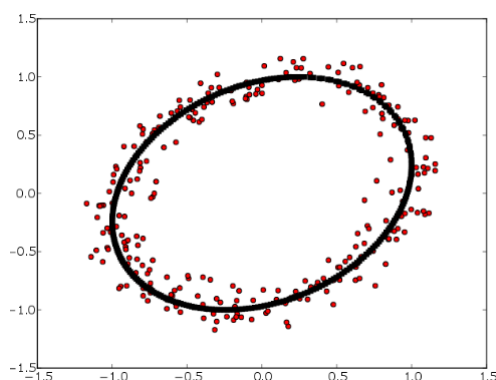


Figura 3.1: Puede esperarse que al estar en cierta manera cercanas la nube de datos formada por los puntos rojos y negros, también lo estén sus códigos de barras. Imagen de [28]

3.1. Intercalados generalizados

Sean \mathcal{M} y $\mathcal{N} \in \text{Mod}_R^{\mathcal{P}}$. Recordemos que \mathcal{M} y \mathcal{N} son isomorfos si existen morfismos de funtores $\phi \in \text{Hom}_{\text{Mod}_R^{\mathcal{P}}}(\mathcal{M}, \mathcal{N})$ y $\psi \in \text{Hom}_{\text{Mod}_R^{\mathcal{P}}}(\mathcal{N}, \mathcal{M})$ tales que su composición por la izquierda y por la derecha es la identidad. En caso de que exista una descomposición en módulos de intervalo, serán isomorfos si descomponen como la misma suma de módulos de intervalo. Aunque es una noción que permite comparar módulos de persistencia, resulta demasiado fuerte. Esto se debe a que solo admite dos posibilidades, isomorfos o no isomorfos.

Ejemplo 3.1. Sea \mathcal{U} el módulo definido por $\{\mathbb{R}_x, (Id_{\mathbb{R}})_{x \leq y}\}_{x \in \mathbb{Z}}$. Sean $\mathcal{V} = \{\mathbb{C}_x^2, (Id_{\mathbb{C}}^2)_{x \leq y}\}_{x \in \mathbb{Z}}$ y \mathcal{W} con $W_x = U_x \forall x \in \mathbb{Z} \setminus \{0\}$, $W_0 = 0$. Ni \mathcal{V} ni \mathcal{W} son isomorfos a \mathcal{U} . Haciendo un único cambio en V_0 podemos obtener \mathcal{U} pero en el caso de \mathcal{W} debemos cambiar el módulo en todos los índices.

En el ejemplo anterior \mathcal{V} parece más "cerca" de ser isomorfo a \mathcal{U} que \mathcal{W} . En esta sección buscaremos una noción de isomorfismo aproximado que formalice esto.

Usaremos la composición vertical y horizontal de morfismos de funtores de la definiciones 2.7 y 2.8. Además se usa el functor asociado a una aplicación monótona de la Definición 2.17.

Definición 3.2. Sea \mathcal{P} un poset. Diremos que una aplicación monótona $\sigma : \mathcal{P} \rightarrow \mathcal{P}$ es una **traslación** si $\sigma(x) \geq x$ para todo $x \in \mathcal{P}$.

Cualquier traslación $\sigma : \mathcal{P} \rightarrow \mathcal{P}$ define un morfismo de funtores:

$$\eta_{\sigma} : Id_{\mathcal{P}} \rightarrow T_{\sigma}$$

donde, para cada $x \in \mathcal{P}$, $(\eta_{\sigma})_x$ es el único morfismo que hay $x \rightarrow \sigma(x)$.

Observación 3.3. Nótese que para un objeto $\mathcal{M} \in \mathcal{C}^{\mathcal{P}}$, donde \mathcal{C} es una categoría arbitraria, $(Id_{\mathcal{M}} * \eta_{\sigma})_x$ es el morfismo de transición $M_{x \leq \sigma(x)}$. \triangle

Si τ es además de monótona biyectiva, su inversa τ^{-1} es también monótona.

Lema 3.4. Sea \mathcal{P} un poset y $\mathcal{M} = (M_x, M_{x \leq y})$ un objeto \mathcal{P} -persistente.

1. Para $\tau, \sigma : \mathcal{P} \rightarrow \mathcal{P}$ aplicaciones monótonas. Entonces $T_{\tau} \circ T_{\sigma} = T_{\tau \circ \sigma}$. Además $T_{\tau}^*(T_{\sigma}^*(\mathcal{M})) = T_{\tau \circ \sigma}^*(\mathcal{M})$.
2. Sea $\tau : \mathcal{P} \rightarrow \mathcal{P}$ monótona y biyectiva. Entonces $Id_{\mathcal{P}} = T_{\tau} \circ T_{\tau}^{-1}$ y $T_{\tau}^*(T_{\tau}^{-1}^*(\mathcal{M})) = T_{\tau \circ \tau^{-1}}^*(\mathcal{M}) = \mathcal{M}$.

3. La composición de las transformaciones naturales asociadas a aplicaciones monótonas τ, σ verifica que: $\eta_\tau \eta_\sigma = \eta_{\tau \circ \sigma}$.

Este lema implica que cuando tenemos una transformación monótona y biyectiva, τ , el objeto $\mathcal{M} \in \mathcal{C}^{\mathcal{P}}$ se puede recuperar a partir de $T_\tau^*(\mathcal{M})$.

Definición 3.5. Sean $\tau, \sigma : \mathcal{P} \rightarrow \mathcal{P}$ dos aplicaciones monótonas en un poset \mathcal{P} . Decimos que (τ, σ) es una **pareja de traslación** si $x \leq (\tau \circ \sigma)(x)$ y $x \leq (\sigma \circ \tau)(x)$ para todo $x \in \mathcal{P}$.

Proposición 3.6. Sea \mathcal{P} un poset y $\sigma, \tau, \rho, \pi : \mathcal{P} \rightarrow \mathcal{P}$ aplicaciones monótonas. Si (τ, σ) y (ρ, π) son parejas de traslación, $(\rho \circ \tau, \sigma \circ \pi)$ es una pareja de traslación.

Demostración. Probamos que $(\rho \circ \tau \circ \sigma \circ \pi)(x) \geq x$. La otra desigualdad se prueba de forma análoga. Se verifica que $(\tau \circ \sigma)(\pi(x)) \geq \pi(x)$ por ser (τ, σ) una pareja de traslación, y como ρ es monótona $\rho((\tau \circ \sigma)(\pi(x))) \geq \rho(\pi(x))$. Acabamos por ser (ρ, π) una pareja de traslación. \square

Definición 3.7. Sea \mathcal{C} una categoría y $\mathcal{M}, \mathcal{N} \in \mathcal{C}^{\mathcal{P}}$, para un poset \mathcal{P} . Si (τ, σ) es una pareja de traslación, decimos que \mathcal{M} y \mathcal{N} están **(τ, σ) -intercalados** si existe una pareja de morfismos de funtores: $\varphi : \mathcal{M} \rightarrow T_\tau^*(\mathcal{N})$ y $\psi : \mathcal{N} \rightarrow T_\sigma^*(\mathcal{M})$:

$$\begin{array}{ccccc}
 \mathcal{P} & \xrightarrow{T_\tau} & \mathcal{P} & \xrightarrow{T_\sigma} & \mathcal{P} \\
 \downarrow M & \searrow T_\tau^*(\mathcal{N}) & \downarrow N & \searrow T_\sigma^*(\mathcal{M}) & \downarrow M \\
 \mathcal{C} & \xrightarrow{\varphi} & \mathcal{C} & \xrightarrow{\psi} & \mathcal{C}
 \end{array}$$

verificando las siguientes igualdades, llamadas **condiciones de conmutatividad**:

$$(\psi * Id_{T_\tau}) \circ \varphi = Id_{\mathcal{M}} * \eta_{(\sigma \circ \tau)} \quad \text{y} \quad (\varphi * Id_{T_\sigma}) \circ \psi = Id_{\mathcal{N}} * \eta_{(\tau \circ \sigma)}.$$

La pareja (ψ, φ) se llama un **(τ, σ) -intercalado generalizado** entre \mathcal{M} y \mathcal{N} .

