

UNIVERSIDAD DE SALAMANCA

Facultad de Ciencias

GRADO EN MATEMÁTICAS

INTRODUCCIÓN AL ANÁLISIS
TOPOLÓGICO DE DATOS

TRABAJO DE FIN DE GRADO

realizado por

Luis Sánchez Cano

bajo la tutela de:

Darío Sánchez Gómez

(Departamento de Matemáticas)



**VNiVERSiDAD
D SALAMANCA**

CAMPUS DE EXCELENCIA INTERNACIONAL



Salamanca, 6 de julio de 2022

Introducción al Análisis Topológico de Datos

Trabajo de fin de Grado presentado por
Luis Sánchez Cano

Luis Sánchez Cano

Darío Sánchez Gómez, tutor

Salamanca, a 6 de julio de 2022.

Índice general

Introducción	1
1. Complejos simpliciales	3
1.1. Complejos simpliciales geométricos	3
1.2. Complejos simpliciales abstractos	6
1.3. Complejos simpliciales de Čech y Vietoris-Rips	7
1.3.1. Complejo de Čech	8
1.3.2. Complejo de Vietoris-Rips	9
1.4. Otros complejos simpliciales	11
1.4.1. Complejo de Delaunay	11
1.4.2. Complejo Alfa	11
2. Homología y códigos de barras	13
2.1. Homología	13
2.2. Cálculo de la homología	14
2.3. Homología simplicial	20
2.4. Homología persistente	22
2.5. Códigos de barras	26
2.5.1. Descomposición y cómputo del código de barras	28
3. Laplaciano combinatorio	35
3.1. Laplaciano en grafos	35
3.2. Laplaciano en complejos simpliciales	37
3.3. Espectro Laplaciano de un complejo simplicial	46
Conclusiones	52

Bibliografía	55
Glosario de términos	59

Índice de figuras

1.1. n-símplices	4
1.2. Poliedro de K	5
1.3. Complejo de Čech	9
1.4. Complejo de Čech vs complejo de Vietoris-Rips	10
1.5. Diagrama de Voronoi.	12
1.6. Complejo de Delaunay	12
2.1. Filtración de un complejo de Čech	23
2.2. Nacimiento y muerte de α , que nace en K_i y muere entrando en K_j	25
2.3. Diagrama de persistencia	27
2.4. Filtración junto con su código de barras asociado. Imagen sacada de [12].	27
3.1. Ejemplo de grafo.	36
3.2. Complejo simplicial orientado	41
3.3. Diferentes orientaciones en un mismo complejo simplicial	48

Introducción

El Análisis Topológico de Datos (TDA por sus siglas en inglés) comenzó a desarrollarse en la década de los noventa con la introducción de la función de tamaño por parte de Patrizio Frosini [10]. Esta función nos da el número de componentes conexas de un espacio topológico que permanecen a lo largo de una filtración, lo que posteriormente veremos como la homología persistente de nivel cero, $H_0(\mathcal{F})$. En 1999, Vanessa Robins [24] comenzó a exponer (aunque no con ese nombre) lo que más adelante se conocería como *homología persistente*, introduciendo para ello los *números de Betti de persistencia*. Dicha *homología persistente* se definiría para una filtración de un complejo simplicial en 2002 con un trabajo de H. Edelsbrunner, D. Letcher, y A. Zomorodian [8] donde, además de esto, presentaron los *diagramas de persistencia*, gráficos que muestran el nacimiento y muerte de los distintos agujeros n -dimensionales que aparecen a lo largo de una filtración. En ese mismo trabajo se dio una segunda definición de *homología persistente* equivalente, esta será la que consideraremos en la memoria. Poco después, en 2005, Carlsson et al. [4] dieron otro método para la representación de los nacimientos y las muertes, los llamados *códigos de barras*. Esta última definición de *homología persistente* será la que consideraremos en la memoria. Para el desarrollo de todo este campo relativamente nuevo se utilizan como herramientas el Álgebra Conmutativa y la Topología Algebraica.

En la actualidad, la inmensa y desestructurada cantidad de datos almacenados en internet (redes sociales, foros, videojuegos, instituciones públicas...) hace imposible su manejo mediante métodos tradicionales, por ello en los últimos años ha tomado una enorme relevancia el *Big Data*, y con ello el TDA. Este enfoque desde la Topología Algebraica permite dotar de una estructura de espacio topológico a una nube de datos mediante el uso de los complejos simpliciales, buscando con ello inferir patrones, relaciones o características de dichos datos a través de la homología. Hoy en día se utilizan estas técnicas en diversos campos como pueden ser la biología, la química, la clasificación de imágenes, la astrofísica, la medicina... Ejemplos de ello los encontramos en [5]. De hecho, contamos con diversos softwares que permiten realizar cálculos de un modo rápido y efectivo, los cuales iremos mencionando a lo largo de este trabajo.

Esta memoria comienza introduciendo las nociones básicas relacionadas con los complejos simpliciales para, posteriormente, pasar a hablar de la homología simplicial, la homología persistente y los códigos de barras, incluyendo su cómputo. En una “segun-

da” parte, claramente diferenciada, trataremos el Laplaciano combinatorio de un complejo simplicial, su espectro y finalmente su estrecha relación con la homología, uniendo así estos dos enfoques utilizados en el TDA. Este desarrollo lo haremos a lo largo de 3 capítulos:

- Capítulo 1: Se tratarán con detalle las nociones de n -símplice, complejo simplicial y presentaremos distintos tipos de complejos simpliciales. Incluyendo además figuras representativas y ejemplos sencillos.
- Capítulo 2: Entraremos a explicar de forma general los complejos de cadenas de A -módulos y la homología asociada, así como un algoritmo para su cómputo mediante el uso de la forma normal de Smith (cuyo cálculo también se detalla) para, posteriormente, centrarnos en el caso simplicial: complejos de cadenas simpliciales y homología simplicial. Con estas herramientas expuestas pasaremos a hablar de la homología persistente, noción fundamental en el TDA, pues nos permite cuantificar la importancia de los agujeros n -dimensionales que se generan en nuestro complejo simplicial a lo largo de una filtración, discriminando así el ruido de la información útil. Tras esto, hablaremos de los códigos de barras y su cómputo a partir de la descomposición de la homología simplicial como módulo graduado, lo cual nos posibilita recoger la información de los datos de una manera más visual.
- Capítulo 3: Introduciremos la matriz combinatoria Laplaciana de un grafo y generalizaremos dicho concepto a los complejos simpliciales posteriormente, dando la definición de Laplaciano combinatorio de un complejo simplicial y la matriz Laplaciana asociada. Estas matrices representan la adyacencia entre los distintos símplices, y detallaremos cómo calcularlas. Igualmente, daremos la noción de espectro del Laplaciano de un complejo simplicial, así como información relativa a dicho objeto. Finalmente mostraremos una versión discreta del Teorema de Hodge, que establece un isomorfismo entre la homología simplicial y el núcleo del Laplaciano combinatorio.

Capítulo 1

Complejos simpliciales

A lo largo de este capítulo introduciremos una serie de nociones básicas sobre los complejos simpliciales: definición, características y tipos de complejos simpliciales; con el objetivo de dar posteriormente un enfoque relacionado con el Análisis Topológico de Datos. Las referencias principales utilizadas para desarrollar este capítulo son [1, 3] y [23].

1.1. Complejos simpliciales geométricos

En primer lugar, abordaremos el tema desde un punto de vista geométrico para, posteriormente, estudiarlo de manera más abstracta.

Diremos que un conjunto $\{a_0, a_1, \dots, a_n\}$ de $n + 1$ puntos de \mathbb{R}^N es **geoméricamente independiente** si para cualesquiera $t_i \in \mathbb{R}$ tales que $\sum_{i=0}^n t_i = 0$ se tiene que:

$$\sum_{i=0}^n t_i a_i = 0 \Leftrightarrow t_0 = t_1 = \dots = t_n = 0$$

Equivalentemente, desde el punto de vista del álgebra lineal, podemos decir que dicho conjunto es **geoméricamente independiente** si y sólo si los vectores $a_1 - a_0, \dots, a_n - a_0$ son linealmente independientes.

A partir de esto tenemos que, dado un conjunto geoméricamente independiente $\{a_0, \dots, a_n\}$ de puntos de \mathbb{R}^N , definimos el **n-símplice** (o simplemente símplice) σ generado por a_0, \dots, a_n (denotado como $\sigma = \langle a_0, \dots, a_n \rangle$) como el conjunto de puntos de \mathbb{R}^N de la forma:

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

Un **n-símplice** se puede entender, por lo tanto, como la envolvente convexa del conjunto de puntos dado. Los puntos a_i se denominan **vértices** de σ y definimos la **dimensión** de

dicho símplice como $\dim(\sigma) = n$. Además, diremos que cualquier símplice τ generado por un subconjunto de $\{a_0, \dots, a_n\}$ es una **cara** de σ , lo cual se denotará como $\tau \leq \sigma$, y en caso de que dicha cara sea diferente de σ la llamaremos **cara propia**.

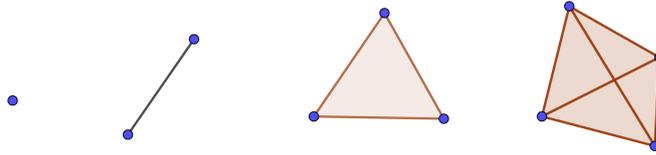


Figura 1.1: n -símplices de distintas dimensiones.

Con esto, llegamos a la siguiente definición:

Definición 1.1. Un **complejo simplicial geométrico** K en \mathbb{R}^N es una colección finita de símplices en \mathbb{R}^N tales que:

1. Si para $\sigma \in K$ tenemos que $\tau \leq \sigma$, entonces $\tau \in K$.
2. Si dos símplices tienen intersección no vacía, entonces dicha intersección es una cara de cada uno de los símplices.

Observación 1.2. En esta memoria nos centraremos en el caso finito. Sin embargo, la definición previa se puede extender al caso infinito.

Además, se dice que L es un **subcomplejo** (simplicial) de K si es una subcolección de K que contiene todas las caras de sus elementos, esto se puede denotar como $L \subseteq K$. Por ejemplo, los **p-esqueletos** $K^{(p)}$ son los subcomplejos de K formados por los símplices de dimensión p o menor. En particular, es fácil ver que $K^{(0)}$ son los vértices de K . Para un complejo simplicial K definimos su **dimensión** como la mayor de las dimensiones de sus símplices.

Otro concepto conocido es el de **poliedro** de un complejo simplicial K , denotado por $|K|$, definido como la unión de los símplices que forman K , i.e., $|K| := \bigcup_{\sigma \in K} \sigma$. Además, se le puede dotar de una estructura de espacio topológico a $|K|$ al considerar la topología inducida por la usual de \mathbb{R}^N , siendo así un espacio compacto.

Ejemplo 1.3. Un complejo simplicial geométrico en \mathbb{R}^2 podría ser:

$$K = \{\langle a_1 \rangle, \langle a_2 \rangle, \langle a_3 \rangle, \langle a_4 \rangle, \langle a_5 \rangle, \langle a_6 \rangle, \langle a_7 \rangle, \langle a_1, a_2 \rangle, \langle a_2, a_3 \rangle, \langle a_2, a_4 \rangle, \langle a_3, a_4 \rangle, \langle a_4, a_5 \rangle,$$

$$\langle a_5, a_6 \rangle, \langle a_5, a_7 \rangle, \langle a_6, a_7 \rangle, \langle a_5, a_6, a_7 \rangle$$

El poliedro asociado a dicho complejo simplicial sería por tanto:

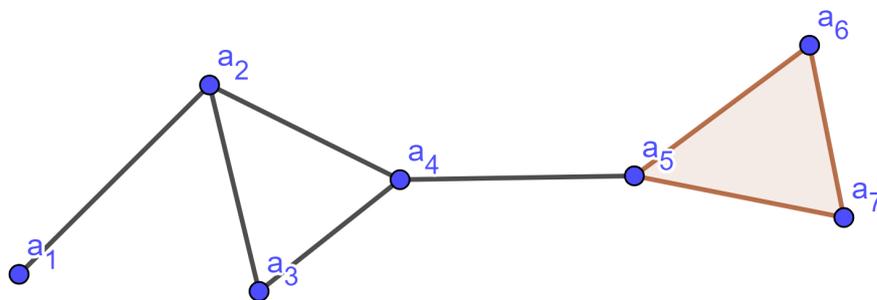


Figura 1.2: Poliedro de K .

Igualmente, es natural definir las aplicaciones entre este tipo de objetos. Dados dos complejos simpliciales K y L , una **aplicación simplicial** $f: K \rightarrow L$ es una aplicación entre sus conjuntos de vértices $f: K^{(0)} \rightarrow L^{(0)}$ de tal modo que para todo símplice $\sigma = \langle a_0, \dots, a_n \rangle \in K$, los vértices $f(a_0), \dots, f(a_n)$ pertenecen a un mismo símplice de L . De manera natural se define la composición de aplicaciones simpliciales, la cual también es una aplicación simplicial. Si además existe otra aplicación simplicial $g: L \rightarrow K$ tal que $g \circ f = 1_K$ y $f \circ g = 1_L$, entonces diremos que f es un **isomorfismo simplicial** y que K y L son **simplicialmente isomorfos**.

Además, toda aplicación simplicial $f: K \rightarrow L$ induce una aplicación continua entre sus poliedros asociados $|f|: |K| \rightarrow |L|$, definida del siguiente modo:

$$|f|(x) := \sum_{i=0}^n t_i f(a_i)$$

para cada $x = \sum_{i=0}^n t_i a_i \in \sigma = \langle a_0, \dots, a_n \rangle$, con $\sigma \in |K|$.

A partir de esta definición es fácil comprobar que $|1_K| = 1_{|K|}$ y $|f \circ g| = |f| \circ |g|$ para cualesquiera aplicaciones simpliciales f y g . Con lo cual, si K y L son simplicialmente isomorfos, se tiene que $|K|$ y $|L|$ son homeomorfos como espacios topológicos.

Por otro lado, diremos que un espacio topológico X es **triangulable** si podemos encontrar un complejo K tal que exista un homeomorfismo $h: |K| \rightarrow X$. En este caso (K, h) recibe el nombre de **triangulación** de X . Así, apoyándonos en lo anterior, si tenemos dos espacios topológicos triangulables X e Y , podemos “reducir” el problema de ver si son homeomorfos probando si los complejos simpliciales de sus triangulaciones son isomorfos.

1.2. Complejos simpliciales abstractos

Debido a las implicaciones que conlleva definir los complejos desde el punto de vista geométrico, pasaremos a definir los complejos simpliciales de un modo más abstracto con el fin de manejar este objeto matemático de una manera más sencilla. De ahora en adelante, si hablamos de complejos simpliciales a secas, nos referiremos al caso abstracto. Además, por comodidad seguiremos utilizando la misma notación que en el caso geométrico.

Definición 1.4. Un **complejo simplicial abstracto** es una colección (finita) K de conjuntos finitos no vacíos, de tal modo que para todo elemento de K también están en K sus subconjuntos no vacíos.

Observación 1.5. De nuevo, nos centraremos en el caso finito, aunque dicha definición se puede extender sin problema al caso infinito. Por tanto, si no se especifica nada, consideraremos el caso finito en el resto de la memoria.

Un elemento X de K recibe el nombre de **símplice** de K y se define su dimensión como el número de sus elementos menos uno. Cada subconjunto no vacío de X es una **cara** de X . La **dimensión** de K es la mayor de las dimensiones de sus elementos. El **conjunto de vértices** V de K (también denotado V_K) se define como la unión de sus elementos con cardinal igual a uno. Las notaciones $v \in V$ y $\{v\} \in K$ representarán el mismo elemento. Además, una subcolección de K que sea también un complejo simplicial recibirá el nombre de **subcomplejo** (simplicial) de K .

Observación 1.6. Observamos que esta nueva definición abstracta de un complejo simplicial es equivalente a la definición geométrica que dimos al principio de la memoria, así como las propiedades que definimos.

Igualmente, podemos definir una **aplicación simplicial abstracta** entre complejos simpliciales abstractos $f: K \rightarrow L$ como una aplicación entre sus respectivos conjuntos de vértices de tal modo que $\{a_0, \dots, a_n\} \in K \Leftrightarrow \{f(a_0), \dots, f(a_n)\} \in L$. Por lo tanto, diremos que dos complejos simpliciales abstractos K y L son **simplicialmente isomorfos** si existe una aplicación simplicial (abstracta) biyectiva entre sus respectivos conjuntos de vértices.

Como ya hemos mencionado, las dos definiciones dadas de complejos simpliciales son equivalentes, veamos entonces cómo están relacionadas. Dado un complejo simplicial geométrico K , definimos el **esquema** de K , y lo denotamos $\mathcal{A}(K)$, como el complejo simplicial abstracto cuyo conjunto de vértices es $K^{(0)}$ y $\{a_0, \dots, a_n\} \in \mathcal{A}(K) \Leftrightarrow \sigma = \langle a_0, \dots, a_n \rangle \in K$. Por otro lado, si tenemos un complejo simplicial abstracto K , llamaremos **realización geométrica** de K a un complejo simplicial geométrico L cuyo esquema asociado sea isomorfo a K , dicha realización geométrica es única salvo isomorfismos.

Teorema 1.7. *Todo complejo simplicial abstracto es isomorfo al esquema de algún complejo simplicial geométrico. Además, dos complejos simpliciales geométricos son simplicialmente isomorfos si y solo si sus esquemas asociados lo son.*

Demostración. Sea K un complejo simplicial abstracto, con $V_K = \{v_0, \dots, v_n\}$ su conjunto de vértices. Construimos una aplicación inyectiva $f: V_K \rightarrow \mathbb{R}^N$, con $N \geq n$, de tal modo que $f(v_i) := a_i$, siendo $\{a_0, \dots, a_n\}$ un conjunto geoméricamente independiente (esta construcción es posible gracias a la elección de N). Con esto, la realización geométrica de K sería un complejo simplicial geométrico L tal que $\langle f(v_{i_0}), \dots, f(v_{i_m}) \rangle \in L$ si y sólo si $\{v_{i_0}, \dots, v_{i_m}\} \in K$.

El segundo resultado se puede comprobar directamente a partir de la construcción del esquema de un complejo simplicial geométrico. \square

Observación 1.8. Es fácil ver entonces que para todo complejo simplicial abstracto podemos encontrar una realización geométrica. Con lo cual, nos será indiferente trabajar con complejos simpliciales geométricos o abstractos. Además, cuando hablemos de poliedro asociado a un complejo simplicial abstracto nos referiremos al poliedro asociado a su realización geométrica. Igualmente, a partir de ahora utilizaremos indistintamente la notación $\sigma = \langle a_0, \dots, a_n \rangle$ o $\{a_0, \dots, a_n\}$ para referirnos a los elementos de un complejo simplicial abstracto.

1.3. Complejos simpliciales de Čech y Vietoris-Rips

Para tratar el siguiente apartado nos apoyaremos en herramientas tanto de la topología como de la topología algebraica. Explicaremos con más detalle dos tipos de complejos simpliciales concretos, el de Čech y el de Vietoris-Rips, los cuales serán objeto de estudio a lo largo de esta memoria.

Dado un espacio topológico X y un recubrimiento finito $\mathcal{U} = \{U_i\}_{i \in I}$ de dicho espacio (podemos suponer que todo conjunto del recubrimiento es distinto del vacío, ya que si alguno fuera igual al vacío se podría suprimir), tenemos la siguiente definición:

Definición 1.9. El **nervio** de \mathcal{U} , denotado por $N(\mathcal{U})$, es el complejo simplicial abstracto cuyo conjunto de vértices es I , de tal modo que el k -símplice $\{i_0, \dots, i_k\} \subseteq I$ pertenece a $N(\mathcal{U})$ si y solo si $\bigcap_{p=0}^k U_{i_p} \neq \emptyset$.

Observación 1.10. Es fácil comprobar que $N(\mathcal{U})$ está bien definido como complejo simplicial abstracto, pues para cualquier subconjunto $S = \{i_0, \dots, i_k\}$ de I , si se tiene que

$\bigcap_{p=0}^k U_{i_p} \neq \emptyset$, también se verificará que la intersección es distinta del vacío cuando consideramos como subíndices cualquier subconjunto de S .