El intercalado mide cómo de lejos están dos módulos de persistencia de ser isomorfos. De forma intuitiva, un (σ, τ) -intercalado entre dos objetos de persistencia \mathcal{M} y \mathcal{N} da una factorización de cada morfismo de transición de manera que los siguientes diagramas sean conmutativos:

$$\begin{array}{ccc}
 M_x & \xrightarrow{M_{x \leq (\sigma \circ \tau)(x)}} & M_{(\sigma \circ \tau)(x)} \\
 \searrow \varphi_x & & \nearrow \psi_{\tau(x)} \\
 & N_{\tau(x)} &
 \end{array}
 \qquad
 \begin{array}{ccc}
 & M_{\sigma(x)} & \\
 \nearrow \psi_x & & \searrow \psi_{\sigma(x)} \\
 N_x & \xrightarrow{N_{x \leq (\tau \circ \sigma)(x)}} & N_{(\tau \circ \sigma)(x)}
 \end{array}$$

La siguiente proposición es una versión de la desigualdad triangular para intercalados generalizados.

Proposición 3.8. *Sea \mathcal{C} una categoría, \mathcal{P} un poset y $\mathcal{U}, \mathcal{V}, \mathcal{W} \in \mathcal{C}^{\mathcal{P}}$. Sean las aplicaciones monótonas $\tau, \sigma, \rho, \pi: \mathcal{P} \rightarrow \mathcal{P}$. Si \mathcal{U}, \mathcal{V} están (τ, σ) -intercalados y \mathcal{V}, \mathcal{W} están (ρ, π) -intercalados, entonces \mathcal{U} y \mathcal{W} están $(\rho \circ \tau, \sigma \circ \pi)$ -intercalados.*

Demostración. Por la Proposición 3.6 $(\rho \circ \tau, \sigma \circ \pi)$ forma una pareja de traslación, luego tiene sentido definir un intercalado.

Como \mathcal{U} y \mathcal{V} están (τ, σ) -intercalados existe un intercalado (ψ_1, φ_1) con $\varphi_1: \mathcal{U} \rightarrow T_{\tau}^*(\mathcal{V})$, $\psi_1: \mathcal{V} \rightarrow T_{\sigma}^*(\mathcal{U})$, cumpliendo las condiciones de conmutatividad:

$$(\psi_1 * Id_{T_{\tau}}) \circ \varphi_1 = Id_{\mathcal{U}} * \eta_{\sigma \circ \tau} \text{ y } (\varphi_1 * Id_{T_{\sigma}}) \circ \psi_1 = Id_{\mathcal{V}} * \eta_{\tau \circ \sigma}.$$

Análogamente para \mathcal{V} y \mathcal{W} tenemos que (ψ_2, φ_2) con $\varphi_2: \mathcal{V} \rightarrow T_{\rho}^*(\mathcal{W})$, $\psi_2: \mathcal{W} \rightarrow T_{\pi}^*(\mathcal{V})$, verificando:

$$(\psi_2 * Id_{T_{\rho}}) \circ \varphi_2 = Id_{\mathcal{V}} * \eta_{\pi \circ \rho} \text{ y } (\varphi_2 * Id_{T_{\pi}}) \circ \psi_2 = Id_{\mathcal{W}} * \eta_{\rho \circ \pi}.$$

Definimos el siguiente morfismo de funtores:

$$\begin{aligned} \varphi &:= (\varphi_2 * Id_{T_{\tau}}) \circ \varphi_1 \\ \psi &:= (\psi_1 * Id_{T_{\pi}}) \circ \psi_2 \end{aligned}$$

Hay que probar que (ψ, φ) es un $(\rho \circ \tau, \sigma \circ \pi)$ -intercalado entre \mathcal{U} y \mathcal{W} . Primero tenemos vemos que ambas transformaciones naturales están bien definidas en el sentido de realizar las transformaciones correctas. Con la composición vertical y horizontal tenemos para φ :

$$\mathcal{U} \xrightarrow{\varphi_1} \mathcal{V}T_{\tau} \xrightarrow{\varphi_2 * Id_{T_{\tau}}} \mathcal{W}T_{\rho}T_{\tau} = \mathcal{W}T_{(\rho \circ \tau)}$$

Análogamente para ψ :

$$\mathcal{W} \xrightarrow{\psi_2} \mathcal{V}T_{\pi} \xrightarrow{\psi_1 * Id_{T_{\pi}}} \mathcal{U}T_{\sigma}T_{\pi} = \mathcal{U}T_{(\sigma \circ \pi)}$$

Falta ver que ambas cumplen las condiciones de conmutatividad:

$$\begin{aligned} (\psi * Id_{T_{(\rho \circ \tau)}}) \circ \varphi &= Id_{\mathcal{U}} * \eta_{(\sigma \circ \pi \circ \rho \circ \tau)} \\ (\varphi * Id_{T_{(\sigma \circ \pi)}}) \circ \psi &= Id_{\mathcal{W}} * \eta_{(\rho \circ \tau \circ \sigma \circ \pi)} \end{aligned}$$

Se da la demostración de una de ellas ya que la otra se prueba de forma análoga. Pensando la transformación natural como la familia de morfismos que la define, basta probar que $\forall x \in \mathcal{P}$:

$$\left((\psi * Id_{T_{(\rho \circ \tau)}}) \circ \varphi \right) (x) = (Id_{\mathcal{U}} * \eta_{(\sigma \circ \pi \circ \rho \circ \tau)}) (x)$$

Usando la observación 3.3 se tiene que $(Id_{\mathcal{U}} * \eta_{(\sigma \circ \pi \circ \rho \circ \tau)})(x) = M_{x \leq (\sigma \circ \pi \circ \rho \circ \tau)(x)}$, así que basta ver que el siguiente diagrama es conmutativo:

$$\begin{array}{ccccc}
 U_x & \xrightarrow{U_{x \leq (\sigma \circ \tau)(x)}} & U_{(\sigma \circ \tau)(x)} & \xrightarrow{U_{(\sigma \circ \tau)(x) \leq (\sigma \circ \pi \circ \rho \circ \tau)(x)}} & U_{(\sigma \circ \pi \circ \rho \circ \tau)(x)} \\
 (\varphi_1)_x \searrow & & \nearrow (\psi_1)_{\tau(x)} & & \nearrow (\psi_1)_{(\pi \circ \rho \circ \tau)(x)} \\
 & & V_{\tau(x)} & \xrightarrow{V_{\tau(x) \leq (\pi \circ \rho \circ \tau)(x)}} & V_{(\pi \circ \rho \circ \tau)(x)} \\
 & & \searrow (\varphi_2)_{\tau(x)} & & \nearrow (\psi_2)_{(\rho \circ \tau)(x)} \\
 & & & & W_{(\rho \circ \tau)(x)}
 \end{array}$$

Los dos triángulos son conmutativos por ser (ψ_1, φ_1) y (ψ_2, φ_2) intercalados y cumplir las condiciones de conmutatividad. El paralelogramo también lo es por ser ψ_1 una transformación natural. \square

Un caso de particular importancia y que ocupará el resto de la sección es que el conjunto ordenado sea $T \subseteq \mathbb{R}$ con el orden usual. En este caso, usando la suma tenemos que existe el tipo de pareja de traslación:

$$\tau(x) = x + \epsilon = \sigma(x)$$

Es claro que por ser cada una de ellas una traslación, forman una pareja de traslación luego está bien definida. Para esta pareja de traslación se dice que \mathcal{M} y \mathcal{N} están ϵ -intercalados y se podrá definir un espacio métrico. La notación será T_ϵ para T_σ y η_ϵ para η_σ , cuando la aplicación monótona sea $\sigma(x) = x + \epsilon$.

Una **distancia** en un conjunto X es una función $d : X \times X \rightarrow \mathbb{R}$ verificando:

1. $d(x, y) = 0 \Leftrightarrow x = y \forall x, y \in X$.
2. $d(x, y) = d(y, x) \forall x, y \in X$.
3. $d(x, z) \leq d(x, y) + d(y, z) \forall x, y, z \in X$ (**.desigualdad triangular**)

Será una **distancia extendida** si verifica los tres puntos anteriores pero valora en $\mathbb{R} \cup \{+\infty\}$. Será una **pseudodistancia extendida** si valora en $\mathbb{R} \cup \{\infty\}$, cumple los axiomas 2 y 3 de distancia y el primero se cambia por $d(x, x) = 0$ y $d(x, y) \geq 0 \forall x, y \in X$.