Antes de continuar, es importante explicar varias nociones relacionadas con la **homotopía**. Dados dos espacios topológicos X e Y , dos aplicaciones continuas $f, g: X \rightarrow Y$ se dice que son **homótopas** (y se denota $f \simeq g$, de hecho, estamos ante una relación de equivalencia en el conjunto de las aplicaciones continuas de X en Y) si existe otra aplicación continua $H: X \times [0, 1] \rightarrow Y$ tal que $H(x, 0) = f(x)$ y $H(x, 1) = g(x)$. Por otro lado, una aplicación continua $f: X \rightarrow Y$ se dice que es una **equivalencia de homotopía** si existe otra aplicación continua $g: Y \rightarrow X$ tal que $g \circ f \simeq Id_X$ y $f \circ g \simeq Id_Y$, en ese caso diremos que X y Y tienen el mismo tipo de homotopía, o que son **homotópicamente equivalentes**. Además, si un espacio topológico X y el espacio unipuntual son homotópicamente equivalentes, se dice que X es **contráctil**.

Teorema 1.11 (Teorema del nervio). *Sean X un espacio topológico paracompacto y $\mathcal{U} = \{U_i\}_{i \in I}$ un recubrimiento abierto de X finito. Si además, para todo subconjunto de $S \subseteq I$ distinto del vacío, se verifica que $\bigcap_{s \in S} U_s$ es contráctil o vacía, entonces se cumple que X y $|N(\mathcal{U})|$ son homotópicamente equivalentes.*

Demostración. Puede consultarse la demostración del teorema en [15]. □

Dicho teorema se puede generalizar más, pero para el contenido tratado en esta memoria nos valdrá con considerarlo así. La importancia de este teorema radica en su uso en la construcción de diversos complejos, como veremos a continuación.

1.3.1. Complejo de Čech

Dados un subconjunto finito $X \subset \mathbb{R}^N$ (el cual sería nuestra nube de puntos, los datos sobre los que queremos trabajar) y el recubrimiento $\mathcal{U} = \{B_\epsilon(x)\}_{x \in X}$ de X , donde $B_\epsilon(x)$ denota la bola abierta de radio ϵ centrada en x (es posible considerar bolas ya que \mathbb{R}^N es un espacio métrico con la distancia usual), el **complejo simplicial de Čech** (asociado a estos datos) se construye como $\check{C}_\epsilon(X) := N(\mathcal{U})$.

Es decir, dado un conjunto finito de puntos (nuestros datos) en un espacio métrico, podemos considerar dichas bolas abiertas de radio ϵ de tal modo que si cualesquiera $n + 1$ bolas tienen intersección no vacía, entonces sus centros formarán un n -símplice.

Como la intersección de las bolas abiertas es vacía o convexa (por tanto contráctil), apoyándonos en el Teorema del nervio sabemos que el poliedro asociado al complejo de

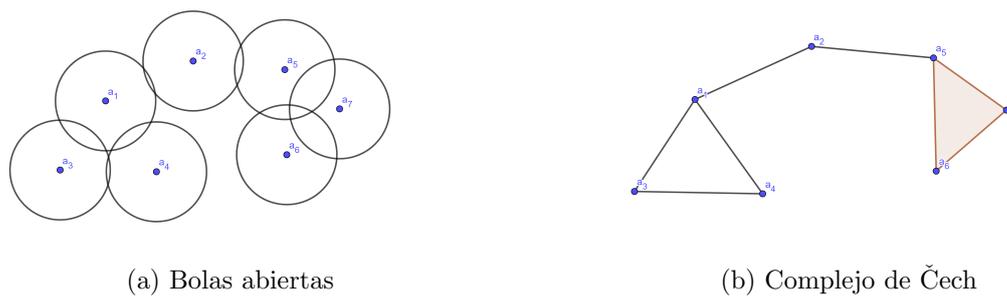


Figura 1.3: Construcción de un complejo de Čech.

Čech, $|N(\mathcal{U})|$, nos permite capturar la topología de X . Sin embargo, este complejo tiene una gran desventaja, pues su construcción es bastante costosa a nivel computacional debido a que se tendrían que comprobar las intersecciones para cada bola.

Observación 1.12. Si consideramos $\epsilon = 0$, por convenio tendremos que $\check{C}_0(X) = \emptyset$. Por otro lado, $\check{C}_\infty(X)$ será un $(|X| - 1)$ -símplice, de hecho a partir de cierto valor d , $\check{C}_p(X)$ con $p \geq d$ siempre será un símplice de dimensión $(|X| - 1)$.

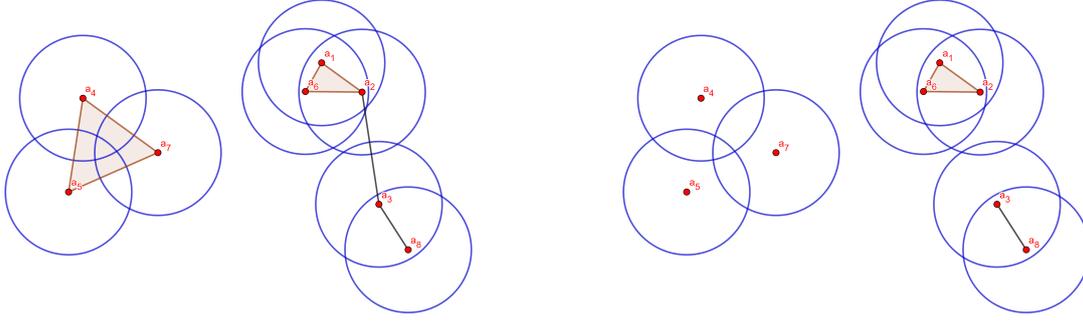
1.3.2. Complejo de Vietoris-Rips

Como posible solución al alto coste computacional en la construcción del complejo de Čech encontramos el complejo de Vietoris-Rips, cuyo cómputo es menos costoso.

Dado un subconjunto finito $X \subset \mathbb{R}^N$ y un parámetro $\epsilon > 0$, definimos el **complejo simplicial de Vietoris-Rips** (asociado a estos datos), denotado por $VR_\epsilon(X)$, como el complejo cuyo conjunto de vértices es X y $\{x_0, \dots, x_n\} \subseteq X$ forman un símplice si $d(x_i, x_j) \leq \epsilon \forall i, j \in \{0, 1, \dots, n\}$. Es fácil ver que se trata de un complejo simplicial, pues si $A = \{x_0, \dots, x_n\} \in VR_\epsilon(X)$, entonces necesariamente los elementos de cualquier subconjunto B de A van a estar a una distancia menor o igual que ϵ unos de otros, por tanto $B \in VR_\epsilon(X)$.

Observación 1.13. Cabe remarcar que en la construcción del complejo de Čech el parámetro ϵ hace referencia al radio de las bolas, sin embargo, en el caso del complejo de Vietoris-Rips observamos que ϵ hace alusión a la distancia entre los vértices. Por lo tanto, si consideramos bolas de radio ϵ centradas en los vértices, en la construcción del complejo de Vietoris-Rips para unir dos vértices necesitamos que dichos vértices estén dentro de la intersección de sus bolas, y éstas han de ser cerradas.

Siguiendo la línea de esta pequeña observación, vamos a mostrar una proposición que



(a) Complejo de Čech

(b) Complejo de Vietoris-Rips

Figura 1.4: Comparación de estos dos tipos de complejos con el mismo ϵ .

nos permite relacionar ambos complejos:

Proposición 1.14. Sean $X \subset \mathbb{R}^N$ finito y $\epsilon > 0$. Entonces se tienen las siguientes inclusiones:

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

Demostración. Primero, tomemos $\sigma = \{x_0, \dots, x_p\} \in \check{C}_\epsilon(X)$. Por definición, sabemos que existe $\hat{x} \in B_\epsilon(x_i) \cap B_\epsilon(x_j)$ para cualesquiera $i, j \in \{0, \dots, p\}$, y por la desigualdad triangular llegamos a que $d(x_i, x_j) \leq d(x_i, \hat{x}) + d(\hat{x}, x_j) \leq \epsilon + \epsilon = 2\epsilon$. Por lo tanto, como esto se verifica para todos los pares de vértices de σ , se cumple que $\sigma \in VR_{2\epsilon}(X)$.

Ahora probemos la otra inclusión. Supongamos que $\tau = \{x_0, \dots, x_q\} \in VR_{2\epsilon}(X)$. Por definición, $d(x_i, x_j) \leq 2\epsilon \forall i, j \in \{0, \dots, q\}$, entonces si consideramos bolas abiertas de radio 2ϵ centradas en x_i y en x_j , dichas bolas tendrán intersección no vacía, puesto que ya de por sí $x_i \in \bar{B}_{2\epsilon}(x_j)$ y $x_j \in \bar{B}_{2\epsilon}(x_i)$ (aunque la construcción del complejo de Čech se haga utilizando bolas abiertas y en este caso hablemos de bolas cerradas, si x_i o x_j estuvieran en la frontera, como $\epsilon > 0$, necesariamente la intersección de las bolas abiertas también sería no vacía). Como esto sucede para todos los pares de vértices de τ , tenemos que $\tau \in \check{C}_{2\epsilon}(X)$. \square

Queda demostrado así que el complejo de Vietoris-Rips es un subcomplejo del complejo de Čech para el mismo ϵ , como pudimos observar en el ejemplo de la figura 1.4.

Observación 1.15. Si nos fijamos, cuando se considera un ϵ lo suficientemente grande, se pueden llegar a formar k -símplices con $k > N$ pese a que nuestra nube de puntos está contenida en \mathbb{R}^N .

1.4. Otros complejos simpliciales

Además de los dos tipos de complejos simpliciales ya explicados, podemos encontrar distintas alternativas a dichos complejos. Esta variedad es útil puesto que siempre se buscará la manera más eficiente de representar los datos y que menos información redundante nos dé. Veremos a lo largo de esta sección un par de ejemplos que también tienen su importancia en el Análisis Topológico de Datos.

1.4.1. Complejo de Delaunay

La idea que subyace en este tipo de complejo simplicial es facilitar el cómputo mediante la reducción de un gran cúmulo de datos sin modificar la estructura topológica. Esta reducción se hace considerando representantes, pensando así cada región del espacio \mathbb{R}^N como un único punto. Para ello se necesita la siguiente definición.

Definición 1.16. Sean $X \subset \mathbb{R}^N$ nuestra nube de datos y $\lambda \in X$ un punto de referencia (representante), llamamos **célula de Voronoi** asociada a λ a

$$V_\lambda = \{x \in \mathbb{R}^N : d(x, \lambda) \leq d(x, \lambda') \forall \lambda' \in X\}.$$

Vemos entonces que dichas células dividen \mathbb{R}^N en regiones “representadas” por un único punto, discretizando así un espacio continuo. Esto es posible ya que $\{V_\lambda\}_{\lambda \in X}$ forman un recubrimiento de \mathbb{R}^N , puesto que para cualquier elemento de \mathbb{R}^N siempre podemos calcular la distancia a los distintos $\lambda \in X$. Dicha colección de las distintas células de Voronoi recibe el nombre de **diagrama de Voronoi** de X .

Tras esto, podemos dar la siguiente definición:

Definición 1.17. El nervio asociado al diagrama de Voronoi es un complejo simplicial que recibe el nombre de **complejo de Delaunay**:

$$\mathcal{D}(X) = \{\sigma \subseteq X : \bigcap_{\lambda \in \sigma} V_\lambda \neq \emptyset\}.$$

1.4.2. Complejo Alfa

Este complejo aparece como una forma de simplificar el complejo de Delaunay, para disminuir así su coste computacional. La idea es tomar la intersección de cada célula de Voronoi V_λ con una bola abierta $B_\epsilon(\lambda)$, obteniendo de este modo un espacio contráctil.

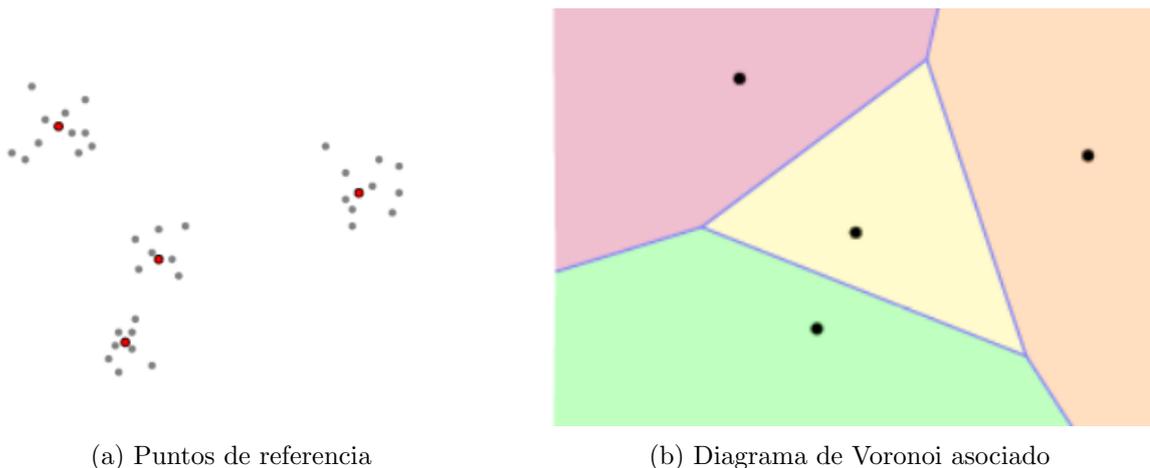


Figura 1.5: Diagrama de Voronoi.

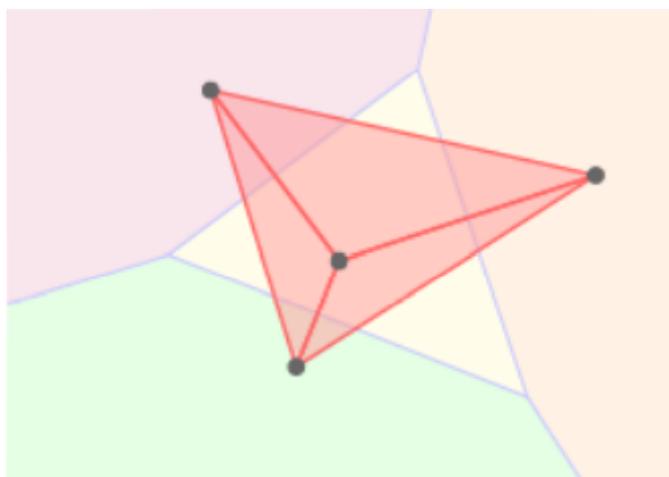


Figura 1.6: Complejo de Delaunay asociado a la figura 1.5.

Por lo tanto, apoyándonos en el Teorema del nervio, intersecando estos subespacios podemos elaborar un complejo simplicial conservando el tipo de homotopía, pues estas intersecciones, $R_\epsilon(\lambda) := V_\lambda \cap B_\epsilon(\lambda)$ para todo $\lambda \in X$, forman un recubrimiento de nuestra nube de puntos X . Veamos cómo construir dicho complejo:

Definición 1.18. Dado $\epsilon > 0$, definimos el **complejo Alfa** como el nervio del recubrimiento dado por $\{R_\epsilon(\lambda)\}_{\lambda \in X}$.

$$\mathcal{A}_\epsilon(X) = \left\{ \sigma \subseteq X : \bigcap_{\lambda \in \sigma} R_\epsilon(\lambda) \neq \emptyset \right\}.$$

Este tipo de complejos nace a partir de las formas Alfa, una definición introducida en [7]. Tanto los complejos Alfa como las formas Alfa han tenido una gran importancia en el desarrollo de modelos de moléculas biológicas como las proteínas o el ADN [14].

Capítulo 2

Homología y códigos de barras

En este segundo capítulo nos centraremos en desarrollar de manera rigurosa las herramientas de la topología algebraica utilizadas en el Análisis Topológico de Datos, así como su interpretación de un modo más “tangible”. Las principales referencias utilizadas en este capítulo son [11] y [27], siendo la primera más introductoria y la segunda está centrada en la homología persistente. Debemos tener en cuenta que en esta memoria siempre trabajaremos con anillos conmutativos con unidad, salvo que se especifique lo contrario.

2.1. Homología

A lo largo de la siguiente sección hablaremos de los complejos de cadenas y definiremos la homología asociada a un complejo de cadenas, herramienta principal de la topología algebraica, que nos permitirá introducir posteriormente la homología persistente. Para hablar de la homología simplicial en secciones posteriores primero daremos una definición genérica y ulteriormente trataremos el caso simplicial.

Definición 2.1. Un **complejo de cadenas** (de A -módulos) es una familia $C = \{C_n, \delta_n\}$, indexada por los enteros ($n \in \mathbb{Z}$), donde los C_n son A -módulos y sus elementos se denominan **n-cadenas**, y $\delta_n: C_n \rightarrow C_{n-1}$ son homomorfismos de A -módulos verificando $\delta_{n-1} \circ \delta_n = 0 \forall n \in \mathbb{Z}$, dichos homomorfismos reciben el nombre de **operador borde** de dimensión n . Podemos representarlo de la siguiente manera:

$$\dots \xrightarrow{\delta_{n+1}} C_n \xrightarrow{\delta_n} C_{n-1} \xrightarrow{\delta_{n-1}} \dots \xrightarrow{\delta_1} C_0 \xrightarrow{\delta_0} C_{-1} \xrightarrow{\delta_{-1}} \dots$$

A partir de un complejo de cadenas podemos definir los siguientes A -submódulos de $C_n \forall n \in \mathbb{Z}$:

$$\begin{aligned} Z_n(C) &:= \text{Ker}(\delta_n), \text{ el } A\text{-módulo de } \mathbf{n}\text{-ciclos.} \\ B_n(C) &:= \text{Im}(\delta_{n+1}), \text{ el } A\text{-módulo de } \mathbf{n}\text{-bordes.} \end{aligned}$$

Como $\delta_{n-1} \circ \delta_n = 0 \forall n \in \mathbb{Z}$, $B_n(C)$ es un A -submódulo de $Z_n(C)$. Con esto podemos dar la siguiente definición:

Definición 2.2. El **A -módulo de homología de dimensión n** asociado a C se define como el cociente

$$H_n(C) := Z_n(C)/B_n(C).$$

Dados dos complejos de cadenas C y C' , llamaremos **homomorfismo de complejos de cadenas**, y lo denotamos por $f: C \rightarrow C'$, a toda colección de homomorfismos de A -módulos $\{f_n: C_n \rightarrow C'_n\}_{n \in \mathbb{Z}}$ verificando para cada $n \in \mathbb{Z}$ que el siguiente diagrama es conmutativo:

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

Además, todo homomorfismo de complejos de cadenas $f: C \rightarrow C'$ induce homomorfismos entre los A -módulos de homología asociados: para cada $n \in \mathbb{Z}$ se tiene

$$\begin{aligned} H_n(f): H_n(C) &\longrightarrow H_n(C') \\ [z] &\longmapsto [f_n(z)] \end{aligned} \tag{2.1}$$

Es inmediato ver que $f_n(z) \in \text{Ker}(\delta'_n)$: si $z \in \text{Ker}(\delta_n)$, entonces $f_{n-1}(\delta_n(z)) = f_{n-1}(0) = 0$ y como por definición $f_{n-1} \circ \delta_n = \delta'_n \circ f_n$, tenemos que $\delta'_n(f_n(z)) = 0$, luego $f_n(z) \in \text{Ker}(\delta'_n)$. Para probar que está bien definida basta con ver que $f_n(B_n(C)) \subseteq B_n(C')$: tomemos $b \in B_n(C)$, entonces $b = \delta_{n+1}(x)$ para algún $x \in C_{n+1}$ y por tanto $f_n(b) = f_n(\delta_{n+1}(x)) = \delta'_{n+1}(f_{n+1}(x)) \in \text{Im}(\delta'_{n+1}) = B_n(C')$, puesto que $f_n \circ \delta_{n+1} = \delta'_{n+1} \circ f_{n+1}$. Así, si $z \sim w$, sabemos que $z - w = b$ para algún $b \in B_n(C)$ y entonces $f_n(z - w) = f_n(z) - f_n(w) = f_n(b) \in B_n(C')$, con lo cual se obtiene que $f_n(z) \sim f_n(w)$.