Definición 3.9. Sean dos objetos \mathbb{N} -persistentes \mathcal{M} y \mathcal{N} . Se define su **distancia de intercalado** como:

$$d_I(\mathcal{M}, \mathcal{N}) := \inf\{\epsilon \geq 0 \text{ tal que } \mathcal{M} \text{ y } \mathcal{N} \text{ están } \epsilon\text{-intercalados}\}$$

Si no existe ningún valor ϵ finito para el cual exista un ϵ -intercalado, se fija que $d_I(\mathcal{M}, \mathcal{N}) = \infty$

Proposición 3.10. *Sea \mathcal{C} una categoría. La distancia de intercalado define una distancia extendida en $\mathcal{C}^{\mathbb{N}}$.*

Demostración. Es claro que $d(\mathcal{M}, \mathcal{M}) = 0$. Si $d(\mathcal{M}, \mathcal{N}) = 0$ tenemos (por estar considerando módulos de persistencia sobre los números naturales) que el ínfimo se alcanza, luego son isomorfos pues el 0-intercalado nos da la conmutatividad del diagrama:

$$\begin{array}{ccccc} M_x & \xrightarrow{Id} & M_x & \xrightarrow{Id} & M_x \\ \varphi \downarrow & & \psi \uparrow & & \downarrow \varphi \\ N_x & \xrightarrow{Id} & N_x & \xrightarrow{Id} & N_x \end{array}$$

Por tanto, $\varphi \circ \psi = Id$ y $\psi \circ \varphi = Id$.

La simetría es inmediata por la definición y la desigualdad triangular se obtiene de la Proposición 3.8. \square

La razón por la cual la definimos como distancia extendida es que existen categorías con objetos de persistencia para los cuales no existe un ϵ -intercalado para ningún ϵ finito. Un ejemplo de ello es $Mod^{\mathbb{N}}$. En efecto, dados los módulos de persistencia \mathcal{M} y \mathcal{N} definidos por:

$$\begin{array}{ccccccc} \dots & \rightarrow & \mathbb{R}^2 & \rightarrow & \mathbb{R}^2 & \rightarrow & \mathbb{R}^2 & \rightarrow & \dots \\ \dots & \rightarrow & \mathbb{R}^3 & \rightarrow & \mathbb{R}^3 & \rightarrow & \mathbb{R}^3 & \rightarrow & \dots \end{array}$$

se tiene que $d(\mathcal{M}, \mathcal{N}) = \infty$. Si existiera un ϵ -intercalado (ψ, φ) entre esos módulos de persistencia debería ocurrir, por las condiciones de conmutatividad, que $\varphi \circ \psi = Id_{\mathbb{R}^3}$, pero esto es imposible porque $\psi : \mathbb{R}^3 \rightarrow \mathbb{R}^2$ no es inyectivo y, por tanto, tampoco la composición.

Por otra parte, en el caso de objetos $\mathcal{M}, \mathcal{N} \in \mathcal{C}^{\mathbb{R}}$ al tomar ínfimo no tenemos garantizado que si $d_I(\mathcal{M}, \mathcal{N}) = 0$ se cumpla que $\mathcal{M} \cong \mathcal{N}$ como se muestra en el siguiente ejemplo:

Ejemplo 3.11. Sean los objetos $\mathcal{M}, \mathcal{N} \in \mathcal{C}^{\mathbb{R}}$ donde $\mathcal{M} = 0$, es decir $M_x = 0 \forall x \in \mathbb{R}$, y

$$\mathcal{N} = \begin{cases} \mathbb{k} & \text{si } x = 0 \\ 0 & \text{en los demás casos} \end{cases}$$

En ambos sus morfismos de transición son siempre el morfismo 0. Entonces para todo $\epsilon > 0$ podemos encontrar una pareja de ϵ -intercalado, luego el ínfimo de los valores para los cuales existe un intercalado es 0. No obstante no son isomorfos pues en su descomposición como módulos de intervalo, la cual es posible por tener un número finito de valores críticos, \mathcal{N} descompone como $[0, 0]$ y \mathcal{M} como el conjunto vacío.

Si es cierto, no obstante, que $d_I(\mathcal{M}, \mathcal{N}) = 0$ si $\mathcal{M} \cong \mathcal{N}$, pues el propio isomorfismo nos da un 0-intercalado. Esto nos permitirá definir una pseudodistancia extendida. Para ello necesitamos el siguiente lema que muestra que los valores para los cuales los objetos de $\mathcal{C}^{\mathbb{R}}$ están intercalados forman un rayo de la recta real.

Lema 3.12. *Si dos objetos de $\mathcal{C}^{\mathbb{R}}$ están ϵ -intercalados, entonces están $\epsilon + \delta$ -intercalados para todo $\delta > 0$.*

Demostración. Sean \mathcal{M} y \mathcal{N} dos objetos ϵ -intercalados y sea (ψ, φ) un ϵ -intercalado. Fijamos un $\delta > 0$. Entonces tenemos la transformación natural

$$\eta_\delta * Id_{T_\epsilon} : T_\epsilon \rightarrow T_\delta T_\epsilon = T_{\epsilon+\delta}$$

luego $Id_{\mathcal{N}} * (\eta_\delta * Id_{T_\epsilon}) : \mathcal{N}T_\epsilon \rightarrow \mathcal{N}T_{\epsilon+\delta}$. Definimos $\hat{\varphi} = (Id_{\mathcal{N}} * (\eta_\delta * Id_{T_\epsilon})) \circ \varphi$. Análogamente definimos $\hat{\psi} = (Id_{\mathcal{M}} * (\eta_\delta * Id_{T_\epsilon})) \circ \psi$. Así definidos se cumplen las condiciones de conmutatividad $(\hat{\varphi} * Id_{T_{\epsilon+\delta}}) \circ \varphi = Id_{\mathcal{M}} * \eta_{2(\epsilon+\delta)}$ y $(\hat{\psi} * Id_{T_{\epsilon+\delta}}) \circ \psi = Id_{\mathcal{N}} * \eta_{2(\epsilon+\delta)}$. \square

Proposición 3.13. *El operador d_I define una pseudodistancia extendida en $\mathcal{C}^{\mathbb{R}}$.*

Demostración. La simetría y su no negatividad son inmediatas de la definición.

Necesitamos demostrar la desigualdad triangular. Sean $\mathcal{U}, \mathcal{V}, \mathcal{W} \in \mathcal{C}^{\mathbb{R}}$ de manera que $d_I(\mathcal{U}, \mathcal{V}) = \epsilon$ y $d_I(\mathcal{V}, \mathcal{W}) = \epsilon'$. Por el Lema 3.12 tenemos que para todo $\delta > 0$ existe un $\epsilon + \delta$ -intercalado de \mathcal{U} y \mathcal{V} , y un $\epsilon' + \delta$ -intercalado de \mathcal{V} y \mathcal{W} . Por la Proposición 3.8 existe un $\epsilon + \epsilon' + 2\delta$ -intercalado de \mathcal{U} y \mathcal{W} , como podemos hacer δ tan pequeño como queramos se tiene que $d_I(\mathcal{U}, \mathcal{W}) \leq \epsilon + \epsilon'$. \square

Corolario 3.14. *Sea la relación de equivalencia en $\text{Obj}(\mathcal{C}^{\mathbb{R}})$*

$$M \sim N \Leftrightarrow d_I(M, N) = 0.$$

En el conjunto cociente definido por esta relación la distancia de intercalado define una distancia generalizada.

Demostración. Para dos clases de equivalencia $[\mathcal{M}]$ y $[\mathcal{N}]$ se define

$$d_I([\mathcal{M}], [\mathcal{N}]) = d_I(\mathcal{M}, \mathcal{N})$$

Hay que ver que está bien definido. Si $\mathcal{M}' \in [\mathcal{M}]$ tenemos por la desigualdad triangular:

$$\begin{aligned} d_I(\mathcal{M}, \mathcal{N}) &\leq d_I(\mathcal{M}, \mathcal{M}') + d_I(\mathcal{M}', \mathcal{N}) = d_I(\mathcal{M}', \mathcal{N}) \\ d_I(\mathcal{M}', \mathcal{N}) &\leq d_I(\mathcal{M}', \mathcal{M}) + d_I(\mathcal{M}, \mathcal{N}) = d_I(\mathcal{M}, \mathcal{N}) \end{aligned}$$

Por simetría se ve el mismo resultado para \mathcal{N} . Los demás axiomas de distancia son inmediatos de la proposición anterior. \square

La siguiente proposición nos muestra que al pasar por un functor dos objetos persistentes, la distancia de intercalado entre sus imágenes queda acotada superiormente por la distancia de intercalado entre los objetos. Esto será necesario para un resultado posterior, consecuencia del teorema de estabilidad.

Proposición 3.15. *Sean \mathcal{C} y \mathcal{D} dos categorías y $H : \mathcal{C} \rightarrow \mathcal{D}$ un functor. Para dos objetos $\mathcal{M}, \mathcal{N} \in \mathcal{C}^{\mathbb{R}}$ se verifica:*

$$d_I(H\mathcal{M}, H\mathcal{N}) \leq d_I(\mathcal{M}, \mathcal{N}).$$

Luego, en particular, si \mathcal{M} y \mathcal{N} están ϵ -intercalados también lo estarán $H\mathcal{M}$ y $H\mathcal{N}$.

Demostración. Si $d_I(\mathcal{M}, \mathcal{N}) = \infty$ es inmediato. Supongamos que $d_I(\mathcal{M}, \mathcal{N}) < \infty$. Si \mathcal{M} y \mathcal{N} están ϵ -intercalados existen morfismos de funtores $\varphi : \mathcal{M} \rightarrow \mathcal{N}T_\epsilon$ y $\psi : \mathcal{N} \rightarrow \mathcal{M}T_\epsilon$ que forman un de ϵ -intercalado con las correspondientes condiciones de conmutatividad:

$$(\psi * Id_{T_\epsilon}) \circ \varphi = Id_{\mathcal{M}} * \eta_{2\epsilon} \text{ y } (\varphi * Id_{T_\epsilon}) \circ \psi = Id_{\mathcal{N}} * \eta_{2\epsilon}.$$

Componiendo con H se tiene $Id_H * \varphi : H\mathcal{M} \rightarrow H\mathcal{N}T_\epsilon$ y $Id_H * \psi : H\mathcal{N} \rightarrow H\mathcal{M}T_\epsilon$ que deberán cumplir:

$$(Id_H * \psi * T_\epsilon) \circ (Id_H * \varphi) = Id_{(H\mathcal{M})} * \eta_{2\epsilon} \text{ y } (Id_H * \varphi * T_\epsilon) \circ (Id_H * \psi) = Id_{(H\mathcal{N})} * \eta_{2\epsilon}.$$

Demostramos la igualdad de la izquierda y la otra se prueba de forma análoga. Por la observación 3.3 tenemos que para todo $x \in \mathbb{R}$:

$$(Id_{(H\mathcal{M})} * \eta_{2\epsilon})_x = (H\mathcal{M})_{x \leq x+2} = H(M_{x \leq x+2})$$

Por otra parte, usando la definición de composición horizontal:

$$\begin{aligned} (Id_H * \psi * Id_{T_\epsilon})_x \circ (Id_H * \varphi)_x &= (Id_H)_{(\mathcal{M}T_\epsilon)_x} \circ H((\psi * Id_{T_\epsilon})_x) \circ (Id_H)_{(\mathcal{N}T_\epsilon)_x} \circ H(\varphi_x) \\ &= (Id_H)_{(M_{x+\epsilon})} \circ H((\psi * Id_{T_\epsilon})_x) \circ (Id_H)_{(N_{x+\epsilon})} \circ H(\varphi_x) \\ &= H((\psi * Id_{T_\epsilon})_x) \circ H(\varphi_x) \\ &= H(\psi * Id_{T_\epsilon})_x \circ \varphi_x \\ &= H(M_{x \leq x+2}) \end{aligned}$$

donde la última igualdad se deduce de las condiciones de conmutatividad de (ψ, φ) . \square

3.2. Distancia de cuello de botella

Una vez explicada la distancia de intercalado, que permite comparar los módulos de persistencia (en general cualquier objeto persistente), buscamos una forma de comparar

sus códigos de barras asociados, para lo que desarrollaremos la distancia de cuello de botella. Además, se esperará que si dos módulos de persistencia están próximos, también lo estén sus códigos de barras asociados.

Denotemos por $\langle a, b \rangle$ a cualquiera de los intervalos (a, b) , $(a, b]$, $[a, b)$, $[a, b]$ de la recta real extendida $(\mathbb{R} \cup \{\pm\infty\})$, con $a \leq b$. Si $a \in \mathbb{R}$ se fijan los convenios de $\mathbb{R} \cup \{\pm\infty\}$: $\infty - a = \infty$, $-\infty + a = -\infty$, $\infty - \infty = 0$, $-\infty - (-\infty) = 0$ y $-\infty < a < \infty$. En esta sección trabajaremos con multiconjuntos de elementos de la forma $\langle a, b \rangle$, a los que llamaremos multiconjuntos de intervalos de $\mathbb{R} \cup \{\pm\infty\}$. Únicamente serán multiconjuntos de cardinalidad finita, aunque la mayoría de los detalles pueden adaptarse para el caso de infinitos intervalos y, en general, a cualquier categoría de Krull-Schmidt [2, Definition 2.1.16].

Definición 3.16. Sean B y B' dos multiconjuntos de intervalos de $\mathbb{R} \cup \{\pm\infty\}$. Una **coincidencia parcial** entre B y B' es un subconjunto (incluyendo el conjunto vacío \emptyset) de $B \times B'$:

$$\Omega = \{(I, I') \in B \times B'\}$$

donde cada elemento de B y B' se encuentra como máximo en una pareja de Ω .

Para $I \in B$ y $I' \in B'$ con $(I, I') \in \Omega$ decimos que I coincide con I' . Si un intervalo de B o B' no está en Ω se dice que no coincide con ninguno. El subconjunto de $B \sqcup B'$ formado por intervalos que no coincidan con ninguno se denotará Ω^C .

Ejemplo 3.17. Sean $B = \{I_1, I_2, I_2, I_3, I_4\}$ y $B' = \{J_1, J_1, J_2, J_3\}$ dos multiconjuntos de intervalos de $\mathbb{R} \cup \{\pm\infty\}$. El subconjunto $\Omega_1 = \{(I_1, J_1), (I_2, J_2), (I_3, J_3)\}$ es una coincidencia parcial pues cada elemento de B y B' aparece máximo en una pareja. Por otro lado, $\Omega_2 = \{(I_1, J_1), (I_1, J_1), (I_3, J_3)\}$ no es una coincidencia parcial porque solo hay una copia de I_1 en B y aparece en dos parejas. $\Omega_3 = \{(I_1, J_1), (I_2, J_1)\}$ también es una coincidencia parcial pues J_1 aparece dos veces en B' . Este último ejemplo muestra que aunque formalmente, como intervalos en la recta real extendida, dos elementos de un multiconjunto sean el mismo, se consideran dos elementos distintos del conjunto.

Si B y B' son dos multiconjuntos de intervalos de $\mathbb{R} \cup \{\pm\infty\}$, una forma equivalente de entender una coincidencia parcial es como una biyección entre un subconjunto de B y un subconjunto de B' .

Definición 3.18. Sea Ω una coincidencia parcial entre B y B' , se definen

$$\text{dom}(\Omega) = \{I \in B \text{ tales que existe } I' \in B' \text{ con } (I, I') \in \Omega\}$$

$$\text{Im}(\Omega) = \{I' \in B' \text{ tales que existe } I \in B \text{ con } (I, I') \in \Omega\}$$

Equivalentemente, una coincidencia parcial entre dos multiconjuntos B y B' es una biyección $\Omega : \text{dom}(\Omega) \xrightarrow{\sim} \text{Im}(\Omega)$ donde si $I \in \text{dom}(\Omega)$, $\Omega(I)$ es el único $I' \in \text{Im}(\Omega)$ tal que $(I, I') \in \Omega$. Por simplicidad, esta biyección se denota $\Omega : B \rightleftharpoons B'$.

Definición 3.19. Sean B y B' dos multiconjuntos de intervalos de $\mathbb{R} \cup \{\pm\infty\}$ y Ω una coincidencia parcial entre ellos. Sean $\langle a, b \rangle = I \in B$ y $\langle c, d \rangle = I' \in B'$ con $(I, I') \in \Omega$ y $\langle e, f \rangle = I'' \in \Omega^C$, definimos:

$$c_{\Omega}(I, I') = \max(|c - a|, |d - b|) \quad (3.1)$$

$$c_{\Omega}(I'') = (f - e) / 2 \quad (3.2)$$

Para la coincidencia parcial Ω definimos su función de **coste** como:

$$c(\Omega) = \max \left(\sup_{(I, I') \in \Omega} c_{\Omega}(I, I'), \sup_{I'' \in \Omega^C} c_{\Omega}(I'') \right) \quad (3.3)$$

Para $\epsilon \in \mathbb{R} \cup \{\pm\infty\}$ decimos que una coincidencia parcial Ω es una ϵ -**coincidencia** si $c(\Omega) \leq \epsilon$.