2.2. Cálculo de la homología

En esta sección mostraremos un algoritmo para el cálculo de la homología asociada a un complejo de cadenas basado en el uso de la forma normal de Smith. Para ello, primero introduciremos una serie de nociones y resultados.

Observación 2.3. A partir de este punto consideraremos que A es un dominio de ideales principales (DIP) y los A -módulos C_n de cualquier complejo de cadenas C son libres y

finito-generados. Por tanto, será para este caso para el que mostraremos el algoritmo mencionado previamente.

Definición 2.4. Sea \mathcal{A} una matriz de orden $m \times n$ con coeficientes en un dominio de ideales principales A , la **forma normal de Smith** asociada a dicha matriz es otra matriz $S \in M_{m \times n}(A)$ tal que $S = PAQ$, donde $P \in M_m(A)$ y $Q \in M_n(A)$ son invertibles, y con la forma

$$S = \left(\begin{array}{ccc|c} s_1 & \cdots & 0 & \\ \vdots & \ddots & \vdots & 0 \\ 0 & \cdots & s_r & \\ \hline & & 0 & 0 \end{array} \right)$$

donde $s_i \neq 0$ para todo $i \in \{1, \dots, r\}$ y cada s_i divide a s_{i+1} para $1 \leq i < r$. Además, estos s_i son únicos salvo multiplicación por unidades (elementos invertibles) y reciben el nombre de **factores invariantes**.

Es conocido el procedimiento para hallar la forma normal de Smith de cualquier matriz \mathcal{A} con coeficientes en un *DIP* (véase [19]). Sin embargo, como a partir de este punto nos centraremos en el caso en el que A sea un **dominio euclídeo** si no se especifica lo contrario, concretamente \mathbb{Z} , presentaremos el algoritmo para hallar la forma normal de Smith en dicha situación, en la cual sabemos que existe un algoritmo de la división.

A continuación, denotaremos la fila i -ésima por f_i y la columna j -ésima por c_j .

Algoritmo para el cálculo de la forma normal de Smith

1. Seleccionamos el pivote a_{ij} con menor norma de \mathcal{A} .
2. “Mejoramos” el pivote según el caso:
 - a) Si el pivote divide a todos los elementos de \mathcal{A} nos movemos al siguiente paso.
 - b) Si el pivote no divide a algún elemento a_{ik} de su fila, entonces reemplazamos c_k por $c_k - q \cdot c_i$, siendo $q \in A$ tal que $r = a_{ik} - q \cdot a_{ij}$ es el resto de dividir a_{ik} entre a_{ij} . Entonces r sería nuestro nuevo pivote y le aplicaríamos este segundo paso.
 - c) Si el pivote no divide a algún elemento a_{kj} de su columna, entonces reemplazamos f_k por $f_k - q \cdot f_j$, siendo $q \in A$ tal que $r = a_{kj} - q \cdot a_{ij}$ es el resto al dividir a_{kj} entre a_{ij} . Entonces r sería nuestro nuevo pivote y le aplicaríamos este segundo paso.

- d) Si el pivote divide a todos los elementos de su fila y columna, pero no a algún elemento a_{st} con $s \neq i$ y $t \neq j$, entonces reemplazamos f_s por $f_s - q \cdot f_j$, siendo $q \in A$ tal que $a_{sj} - q \cdot a_{ij} = 0$. Posteriormente, en la matriz que obtenemos reemplazamos f_j por $f_j + f_s$. De nuevo, realizamos el paso 2 con a_{ij} como pivote, que no ha sido modificado.
3. Trasladamos nuestro pivote a la posición $(1, 1)$ aplicando intercambios de filas y columnas.
 4. Reemplazamos las filas y columnas por filas y columnas convenientes de tal modo que se anulen todos los elementos de la fila y columna de nuestro pivote.
 5. Volvemos al primer paso considerando ahora el menor resultante de eliminar la fila y columna de nuestro pivote. Si el menor es nulo o no existe, acabamos.

En este algoritmo es fácil ver que cada pivote (que será el factor invariante s_i) dividirá al siguiente pivote, pues es elegido de tal modo que divide a todo elemento de la matriz considerada y por lo tanto seguirá dividiendo a todo elemento de dicha matriz tras aplicar operaciones elementales a filas y columnas. En cuanto a la unicidad de estos factores invariantes, véase [19].

Ejemplo 2.5. Tomemos por ejemplo la siguiente matriz con coeficientes en \mathbb{Z} :

$$\mathcal{A} = \begin{pmatrix} 2 & 5 & 1 & 6 \\ 3 & 2 & 3 & 4 \end{pmatrix}$$

Aplicando el algoritmo tenemos los siguientes pasos:

$$\begin{aligned} \begin{pmatrix} 2 & 5 & 1 & 6 \\ 3 & 2 & 3 & 4 \end{pmatrix} &\xrightarrow{f_3 \leftrightarrow f_1} \begin{pmatrix} 1 & 5 & 2 & 6 \\ 3 & 2 & 3 & 4 \end{pmatrix} \xrightarrow{c_2 - 5c_1, c_3 - 2c_1, c_4 - 6c_1} \begin{pmatrix} 1 & 0 & 0 & 0 \\ 3 & -13 & -3 & -14 \end{pmatrix} \xrightarrow{f_2 - 3f_1} \\ &\xrightarrow{f_2 - 3f_1} \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & -13 & -3 & -14 \end{pmatrix} \xrightarrow{c_2 - 4c_3} \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & -1 & -3 & -14 \end{pmatrix} \xrightarrow{c_3 - 3c_2, c_4 - 14c_2} \\ &\xrightarrow{c_3 - 3c_2, c_4 - 14c_2} \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \end{pmatrix} = S \end{aligned}$$

Obtenemos que $s_1 = 1$ y $s_2 = -1$, luego s_1 divide a s_2 . Como hemos comentado previamente, estos son únicos salvo multiplicación por unidades, i.e., podríamos haber tenido $s_1 = 1 = s_2$.

Observación 2.6. Cabe mencionar que también existen softwares que permiten calcular la forma normal de Smith. Por ejemplo, MATLAB proporciona la función $smithForm()$, que devuelve la forma normal de Smith de una matriz cuadrada invertible con coeficientes enteros o polinomios. Además, otros usuarios han desarrollado una función llamada $MNsmithForm()$, la cual calcula la forma normal de Smith de una matriz de polinomios de orden $m \times n$.

A continuación veremos una serie de resultados que nos permitirán desarrollar finalmente el algoritmo para el cómputo de la homología en el caso que mencionamos al principio de esta sección. También nos apoyaremos, en esta y en secciones posteriores, en el siguiente resultado:

Teorema 2.7 (Teorema de descomposición de A -módulos). *Sea M un A -módulo finitamente generado por n elementos sobre un DIP, entonces existen $s_1, \dots, s_r \in A$ con $0 \leq r \leq n$ verificando:*

$$M \simeq A/(s_1) \oplus \dots \oplus A/(s_r) \oplus A^{n-r}$$

donde $(s_r) \subseteq (s_{r-1}) \subseteq \dots \subseteq (s_1)$, y cada s_i es único salvo multiplicación por unidades.

En esta descomposición, se denomina **parte libre** de M (y se denota L_M) a A^{n-r} y $\beta := n-r$ recibe el nombre de **número de Betti** de M . Por otro lado, $A/(s_1) \oplus \dots \oplus A/(s_r)$ es la **parte de torsión** de M , denotada por T_M .

También necesitaremos la siguiente definición:

Definición 2.8. Sea C un complejo de cadenas, llamamos A -módulo de los **n -bordes débiles** al conjunto

$$W_n := \{c \in C_n : \exists \lambda \in A \setminus \{0\} \text{ tales que } \lambda \cdot c \in B_n(C)\}.$$

Es fácil ver que se trata de un A -submódulo de Z_n .

Observación 2.9. En los próximos resultados debemos tener en cuenta que al considerar que A es un DIP se verifica que todo submódulo de un A -módulo libre también es libre. Véase [17] (Teorema 5.1).

Por brevedad pasaremos a usar la notación Z_n, B_n y H_n para $Z_n(C), B_n(C)$ y $H_n(C)$ respectivamente.

Lema 2.10. *Dadas las bases $B_1 = \{[c_1], \dots, [c_k]\}$ de Z_n/W_n y $B_2 = \{d_1, \dots, d_l\}$ de W_n . Entonces, $B = \{c_1, \dots, c_k, d_1, \dots, d_l\}$ es base de Z_n .*

Teorema 2.11. *Sea C un complejo de cadenas de A -módulos libres y finitamente generados sobre un DIP. Tenemos que, para cada $n \in \mathbb{Z}$, existen A -submódulos $U_n, V_n, W_n \subseteq C_n$ tales que*

$$C_n \simeq U_n \oplus V_n \oplus W_n$$

$$Z_n \simeq V_n \oplus W_n$$

Además, las partes libre L_n y de torsión T_n del A -módulo de homología H_n se identifican con $L_n \simeq Z_n/W_n$ y $T_n \simeq W_n/B_n$.

Demostración. Primero consideremos la siguiente sucesión exacta de A -módulos:

$$0 \longrightarrow Z_n \xrightarrow{i_n} C_n \xrightarrow{\delta_n} B_{n-1} \longrightarrow 0$$

A continuación queremos construir un homomorfismo $s_n: B_{n-1} \rightarrow C_n$ verificando que $C_n = Z_n \oplus \text{Im}(s_n)$. Vamos a poder construir un homomorfismo s_n tal que $\delta_n \circ s_n = 1_{B_{n-1}}$. Como B_{n-1} es libre y δ_n epiyectiva, para cada elemento e de la base de B_{n-1} sabemos que existe un $c_e \in C_n$ tal que $\delta_n(c_e) = e$. Podemos definir entonces $s_n: B_{n-1} \rightarrow C_n$ extendiendo linealmente que $s_n(e) := c_e$. Así, se verifica $\delta_n(s_n(e)) = e = 1_{B_{n-1}}(e)$, i.e., $\delta_n \circ s_n = 1_{B_{n-1}}$. Con esta igualdad podemos probar que $C_n = Z_n \oplus \text{Im}(s_n)$: tomemos $c \in C_n$, se cumple $c = c - s_n(\delta_n(c)) + s_n(\delta_n(c))$ y es fácil ver que $c - s_n(\delta_n(c)) \in Z_n$ y $s_n(\delta_n(c)) \in \text{Im}(s_n)$. Además, $Z_n \cap \text{Im}(s_n) = 0$: supongamos que $c \in Z_n \cap \text{Im}(s_n)$, entonces $c = s_n(b)$ para algún $b \in B_{n-1}$, luego $\delta_n(c) = \delta_n(s_n(b)) = 1_{B_{n-1}}(b) = b = 0$ puesto que $c \in Z_n$, y si $b = 0$, necesariamente $c = 0$ por ser s_n homomorfismo.

Ahora, definimos $U_n := \text{Im}(s_n)$ y $V_n := \langle c_1, \dots, c_k \rangle$ con $\{[c_1], \dots, [c_k]\}$ base de Z_n/W_n . Por el lema 2.10, se verifica que $Z_n \simeq V_n \oplus W_n$, y con la igualdad vista anteriormente llegamos a $C_n \simeq U_n \oplus V_n \oplus W_n$.

Para probar la segunda parte del enunciado consideremos primero los homomorfismos canónicos:

$$\pi_1: Z_n \longrightarrow H_n, \pi_1(z) = [z]$$

$$\pi_2: H_n \longrightarrow H_n/T_n, \pi_2([z]) = [\bar{z}]$$

y consideremos también el isomorfismo $h: H_n/T_n \rightarrow L_n$, que sabemos que existe puesto que $H_n = T_n \oplus L_n$. Tomemos ahora el homomorfismo epiyectivo $g := h \circ \pi_2 \circ \pi_1$, donde aplicando el Primer Teorema de Isomorfía obtenemos que $Z_n/\text{Ker}(g) \simeq L_n$. Probemos que $\text{Ker}(g) = W_n$: si $z \in \text{Ker}(g)$, $h([\bar{z}]) = 0$ y como h es un isomorfismo se verifica que $[\bar{z}] = [\bar{0}]$, luego $[z] \in T_n$. Por tanto, por la definición de torsión existe $\lambda \in A \setminus \{0\}$ tal que $\lambda[z] = [\lambda z] = [0]$, es decir, $\lambda z \in B_n \Rightarrow z \in W_n$. De manera análoga se puede probar que

W_n está contenido en $\text{Ker}(g)$. Sustituyendo entonces, tenemos que $Z_n/W_n \simeq L_n$. Para probar el isomorfismo de la parte de torsión basta con hacer el siguiente desarrollo:

$$H_n = Z_n/B_n \simeq (V_n \oplus W_n)/B_n \underset{B_n \subseteq W_n}{\simeq} V_n \oplus (W_n/B_n) \simeq (Z_n/W_n) \oplus (W_n/B_n) \simeq L_n \oplus (W_n/B_n).$$

Como la descomposición $H_n = L_n \oplus T_n$ es única, concluimos entonces que necesariamente $T_n \simeq W_n/B_n$. \square

Proposición 2.12. *Sea $M(\delta_n)$ la matriz asociada al operador borde δ_n respecto de las bases $\{e_1, \dots, e_q\}$ y $\{e'_1, \dots, e'_q\}$ de C_n y C_{n-1} respectivamente, tales que $M(\delta_n)$ está en la forma normal de Smith (y la denotamos $S(\delta_n)$), cuyos factores invariantes son s_1, \dots, s_k . Entonces:*

1. $\{e_{k+1}, \dots, e_q\}$ es una base de Z_n .
2. $\{s_1 e'_1, \dots, s_k e'_k\}$ es una base de B_{n-1} .
3. $\{e'_1, \dots, e'_k\}$ es una base de W_{n-1} .

Demostración. Tomemos $c = \sum_{i=1}^q \lambda_i e_i \in C_n$, entonces $\delta_n(c) = \sum_{i=1}^q \lambda_i \delta_n(e_i)$. Como la matriz asociada está en la forma normal de Smith, tenemos que

$$\delta_n(e_i) = \begin{cases} s_i e'_i & \text{si } 1 \leq i \leq k \\ 0 & \text{si } k+1 \leq i \leq q \end{cases}$$

Por lo tanto, $\delta_n(c) = \sum_{i=1}^k \lambda_i s_i e'_i$. Probemos ahora los distintos puntos.

1. Sea $c \in Z_n$, entonces $\delta_n(c) = 0$, lo cual se verifica si y solo si $\lambda_i s_i = 0$ para todo $i \in \{1, \dots, k\}$ puesto que $\{e'_i\}_{i=1}^k$ son linealmente independientes, y como $s_i \neq 0$ y A se considera anillo íntegro, necesariamente $\lambda_i = 0$ para todo $1 \leq i \leq k$. Por consiguiente, $c = \sum_{i=k+1}^q \lambda_i e_i$, es decir, $\{e_{k+1}, \dots, e_q\}$ es un sistema generador. También son linealmente independientes por ser un subconjunto de la base de C_n , luego forman una base de Z_n .
2. Sea $b \in B_{n-1}$, entonces existe $c \in C_n$ tal que $b = \delta_n(c) = \sum_{i=1}^k \lambda_i s_i e'_i$, i.e., $\{s_1 e'_1, \dots, s_k e'_k\}$ son sistema generador. Como hemos visto previamente, $\sum_{i=1}^k \lambda_i s_i e'_i = 0$ si y solo si $\lambda_i = 0$ para $1 \leq i \leq k$, con lo que concluimos que $\{s_i e'_i\}_{i=1}^k$ son base de B_{n-1} .

3. Como $\delta_n(e_i) = s_i e'_i \in B_{n-1}$ para $1 \leq i \leq k$, por definición $e'_i \in W_{n-1}$ para dichos valores de i . Basta con probar que $\{e'_1, \dots, e'_k\}$ son sistema generador. Sea $c \in W_{n-1}$, entonces existe $\lambda \in A \setminus \{0\}$ tal que $\lambda \cdot c \in B_{n-1}$. Ahora, por el segundo punto $\lambda \cdot c = \sum_{i=1}^k \lambda_i s_i e'_i$. Como $c \in C_{n-1}$, $c = \sum_{i=1}^{q'} \mu_i e'_i$, luego $\lambda \cdot c = \sum_{i=1}^{q'} \lambda \mu_i e'_i$. Comparando estas ecuaciones de $\lambda \cdot c$ llegamos a que $\lambda \mu_i = 0$ para $k+1 \leq i \leq q'$, y como $\lambda \neq 0$ y A es anillo íntegro, $\mu_i = 0$ para $i \in \{k+1, \dots, q'\}$. Por lo tanto, $c = \sum_{i=1}^k \mu_i e'_i$ si $c \in W_{n-1}$, luego $\{e'_i\}_{i=1}^k$ es sistema generador de W_{n-1} y además base, ya que son linealmente independientes por ser subconjunto de una base de C_{n-1} .

□

Tras esto, es fácil ver que con conocer $S(\delta_n)$ y $S(\delta_{n+1})$ podemos obtener el A -módulo de homología H_n . Un algoritmo que nos permite realizar dicho cálculo es el siguiente.

Algoritmo para el cómputo de H_n

1. Tomamos $M(\delta_n)$ y $M(\delta_{n+1})$ respecto de bases arbitrarias.
2. Calculamos $S(\delta_n)$ y $S(\delta_{n+1})$, y se obtiene que:
 - a) El número de Betti β_n de H_n es $\beta_n = \text{rang}(Z_n/W_n) = \text{rang}(Z_n) - \text{rang}(W_n) = |\text{columnas nulas de } S(\delta_n)| - |\text{filas no nulas de } S(\delta_{n+1})|$
 - b) Los elementos s_1, \dots, s_k no nulos y distintos de las unidades en $S(\delta_{n+1})$ son los coeficientes de torsión.
3. Por tanto, $H_n \simeq A^{\beta_n} \oplus A/(s_1) \oplus \dots \oplus A/(s_k)$.

2.3. Homología simplicial

En esta sección expondremos la homología simplicial, una de las herramientas claves para nuestro objetivo final en esta primera parte de la memoria: entender los **códigos de barras**.

Para poder hablar de homología simplicial primero debemos definir la **orientación** de un símplice $\sigma = \{v_0, \dots, v_n\}$, que consiste en tomar clases de equivalencia en el conjunto de las ordenaciones de sus vértices. Diremos que dos símplices $\{v_0, \dots, v_n\}$ y $\{v_{\pi(0)}, \dots, v_{\pi(n)}\}$ están relacionados (i.e., tienen la misma orientación) si y solo si π es una permutación par

de sus vértices. Con esto, vemos que para cada símplice σ solo existen dos orientaciones posibles, excepto para $n = 0$, donde solo tendríamos una orientación. Si tenemos una orientación, σ recibe el nombre de **símplice orientado**, denotado por $[\sigma] = [v_0, \dots, v_n]$. Para considerar la orientación opuesta escribimos $-[\sigma]$.

Cuando un complejo simplicial K tiene todos sus símplices orientados, recibe el nombre de **complejo simplicial orientado**. Con esto, podemos definir el A -módulo $C_n(K; A) \equiv C_n(K)$ como las combinaciones A -lineales (finitas) de los n -símplices orientados de K , cuyos elementos reciben el nombre de **n-cadenas**. Una base de $C_n(K)$ la forman los n -símplices orientados de K , las llamadas **n-cadenas elementales**. Y por lo tanto, dado un complejo simplicial orientado K , definimos un **complejo de cadenas simpliciales** como el complejo $C(K; A) \equiv C(K) = \{C_n(K), \delta_n\}_{n \in \mathbb{Z}}$, donde el operador borde

$$\delta_n: C_n(K) \longrightarrow C_{n-1}(K) \tag{2.2}$$

viene dado por la extensión lineal de

$$\delta_n([v_0, \dots, v_n]) := \sum_{i=0}^n (-1)^i [v_0, \dots, \hat{v}_i, \dots, v_n]$$

y se puede comprobar sin mucha dificultad que $\delta_{n-1} \circ \delta_n = 0$ (véase [23], Lema 5.3). Además, observamos que tanto para $n < 0$ como para $n > \dim(K)$ tenemos $C_n(K) = 0$, es decir, podríamos escribir $C(K) = \{C_n(K), \delta_n\}_{n \in \{0, 1, \dots, \dim(K)\}}$.