En el caso de que los multiconjuntos B y B' sean códigos de barras, una ϵ -coincidencia Ω puede interpretarse geoméricamente. Si representamos el diagrama de persistencia asociado a ese código de barras, conectamos los puntos que están en Ω y por tanto se encuentran más próximos que ϵ . Los puntos restantes están a una distancia de la diagonal menor que ϵ . En el caso de 3.2 los tres puntos no conectados a otros puntos se encuentran a menos de ϵ de diagonal.

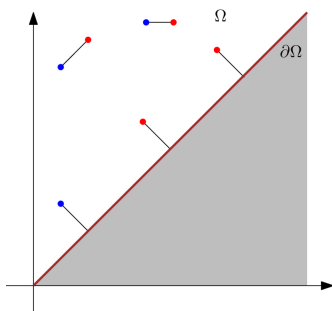


Figura 3.2: interpretación gráfica de la distancia de cuello de botella asociada a un código de barras. Esta imagen está extraída de [16]

Definición 3.20. Sean B, B' y B'' multiconjuntos de intervalos de $\mathbb{R} \cup \{\pm\infty\}$. Sea Ω una coincidencia parcial entre B y B' y Ω' una coincidencia parcial entre B' y B'' . Podemos

definir una coincidencia parcial entre B y B'' llamada **composición**, denotada $\Omega \circ \Omega'$, como el subconjunto de $B \times B''$:

$$\Omega' \circ \Omega = \begin{cases} \{(I, I'') \in B \times B'' \text{ tales que } \Omega(I) = \Omega'^{-1}(I'')\} \\ \emptyset \text{ si el anterior conjunto es vacío} \end{cases}$$

Nótese que para $I \in B$ y $J'' \in B''$:

$$\begin{aligned} I \in \text{dom}(\Omega' \circ \Omega) &\Leftrightarrow \text{existe } I'' \in B'' \text{ con } \Omega(I) = \Omega'^{-1}(I'') \\ J'' \in \text{Im}(\Omega' \circ \Omega) &\Leftrightarrow \text{existe } J \in B \text{ con } \Omega(J) = \Omega'^{-1}(J'') \end{aligned}$$

En el primer caso $(\Omega' \circ \Omega)(I) = I''$ y en el segundo $(\Omega' \circ \Omega)^{-1}(J'') = J$.

Proposición 3.21. *Sean B, B' y B'' multiconjuntos de intervalos de \mathbb{R} . Si Ω es una ϵ -coincidencia entre B y B' y Ω' es una ϵ' -coincidencia entre B' y B'' ; su composición es una $\epsilon + \epsilon'$ -coincidencia entre B y B'' .*

Demostración. Sean $I = \langle a, b \rangle \in B, I' = \langle a', b' \rangle \in B'$ y $I'' = \langle a'', b'' \rangle \in B''$. Tenemos tres posibilidades:

1. En primer lugar supongamos que I coincide con I' y que I' coincide con I'' . En este caso se tiene que

$$\begin{aligned} |a'' - a| &\leq |a'' - a'| + |a' - a| \leq \epsilon + \epsilon' \\ |b'' - b| &\leq |b'' - b'| + |b' - b| \leq \epsilon + \epsilon'. \end{aligned}$$

2. En segundo lugar supongamos que I coincide con I' en Ω pero $I' \in B' \setminus \text{dom}(\Omega')$. En este caso tenemos que $b' \leq a' + 2\epsilon'$ por la definición de $c_{\Omega'}(I')$ luego se verifica que $a + 2(\epsilon + \epsilon') \geq a + \epsilon + 2\epsilon' \geq b' + \epsilon \geq b$, lo que implica que $c_{\Omega' \circ \Omega}(I) \leq \epsilon + \epsilon'$. Además si I'' no coincide con otro elemento de B' , como Ω' era una ϵ' -coincidencia se verifica que $c_{\Omega' \circ \Omega}(I'') = c_{\Omega'}(I'') \leq \epsilon \leq \epsilon + \epsilon'$.
3. Por último si I' coincide con I'' en Ω' pero $I' \in B' \setminus \text{Im}(\Omega)$ puede usarse un razonamiento análogo para ver que $c_{\Omega' \circ \Omega}(I'') \leq \epsilon + \epsilon'$.

□

Estamos ahora ya en condiciones de definir una distancia entre multiconjuntos de intervalos de $\mathbb{R} \cup \{\pm\infty\}$.

Definición 3.22. Sean B y B' dos multiconjuntos de intervalos de $\mathbb{R} \cup \{\pm\infty\}$. La **distancia de cuello de botella** entre B y B' es:

$$d_B(B, B') = \inf\{c(\Omega) \text{ para cada } \Omega \text{ coincidencia parcial entre } B \text{ y } B'\}.$$

Sea \mathbb{B} el conjunto de todos los posibles multiconjuntos de intervalos finitos.

Lema 3.23. *La distancia de cuello de botella define una pseudodistancia extendida en \mathbb{B}*

Demostración. La definición y la no negatividad de la función de coste muestra la no negatividad de la distancia. La simetría se obtiene tomando la aplicación inversa de la coincidencia parcial pensada como una biyección. Claramente $d_B(B, B) = 0$ y la desigualdad triangular es la Proposición 3.21. \square

Nótese que los intervalos **puntuales** del tipo $[a, a]$, es decir puntos en la diagonal del plano, no contribuyen a las funciones de coste, luego no es cierto en general que $d_B(B, B') = 0$ implique que $B = B'$.

Ejemplo 3.24. Sean $B = \{\langle 5, 6 \rangle, [4, 4]\}$ y $B' = \{\langle 5, 6 \rangle\}$ que son multiconjuntos distintos. Tomando la coincidencia parcial $\Omega = \{(\langle 5, 6 \rangle, \langle 5, 6 \rangle)\}$ tenemos que $c_\Omega(\langle 5, 6 \rangle, \langle 5, 6 \rangle) = 0$ y $c_\Omega([4, 4]) = (4 - 4)/2 = 0$ luego

$$d_B(B, B') = c(\Omega) = \max(c_\Omega(\langle 5, 6 \rangle, \langle 5, 6 \rangle), c_\Omega([4, 4])) = 0$$

Lema 3.25. *Sean $B, B' \in \mathbb{B}$:*

$$d_B(B, B') = 0 \Leftrightarrow \text{se diferencian en un número finito de intervalos del tipo } [a, a].$$

Demostración. La implicación \Leftarrow es se deduce del ejemplo anterior. Probamos \Rightarrow . Si $d_B(B, B') = 0$ existe una coincidencia parcial para la cual $\max(|c - a|, |d - b|) = 0$ para todos los intervalos $(\langle a, b \rangle, \langle c, d \rangle) \in \Omega$. Para todos los intervalos $\langle e, f \rangle \in \Omega^C$ debe cumplirse $(f - e)/2 = 0$ luego todos los intervalos que no sean un punto deben estar en una pareja de Ω y por tanto $B = B'$ salvo por un número finito de intervalos puntuales. \square

Definimos en \mathbb{B} la relación de equivalencia $B \sim B'$ si y sólo si se diferencian en un conjunto finito de intervalos puntuales.

Proposición 3.26. *La distancia de cuello de botella define una distancia extendida en el conjunto cociente de \mathbb{B} por la anterior relación de equivalencia.*

Demostración. Definimos, para dos clases $[B], [B']$ cualesquiera, la distancia

$$d_B([B], [B']) = d_B(B, B').$$

Su correcta definición se prueba de forma similar al Corolario 3.14 usando el Lema 3.25. Las propiedades de distancia extendida se deducen de que d_B es una pseudodistancia extendida. \square

Observación 3.27. La distancia de cuello de botella es un caso particular de la **distancia de Wasserstein** de parámetro q para multiconjuntos B y B' de intervalos de \mathbb{R} :

$$W_q(B, B') = \inf_{\Omega=B'} \left(\sum_{I \in \text{dom}(f)} c(I, f(I))^q + \sum_{I \in \text{dom}(f)^c} c(I)^q \right)^{1/q}$$

tomando $q = \infty$, es decir la norma del supremo. [6].

3.3. Teorema de estabilidad de las distancias

Comenzamos esta sección dando dos lemas que son necesarios para la demostración de una de las implicaciones del teorema de estabilidad. Recordemos que χ^I denotaba al módulo de intervalo de I (Definición 2.40) y $B_{\mathcal{M}}$ denotaba al código de barras asociado a un módulo de persistencia \mathcal{M} . Para los módulos de intervalo tendremos entonces que su código de barras asociado es: $B_{\chi^I} = \{I\}$.

Lema 3.28. [5, Lemma 4.14-4.15] Sean $I, I' \subset \mathbb{R}$ intervalos finitos y h y h' la mitad de sus longitudes respectivamente. Sea m el punto medio de I . Entonces $d_I(\chi^I, \chi^{I'}) \leq \max(h, h')$ y, si además $m \notin I'$, entonces $d_I(\chi^I, \chi^{I'}) \geq h$.

Lema 3.29. Sean I e I' intervalos en $\mathbb{R} \cup \{\pm\infty\}$. Entonces se tienen las siguientes igualdades:

$$\begin{aligned} d_B(\{I\}, \{I'\}) &= d_I(\chi^I, \chi^{I'}) \\ d_B(\{I\}, \emptyset) &= d_I(\chi^I, 0) \end{aligned}$$

Demostración. Damos la idea de la demostración para la primera de ellas en el caso de intervalos finitos, pues los demás casos se pueden probar de forma similar.

Sean $I = \langle a, b \rangle$ e $I' = \langle c, d \rangle$ dos intervalos finitos de \mathbb{R} . Sólo existen dos posibles coincidencias parciales $\Omega_1 = \{(I, I')\}$ y $\Omega_2 = \emptyset$ y $d_B(\{I\}, \{I'\}) = \min(c(\Omega_1), c(\Omega_2))$. Estudiando la posición relativa de un intervalo a otro, el mínimo anterior vendrá dado por Ω_2 en los siguientes casos:

1. $a \leq b \leq c \leq d$
2. $a \leq c \leq d \leq (b+a)/2 \leq b$
3. $a \leq c \leq b \leq (d+c)/2 \leq d$
4. $a \leq (b+a)/2 \leq c \leq (d+c)/2 \leq b \leq d$

donde usando el lema anterior llegamos a que $d_B(\{I\}, \{I'\}) = c(\Omega_2) = d_I(\chi^I, \chi^{I'})$.

En los casos $a \leq c \leq (b+a)/2 \leq d \leq b$ y $a \leq c \leq (b+a)/2 \leq (d+c)/2 \leq b \leq d$ tenemos que $d_B(\{I\}, \{I'\}) = c(\Omega_1) = \epsilon$. En estos casos podemos suponer que $d-b = \epsilon$ y construir un ϵ -intercalado (ψ, φ) . Demostramos el caso $a \leq c \leq (b+a)/2 \leq d \leq b$. Sean $\varphi: \chi^I \rightarrow \chi^{I'}T_\epsilon$ y $\psi: \chi^{I'} \rightarrow \chi^IT_\epsilon$ definidos como:

$$\varphi_x = \begin{cases} Id_{\mathbb{k}} & \text{si } a \leq x \leq d - \epsilon \\ 0 & \text{en los demás casos} \end{cases} \quad \text{y} \quad \psi_x = \begin{cases} Id_{\mathbb{k}} & \text{si } a \leq x \leq b \\ 0 & \text{en los demás casos} \end{cases}$$

Si existiera un $\epsilon' < \epsilon$ tal que existe un ϵ' -intercalado, podríamos factorizar la identidad por 0, lo cual es imposible 3.1. \square

Lema 3.30. Sean $\{\mathcal{M}\}_{j \in J}$ y $\{\mathcal{N}\}_{j \in J}$ dos familias de módulos de persistencia indexadas por el mismo conjunto finito J . Se verifica:

$$d_I\left(\bigoplus_{j \in J} \mathcal{M}_j, \bigoplus_{j \in J} \mathcal{N}_j\right) \leq \sup_{j \in J} \{d_I(\mathcal{M}_j, \mathcal{N}_j)\}.$$

Demostración. Si para algún $j \in J$ los módulos $\mathcal{M}_j, \mathcal{N}_j$ no están ϵ -intercalados para algún ϵ finito, entonces $d_I(\mathcal{M}_j, \mathcal{N}_j) = \infty$ y la desigualdad se mantiene. Supongamos ahora que para todo índice $j \in J$ están ϵ_j -intercalados, con ϵ_j finito. El conjunto $\{\epsilon_j\}_{j \in J}$ tendrá un máximo ϵ . Por el Lema 3.12 sabemos que existe una familia de ϵ -intercalados $\{\vartheta_j, \varsigma_j\}_{j \in J}$. La suma directa $\vartheta := \bigoplus_{j \in J} \vartheta_j$ y $\varsigma := \bigoplus_{j \in J} \varsigma_j$ da un ϵ -intercalado de M y N . Por ello cualquier

cota superior de $d_I(\mathcal{M}_j, \mathcal{N}_j)$ es una cota superior de $d_I\left(\bigoplus_{j \in J} \mathcal{M}_j, \bigoplus_{j \in J} \mathcal{N}_j\right)$, con lo que acabamos. \square

Teorema 3.31. Sean \mathcal{M} y \mathcal{N} dos \mathbb{R} -módulos de persistencia sobre un cuerpo, moderados y puntualmente finito generados. Entonces se tiene que:

$$d_B(B_{\mathcal{M}}, B_{\mathcal{N}}) = d_I(\mathcal{M}, \mathcal{N})$$

Demostración. La demostración consta de dos desigualdades.

Veamos $d_B(B_{\mathcal{M}}, B_{\mathcal{N}}) \geq d_I(\mathcal{M}, \mathcal{N})$. Sea $\Omega B_{\mathcal{M}} \rightleftharpoons B_{\mathcal{N}}$ una coincidencia parcial con coste $c(\Omega) = \epsilon$. Sea I un intervalo de $B_{\mathcal{M}}$ que coincide con un intervalo I' de $B_{\mathcal{N}}$, es decir $I \in \text{dom}(\Omega)$. Entonces se tiene:

$$d_B(\{I\}, \{I'\}) \leq c(I, I') \leq \epsilon$$

Por el lema 3.29 tenemos que para todo $\delta > 0$ existe un $\epsilon + \delta$ -intercalado entre χ^I y $\chi^{I'}$.

Análogamente, si $I'' \in (B \setminus \text{dom}(\Omega)) \sqcup (B' \setminus \text{Im}(\Omega))$, es decir, I'' no coincide con ninguno, se tiene por 3.29 que $\chi^{I''}$ está $\epsilon + \delta$ -intercalado con el módulo 0 para todo δ arbitrariamente pequeño. Tenemos las siguientes descomposiciones para \mathcal{M} y \mathcal{N} :

$$\begin{aligned} \mathcal{M} &\cong \left(\bigoplus_{I \in \text{dom}(\Omega)} \chi^I \right) \oplus \left(\bigoplus_{I \in B \setminus \text{dom}(\Omega)} \chi^I \right) \oplus \left(\bigoplus 0 \right) \\ \mathcal{N} &\cong \left(\bigoplus_{I \in \text{Im}(\Omega)} \chi^I \right) \oplus \left(\bigoplus_{I \in B' \setminus \text{Im}(\Omega)} \chi^I \right) \oplus \left(\bigoplus 0 \right) \end{aligned}$$

donde el 0 aparece tantas veces como sea necesario para que ambas sumas tengan el mismo número de elementos. Por el lema 3.30 tenemos que para todo $\delta > 0$ los módulos de persistencia \mathcal{M} y \mathcal{N} están $\epsilon + \delta$ -intercalados. Luego $d_I(\mathcal{M}, \mathcal{N}) \leq \epsilon$. Como esto era cierto para cualquier coincidencia parcial tenemos que $d_B(B_{\mathcal{M}}, B_{\mathcal{N}}) \geq d_I(\mathcal{M}, \mathcal{N})$.