A partir de esto, podemos definir la **homología simplicial** (con coeficientes en A) del complejo simplicial orientado K como la homología asociada al complejo de cadenas simpliciales visto, denotada por

$$H_n(K; A) \equiv H_n(K) := H_n(C(K)).$$

La homología simplicial es independiente de la orientación de K , puesto que gracias a la linealidad de δ_n y δ_{n+1} , si un elemento está en $\text{Ker}(\delta_n)$ o en $\text{Im}(\delta_{n+1})$ también lo estará su opuesto.

Observación 2.13. En la actualidad existen diversos softwares que permiten calcular la homología simplicial de un complejo simplicial finito cualquiera con coeficientes en \mathbb{Z} . Uno de ellos es el paquete “Simplicial Homology” para GAP, que se apoya en la forma normal de Smith.

Observación 2.14. El cociente Z_n/B_n se puede entender como eliminar del conjunto de subcomplejos n -dimensionales de K que tienen “forma de ciclo” aquellos que son borde de un $(n+1)$ -símplice, es decir, la homología simplicial $H_n(K)$ nos proporciona la información acerca de los agujeros n -dimensionales presentes en el complejo simplicial K . Veremos

cómo la homología persistente nos ayuda a conocer la trascendencia de cada uno de estos agujeros, con lo que podremos limpiar de ruido nuestros datos.

Antes de continuar, veamos cómo una aplicación simplicial induce un homomorfismo entre las homología simpliciales asociadas. Consideremos una aplicación simplicial $f: K \rightarrow L$ cualquiera, con K y L complejos simpliciales orientados. Esta induce un homomorfismo entre los complejos de cadenas simpliciales asociados

$$f_*: C(K) \rightarrow C(L)$$

donde, para cada $n \geq 0$, se define $(f_*)_n: C_n(K) \rightarrow C_n(L)$ como la extensión lineal de

$$(f_*)_n([v_0, \dots, v_n]) := \begin{cases} [f(v_0), \dots, f(v_n)] & \text{si } f(v_i) \neq f(v_j) \text{ para todo } i \neq j \\ 0 & \text{en cualquier otro caso} \end{cases} \quad (2.3)$$

Observamos que en particular $(f_*)_0(v) = f(v)$ para todo vértice $v \in K^{(0)}$.

Luego, por lo visto en la Sección 2.1, sabemos que dicho homomorfismo entre los complejos de cadenas $f_*: C(K) \rightarrow C(L)$ induce otro homomorfismo (véase 2.1) entre los A -módulos de homología simplicial

$$H_n(f) := H_n(f_*): H_n(K) \rightarrow H_n(L).$$

Observación 2.15. A partir de aquí utilizaremos la notación σ para símlices orientados también, reservando la notación $[\sigma]$ para considerar la clase de equivalencia en la homología.

2.4. Homología persistente

Cuando tratamos con complejos de Čech o de Vietoris-Rips, el TDA se centra en la búsqueda de un valor óptimo de ϵ que permita representar de manera fiel la realidad y la información relevante. Para llevar a cabo dicha búsqueda se trabaja con filtraciones de complejos simpliciales, noción sobre la que se construye la homología persistente. Veamos de qué se trata.

Definición 2.16. Una **filtración** de un complejo simplicial K , denotada por $\mathcal{F} = \{K_i\}_{i \in I}$ con I totalmente ordenado, es una familia de subcomplejos de K tal que se verifica que $K_i \subseteq K_j$ para todo $i \leq j$ y $\emptyset, K \in \mathcal{F}$.

Observación 2.17. Al tratar siempre, a lo largo de esta memoria, con complejos simpliciales finitos, el número de subcomplejos distintos también será finito. Por lo tanto, podemos considerar las filtraciones incluyendo únicamente los subcomplejos K_i diferentes entre sí, i.e., tendríamos que $\emptyset = K_0 \subset K_1 \subset \dots \subset K_n = K$, con $I = \{0, 1, \dots, n\}$.

Ejemplo 2.18. Podemos ver un ejemplo sencillo considerando el complejo de Čech asociado a una nube de puntos. Dado un conjunto $X \subset \mathbb{R}^N$, podemos variar nuestro ϵ e ir tomando los distintos complejos de Čech resultantes, es decir, la filtración que obtenemos sería $\mathcal{F} = \{\check{C}_{\epsilon_i}(X)\}_{i \in I}$ con $\epsilon_i < \epsilon_j$ para todo $i < j$. Como ya hemos visto, por convenio $\check{C}_0(X) = \emptyset$, y además a partir de un ϵ_n suficientemente grande todos nuestros puntos estarían unidos. Análogamente se pueden construir filtraciones para el caso de los complejos de Vietoris-Rips.

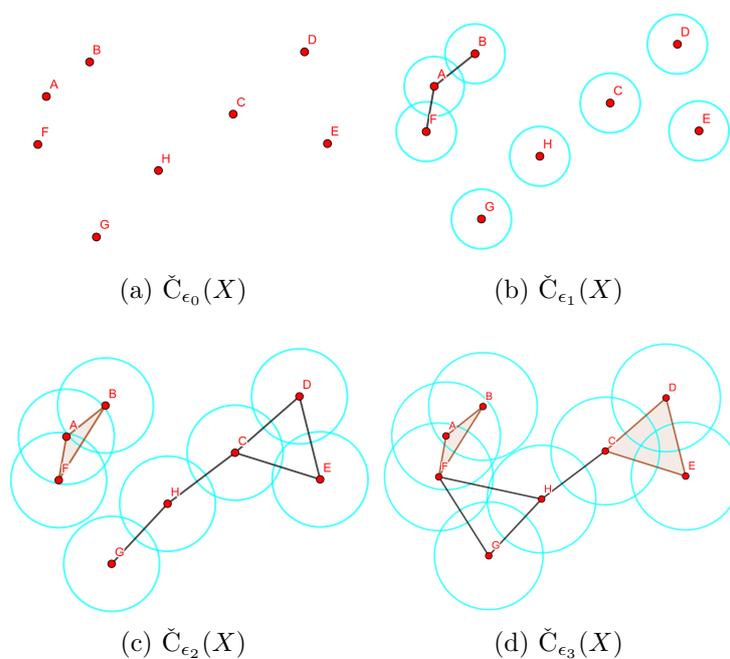


Figura 2.1: Filtración del complejo de Čech $K = \check{C}_{\epsilon_3}(X)$.

Para hablar de homología persistente primero necesitaremos una filtración de un complejo simplicial orientado K . Apoyándonos en las definiciones de complejo de Čech o Vietoris-Rips vistas en la Sección 1.3, podemos dar fácilmente una filtración como hemos detallado en el ejemplo 2.18. Obtendríamos entonces:

$$\mathcal{F} \equiv \emptyset = K_0 \subset K_1 \subset \dots \subset K_n = K$$

A partir de dicha filtración podemos considerar los homomorfismos entre los A -módulos de homología simplicial inducidos por las inclusiones entre los subcomplejos $K_i \subseteq K_j$. Es

decir, para para $p \in \mathbb{N}$ y $0 \leq i \leq j \leq n$ tenemos

$$\begin{aligned} f_p^{i,j}: H_p(K_i) &\longrightarrow H_p(K_j) \\ [z] &\longmapsto [(\rho_{i,j_*})_p(z)] \end{aligned}$$

siendo $\rho_{i,j}: K_i \hookrightarrow K_j$, $z \in Z_p(K_i) \equiv Z_p(C(K_i))$ y $f_p^{i,j} = H_p(\rho_{i,j_*})$ siguiendo la notación de 2.1 y 2.3.

Observación 2.19. En esta situación i y j no tienen por qué ser inmediatamente seguidos. En caso de no serlo simplemente se aplica la composición.

Definición 2.20. Para $p \in \mathbb{N}$ y $i, j \in \{0, 1, \dots, n\}$, definimos el (i, j) -ésimo **A-módulo de homología persistente** de nivel p asociado a la filtración \mathcal{F} como:

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

Asimismo, el (i, j) -ésimo **número de Betti de persistencia** de nivel p de \mathcal{F} es

$$\beta_p^{i,j} = \text{rang}(H_p^{i,j}(\mathcal{F})).$$

Nótese que cuando se considera (i, i) coincide con el número de Betti de $H_p(K_i)$, pues $H_p^{i,i}(\mathcal{F}) = H_p(K_i)$. Este número de Betti $\beta_p^{i,j}$ nos indica el número de clases (agujeros p -dimensionales) que han sobrevivido desde K_i a K_j en la filtración, esto se entenderá mejor posteriormente.

Otra forma de entender la homología persistente, y la razón por la que se denomina “homología”, es mediante el siguiente resultado. Este resultado es de hecho la definición original de homología persistente.

Proposición 2.21. *Dada una filtración \mathcal{F} de un complejo simplicial, tenemos el siguiente isomorfismo de A-módulos:*

$$H_p^{i,j}(\mathcal{F}) \simeq Z_p(K_i)/(B_p(K_j) \cap Z_p(K_i))$$

Demostración. Primero, sabemos que tanto $Z_p(K_i)$ como $B_p(K_j)$ son submódulos de $C_p(K_j)$. Luego, en particular $B_p(K_j) \cap Z_p(K_i)$ es un submódulo de $Z_p(K_i)$, por lo que el cociente del enunciado está bien definido. Ahora, consideremos el siguiente homomorfismo de A-módulos:

$$\begin{aligned} \phi_i: Z_p(K_i) &\longrightarrow H_p(K_j) \\ z &\longmapsto [z] \end{aligned}$$

Y aplicando el Primer Teorema de Isomorfía se obtiene que $Z_p(K_i)/\text{Ker}(\phi_i) \simeq \text{Im}(\phi_i)$.

Observamos que $\text{Ker}(\phi_i) = \{z \in Z_p(K_i) : \phi_i(z) = [0]\} = \{z \in Z_p(K_i) : [z] = [0]\} = \{z \in Z_p(K_i) : z \in B_p(K_j)\} = B_p(K_j) \cap Z_p(K_i)$. Por otro lado, $H_p^{i,j}(\mathcal{F}) = \text{Im}(f_p^{i,j}) = \{f_p^{i,j}([z]) \text{ con } [z] \in H_p(K_i)\} = \{[(\rho_{i,j})_p(z)] \text{ con } z \in Z_p(K_i)\} = \{\phi_i(z) \text{ con } z \in Z_p(K_i)\} = \text{Im}(\phi_i)$. Con esto concluimos que $Z_p(K_i)/(B_p(K_j) \cap Z_p(K_i)) \simeq H_p^{i,j}(\mathcal{F})$. \square

Observación 2.22. Para poder comprender las diferencias entre homología simplicial y homología persistente es importante tener en mente que la homología simplicial está asociada a un complejo simplicial K fijado y la homología persistente se obtiene a partir de una filtración de dicho complejo simplicial.

Así, la homología persistente nos permitirá obtener información cuantitativa acerca de la relevancia de los agujeros p -dimensionales que aparecen a medida que avanzamos por la filtración, i.e., podremos detectar los cambios en la homología del complejo simplicial K a lo largo de \mathcal{F} y de este modo conocer cuándo aparecen y desaparecen dichos agujeros. Veamos cómo.

Definición 2.23. Dada una filtración \mathcal{F} , diremos que un elemento no nulo $\alpha \in H_p(K_i)$ **nace** en K_i si $\alpha \notin H_p^{i-1,i}(\mathcal{F})$. Además, se dice que α **muerde** entrando en K_j si se fusiona con una clase proveniente de un nivel anterior cuando nos movemos desde K_{j-1} a K_j , i.e., si se verifica $f_p^{i,j-1}(\alpha) \notin H_p^{i-1,j-1}(\mathcal{F})$ y $f_p^{i,j}(\alpha) \in H_p^{i-1,j}(\mathcal{F})$.

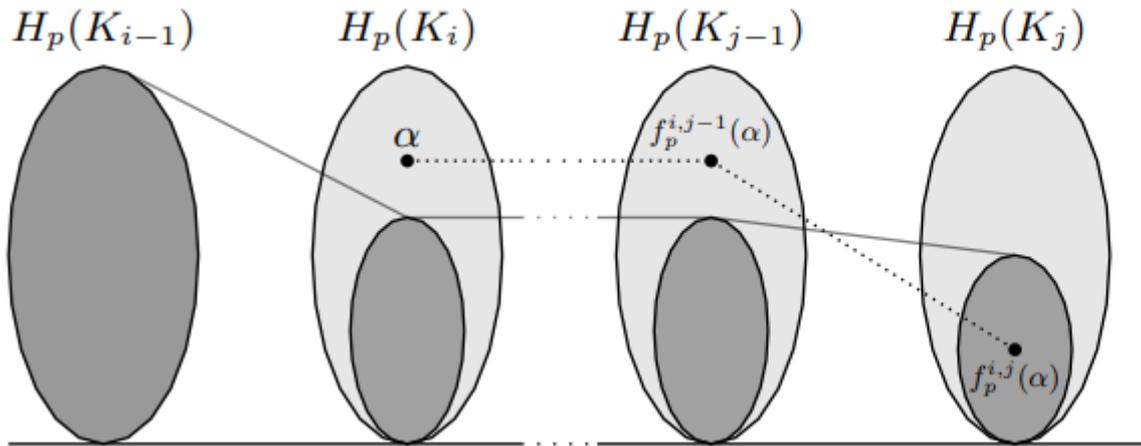


Figura 2.2: Nacimiento y muerte de α , que nace en K_i y muere entrando en K_j .

Cuando esto sucede, decimos que el **índice de persistencia** de α es $\text{pers}(\alpha) = j - i$. Igualmente, puede suceder que α nunca muera, en ese caso diremos que su índice de persistencia es infinito, $\text{pers}(\alpha) = \infty$.

2.5. Códigos de barras

En esta sección trataremos la representación gráfica de la homología persistente, la cual inicialmente se llevaba a cabo utilizando los diagramas de persistencia (introducidos por Edelsbrunner et al.) [8], pero posteriormente Zomorodian y Carlsson desarrollaron los códigos de barras junto con un algoritmo para obtenerlos [4]. Estos últimos son los más empleados hoy en día, por lo que nos centraremos en ellos. Otro método de representación fue introducido por Peter Bubenik, los llamados paisajes de persistencia, pero nosotros no trataremos ese caso en esta memoria, para más detalles véase [2].

Antes de proceder, aclarar que puede haber más de una clase generadora $\alpha \in H_p(K_i)$ que nace en i y muere entrando en j , es decir, podemos hablar de una multiplicidad $\mu_p^{i,j}$, que representa el número de clases generadoras independientes naciendo en i y muriendo en j . Además, se verifica lo siguiente:

$$\mu_p^{i,j} = (\beta_p^{i,j-1} - \beta_p^{i,j}) - (\beta_p^{i-1,j-1} - \beta_p^{i-1,j}). \quad (2.4)$$

Donde notamos que el primer paréntesis nos da el número de clases generadoras independientes que nacen en K_i o antes y mueren entrando en K_j , y el segundo paréntesis nos dice el número de clases generadoras independientes que nacen en K_{i-1} o antes y mueren entrando en K_j . Análogamente, se tiene $\mu_p^{i,\infty} = \beta_p^{i,\infty} - \beta_p^{i-1,\infty}$. Para más detalle, véase [9].

La primera herramienta que apareció para la representación de estas clases fueron los **diagramas de persistencia**, donde cada clase $\alpha \in H_p(\bullet)$ está representada por un punto en el plano, que viene dado por las coordenadas $(i, j) \in \mathbb{N} \times \mathbb{N} \cup \{+\infty\}$ si α nace en i y muere entrando en j .

Supongamos que tenemos una filtración (de 6 pasos) cualquiera de un complejo simplicial K , entonces podríamos obtener el diagrama de persistencia de la figura 2.3 para $H_p(K)$.

Observación 2.24. Como hemos considerado una filtración de 6 pasos, vamos a suponer que la recta $y = 7$ representa el valor ∞ . Por otro lado, la multiplicidad de las clases generadoras se puede representar con un número en rojo al lado de cada punto (si se omite, se considera multiplicidad 1). Dicho esto, en la figura (a) podemos ver cuándo nacen y mueren los distintos agujeros p -dimensionales. Y en la figura (b) se encuentra señalada la región que nos da el número de Betti $\beta_p^{3,4}$, en concreto será el número de puntos dentro de dicha región (contando multiplicidades), i.e., $\beta_p^{3,4} = 3$. De forma general, $\beta_p^{i,j}$ viene dado por el número de puntos que hay en el diagrama de persistencia de H_p dentro de la región $\{(x, y) \in \mathbb{R}^2 : x \leq i, y > j\}$. Por otro lado, una de las desventajas que observamos es que

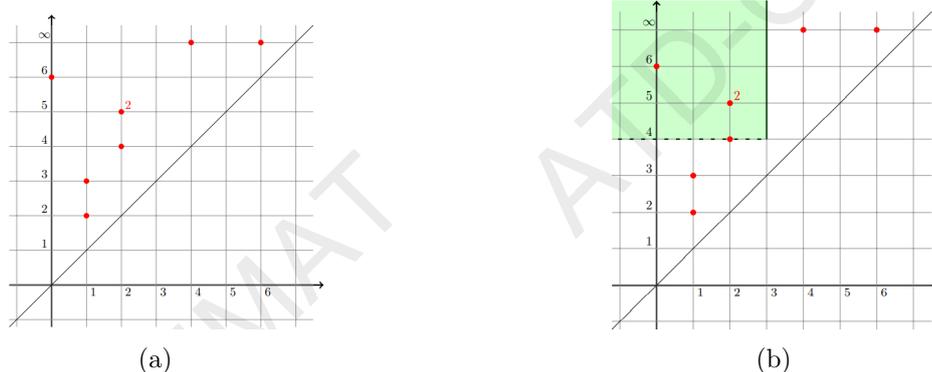


Figura 2.3: Diagrama de persistencia para H_p . Gráficos tomados de [9].

necesitamos un diagrama para cada dimensión.

Como ya hemos mencionado, más adelante se desarrolló una forma de representar la homología persistente mediante el uso de los conocidos **códigos de barras**, donde se muestran los índices de persistencia de las clases generadoras mediante segmentos. Es decir, quedan claramente representadas las etapas de nacimiento-muerte de cada agujero p -dimensional, pudiendo examinar también la multiplicidad en cada caso. Y todo esto en un único gráfico, a diferencia de lo que sucedía con los diagramas de persistencia. Daremos su definición rigurosa a posteriori.

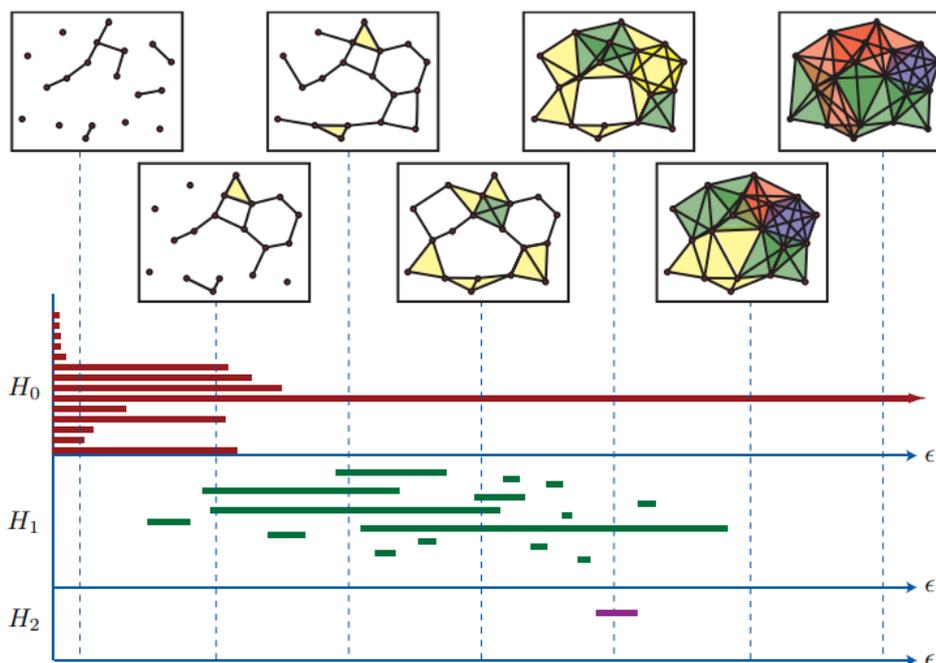


Figura 2.4: Filtración junto con su código de barras asociado. Imagen sacada de [12].

En la figura 2.4 se observa que los segmentos con mayor longitud son las clases que persisten más a lo largo de la filtración, esto se puede entender como los agujeros p -dimensionales que más tiempo permanecen en dicha filtración (cuando alguno de los segmentos aparece con una flecha en un extremo representa una clase que muere en el infinito). Por otro lado, los segmentos de escasa longitud pueden interpretarse como ruido. En particular, el código de barras para H_0 nos dice cuántas componentes conexas tenemos. Asimismo, el rango de $H_p(K_i)$ es el número de segmentos en H_p que intersecan con la línea vertical que representa el paso i -ésimo de la filtración.

2.5.1. Descomposición y cómputo del código de barras

A continuación, veremos la estructura algebraica que subyace en los códigos de barras, los módulos de persistencia. Hablaremos de su descomposición y la relación directa con los índices de persistencia. Para ello, primero daremos un lema y una serie de definiciones.

Lema 2.25. *Un A -módulo M está finitamente generado por n elementos si y sólo si existe un homomorfismo epiyectivo $\phi: A^n \rightarrow M$.*

Demostración. Supongamos primero que M está generado por $\{m_1, \dots, m_n\}$. Tomemos el homomorfismo $\phi: A^n \rightarrow M$ definido por $\phi(a_1, \dots, a_n) := \sum_{i=1}^n a_i m_i$. Es claro que ϕ es epiyectivo.

Recíprocamente, si existe $\phi: A^n \rightarrow M$ epiyectivo, entonces todo elemento $m \in M$ se puede expresar como $m = \phi(a_1, \dots, a_n) = \phi\left(\sum_{i=1}^n a_i e_i\right) = \sum_{i=1}^n \phi(a_i e_i) = \sum_{i=1}^n a_i \phi(e_i)$, para ciertos $a_i \in A$ y siendo $e_i = (0, \dots, \overbrace{1_A}^i, \dots, 0)$, por lo que M estaría generado por $\{\phi(e_i)\}_{i=1}^n$. \square

Definición 2.26. Un A -módulo M se dice que está **finitamente presentado** (o que es de presentación finita) cuando está finitamente generado y el $\text{Ker}(\phi)$ también lo está.

Definición 2.27. Un **módulo de persistencia** sobre A es una familia $\{M_i\}_{i=0}^\infty$ de A -módulos, de manera que para cada $i \leq j$ se tiene un homomorfismo de A -módulos $f_{i,j}: M_i \rightarrow M_j$ verificando:

1. $f_{i,i} = 1_{M_i}$ para $i = 0, 1, 2, \dots$
2. $f_{j,k} \circ f_{i,j} = f_{i,k}$ para todo $i \leq j \leq k$.

Diremos que un módulo de persistencia es de **tipo finito** si existe un $n \in \mathbb{N}$ tal que para todo $n \leq i \leq j$, $f_{i,j}$ es un isomorfismo de A -módulos. Además, se dice que es de **tipo finitamente generado** si es de tipo finito y cada M_i es finitamente generado. Cuando cada M_i sea finitamente presentado diremos que el módulo de persistencia es de **tipo finitamente presentado**.

Definición 2.28. Un anillo A diremos que es **graduado** (o \mathbb{N} -graduado) si se puede descomponer como:

$$A = \bigoplus_{i=0}^{\infty} A_i$$

donde cada A_i es un grupo abeliano y además $a_i \cdot a_j \in A_{i+j}$ para todo $a_i \in A_i$ y $a_j \in A_j$. Cada elemento $a_i \in A_i$ recibe el nombre de **elemento homogéneo** de grado i .

Los ideales generados por elementos homogéneos se llaman **ideales homogéneos o graduados**.

Definición 2.29. Dado un anillo A graduado, un A -módulo M se dice que es un **A -módulo graduado** si puede descomponerse en grupos abelianos:

$$M = \bigoplus_{i=0}^{\infty} M_i$$

de modo que para todo $m_i \in M_i$ y $a_j \in A_j$ se verifica que $a_j \cdot m_i \in M_{i+j}$. A los elementos de M_i se les denomina **elementos homogéneos** de grado i .

Diremos que un A -módulo graduado M es de **tipo finitamente generado** si lo es como A -módulo (todo A -módulo graduado es en particular un A -módulo). Lo mismo sucede con el caso **finitamente presentado**.

A continuación, veremos cómo se relacionan los módulos de persistencia sobre A y los $A[t]$ -módulos graduados. Inicialmente, Zomorodian y Carlsson [27] dieron un resultado que relaciona los A -módulos de persistencia finitamente generados con los $A[t]$ -módulos graduados finitamente generados, pero de manera errónea, puesto que el anillo A requiere más propiedades para que esto se dé. Ante esto, Corbet y Kerber [6] dan un enunciado considerando módulos finitamente presentados, sobre un anillo A con unidad, en lugar de finitamente generados:

Teorema 2.30. *Sea A un anillo, entonces los A -módulos de persistencia finitamente presentados están en correspondencia biunívoca con los $A[t]$ -módulos graduados finitamente presentados.*

Observación 2.31. La acción del anillo de polinomios $A[t]$ dada a partir del A -módulo de persistencia $\mathcal{M} = (\{M_i\}_{i=0}^\infty, \{f_{i,j}\}_{0 \leq i \leq j})$ se define como $t \cdot m = t \cdot (m_i)_{i=0}^\infty := (f_{i,i+1}(m_i))_{i=0}^\infty$ para $t \in A[t]$ y $m \in M = \bigoplus_{i=0}^\infty M_i$. Recíprocamente, si tenemos un $A[t]$ -módulo graduado $M = \bigoplus_{i=0}^\infty M_i$, los homomorfismos $f_{i,i+1}$ del A -módulo de persistencia correspondiente $\mathcal{M} = \{M_i\}_{i=0}^\infty$ vienen definidos por $f_{i,i+1}(m_i) := t \cdot m_i$, con $t \in A[t]$ y $m_i \in M_i$, con lo que podemos definir los $f_{i,j}$ para $0 \leq i \leq j$ simplemente considerando la composición.

Este resultado será suficiente para nosotros, ya que a partir de aquí consideraremos A como un cuerpo, denotándolo F , por lo que será unitario y además todo F -módulo finitamente generado será finitamente presentado y viceversa. Por tanto, bajo estas condiciones es equivalente trabajar con unos módulos u otros. Cabe mencionar también que esta correspondencia uno a uno entre módulos de persistencia y módulos graduados es a nivel categorial, es decir, existe también para los morfismos entre dichos objetos (noción que no hemos definido). Sin embargo, no entraremos en esto ya que para nuestros objetivos es suficiente con saber qué ocurre entre los módulos.

Tras esto, veamos uno de los resultados principales que nos permitirá entender la relación directa de los módulos graduados con los códigos de barras.

Diremos que dos módulos graduados son isomorfos si existe entre ellos un isomorfismo de módulos que preserva la graduación.

Teorema 2.32 (Teorema de descomposición para módulos graduados). *Sea M un $F[t]$ -módulo graduado finitamente generado. Entonces M se puede descomponer de forma única, salvo isomorfismos, como:*

$$M \simeq \left(\bigoplus_{i=1}^{n-m} t^{a_i} \cdot F[t] \right) \oplus \left(\bigoplus_{j=1}^m t^{b_j} \cdot \frac{F[t]}{(t^{c_j})} \right)$$

con $a_i, b_j \in \mathbb{N}$, $c_j \in \mathbb{N}$ y los t^{c_j} son elementos homogéneos tales que t^{c_j} divide a $t^{c_{j+1}}$.

Demostración. Como F es un cuerpo, $F[t]$ es un dominio euclídeo y en particular un DIP. Luego, estamos ante un caso similar al del Teorema de descomposición de A -módulos. La idea de la demostración consiste en aplicar a M el Teorema de descomposición de A -módulos, obteniendo así una parte libre con copias de $F[t]$ y una parte de torsión con copias de $F[t]/(s_j)$, y como (s_j) es un ideal graduado, $(s_j) = (t^{c_j})$ para algún $c_j \in \mathbb{N}$. Finalmente, para mantener la graduación bajo el isomorfismo se necesita ajustar los grados, por ello necesitamos los elementos t^{a_i} y t^{b_j} . Los detalles de la demostración se siguen de [26] \square

Siguiendo esta línea, veremos cómo a partir de una filtración de un complejo simplicial podemos llegar a dicha descomposición.

Supongamos que contamos con una filtración \mathcal{F} de K , para cada uno de los subcomplejos (orientados) de la filtración podemos considerar el complejo de cadenas simpliciales asociado y entre dichos complejos de cadenas los homomorfismos inclusión. Así, observamos que se tiene la siguiente estructura (donde $C_j^i \equiv C_j(K_i)$):

$$\begin{array}{ccccccc}
 \vdots & & \vdots & & \vdots & & \\
 \delta_3^0 \downarrow & & \delta_3^1 \downarrow & & \delta_3^2 \downarrow & & \\
 C_2^0 & \xrightarrow{f_2^0} & C_2^1 & \xrightarrow{f_2^1} & C_2^2 & \xrightarrow{f_2^2} & \dots \\
 \delta_2^0 \downarrow & & \delta_2^1 \downarrow & & \delta_2^2 \downarrow & & \\
 C_1^0 & \xrightarrow{f_1^0} & C_1^1 & \xrightarrow{f_1^1} & C_1^2 & \xrightarrow{f_1^2} & \dots \\
 \delta_1^0 \downarrow & & \delta_1^1 \downarrow & & \delta_1^2 \downarrow & & \\
 C_0^0 & \xrightarrow{f_0^0} & C_0^1 & \xrightarrow{f_0^1} & C_0^2 & \xrightarrow{f_0^2} & \dots \\
 \delta_0^0 \downarrow & & \delta_0^1 \downarrow & & \delta_0^2 \downarrow & & \\
 \vdots & & \vdots & & \vdots & &
 \end{array}$$

Proseguiremos desarrollando estas ideas de forma generalizada, pero sin olvidar que lo explicado se podrá extrapolar al caso de una filtración de un complejo simplicial.

Observamos que al fijar el subíndice j se tiene una estructura de módulo de persistencia:

$$C_j^\bullet \equiv C_j^0 \xrightarrow{f_j^0} C_j^1 \xrightarrow{f_j^1} C_j^2 \xrightarrow{f_j^2} \dots$$

Por el Teorema 2.30, dicho modulo se corresponde con un único $F[t]$ -módulo graduado, que denotaremos por $\alpha(C_j^\bullet)$. De hecho, la familia formada por los $\alpha(C_j^\bullet)$, denotada $\alpha(C)$, constituye un complejo de cadenas de $F[t]$ -módulos graduados donde el nivel p se define como $\alpha(C)_p := \alpha(C_p^\bullet)$.

Finalmente, vemos también que se obtiene otro módulo de persistencia considerando las homologías de dimensión p de cada complejo de cadenas, donde $\tilde{f}_p^i := H_p(f^i)$:

$$H_p(K) \equiv H_p(K_0) \xrightarrow{\tilde{f}_p^0} H_p(K_1) \xrightarrow{\tilde{f}_p^1} H_p(K_2) \xrightarrow{\tilde{f}_p^2} \dots$$

De nuevo, por el Teorema 2.30 se obtiene un módulo graduado, $\alpha(H_p(K))$. Éste será nuestro objeto de estudio a continuación.

Gracias a la correspondencia biyectiva dada por el Teorema 2.30, podemos tratar directamente con $\alpha(H_p(K))$ y con su descomposición como módulo graduado con el fin de recabar la información deseada acerca de $H_p(K)$, i.e., a partir del Teorema 2.32 podremos

conocer el nacimiento y la muerte de las clases generadoras de la homología mediante el grado de dichos elementos y si tienen o no torsión.

Considerando entonces una filtración de un complejo simplicial, tendríamos que las \tilde{f}_p^i son las $f_p^{i,i+1}$ definida en la Sección 2.4, conectando así la homología persistente con esta explicación.

Por lo tanto, apoyándonos en lo recién explicado y abusando de la notación, tendríamos en virtud del Teorema 2.32 el siguiente isomorfismo:

$$H_p(K) \simeq \left(\bigoplus_{i=1}^{n-m} t^{a_i} \cdot F[t] \right) \oplus \left(\bigoplus_{j=1}^m t^{b_j} \cdot \frac{F[t]}{(t^{c_j})} \right).$$

En la descomposición mostrada tenemos la siguiente información: la parte libre se corresponde con las clases generadoras que nacen en la etapa a_i y persisten hasta el infinito, mientras que los elementos de la parte de torsión nos muestran las clases que nacen en b_j y mueren en la etapa $b_j + c_j$, información suficiente para obtener por tanto el código de barras. Para más detalle véase [12] y [27].

Ahora sí podemos dar la siguiente definición:

Definición 2.33. Un **código de barras** es un multiconjunto finito de intervalos acotados inferiormente, los cuales vienen dados por los índices a_i, b_j y c_j vistos anteriormente. Dichos intervalos serían los $[a_i, \infty)$ y los $[b_j, b_j + c_j]$.

Observación 2.34. En la actualidad son conocidas diversas librerías en lenguajes de programación que permiten calcular la homología persistente (y con ello los códigos de barras). Entre ellas podemos encontrar la librería CRAN en R; así como las librerías DIONYSUS, GUDHI, y PHAT, disponibles en C++.

Tras ver cómo se construye un código de barras a partir de una filtración de un complejo simplicial, surge una pregunta de manera natural: ¿se puede realizar el proceso inverso? Es decir, a partir de un código de barras dado, ¿cuáles son las filtraciones de complejos simpliciales que llevan ese código de barras asociado? Hoy en día se está estudiando dicho proceso, el cual resulta claramente más complejo.

Existen algunas situaciones para las que se conoce la filtración de un complejo simplicial (o la familia de filtraciones de complejos simpliciales) cuyo código de barras es el dado, pero para la gran mayoría de casos se sabe muy poco. De hecho, no es fácil probar que para cierto código de barras prefijado exista alguna filtración de un complejo simplicial con dicho código de barras asociado. Para el próspero desarrollo de esta teoría sería altamente útil conseguir un algoritmo que permita realizar estos cálculos dado un código

de barras. Un acercamiento a esto lo proponen J. Leygonie y G. Henselman-Petrusek en [16].

Capítulo 3

Laplaciano combinatorio

En este capítulo introduciremos la noción del Laplaciano combinatorio, así como cómputo matricial, para posteriormente relacionarlo con la homología de un complejo simplicial finito. Para ello, nos apoyaremos principalmente en [13].

3.1. Laplaciano en grafos

Es bien conocida la noción de **grafo**, la cual es un par $G = (V(G), E(G))$, donde $V(G)$ denota el conjunto de vértices del grafo y $E(G)$ las aristas de dicho grafo. En esta sección introduciremos brevemente la visión del problema que se trata en la memoria, pero desde el punto de vista de la teoría de grafos. Antes de comenzar, $v_i \sim v_j$ quiere decir que los vértices v_i y v_j son adyacentes y k_i denota el grado de v_i .

Definición 3.1. Sea un grafo G finito y simple, con $V(G) = \{v_1, \dots, v_n\}$. Se define la **matriz (combinatoria) Laplaciana** de G , denotada \mathcal{L}_G , como la matriz $n \times n$ dada por:

$$(\mathcal{L}_G)_{i,j} = \begin{cases} k_i & \text{si } i = j \\ -1 & \text{si } i \neq j \text{ y } v_i \sim v_j \\ 0 & \text{si } i \neq j \text{ y } v_i, v_j \text{ no son adyacentes} \end{cases}$$

para todo $i, j \in \{1, \dots, n\}$.

Es decir, $\mathcal{L}_G = D_G - A_G$ siendo A_G la matrix de adyacencia del grafo ($a_{ij} = 1$ si los vértices v_i y v_j son adyacentes o 0 en caso de no serlo) y D_G es la matriz diagonal con el grado de los vértices en la diagonal ($d_{ii} = k_i$).

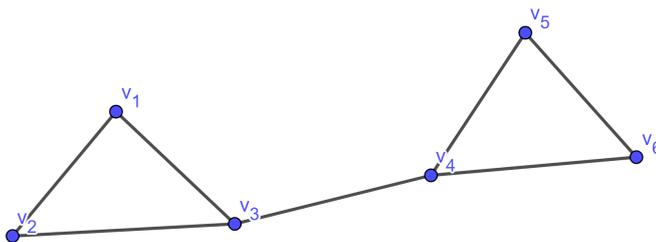


Figura 3.1: Ejemplo de grafo.

Ejemplo 3.2. La matriz combinatoria Laplaciana del grafo de la figura 3.1 sería entonces:

$$\begin{pmatrix} 2 & -1 & -1 & 0 & 0 & 0 \\ -1 & 2 & -1 & 0 & 0 & 0 \\ -1 & -1 & 3 & -1 & 0 & 0 \\ 0 & 0 & -1 & 3 & -1 & -1 \\ 0 & 0 & 0 & -1 & 2 & -1 \\ 0 & 0 & 0 & -1 & -1 & 2 \end{pmatrix}$$

Esta matriz está estrechamente relacionada con la matriz del operador borde δ_1 asociado al grafo G si le dotamos de una orientación y lo consideramos como un complejo simplicial 1-dimensional. El siguiente lema nos dice cómo.

Lema 3.3. *Sea G un grafo finito y simple, con $V(G) = \{v_1, \dots, v_n\}$ y $E(G) = \{e_1, \dots, e_m\}$. Consideremos una orientación arbitraria de las aristas de G y sea \mathcal{B}_1 la matriz del operador borde $\delta_1: C_1(G; \mathbb{R}) \rightarrow C_0(G; \mathbb{R})$ sobre \mathbb{R} . Entonces $\mathcal{L}_G = \mathcal{B}_1 \mathcal{B}_1^T$.*

Demostración. Denotando $b_{ij} = (\mathcal{B}_1)_{ij}$ tenemos que

$$(\mathcal{B}_1)_{ij} = \begin{cases} -1 & \text{si la arista } e_i \text{ comienza en } v_j \\ 0 & \text{si } v_j \notin e_i \\ 1 & \text{si la arista } e_i \text{ acaba en } v_j \end{cases}$$

luego

$$(\mathcal{B}_1 \mathcal{B}_1^T)_{ij} = \sum_{k=1}^m b_{ik} b_{jk} = \begin{cases} k_i & \text{si } i = j \\ -1 & \text{si } i \neq j \text{ y } v_i \sim v_j \\ 0 & \text{si } i \neq j \text{ y } v_i, v_j \text{ no son adyacentes} \end{cases}$$

□

Posteriormente veremos un resultado generalizado de este lema. Recordemos, para esta y próximas secciones, que el operador borde viene definido en 2.2.

3.2. Laplaciano en complejos simpliciales

Comenzamos esta sección dando algunas nociones y propiedades relacionadas con la adyacencia entre símlices en un complejo simplicial, que posteriormente aparecerán reflejadas en el Laplaciano del complejo.

A partir de este punto consideraremos cadenas sobre \mathbb{R} por conveniencia. También hablaremos del **f-vector** de un complejo simplicial K de dimensión n , el cual es un vector de dimensión $n + 1$ de la forma $(f_0(K), \dots, f_n(K))^T$, donde $f_i(K)$ es el número de i -símlices que hay en K .