La desigualdad $d_B(B_{\mathcal{M}}, B_{\mathcal{M}}) \leq d_I(\mathcal{M}, \mathcal{N})$ está demostrada en [12, Theorem 4.4] y se omite su demostración por motivos de extensión. \square

Este resultado es de un enorme significado en para el campo del análisis topológico de datos y la homología persistente. Bajo ciertas condiciones permite computar la proximidad entre dos módulos de persistencia a través de las distancias entre sus códigos de barras, lo cual evita la búsqueda de los morfismos de funtores del intercalado. Además como corolarios de este teorema mostramos que es posible aplicar la homología persistente a nubes de datos de forma estable bajo ligeras variaciones.

Sea S un conjunto finito de puntos, $\text{VR}(S)$ el complejo de Vietoris-Rips explicado en el Capítulo 1 de la memoria y H_i la homología simplicial en nivel i .

Corolario 3.32. Sean P y Q dos conjuntos finitos de puntos en \mathbb{R}^n y sea $\pi : P \rightarrow Q$ una biyección entre ellos de forma que $\|p - \pi(p)\| \leq \epsilon \forall p \in P$. Entonces para cualquier natural i :

$$d_B(B_{H_i(\text{VR}(P))}, B_{H_i(\text{VR}(Q))}) \leq \epsilon$$

Demostración. Sean $P = \{p_1, \dots, p_m\}$ y $Q = \{q_1, \dots, q_m\}$ ordenados de manera que $\|p_i - q_i\| \leq \epsilon$ para todo $1 \leq i \leq m$ (lo cual es posible tomando $\pi(p_i) = q_i$). Si tenemos un símlice $\sigma = \{p_{i_1}, \dots, p_{i_l}\}$ de $\text{VR}_r(P)$ tenemos que $\max_{x, y \in \sigma} \|x - y\|$ es menor o igual a r . La imagen de σ por π entonces cumple que la máxima distancia entre dos cualesquiera de sus puntos es $r + \epsilon$. Se tiene así la inclusión definida por σ : $\text{VR}_r(P) \subseteq \text{VR}_{r+\epsilon}(Q)$.

Haciendo un proceso simétrico con la inversa de σ llegamos a que $\text{VR}_r(Q) \subseteq \text{VR}_{r+\epsilon}(P)$.
 Luego juntando ambos resultados aparecen las siguientes inclusiones:

$$\begin{aligned} \text{VR}_r(P) &\subseteq \text{VR}_{r+\epsilon}(Q) \subseteq \text{VR}_{r+2\epsilon}(P) \\ \text{VR}_r(Q) &\subseteq \text{VR}_{r+\epsilon}(P) \subseteq \text{VR}_{r+2\epsilon}(Q) \end{aligned}$$

Tomando la homología en grado i de la filtración dada por los valores de r , tenemos que esas inclusiones descienden al pasarlas por el functor H_i a una pareja de ϵ -intercalado $(H_i(\sigma^{-1}), H_i(\sigma))$:

$$\begin{array}{ccccccc} \dots & \longrightarrow & H_i(\text{VR}_r(P)) & \longrightarrow & H_i(\text{VR}_{r+\epsilon}(P)) & \longrightarrow & H_i(\text{VR}_{r+2\epsilon}(P)) & \longrightarrow & \dots \\ & \searrow & \nearrow & & \searrow & \nearrow & \searrow & \nearrow & \\ \dots & \longrightarrow & H_i(\text{VR}_r(Q)) & \longrightarrow & H_i(\text{VR}_{r+\epsilon}(Q)) & \longrightarrow & H_i(\text{VR}_{r+2\epsilon}(Q)) & \longrightarrow & \dots \end{array}$$

donde hemos usado la Proposición 3.15 para que se preserven los intercalados.

Como $\text{Card}(P) = \text{Card}(Q) < \infty$ tenemos que $\{H_i(\text{VR}_r(P))\}$ y $\{H_i(\text{VR}_r(Q))\}$ son moderados pues solo tienen un número finito de valores críticos y por tanto tendrán asociado un código de barras finito. Se tiene entonces:

$$d_B(B_{H_i(\text{VR}(P))}, B_{H_i(\text{VR}(Q))}) = d_I(H_i(\text{VR}(P)), H_i(\text{VR}(Q))) \leq \epsilon$$

usando el Teorema 3.31. □

Conclusiones

En esta memoria se ha tratado de realizar un estudio general de la estructura de los módulos de persistencia, con especial atención a los detalles para los módulos de persistencia uniparamétricos que surgen en el análisis topológico de datos, es decir, los moderados puntualmente finito generados. El capítulo 1 ha servido para justificar por qué estos módulos de persistencia son una estructura relevante en el análisis topológico de datos.

El capítulo 2 se ha estudiado las propiedades de los módulos de persistencia uniparamétrica como estructura algebraica. En particular, aportar un análisis detallado de todos los elementos necesarios para clasificar espacios vectoriales de persistencia moderados puntualmente finito generados como suma de módulos de intervalo. Tras un breve repaso de la estructura de módulo graduado sobre un anillo graduado (Sección 2.3) se ha mostrado cómo un módulo de \mathbb{N} -persistencia es esencialmente un módulo graduado. Para esto se ha dado una equivalencia de categorías en el Lema 2.34. Se ha mostrado además cómo mantiene esta equivalencia de categorías las condiciones de finitud. Por construcción es claro que la presentación/generación finita se mantiene en cada componente homogénea. Hemos utilizado los Lemas 2.35, 2.36 y 2.37 para ver cómo añadiendo la condición de tipo finito se podía lograr que el módulo graduado obtenido fuera también finito generado pudiendo usar el teorema de la estructura 2.31 para descomponerlo de modo único. En la Sección 2.5 se da la clasificación de módulos de \mathbb{N} -persistencia como módulos de intervalo. Esta clasificación se extiende a otros conjuntos totalmente ordenados en la Sección 2.6, incluyendo el caso de los módulos moderados. Se termina dado las representaciones por códigos de barras o diagramas de persistencia.

Como futuras líneas de trabajo surge la búsqueda de la generalización de los resultados mostrados en este capítulo, para persistencia uniparamétrica, a módulos multipersistentes, es decir dependientes de varios parámetros. El caso más relevante es el poset \mathbb{R}^n donde $(a_1, \dots, a_n) \leq (b_1, \dots, b_n)$ si $a_i \leq b_i$ para todo $1 \leq i \leq n$. Una línea de trabajo es tratar de clasificarlos como suma de módulos de bloque (producto de intervalos) que son el análogo n -dimensional de los módulos de intervalo, aunque muestra grandes problemas en su computación [9]. El resultado más fuerte en esta línea se debe a Crawley y Botnan en [3] y da la clasificación de módulos \mathbb{R}^2 -persistentes. Este tipo de módulos aparecen en el análisis topológico de datos de diversas formas como en el análisis de muestras de datos

ruidosas [10]. Otras referencias sobre multipersistencia y sus invariantes son [27], [33] y [22].

En capítulo tres estudia la teoría de las distancias entre módulos de persistencia y sus códigos de barras centrándose en módulos de persistencia sobre \mathbb{R} con el objetivo de agrupar todos los resultados conocidos hasta la fecha y mostrar su relevancia en el análisis topológico de datos. En la Sección 3.1 se ha desarrollado la distancia de intercalado y mostrado cómo, en efecto, define una distancia en un cociente de $C^{\mathbb{R}}$. En la Sección 3.2 se estudia la distancia de cuello de botella y se demuestra que define una distancia. Finalmente en la Sección 3.3 se da la equivalencia entre ambas y se utiliza para demostrar la estabilidad de la homología persistente de complejos de Vietoris-Rips.

Aunque los detalles se dan para el caso de un número finito de intervalos, el teorema de estabilidad se cumple en general como muestra Lesnick [24]. La teoría de distancia de intercalado se puede generalizar a la multipersistencia [23] donde se buscan similares resultados de estabilidad.