Como ya vimos, C_i denota el conjunto de combinaciones de i -cadenas de K , y además tiene estructura de \mathbb{R} -espacio vectorial. Una base de C_i viene dada por las i -cadenas elementales de K , luego su dimensión es $f_i(K)$. Por otro lado, si K tiene dimensión n , tanto C_{-1} como todo C_p para $p > n$ son espacios vectoriales formados por el cero únicamente.

A continuación, expondremos una serie de nuevas definiciones:

Definición 3.4. Dado un complejo simplicial finito K , diremos que dos n -símlices $\sigma_1, \sigma_2 \in K$ son **superiormente adyacentes**, y se denota $\sigma_1 \sim_U \sigma_2$, si ambos son caras del mismo $(n + 1)$ -símplice $\tau \in K$, el cual recibe el nombre de **símplice superior común**. También llamaremos **grado superior** de un n -símplice $\sigma \in K$, y lo denotaremos $\deg_U(\sigma)$, al número de $(n + 1)$ -símlices de K de los que σ es una cara.

Además, si dotamos a K de una orientación y tenemos dos n -símlices superiormente adyacentes σ_1, σ_2 con τ su símplice superior común, observando los signos de dichos n -símlices en $\delta_{n+1}(\tau)$ podemos decir que σ_1 y σ_2 tienen **orientación similar** con respecto a τ si los signos son iguales. En caso contrario diremos que tienen **orientación disímil**.

Observación 3.5. Observamos que esta similitud o desigualdad en los signos de los símlices superiormente adyacentes en $\delta_{n+1}(\tau)$ no depende de la orientación de τ , es decir, que tengan o no orientación similar solamente depende de la orientación de dichos símlices superiormente adyacentes.

Lema 3.6 (Unicidad del símplice superior común). *Sea K un complejo simplicial finito y sean $\sigma_1, \sigma_2 \in K$ dos n -símlices distintos. Si éstos son superiormente adyacentes, entonces su símplice superior común es único.*

Demostración. Supongamos que τ_1, τ_2 son dos $(n+1)$ -símlices de K y símlices superiores comunes de σ_1 y σ_2 . Se tiene entonces que $\sigma_1 \cup \sigma_2 \subseteq \tau_1 \cap \tau_2$. Además, por la definición

de s mplice, $\tau_1 \cap \tau_2$ (que sabemos que no son disjuntos puesto que tienen σ_1, σ_2 como caras ambos) es una cara de τ_1 y de τ_2 . Por otro lado, como σ_1 y σ_2 son dos n -s mplices distintos, la  nica cara de τ_1 que los contiene es τ_1 , luego $\tau_1 \cap \tau_2 = \tau_1$. As , vemos que τ_1 es una cara de τ_2 , pero al ser ambos $(n + 1)$ -s mplices no queda otra que $\tau_1 = \tau_2$. \square

Definici n 3.7. Dado un complejo simplicial finito K , diremos que dos n -s mplices $\sigma_1, \sigma_2 \in K$ son **inferiormente adyacentes**, y se denota $\sigma_1 \sim_L \sigma_2$, si ambos tienen un $(n - 1)$ -s mplice com n no vac o $\eta \in K$ como cara, el cual recibe el nombre de **s mplice inferior com n**. Tambi n llamaremos **grado inferior** de un n -s mplice $\sigma \in K$, y lo denotaremos $deg_L(\sigma)$, al n mero de $(n - 1)$ -s mplices no vac os de K que son caras de σ .

De nuevo, si dotamos a K de una orientaci n y tenemos dos n -s mplices inferiormente adyacentes σ_1, σ_2 con η su s mplice inferior com n, observando los signos de η en $\delta_n(\sigma_1)$ y en $\delta_n(\sigma_2)$ podemos decir que η es un **s mplice inferior com n similar** si ambos signos son iguales. En caso contrario diremos que se trata de un **s mplice inferior com n dis mil**.

Observaci n 3.8. De manera an loga al caso de los s mplices superiormente adyacentes, en esta situaci n observamos que si el s mplice inferior com n es o no similar solo depende de la orientaci n de los s mplices inferiormente adyacentes y no de su propia orientaci n.

Lema 3.9 (Unicidad del s mplice inferior com n). *Sea K un complejo simplicial finito y sean $\sigma_1, \sigma_2 \in K$ dos n -s mplices distintos. Si  stos son inferiormente adyacentes, entonces su s mplice inferior com n es la intersecci n de los dos s mplices. Por tanto, dicho s mplice inferior com n es  nico en caso de existir.*

Demostraci n. Supongamos que η es un s mplice inferior com n de σ_1 y σ_2 , en ese caso se tiene que $\eta \neq \emptyset$ y $n > 0$ (pues si $n = 0$, σ_1, σ_2 son v rtices y  stos no pueden ser inferiormente adyacentes). As , tenemos que $\eta \subseteq \sigma_1 \cap \sigma_2$, pero $\sigma_1 \cap \sigma_2$ es una cara de σ_1 conteniendo a η y como η es un $(n - 1)$ -s mplice, las  nicas caras de σ_1 que contienen a η son σ_1 y η . Si fuese σ_1 , entonces tendr amos que $\sigma_1 = \sigma_1 \cap \sigma_2$, luego σ_1 ser a una cara de σ_2 y por dimensiones $\sigma_1 = \sigma_2$, lo cual ser a una contradicci n del enunciado. Por lo tanto, $\eta = \sigma_1 \cap \sigma_2$, con lo cual podemos ver que el s mplice inferior com n de dos n -s mplices es la intersecci n de ambos y por consiguiente  nico. \square

Corolario 3.10. *Dos n -s mplices distintos σ_1, σ_2 de un complejo simplicial finito K son inferiormente adyacentes si y s lo si $\sigma_1 \cap \sigma_2$ es un $(n - 1)$ -s mplice no vac o de K .*

Lema 3.11. *Sea K un complejo simplicial finito orientado y sea n un entero tal que $0 < n < dim(K)$. Dados dos n -s mplices distintos $\sigma_1, \sigma_2 \in K$ superiormente adyacentes*

con τ su símplice superior común, σ_1 y σ_2 tienen orientación similar con respecto a τ si y sólo si tienen un símplice inferior común disímil.

Demostración. Supongamos, sin pérdida de generalidad, que $\tau = \langle v_0, v_1, \dots, v_{n+1} \rangle$ y $\sigma_1 = c_1 \langle v_1, v_2, \dots, v_{n+1} \rangle$, $\sigma_2 = c_2 \langle v_0, v_2, \dots, v_{n+1} \rangle$, con c_1, c_2 iguales a 1 ó -1 según la orientación. Como sabemos que $n > 0$, entonces $n + 1 \geq 2$ y por tanto $\eta = k \langle v_2, v_3, \dots, v_{n+1} \rangle$ es el símplice inferior común de σ_1, σ_2 , con k igual a 1 ó -1 .

Ahora, dado que k, c_1, c_2 valdrán 1 ó -1 , se tiene que

$$\begin{aligned} \langle v_1, v_2, \dots, v_{n+1} \rangle &= c_1 \sigma_1 \\ \langle v_0, v_2, \dots, v_{n+1} \rangle &= c_2 \sigma_2 \\ \langle v_2, v_3, \dots, v_{n+1} \rangle &= k \eta. \end{aligned}$$

Por lo tanto,

$$\begin{aligned} \delta_{n+1}(\tau) &= \delta_{n+1}(\langle v_0, v_1, \dots, v_{n+1} \rangle) = \\ &= \langle v_1, v_2, \dots, v_{n+1} \rangle - \langle v_0, v_2, \dots, v_{n+1} \rangle + R = \\ &= c_1 \sigma_1 - c_2 \sigma_2 + R. \end{aligned}$$

siendo R el resto de sumandos, los cuales no son de interés.

Del mismo modo tenemos:

$$\begin{aligned} \delta_n(\sigma_1) &= \delta_n(c_1 \langle v_1, v_2, \dots, v_{n+1} \rangle) = c_1 \langle v_2, \dots, v_{n+1} \rangle + R_1 = c_1 k \eta + R_1 \\ \delta_n(\sigma_2) &= \delta_n(c_2 \langle v_0, v_2, \dots, v_{n+1} \rangle) = c_2 \langle v_2, \dots, v_{n+1} \rangle + R_2 = c_2 k \eta + R_2 \end{aligned}$$

donde R_1 y R_2 representan el resto de los sumandos en cada caso.

Tras esto, podemos pasar a la demostración.

\Rightarrow) Supongamos que σ_1 y σ_2 tienen orientación similar con respecto a τ , entonces sabemos que $c_1 = -c_2$. Luego, los coeficientes de η en $\delta_n(\sigma_1)$ y $\delta_n(\sigma_2)$, que son $c_1 k$ y $c_2 k$ respectivamente, tienen signos contrarios. Esto significa que η es un símplice inferior común disímil de σ_1 y σ_2 .

\Leftarrow) Supongamos que η es un símplice inferior común disímil de σ_1 y σ_2 . Por definición, tendríamos que $c_1 k$ y $c_2 k$ tienen signos contrarios, i.e., $c_1 k = -c_2 k$, por tanto $c_1 = -c_2$. Con esto, tendríamos que σ_1 y σ_2 aparecen con el mismo signo en $\delta_{n+1}(\tau)$, lo cual implica que ambos tienen orientación similar con respecto a τ . \square

Visto esto, se tiene también el siguiente corolario:

Corolario 3.12. *Sea $n > 0$ un entero. Si dos n -símplices distintos de un complejo simplicial finito son superiormente adyacentes, entonces también son inferiormente adyacentes.*

En lo que sigue estudiaremos los Laplacianos en el caso de los complejos simpliciales. Como estamos considerando nuestro anillo como un cuerpo, pasaremos a hablar de espacios vectoriales de ahora en adelante.

Observación 3.13. Llegados a este punto consideraremos una métrica euclídea $\langle -, - \rangle$ para todos los desarrollos.

Antes de nada necesitaremos definir el producto escalar en $C_k(K)$ para poder hablar también de ortogonalidad posteriormente. Consideremos que $\{\sigma_1, \dots, \sigma_n\}$ forman una base de $C_k(K)$. Dados $c_1 = \lambda_1\sigma_1 + \dots + \lambda_n\sigma_n$ y $c_2 = \mu_1\sigma_1 + \dots + \mu_n\sigma_n$ pertenecientes a $C_k(K)$, el producto escalar se define como:

$$\langle c_1, c_2 \rangle = \sum_{i=1}^n \lambda_i \mu_i$$

Para los operadores borde $\delta_n: C_n \rightarrow C_{n-1}$ denotaremos como \mathcal{B}_n a la matriz asociada considerando las bases estándar de C_n y C_{n-1} con un orden fijado. Además, en esta sección haremos uso de la adjunta de δ_n , $\delta_n^*: C_{n-1} \rightarrow C_n$, cuya matriz asociada respecto de las mismas bases sabemos que es \mathcal{B}_n^T . Recordemos su definición:

Definición 3.14. Dada una aplicación lineal $T: V \rightarrow W$, llamamos **adjunta** de T a la aplicación lineal $T^*: W \rightarrow V$ (que es única) verificando que:

$$\langle T(v), w \rangle = \langle v, T^*(w) \rangle$$

para cualesquiera $v \in V$ y $w \in W$.

Definición 3.15. Sea K un complejo simplicial finito orientado. Dado un entero $n \geq 0$, el **n-ésimo Laplaciano combinatorio** es la aplicación lineal $\Delta_n: C_n \rightarrow C_n$ definida por

$$\Delta_n = \delta_{n+1} \circ \delta_{n+1}^* + \delta_n^* \circ \delta_n.$$

Para facilitar la escritura usaremos la siguiente notación:

$$\begin{aligned} \Delta_n^{UP} &= \delta_{n+1} \circ \delta_{n+1}^* \\ \Delta_n^{DN} &= \delta_n^* \circ \delta_n \end{aligned}$$

Definición 3.16. La **n-ésima matriz Laplaciana** de K respecto de las bases estándar ordenadas de C_n y C_{n-1} , denotada por \mathcal{L}_n , es la matriz asociada a Δ_n . Por lo tanto,

$$\mathcal{L}_n = \mathcal{B}_{n+1} \mathcal{B}_{n+1}^T + \mathcal{B}_n^T \mathcal{B}_n$$

De nuevo, usaremos la notación $\mathcal{L}_n^{UP} = \mathcal{B}_{n+1} \mathcal{B}_{n+1}^T$ y $\mathcal{L}_n^{DN} = \mathcal{B}_n^T \mathcal{B}_n$.

Ejemplo 3.17. La matriz Laplaciana \mathcal{L}_1 de la figura 3.2 es la siguiente:

$$\mathcal{B}_1 = \begin{pmatrix} -1 & 0 & -1 & 0 \\ 1 & -1 & 0 & 0 \\ 0 & 1 & 1 & -1 \\ 0 & 0 & 0 & 1 \end{pmatrix} \quad \text{y} \quad \mathcal{B}_2 = \begin{pmatrix} 1 \\ 1 \\ -1 \\ 0 \end{pmatrix}$$

por tanto

$$\begin{aligned} \mathcal{L}_1 &= \begin{pmatrix} 1 \\ 1 \\ -1 \\ 0 \end{pmatrix} \begin{pmatrix} 1 & 1 & -1 & 0 \end{pmatrix} + \begin{pmatrix} -1 & 1 & 0 & 0 \\ 0 & -1 & 1 & 0 \\ -1 & 0 & 1 & 0 \\ 0 & 0 & -1 & 1 \end{pmatrix} \begin{pmatrix} -1 & 0 & -1 & 0 \\ 1 & -1 & 0 & 0 \\ 0 & 1 & 1 & -1 \\ 0 & 0 & 0 & 1 \end{pmatrix} = \\ &= \begin{pmatrix} 3 & 0 & 0 & 0 \\ 0 & 3 & 0 & -1 \\ 0 & 0 & 3 & -1 \\ 0 & -1 & -1 & 2 \end{pmatrix}. \end{aligned}$$

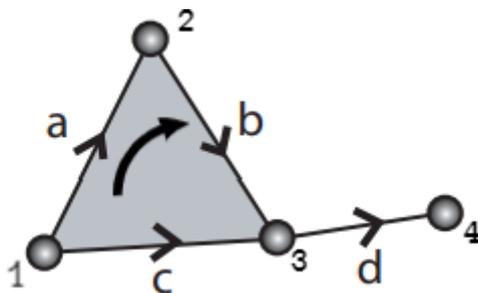


Figura 3.2: Complejo simplicial orientado.

Observación 3.18. La n -ésima matriz Laplaciana nos da información relativa a la adyacencia entre símplices y su orientación. En este ejemplo, la diagonal nos indica el número de 2-símplices adyacentes a cada 1-símplice más 2. Los elementos (i, j) -ésimos fuera de la diagonal nos dicen en qué situación se encuentra cada par de 1-símplices de entre varias posibles respecto a su adyacencia. Esto se detallará y veremos por qué es así con los próximos resultados.

Volviendo al Lema 3.3 enunciado al principio de este capítulo podemos ver la relación entre la matriz Laplaciana de un complejo simplicial y la matriz combinatoria Laplaciana de un grafo, pues la primera es una generalización de la segunda, ya que los grafos finitos y simples se pueden pensar como complejos simpliciales de dimensión 1 inmersos en un espacio Euclídeo. En concreto, $\mathcal{L}_G = \mathcal{L}_0$ para un grafo G finito y simple.

A continuación procederemos a explicar varios resultados que nos facilitan el cómputo de la matriz Laplaciana de un complejo simplicial.

Proposición 3.19. Sean K un complejo simplicial finito orientado y d un entero tal que $0 \leq d \leq \dim(K)$. Sean $\{\sigma_1, \dots, \sigma_n\}$ los d -símplices de K y sean $\{\tau_1, \dots, \tau_m\}$ los $(d+1)$ -símplices de K . Entonces, para $i, j \in \{1, \dots, n\}$, tenemos

$$(\mathcal{L}_d^{UP})_{ij} = \begin{cases} \deg_U(\sigma_i) & \text{si } i = j \\ 1 & \text{si } i \neq j, \sigma_i \sim_U \sigma_j \text{ y tienen orientación similar} \\ -1 & \text{si } i \neq j, \sigma_i \sim_U \sigma_j \text{ y tienen orientación disímil} \\ 0 & \text{si } i \neq j \text{ y } \sigma_i, \sigma_j \text{ no son superiormente adyacentes} \end{cases}$$

Demostración. Es inmediato ver que si $d = \dim(K)$, entonces $\delta_{d+1}: C_{d+1} \rightarrow C_d$ es la aplicación nula, pues C_{d+1} es el espacio vectorial trivial, por lo tanto \mathcal{L}_d^{UP} es una matriz nula. En este caso se verifica el enunciado, ya que no existen $(d+1)$ -símplices en K y por tanto, no podemos encontrar d -símplices superiormente adyacentes.

A continuación, consideremos $0 \leq d < \dim(K)$. Como $\mathcal{L}_d^{UP} = \mathcal{B}_{d+1} \mathcal{B}_{d+1}^T$, la componente ij -ésima de esta matriz viene dada por el producto escalar de la fila i -ésima y la fila j -ésima de \mathcal{B}_{d+1} . Podemos denotar como X e Y a dichas filas respectivamente, y a cada uno de los productos de las componentes de las filas mencionadas los denominaremos sumandos.

Supongamos que $i = j$ y sea τ_k un $(d+1)$ -símplice en K . Si σ_i es una cara de τ_k , esto se traduce en que la componente k -ésima de X es 1 ó -1 dependiendo de la orientación de σ_i . Así, el k -ésimo sumando del producto $X \cdot X$ es 1. Por otro lado, si σ_i no es una cara de τ_k , la componente k -ésima de X sería 0, con lo cual el k -ésimo sumando de $X \cdot X$ sería 0. Luego, $X \cdot X$ nos está dando el número de $(d+1)$ -símplices de los que σ_i es una cara, es decir, $(\mathcal{L}_d^{UP})_{ii} = X \cdot X = \deg_U(\sigma_i)$.

Ahora supongamos que $i \neq j$ y sea τ_k un $(d+1)$ -símplice en K . Sean σ_i, σ_j dos caras de τ_k . Si tienen orientación similar, el k -ésimo sumando de $X \cdot Y$ será 1. Si tienen orientación disímil, dicho sumando será -1 . En caso de que σ_i o σ_j no sea cara de τ_k es inmediato ver que el k -ésimo sumando de $X \cdot Y$ es 0.

Además, por la unicidad del símplice superior común sabemos que como mucho existe un $(d+1)$ -símplice del que σ_i y σ_j son caras. En consecuencia, si σ_i y σ_j son superiormente adyacentes, un único sumando de $X \cdot Y$ será 1 ó -1 y el resto serán todos 0, luego $(\mathcal{L}_d^{UP})_{ij} = X \cdot Y$ es 1 ó -1 en función de las orientaciones de σ_i y σ_j . Si no fueran superiormente adyacentes, $X \cdot Y$ sería 0 y $(\mathcal{L}_d^{UP})_{ij} = X \cdot Y = 0$, probando así el enunciado. \square

Proposición 3.20. Sean K un complejo simplicial finito orientado y d un entero tal que $0 \leq d \leq \dim(K)$. Sean $\{\sigma_1, \dots, \sigma_n\}$ los d -símplices de K . Entonces, para $i, j \in \{1, \dots, n\}$, tenemos

$$(\mathcal{L}_d^{DN})_{ij} = \begin{cases} \deg_L(\sigma_i) & \text{si } i = j \\ 1 & \text{si } i \neq j \text{ y } \sigma_i, \sigma_j \text{ tienen un símplice inferior común similar} \\ -1 & \text{si } i \neq j \text{ y } \sigma_i, \sigma_j \text{ tienen un símplice inferior común disímil} \\ 0 & \text{si } i \neq j \text{ y } \sigma_i, \sigma_j \text{ no son inferiormente adyacentes} \end{cases}$$

Demostración. En el caso $d = 0$ es inmediato ver que el operador $\delta_0: C_0 \rightarrow C_{-1}$ es la aplicación nula, pues C_{-1} es el espacio vectorial trivial. Por consiguiente, \mathcal{L}_0^{DN} es una matriz nula. Como los vértices de un complejo simplicial (es decir, los 0-símplices) no pueden ser inferiormente adyacentes, se verifica el enunciado.

Supongamos ahora que $0 < d \leq \dim(K)$. Como $\mathcal{L}_d^{DN} = \mathcal{B}_d^T \mathcal{B}_d$, la componente ij -ésima de esta matriz viene dada por el producto escalar de la columna i -ésima y la columna j -ésima de \mathcal{B}_d . Denotemos a dichas columnas por X e Y respectivamente, y sean $\{\mu_1, \dots, \mu_m\}$ los $(d-1)$ -símplices de K .

Primero, supongamos que $i = j$ y sea μ_k un $(d-1)$ -símplice de K . Si μ_k es una cara de σ_i , la componente k -ésima de X será 1 ó -1 en función de la orientación de μ_k . Así, el sumando k -ésimo de $X \cdot X$ será 1. Si μ_k no fuera una cara de σ_i , el sumando k -ésimo de $X \cdot X$ sería entonces 0. Por lo tanto, $(\mathcal{L}_d^{DN})_{ii} = X \cdot X = \deg_L(\sigma_i)$, pues $X \cdot X$ nos está dando el número de $(d-1)$ -símplices de K que son caras de σ_i .

A continuación, supongamos que $i \neq j$ y sea μ_k un $(d-1)$ -símplice de K . Si μ_k es una cara de σ_i y de σ_j y es un símplice inferior común similar, entonces la componente k -ésima de X y de Y será 1 ó -1 en ambos casos, por lo que el k -ésimo sumando de $X \cdot Y$ será 1. Si μ_k fuese en cambio un símplice inferior común disímil, el k -ésimo sumando de $X \cdot Y$ sería -1 . Por otro lado, si μ_k no es una cara de σ_i o de σ_j , es fácil ver que el k -ésimo sumando de $X \cdot Y$ sería 0.

Apoyándonos en la unicidad del símplice inferior común sabemos que como máximo σ_i y σ_j tendrán un $(d-1)$ -símplice inferior común, que de hecho sería $\sigma_i \cap \sigma_j$. Con lo cual, si tienen un símplice inferior común similar, solamente uno de los sumandos de $X \cdot Y$ será 1 y el resto 0, i.e., $(\mathcal{L}_d^{DN})_{ij} = X \cdot Y = 1$. Si en cambio se tratase de un símplice inferior común disímil, se tendría $(\mathcal{L}_d^{DN})_{ij} = X \cdot Y = -1$. Y si no tuviesen ninguna cara en común, $(\mathcal{L}_d^{DN})_{ij} = X \cdot Y = 0$, probando así el enunciado. \square

Finalmente, haciendo uso de los resultados previos podemos demostrar el siguiente

teorema, que nos dará el modo de calcular la matriz Laplaciana de un complejo simplicial.

Teorema 3.21 (Teorema de la matriz Laplaciana). *Sean K un complejo simplicial finito orientado, d un entero tal que $0 \leq d \leq \dim(K)$ y $\{\sigma_1, \dots, \sigma_n\}$ los d -símplices de K . Entonces, para $i, j \in \{1, \dots, n\}$, tenemos*

1. Si $d = 0$

$$(\mathcal{L}_d)_{ij} = \begin{cases} \deg_U(\sigma_i) & \text{si } i = j \\ -1 & \text{si } \sigma_i, \sigma_j \text{ son distintos y superiormente adyacentes} \\ 0 & \text{si } \sigma_i, \sigma_j \text{ son distintos y no son superiormente adyacentes} \end{cases}$$

2. Si $0 < d \leq \dim(K)$

$$(\mathcal{L}_d)_{ij} = \begin{cases} \deg_U(\sigma_i) + d + 1 & \text{si } i = j \\ 1 & \text{si } i \neq j \text{ y } \sigma_i, \sigma_j \text{ no son superiormente adyacentes} \\ & \text{pero tienen un símplice inferior común similar} \\ -1 & \text{si } i \neq j \text{ y } \sigma_i, \sigma_j \text{ no son superiormente adyacentes} \\ & \text{pero tienen un símplice inferior común disímil} \\ 0 & \text{si } i \neq j \text{ y } \sigma_i, \sigma_j \text{ son superiormente adyacentes} \\ & \text{o no son inferiormente adyacentes} \end{cases}$$

Demostración. 1. Consideremos $d = 0$. Sabemos que $\mathcal{L}_0 = \mathcal{B}_1 \mathcal{B}_1^T$ y cabe mencionar que para los vértices se considera una única orientación, con lo cual dos vértices cualesquiera superiormente adyacentes siempre tienen orientación disímil. Tras esto, vemos que esta parte se demuestra directamente a partir de la Proposición 3.19.

2. Consideremos ahora $0 < d \leq \dim(K)$. Si $i = j$, a partir de las Proposiciones 3.19 y 3.20 sabemos que $(\mathcal{L}_d)_{ii} = (\mathcal{L}_d^{UP})_{ii} + (\mathcal{L}_d^{DN})_{ii} = \deg_U(\sigma_i) + \deg_L(\sigma_i)$. Además, para $d > 0$ se tiene que el número de caras de dimensión $(d - 1)$ de σ_i es siempre $d + 1$, luego podemos afirmar que $(\mathcal{L}_d)_{ii} = \deg_U(\sigma_i) + d + 1$.

De modo análogo, apoyándonos en las proposiciones mencionadas, podemos tratar fácilmente el caso $i \neq j$. Si σ_i y σ_j no son superiormente adyacentes pero tienen un símplice inferior común similar, entonces $(\mathcal{L}_d)_{ij} = 0 + 1 = 1$. Si σ_i y σ_j no son superiormente adyacentes pero tienen un símplice inferior común disímil, se obtiene que $(\mathcal{L}_d)_{ij} = 0 - 1 = -1$.

Si σ_i, σ_j son superiormente adyacentes, por el Lema 3.11 sabemos que si tienen orientación similar, entonces también tienen un símplice inferior común disímil, luego $(\mathcal{L}_d)_{ij} = 1 - 1 = 0$. Si en cambio tienen orientación disímil, apoyándonos en dicho lema sabemos que $(\mathcal{L}_d)_{ij} = -1 + 1 = 0$.

Por otra parte, si σ_i, σ_j no son inferiormente adyacentes, entonces tampoco pueden ser superiormente adyacentes por el Corolario 3.12. Con esto, es fácil ver que $(\mathcal{L}_d)_{ij} = 0 + 0 = 0$.

□

Observación 3.22. Apoyándonos de nuevo en el hecho de que un grafo G se puede entender como un complejo simplicial de dimensión 1, el Lema 3.3 es un corolario de este último teorema, ya que a partir de él podemos ver que la definición de \mathcal{L}_G coincide con \mathcal{L}_0 de G como complejo simplicial.

Observación 3.23. Cabe mencionar que la d -ésima matriz Laplaciana de un complejo simplicial K depende únicamente de las orientaciones de los d -símplices de K , como se observa en el teorema previo.

Corolario 3.24. *Sea K un complejo simplicial finito orientado.*

1. Sean v_1, \dots, v_m los vértices de K . Se verifica para $i \in \{1, \dots, m\}$:

$$\Delta_0(v_i) = \sum_{v_j \sim_U v_i} (v_i - v_j)$$

2. Sean d un entero tal que $0 < d \leq \dim(K)$ y $\sigma_1, \dots, \sigma_n$ los d -símplices orientados de K . Se verifica para $i \in \{1, \dots, n\}$:

$$\Delta_d(\sigma_i) = \sum_{\sigma_j \sim_L \sigma_i} (\sigma_i + s_{ij}\sigma_j) + \sum_{\sigma_k \sim_U \sigma_i} (\sigma_i - s_{ik}\sigma_k)$$

donde s_{ij} es 1 ó -1 dependiendo de si σ_i y σ_j tienen un símplex inferior común similar o disímil respectivamente. Del mismo modo, s_{ik} es 1 ó -1 en función de si σ_i y σ_k tienen un símplex superior común similar o disímil respectivamente.

Demostración. 1. Por el teorema previo sabemos qué forma tiene \mathcal{L}_0 . Así, si representamos cada vértice v_i con el vector i -ésimo de la base estándar de \mathbb{R}^m , denotado e_i , entonces $\Delta_0(v_i)$ es exactamente $\mathcal{L}_0 e_i$. Es decir, la i -ésima columna de \mathcal{L}_0 :

$$\Delta_0(v_i) = \deg_U(v_i)v_i - \sum_{v_j \sim_U v_i} v_j$$

Como el número de vértices v_j , con $j \in \{1, \dots, m\} - \{i\}$, que son superiormente adyacentes a v_i es de hecho $\deg_U(v_i)$, se tiene:

$$\Delta_0(v_i) = \sum_{v_j \sim_U v_i} (v_i - v_j)$$

2. Análogamente, podemos representar cada σ_i con el i -ésimo elemento de la base estándar de \mathbb{R}^n , denotado e_i . Por tanto, $\Delta_d(\sigma_i) = \mathcal{L}_d e_i$, es decir, la i -ésima columna de \mathcal{L}_d . Apoyándonos en el teorema anterior tenemos:

$$\Delta_d(\sigma_i) = (\deg_U(\sigma_i) + d + 1)\sigma_i + \sum_{\sigma_j \sim_L \sigma_i} s_{ij}\sigma_j - \sum_{\sigma_k \sim_U \sigma_i} s_{ik}\sigma_k$$

Sabemos que dos símlices superiormente adyacentes también son inferiormente adyacentes, luego la resta de esos dos sumatorios hace que se cancelen los términos que tengan un símlice σ_k superiormente adyacente a σ_i , que es lo que se busca, ya que el elemento k -ésimo de la i -ésima columna de \mathcal{L}_d es 0 cuando σ_i y σ_k son superiormente adyacentes. Los s_{ij} representan los signos dependiendo de si los símlices inferiores comunes son similares o no e igualmente, los s_{ik} son los signos en función de si los símlices superiores comunes son similares o no. Finalmente, como sabemos que el número de símlices inferiormente adyacentes a σ_i es $d + 1$ y el número de símlices superiormente adyacentes a σ_i es $\deg_U(\sigma_i)$, podemos concluir que:

$$\Delta_d(\sigma_i) = \sum_{\sigma_j \sim_L \sigma_i} (\sigma_i + s_{ij}\sigma_j) + \sum_{\sigma_k \sim_U \sigma_i} (\sigma_i - s_{ik}\sigma_k)$$

□

3.3. Espectro Laplaciano de un complejo simplicial

En esta sección trataremos brevemente la noción de espectro de un laplaciano y finalmente relacionaremos la homología de la que hablamos en capítulos anteriores con el laplaciano combinatorio de un complejo simplicial K . Para ello trabajaremos con los valores propios de la matriz laplaciana, en particular con el valor propio cero, como veremos más adelante. Se seguirán principalmente las referencias [13], [18], [20] y [21].

Definición 3.25. Sea K un complejo simplicial finito orientado y sea d un entero tal que $0 \leq d \leq \dim(K)$. Llamamos **espectro del d -ésimo Laplaciano** de K al conjunto de los valores propios de Δ_d de K , y lo denotamos por $\text{Spec}(\Delta_d(K))$.

Teorema 3.26. *Sea K un complejo simplicial finito y sea d un entero tal que $0 \leq d \leq \dim(K)$. Entonces $\text{Spec}(\Delta_d(K))$ no depende de la orientación considerada para los d -símlices de dicho complejo simplicial.*

Demostración. A partir del teorema de la matriz Laplaciana podemos deducir fácilmente que los elementos de la diagonal de dicha matriz no dependen de la orientación de los

símplices. Igualmente, que los elementos fuera de la diagonal sean cero o distintos de cero tampoco depende de la orientación. Con esto, vemos que lo único que depende de la orientación de los d -símplices es el signo de los elementos fuera de la diagonal que son distintos de cero.

Si $d = 0$ no es necesario probar nada, pues los elementos fuera de la diagonal serán 0 ó -1 .

Supongamos que $0 < d \leq \dim(K)$. Denotemos a los d -símplices de K por $\{\sigma_1, \dots, \sigma_n\}$. Ahora, consideremos una orientación para los d -símplices de K y denotemos a este complejo simplicial por K_1 . Por otro lado, consideremos otra orientación para dichos d -símplices de tal modo que coincide con la de K_1 excepto para uno de los d -símplices, por ejemplo σ_p (que tendrá la orientación contraria) para $p \in \{1, \dots, n\}$, y llamemos a dicho complejo K_2 . De este modo, hablaríamos de σ_p en K_1 y de $-\sigma_p$ en K_2 .

Sean $\mathcal{L}_d(K_1)$ y $\mathcal{L}_d(K_2)$ las d -ésimas matrices Laplacianas de K_1 y K_2 respectivamente. Supongamos que σ_i y σ_p son inferiormente adyacentes, pero no superiormente adyacentes. Sea μ el símplex inferior común de dichos símplices. Como se verifica que $\delta_d(-\sigma_p) = -\delta_d(\sigma_p)$, sabemos que el signo que acompaña a μ en $\delta_d(\sigma_p)$ es el opuesto al que acompaña a μ en $\delta_d(-\sigma_p)$. Así, si por ejemplo μ aparece con el mismo signo en $\delta_d(\sigma_i)$ y en $\delta_d(\sigma_p)$, entonces aparecerá con diferente signo en $\delta_d(\sigma_i)$ y en $\delta_d(-\sigma_p)$. Por lo tanto, si μ es un símplex inferior común similar en K_1 , entonces es un símplex inferior común disímil en K_2 y viceversa. Se verifica por consiguiente que $(\mathcal{L}_d(K_1))_{ip} = (\mathcal{L}_d(K_1))_{pi}$ tiene signo opuesto que $(\mathcal{L}_d(K_2))_{ip} = (\mathcal{L}_d(K_2))_{pi}$.

Visto esto, sabemos que $\mathcal{L}_d(K_1)$ y $\mathcal{L}_d(K_2)$ son idénticas exceptuando que los signos de los elementos distintos de cero fuera de la diagonal en la fila p o en la columna p son opuestos, pues lo anterior se verifica para cualquier σ_i que sea inferiormente adyacente con σ_p y no superiormente adyacente.

Consideremos ahora una matriz $n \times n$ para cada $x \in \{1, \dots, n\}$, denotada por J_x , tal que:

$$(J_x)_{ij} = \begin{cases} -1 & \text{si } i = x = j \\ 1 & \text{si } x \neq i = j \\ 0 & \text{en el resto de casos} \end{cases}$$

Se observa que J_x es invertible para cada $x \in \{1, \dots, n\}$. Además, $\mathcal{L}_d(K_2) = J_p \mathcal{L}_d(K_1) J_p$, y como $J_x^{-1} = J_x$ para cada $x \in \{1, \dots, n\}$ se tiene que $\mathcal{L}_d(K_1)$ y $\mathcal{L}_d(K_2)$ son matrices semejantes. Sabemos que los valores propios de dos matrices semejantes son los mismos, con lo cual $\text{Spec}(\Delta_d(K_1)) = \text{Spec}(\Delta_d(K_2))$.

Como estamos considerando complejos simpliciales finitos, bastaría con aplicar induc-

ción para terminar la demostración. \square

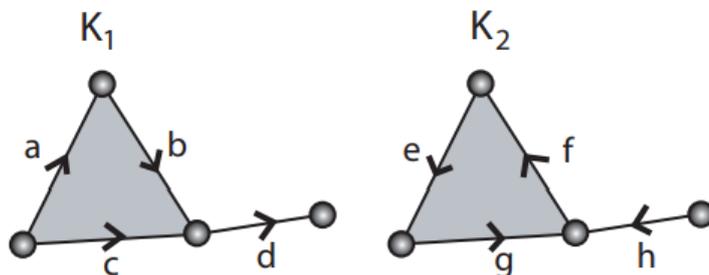


Figura 3.3: Diferentes orientaciones en un mismo complejo simplicial.

Ejemplo 3.27. Si calculamos $\mathcal{L}_1(K_1)$ y $\mathcal{L}_1(K_2)$ para los complejos de la figura 3.3 obtenemos:

$$\mathcal{L}_1(K_1) = \begin{pmatrix} 3 & 0 & 0 & 0 \\ 0 & 3 & 0 & -1 \\ 0 & 0 & 3 & -1 \\ 0 & -1 & -1 & 2 \end{pmatrix} \quad \text{y} \quad \mathcal{L}_1(K_2) = \begin{pmatrix} 3 & 0 & 0 & 0 \\ 0 & 3 & 0 & -1 \\ 0 & 0 & 3 & 1 \\ 0 & -1 & 1 & 2 \end{pmatrix}.$$

Observamos que ambos complejos son idénticos desde una perspectiva geométrica, sin embargo tienen símplices con orientaciones diferentes. Aún así, se tiene que efectivamente $\text{Spec}(\Delta_1(K_1)) = \{1, 3, 3, 4\} = \text{Spec}(\Delta_1(K_2))$.

Observación 3.28. Previamente vimos que la d -ésima matriz Laplaciana de un complejo dependía solo de las orientaciones de los d -símplices. Ahora observamos que los valores propios de estas matrices no dependen siquiera de la orientación, por lo que podemos entonces tratar con el espectro sin fijar una orientación.

A continuación daremos una serie de resultados mostrando ciertas características del espectro del d -ésimo Laplaciano de K . Inicialmente nos centraremos en los valores propios no nulos, denotados $\text{Spec}_{NZ}(\Delta_d(K))$, y posteriormente en el valor propio cero, dando con ello el resultado principal de esta parte de la memoria. En estos resultados \cup_M denota la unión como multiconjuntos, es decir, si hay elementos repetidos en los conjuntos, al considerar la unión se mantienen esas repeticiones.

Teorema 3.29. *Sea E un espacio de dimensión finita sobre un cuerpo con producto escalar y sean T y U operadores lineales autoadjuntos sobre E tales que $T \circ U = 0 = U \circ T$, entonces $\text{Spec}_{NZ}(T + U) = \text{Spec}_{NZ}(T) \cup_M \text{Spec}_{NZ}(U)$.*

Demostración. Véase [13] (Teorema 2.2.5). \square

Corolario 3.30. *Dados un complejo simplicial finito orientado K y $d \in \mathbb{Z}$ tal que $0 \leq d \leq \dim(K)$, se verifica:*

$$\text{Spec}_{NZ}(\Delta_d(K)) = \text{Spec}_{NZ}(\Delta_d^{UP}(K)) \cup_M \text{Spec}_{NZ}(\Delta_d^{DN}(K))$$

Demostración. Se tiene que Δ_d^{UP} y Δ_d^{DN} son autoadjuntos como operadores lineales. Por definición, sabemos que la imagen de un operador borde está contenida en el núcleo del operador borde de una dimensión menor. Sabemos además que el adjunto de la composición es la composición de los adjuntos en orden inverso. Tenemos entonces:

$$\begin{aligned} \Delta_d^{UP} \circ \Delta_d^{DN} &= (\delta_{d+1} \circ \delta_{d+1}^*) \circ (\delta_d^* \circ \delta_d) = \delta_{d+1} \circ (\delta_{d+1}^* \circ \delta_d^*) \circ \delta_d = \delta_{d+1} \circ (\delta_d \circ \delta_{d+1})^* \circ \delta_d = \\ &= \delta_{d+1} \circ 0^* \circ \delta_d = \delta_{d+1} \circ 0 \circ \delta_d = 0 \end{aligned}$$

Análogamente, tenemos que $\Delta_d^{DN} \circ \Delta_d^{UP} = 0$. Luego, dichos operadores son autoadjuntos y además se anulan mutuamente. Apoyándonos entonces en el teorema 3.29 podemos concluir. \square

Teorema 3.31. *Sean F un cuerpo, y $U: F^m \rightarrow F^n$ y $T: F^n \rightarrow F^m$ con $n, m \in \mathbb{N}$ dos aplicaciones lineales, entonces $\text{Spec}_{NZ}(T \circ U) = \text{Spec}_{NZ}(U \circ T)$.*

Demostración. Denotaremos por 0 al vector cero de la dimensión correspondiente y suprimiremos el símbolo de composición \circ por brevedad. Sea $\lambda \in F$ un valor propio no nulo de UT . Tomemos un vector propio de UT , $x \in F^m$, de valor propio λ , entonces $UT(x) = \lambda x$, luego $(TU)T(x) = T(UT)(x) = T(\lambda x) = \lambda T(x)$. Si $T(x) = 0$, entonces $\lambda x = UT(x) = U(0) = 0$, y por ser x un vector propio de valor propio no nulo, $x \neq 0$, luego tendríamos que $\lambda = 0$, llegando a contradicción. Así, $T(x) \neq 0$, por lo que $T(x)$ es un vector propio de TU de valor propio λ . Análogamente se tiene que si $y \in F^n$ es un vector propio de TU de valor propio no nulo λ' , entonces $U(y)$ es un vector propio de UT con el mismo valor propio. Es decir, $\lambda \neq 0$ es un valor propio de TU si y solo si lo es de UT . Además, T manda vectores propios de UT a vectores propios de TU con el mismo valor propio y U manda vectores propios de TU a vectores propios de UT con el mismo valor propio.