Bibliografía

- [1] Michael Atiyah. *Introduction to commutative algebra*. CRC Press, 2018.
- [2] Nicolas Berkouk. «Persistence and Sheaves: from Theory to Applications». Tesis doct. Institut polytechnique de Paris, 2020.
- [3] Magnus Botnan y William Crawley-Boevey. «Decomposition of persistence modules». En: *Proceedings of the American Mathematical Society* 148.11 (2020), págs. 4581-4596.
- [4] Peter Bubenik. *Software-TDA*. English. University of Florida, Department of Mathematics. 2021. URL: <https://people.clas.ufl.edu/peterbubenik/software/>.
- [5] Peter Bubenik y Jonathan A Scott. «Categorification of persistent homology». En: *Discrete & Computational Geometry* 51.3 (2014), págs. 600-627.
- [6] Gunnar Carlsson. «Persistent Homology and Applied Homotopy Theory». En: *arXiv preprint arXiv:2004.00738* (2020).
- [7] Gunnar Carlsson. «Topological pattern recognition for point cloud data». En: *Acta Numerica* 23 (2014), págs. 289-368.
- [8] Gunnar Carlsson. «Topology and data». En: *Bulletin of the American Mathematical Society* 46.2 (2009), págs. 255-308.
- [9] Gunnar Carlsson, Gurjeet Singh y Afra J Zomorodian. «Computing multidimensional persistence». En: *International Symposium on Algorithms and Computation*. Springer. 2009, págs. 730-739.
- [10] Gunnar Carlsson y Afra J Zomorodian. «The theory of multidimensional persistence». En: *Discrete & Computational Geometry* 42.1 (2009), págs. 71-93.
- [11] Gunnar Carlsson y col. «Persistence barcodes for shapes». En: *International Journal of Shape Modeling* 11.02 (2005), págs. 149-187.
- [12] Frédéric Chazal y col. «Proximity of persistence modules and their diagrams». En: *Proceedings of the twenty-fifth annual symposium on Computational geometry*. 2009, págs. 237-246.
- [13] René Corbet y Michael Kerber. «The representation theorem of persistence revisited and generalized». En: *Journal of Applied and Computational Topology* 2.1 (2018), págs. 1-31.

- [14] William Crawley-Boevey. «Decomposition of pointwise finite-dimensional persistence modules». En: *Journal of Algebra and its Applications* 14.05 (2015), pág. 1550066.
- [15] Meryll Dindin, Yuhei Umeda y Frederic Chazal. «Topological data analysis for arrhythmia detection through modular neural networks». En: *Canadian Conference on Artificial Intelligence*. Springer. 2020, págs. 177-188.
- [16] Vincent Divol y Théo Lacombe. «Understanding the topology and the geometry of the space of persistence diagrams via optimal partial transport». En: *Journal of Applied and Computational Topology* 5.1 (2021), págs. 1-53.
- [17] David S Dummit y Richard M Foote. *Abstract algebra*. Vol. 1999. Prentice Hall Englewood Cliffs, NJ, 1991.
- [18] Herbert Edelsbrunner, David Letscher y Afra J Zomorodian. «Topological persistence and simplification». En: *Proceedings 41st annual symposium on foundations of computer science*. IEEE. 2000, págs. 454-463.
- [19] David Eisenbud. *Commutative algebra: with a view toward algebraic geometry*. Vol. 150. Springer Science & Business Media, 2013.
- [20] Marian Gidea y Yuri Katz. «Topological data analysis of financial time series: Landscapes of crashes». En: *Physica A: Statistical Mechanics and its Applications* 491 (2018), págs. 820-834.
- [21] Anubha Goel, Puneet Pasricha y Aparna Mehra. «Topological data analysis in investment decisions». En: *Expert Systems with Applications* 147 (2020), pág. 113222.
- [22] Heather A Harrington y col. «Stratifying multiparameter persistent homology». En: *SIAM Journal on Applied Algebra and Geometry* 3.3 (2019), págs. 439-471.
- [23] Michael Lesnick. «The theory of the interleaving distance on multidimensional persistence modules». En: *Foundations of Computational Mathematics* 15.3 (2015), págs. 613-650.
- [24] Michael Phillip Lesnick. *Multidimensional interleavings and applications to topological inference*. Stanford University, 2012.
- [25] Rachel Levanger. «A Comparison Framework For Interleaved Persistence Modules and Applications of Persistent Homology to Problems in Fluid Dynamics». Tesis doct. Rutgers, The State University of New Jersey, 2017.
- [26] Saunders Mac Lane. *Categories for the working mathematician*. Vol. 5. Springer Science & Business Media, 2013.
- [27] Ezra Miller. «Data structures for real multiparameter persistence modules». En: *arXiv preprint arXiv:1709.08155* (2017).

- [28] Elizabeth Munch. «A users' guide to topological data analysis». En: *Journal of Learning Analytics* 4.2 (2017), págs. 47-61.
- [29] James R Munkres. *Elements of algebraic topology*. The Benjamin/Cummings Publishing Company, Inc., 1984.
- [30] Takenobu Nakamura y col. «Persistent homology and many-body atomic structure for medium-range order in the glass». En: *Nanotechnology* 26.30 (2015), pág. 304001.
- [31] Monica Nicolau, Arnold J Levine y Gunnar Carlsson. «Topology based data analysis identifies a subgroup of breast cancers with a unique mutational profile and excellent survival». En: *Proceedings of the National Academy of Sciences* 108.17 (2011), págs. 7265-7270.
- [32] Dayten Sheffar. «Introductory Topological Data Analysis». En: *arXiv preprint arXiv:2004.04108* (2020).
- [33] Pawin Vongmasa. «Generalized persistence modules and some of their invariants». Tesis doct. 2015.
- [34] Raoul R Wadhwa y col. «A flat persistence diagram for improved visualization of persistent homology». En: *arXiv preprint arXiv:1812.04567* (2018).
- [35] Cary Webb. «Decomposition of graded modules». En: *Proceedings of the American Mathematical Society* 94.4 (1985), págs. 565-571.
- [36] Afra J Zomorodian. *Computing and comprehending topology: Persistence and hierarchical morse complexes*. University of Illinois at Urbana-Champaign, 2001.
- [37] Afra J Zomorodian. *Topology for computing*. Vol. 16. Cambridge university press, 2005.
- [38] Afra J Zomorodian y Gunnar Carlsson. «Computing persistent homology». En: *Discrete & Computational Geometry* 33.2 (2005), págs. 249-274.

Glosario de términos

índice de persistencia, 14

anillo

graduado, 22

Noetheriano, 23

anillo graduado

ideal graduado, 23

aplicación monótona, 20

código de barras, 37

cara, 4

propia, 4

categoría, 15

Krull-Schmidt, 19

abeliana, 19

coproducto, 19

equivalencia de, 17

producto, 19

subcategoría, 16

coincidencia parcial, 47

ϵ -coincidencia, 48

composición, 49

coste, 48

dominio, 47

imagen, 47

complejo de Čech, 11

complejo de cadenas, 7

c-cadena, 7

morfismo de, 8

n-borde, 7

n-ciclo, 7

operador de borde, 7

complejo de Vietoris-Rips, 11

complejo simplicial

abstracto, 5

cara, 5

dimensión, 5

filtrado, 12

geométrico, 4

orientado, 7

poliedro asociado, 4

subcomplejo, 4, 5

diagrama de persistencia, 37

dimensión, 3

distancia, 43

extendida, 43

pseudodistancia, 43

distancia cuello de botella, 50

epimorfismo, 18

esqueleto, 4

esquema de vértices, 6

filtración, 12

functor covariante, 17

composición, 17

quasi-inversa, 17

- homología persistente, 14
- intercalado
 - ϵ -intercalado, 43
 - distancia, 43
 - generalizado, 41
- intervalo, 31
- módulo
 - finito generado, 21
 - finito presentado, 22
 - graduado, 23
 - Noetheriano, 23
- módulo de homología simplicial, 9
- módulo de homología, 8
- módulo de persistencia
 - descomponible, 32
 - indescomponible, 32
 - moderado, 35
- módulo persistencia
 - intervalo, 31
- módulo persistente
 - puntualmente finito generado, 22
 - puntualmente finito presentado, 22
 - tipo finito, 22
- monomorfismo, 18
- morfismo
 - de vértices, 4
 - isomorfismo, 6
 - simplicial, 4
 - simplicial abstracto, 6
- morfismo de funtores, 17
 - composición horizontal, 18
 - composición vertical, 18
- morfismo de transición, 20
- morfismo graduado, 23
- multiconjunto, 37
- número de Betti, 8, 14
- nervio, 10
- objeto, 15
 - cero, 18
 - inicial, 18
 - terminal, 18
- objeto persistente, 20
- orden
 - parcial, 11
 - total, 11
- orientación, 7
- pareja de translación, 41
- poset, 11, 16
- realización geométrica, 6
- símplice, 3, 5
- transformación natural, 17
- translación, 40
- triangulable, 4
- vértice, 3
- valor
 - crítico, 35
 - regular, 35