Con esto, podemos construir las siguientes aplicaciones:

$$\begin{aligned} \phi: E_\lambda(UT) &\rightarrow E_\lambda(TU), \phi(x) = T(x) \\ \varphi: E_\lambda(TU) &\rightarrow E_\lambda(UT), \varphi(y) = \frac{1}{\lambda}U(y) \end{aligned}$$

donde $E_\lambda(\bullet)$ denota el espacio de vectores propios de \bullet de valor propio λ . De este modo, se verifica para todo $x \in E_\lambda(UT)$ y todo $y \in E_\lambda(TU)$ que

$$(\varphi \circ \phi)(x) = \frac{1}{\lambda}UT(x) = \frac{1}{\lambda}\lambda x = x$$

y

$$(\phi \circ \varphi)(y) = T\left(\frac{1}{\lambda}U(y)\right) = \frac{1}{\lambda}TU(y) = \frac{1}{\lambda}\lambda y = y$$

Luego, ϕ y φ son inversas, con lo cual $E_\lambda(UT)$ y $E_\lambda(TU)$ son isomorfos. Finalmente, como la dimensión de los espacios de vectores propios viene dada por la multiplicidad del valor propio y las dimensiones de $E_\lambda(UT)$ y $E_\lambda(TU)$ coinciden, sabemos que todo valor propio $\lambda \neq 0$ tiene la misma multiplicidad en UT y en TU . Es decir, $\text{Spec}_{\mathbb{N}\mathbb{Z}}(T \circ U) = \text{Spec}_{\mathbb{N}\mathbb{Z}}(U \circ T)$. \square

Corolario 3.32. *Dado un complejo simplicial finito orientado K , se verifica:*

$$\text{Spec}_{\mathbb{N}\mathbb{Z}}(\Delta_{d-1}^{UP}(K)) = \text{Spec}_{\mathbb{N}\mathbb{Z}}(\Delta_d^{DN}(K))$$

para todo entero positivo $d \leq \dim(K)$.

Demostración. Por definición, $\Delta_{d-1}^{UP} = \delta_d \circ \delta_d^*$ y $\Delta_d^{DN} = \delta_d^* \circ \delta_d$. A partir del teorema 3.31 se concluye. \square

Teorema 3.33. *Dados un complejo simplicial finito orientado K e $i \in \mathbb{Z}$ tal que $0 \leq i \leq \dim(K)$, entonces $\text{Spec}(\Delta_i(K))$ no contiene racionales no enteros.*

Demostración. Sabemos que los valores propios de $\Delta_i(K)$ vienen dados por las raíces del polinomio característico, $p(x) = |\mathcal{L}_i(K) - xI_{f_i(K)}|$, siendo $I_{f_i(K)}$ la matriz identidad de orden $f_i(K)$. Por el teorema 3.21 conocemos que los coeficientes de $\mathcal{L}_i(K)$ son enteros, por lo que $p(x) \in \mathbb{Z}[x]$. Es conocido que el coeficiente de mayor grado de $p(x)$ será $(-1)^{f_i(K)}$. Igualmente, si $p(x)$ tuviese alguna raíz estrictamente racional, el denominador de dicha raíz dividiría al coeficiente de mayor grado, por lo tanto el denominador solo podría ser ± 1 , es decir, no sería una raíz estrictamente racional, sino entera. Con esto concluimos que las raíces solo pueden ser enteras o irracionales. \square

El próximo resultado nos relaciona las dos partes claramente diferenciadas de esta memoria, la homología y el Laplaciano combinatorio. Así vemos que nos encontramos ante dos métodos útiles para el estudio de los agujeros n -dimensionales de un complejo simplicial.

Teorema 3.34 (Teorema Combinatorio de Hodge para Laplacianos). *Sea K un complejo simplicial finito. Entonces*

$$H_k(K) \simeq \text{Ker}(\Delta_k(K))$$

para cualquier entero k tal que $0 \leq k \leq \dim(K)$, considerando el isomorfismo como espacios vectoriales sobre \mathbb{R} (Eckmann, 1944).

Es decir, sabemos que la multiplicidad del valor propio 0 en el k -ésimo Laplaciano combinatorio de K nos da el número de agujeros k -dimensionales de dicho complejo simplicial, lo que ya conocemos como número de Betti.

Se trata del principal resultado de esta parte de la memoria, para el cual daremos su demostración. Para ello, primero mostraremos una serie de resultados que usaremos para dicha demostración.

Lema 3.35. *Dada una aplicación lineal $T: V \rightarrow W$ y su adjunta $T^*: W \rightarrow V$. Entonces, V y W descomponen ortogonalmente del siguiente modo:*

$$V = \text{Ker}(T) \oplus \text{Im}(T^*) \text{ y } W = \text{Ker}(T^*) \oplus \text{Im}(T)$$

Teorema 3.36 (Teorema de Descomposición Combinatoria de Hodge). $C_k(K; \mathbb{R})$ admite una descomposición ortogonal

$$C_k(K; \mathbb{R}) = \text{Im}(\delta_k^*) \oplus \text{Ker}(\Delta_k) \oplus \text{Im}(\delta_{k+1})$$

donde $\text{Ker}(\Delta_k) = \text{Ker}(\delta_{k+1}^*) \cap \text{Ker}(\delta_k)$.

Demostración. Dado un complejo simplicial K sabemos por el teorema 3.35 que el espacio $C_k(K)$ descompone ortogonalmente del siguiente modo:

$$C_k(K) = \text{Ker}(\delta_{k+1}^*) \oplus \text{Im}(\delta_{k+1})$$

Por lo tanto, apoyándonos en que $\text{Ker}(\delta_k) \subseteq C_k(K)$ y además $\text{Im}(\delta_{k+1}) \subseteq \text{Ker}(\delta_k)$, obtenemos lo siguiente:

$$\begin{aligned} C_k(K) \cap \text{Ker}(\delta_k) &= [\text{Ker}(\delta_{k+1}^*) \oplus \text{Im}(\delta_{k+1})] \cap \text{Ker}(\delta_k) = \\ &= [\text{Ker}(\delta_{k+1}^*) \cap \text{Ker}(\delta_k)] \oplus [\text{Im}(\delta_{k+1}) \cap \text{Ker}(\delta_k)] = [\text{Ker}(\delta_{k+1}^*) \cap \text{Ker}(\delta_k)] \oplus \text{Im}(\delta_{k+1}) \Rightarrow \\ &= \text{Ker}(\delta_k) = [\text{Ker}(\delta_{k+1}^*) \cap \text{Ker}(\delta_k)] \oplus \text{Im}(\delta_{k+1}) \end{aligned}$$

De nuevo, por el lema 3.35 sabemos que $C_k(K)$ descompone también del siguiente modo:

$$C_k(K) = \text{Ker}(\delta_k) \oplus \text{Im}(\delta_k^*)$$

Y por la descomposición de $\text{Ker}(\delta_k)$ que hemos visto previamente se tiene que $C_k(K) = \text{Im}(\delta_k^*) \oplus [\text{Ker}(\delta_{k+1}^*) \cap \text{Ker}(\delta_k)] \oplus \text{Im}(\delta_{k+1})$

A continuación, probemos que $\text{Ker}(\Delta_k) = \text{Ker}(\delta_{k+1}^*) \cap \text{Ker}(\delta_k)$.

Es inmediato ver que $\text{Ker}(\delta_{k+1}^*) \cap \text{Ker}(\delta_k) \subseteq \text{Ker}(\Delta_k)$. Sea $z \in \text{Ker}(\delta_{k+1}^*) \cap \text{Ker}(\delta_k)$, entonces $\Delta_k(z) = (\delta_{k+1} \circ \delta_{k+1}^* + \delta_k^* \circ \delta_k)(z) = \delta_{k+1}(\delta_{k+1}^*(z)) + \delta_k^*(\delta_k(z)) = \delta_{k+1}(0) + \delta_k^*(0) = 0 \Rightarrow z \in \text{Ker}(\Delta_k)$

Ahora, la otra inclusión. Sea $z \in Ker(\Delta_k)$, luego $\Delta_k(z) = 0$. Entonces, $\langle \Delta_k(z), z \rangle = 0 \Rightarrow \langle \delta_{k+1}(\delta_{k+1}^*(z)), z \rangle + \langle \delta_k^*(\delta_k(z)), z \rangle = \langle \delta_{k+1}^*(z), \delta_{k+1}^*(z) \rangle + \langle \delta_k(z), \delta_k(z) \rangle = 0$. Como estamos trabajando con una métrica definido positiva, de hecho euclídea, necesariamente tenemos que

$$\langle \delta_{k+1}^*(z), \delta_{k+1}^*(z) \rangle = 0 \Rightarrow \delta_{k+1}^*(z) = 0 \Rightarrow z \in Ker(\delta_{k+1}^*)$$

$$\langle \delta_k(z), \delta_k(z) \rangle = 0 \Rightarrow \delta_k(z) = 0 \Rightarrow z \in Ker(\delta_k)$$

Es decir, $z \in Ker(\delta_{k+1}^*) \cap Ker(\delta_k)$. Con lo cual, $Ker(\Delta_k) = Ker(\delta_{k+1}^*) \cap Ker(\delta_k)$.

Finalmente, sustituyendo en la descomposición obtenida al principio de la demostración llegamos a $C_k(K) = Im(\delta_k^*) \oplus Ker(\Delta_k) \oplus Im(\delta_{k+1})$. \square

Así, podemos dar la demostración del Teorema Combinatorio de Hodge para Laplacianos como si se tratase de un corolario de este anterior teorema.

Demostración. Por uno de los resultados probados en el teorema previo, junto con la igualdad $Ker(\delta_k) = [Ker(\delta_{k+1}^*) \cap Ker(\delta_k)] \oplus Im(\delta_{k+1})$, podemos afirmar que

$$Ker(\delta_k) = Im(\delta_{k+1}) \oplus Ker(\Delta_k)$$

Con esto, es directo comprobar lo que buscamos:

$$H_k(K) = Ker(\delta_k)/Im(\delta_{k+1}) \simeq Ker(\Delta_k)$$

\square

Además de este resultado principal que nos permite ver la importancia del valor propio cero, el resto de autovalores dan igualmente información relevante acerca de cómo es el complejo simplicial. En el caso de los grafos, ciertos autovalores están relacionados con invariantes como podría ser el *número de independencia* de un grafo. Asimismo, el espectro del Laplaciano está estrechamente relacionado con la matriz de adyacencia de un grafo, para más detalle puede consultarse [22].

Conclusiones

A lo largo de esta memoria hemos detallado dos acercamientos al estudio de los agujeros n -dimensionales existentes en una filtración de un complejo simplicial, mostrando finalmente cómo se relacionan estos dos enfoques. Hemos explicado la matemática teórica subyacente y su visualización de manera gráfica con los *códigos de barras*, consiguiendo plasmar así qué información es relevante.

Las técnicas expuestas resultan bastantes útiles y precisas por el rigor matemático. De este modo, el TDA se convierte en una herramienta valiosa para el análisis de datos. No sería de extrañar que en los próximos años este campo se consiguiera desarrollar de manera notable, pues cada vez más y más empresas requieren de analistas de datos para poder comercializar y mejorar sus productos, además de la importancia creciente que tiene el análisis de datos en campos como la medicina, la química, la psicología... Como ya hemos mencionado, actualmente existen diversos softwares que permiten calcular la forma normal de Smith, la homología, etc. Es decir, a nivel computacional este campo está evolucionando cada vez más, con algoritmos más eficientes y menos costosos computacionalmente, lo cual es completamente necesario para que el TDA sea verdaderamente provechoso en la práctica. Hasta la fecha se han desarrollado varios métodos alternativos para realizar estos cálculos: las *optimizaciones directas*, métodos que mejoran la eficiencia del algoritmo mostrado en la memoria; los *acercamientos distribuidos*, que recuperan la información homológica de un complejo simplicial a través de cálculos paralelos y distribuidos; los métodos basados en *anotaciones*, las cuales son vectores asociados a cada símplice del complejo simplicial codificando de manera compacta la clase de homología a la que pertenece dicho símplice. Otra de las líneas de investigación actuales está centrada en el desarrollo de mecanismos que permitan la visualización, con el fin de hacer más fácil el acceso de personas con menos formación en este campo.

En conclusión, hemos visto un claro ejemplo de cómo incluso las matemáticas más puras pueden tener una aplicación indudable y “tangible” en la vida real, asegurando así rigurosidad y robustez.

Bibliografía

- [1] AYALA, R.; DOMÍNGUEZ, E.; QUINTERO, A., *Elementos de la teoría de homología clásica*, Universidad de Sevilla, Secretariado de publicaciones, 2002.
- [2] BUBENIK, P., *Statistical Topological Data Analysis using Persistence Landscapes*, Journal of Machine Learning Research 16, pp. 77-102, 2015.
- [3] CARLSSON, G., *Topology and Data*, Bulletin Of The American Mathematical Society, vol. 46, number 2, pp. 255–308, 2009.
- [4] CARLSSON, G.; COLLINS, A.; GUIBAS, L. J.; ZOMORODIAN, A., *Persistence barcodes for shapes*, International Journal of Shape Modeling, 11 (2), pp. 149-187, 2005.
- [5] CARLSSON, G.; VEJDEMO-JOHANSSON, M., *Topological Data Analysis with Applications*, Cambridge University Press, 2021.
- [6] CORBET, R.; KERBER, M., *The representation theorem of persistence revisited and generalized*, Journal of Applied and Computational Topology, 2, pp. 1-31, 2017.
- [7] EDELSBRUNNER, H.; KIRKPATRICK, D. G.; SEIDEL, R., *On the shape of a set of points in the plane*, IEEE Transactions on Information Theory vol. IT-29, 1983.
- [8] EDELSBRUNNER, H.; LETSCHER, D.; ZOMORODIAN, A., *Topological Persistence and Simplification*, Discrete Comput. Geom., 28, pp. 511-533, 2002.
- [9] ESPINOSA, M. E., *Homología persistente*, Centro de Investigación en Matemáticas (CIMAT), 2015.
- [10] FROSINI, P., *A distance for similarity classes of submanifolds of a Euclidean space*, Bulletin of the Australian Mathematical Society, 42 (3), pp. 407-415, 1990.
- [11] FUGACCI, U.; SCARAMUCCIA, S.; IURICICH, F.; DE FLORIANI, L., *Persistent homology: a step-by-step introduction for newcomers*, STAG: Smart Tools and Apps in computer Graphics, 2016.
- [12] GHRIST, R., *Barcodes: The persistent topology of data*, Bulletin Of The American Mathematical Society, vol. 45, number 1, pp. 61-75, 2008.

- [13] GOLDBERG, T. E., *Combinatorial Laplacians of Simplicial Complexes*, The Division of Natural Science and Mathematics of Bard College, New York, 2002.
- [14] HARRIS, R. A.; PATEL, M. S.; ROCHE, T. E., *Alpha-Keto Acid Dehydrogenase Complexes*, Birkhäuser, 1996.
- [15] HATCHER, A., *Algebraic Topology*, Cambridge University Press, 2002.
- [16] HENSELMAN-PETRUSEK, G.; LEYGONIE, J., *Algorithmic Reconstruction of the Fiber of Persistent Homology on Cell Complexes*, arXiv preprint arXiv:2110.14676., 2021.
- [17] HILTON, P.J.; STAMMBACH, U., *A Course in Homological Algebra*, Graduate texts in mathematics, Second edition, Springer, 2012.
- [18] HORAK, D.; JOST, J., *Spectra of combinatorial Laplace operators on simplicial complexes*, Advances in Mathematics 244, pp. 303-336, 2013.
- [19] JACOBSON, N., *Basic Algebra I, Second Edition*, W. H. Freeman, San Francisco, 1985.
- [20] JIANG, X.; LIM, L. H.; YAO, Y.; YE, Y., *Statistical ranking and combinatorial Hodge theory*, Mathematical Programming, 127(1), pp. 203-244, 2011.
- [21] MALETIC, S. V., *Simplicial complexes and complex networks: The influence of higher-order (sub)structures on network properties*, Doctoral dissertation, University of Belgrade, 2013.
- [22] MOHAR, B., *The Laplacian spectrum of graphs*, Graph Theory, Combinatorics and Applications, 2, 871, 1991.
- [23] MUNKRES, JAMES R., *Elements of Algebraic Topology*, Addison-Wesley Publishing Company, Menlo Park, CA, 1984.
- [24] ROBINS, V., *Towards computing homology from finite approximations*, Topology Proceedings, vol. 24, pp. 503-532, 1999.
- [25] SIZEMORE, ROBERT K., *Hodgerank: Applying Combinatorial Hodge Theory to Sports Ranking*, North Carolina, 2013.
- [26] C.WEBB, *Decomposition of graded modules*, Proceeding of the American Mathematical Society, vol 94, 4, 1985.

- [27] ZOMORODIAN, A. AND CARLSSON, G., *Computing Persistent Homology*, Discrete Comput. Geom., 33, pp. 249-274, 2005.

Glosario de términos

- A*-módulo
 - de homología, 14
 - de homología persistente, 24
 - de n -bordes, 13
 - de n -ciclos, 13
 - graduado, 29
- Anillo graduado, 29
- Aplicación
 - adjunta, 40
 - simplicial, 6
- Cara de un símplice, 6
- Célula de Voronoi, 11
- Código de barras, 32
- Complejo
 - Alfa, 12
 - de cadenas, 13
 - de cadenas simpliciales, 21
 - de Delaunay, 11
 - simplicial abstracto, 6
 - simplicial de Čech, 8
 - simplicial de Vietoris-Rips, 9
 - simplicial geométrico, 4
 - simplicial orientado, 21
- Dimensión
 - de un complejo simplicial, 6
 - de un símplice, 6
- Elemento homogéneo
 - de un A -módulo graduado, 29
 - de un anillo graduado, 29
- Espacio topológico
 - contráctil, 8
 - triangulable, 5
- Espectro del d -ésimo Laplaciano, 46
- Esquema de un complejo simplicial geométrico, 6
- Filtración, 22
- Finitamente presentado, 28
- Forma normal de Smith, 15
- Grado
 - inferior de un símplice, 38
 - superior de un símplice, 37
- Homología simplicial, 21
- Homomorfismo
 - de complejos de cadenas, 14
 - inducido entre homologías, 14
- Ideal graduado, 29
- Índice de persistencia, 25
- Matriz (combinatoria) Laplaciana, 35
- Módulo de persistencia, 28
- Muerte de una clase, 25
- n -bordes débiles, 17
- n -cadenas, 13
- n -ésima matriz Laplaciana, 40
- n -ésimo Laplaciano combinatorio, 40
- n -símplice, 3

Número de Betti

de persistencia, 24

de un A -módulo, 17

Nacimiento de una clase, 25

Nervio, 7

Operador borde, 13

Orientación de un símlice, 20

Poliedro de un complejo simplicial, 4

Realización geométrica de un complejo simplicial abstracto, 6

Símlice

inferior común, 38

orientado, 21

superior común, 37

Símplices

inferiormente adyacentes, 38

superiormente adyacentes, 37

Subcomplejo simplicial, 6

Vértices de un complejo simplicial, 6