

Sobre la homología persistente en redes neuronales

José Manuel Ros Rodrigo

Facultad de Ciencia y Tecnología
Universidad de La Rioja

Abril 2022

Resumen

(En construcción.)

Índice

1. Introducción	4
2. Preliminares	5
2.1. Complejos simpliciales	5
2.2. Homología. Homología persistente	9
2.3. Homología persistente	12
3. Homología persistente en redes neuronales	13
3.1. Construcción de un complejo simplicial a partir de una red neuronal . . .	13
4. Experimentos	20

1. Introducción

(En construcción.)

2. Preeliminares

A lo largo de este capítulo vamos a ver todas las nociones teóricas necesarias para el uso de la homología persistente en redes neuronales.

2.1. Complejos simpliciales

Comenzamos con el primer concepto fundamental de todo el trabajo, los *complejos simpliciales*. Esta noción admite dos enfoques diferentes, por lo que debemos distinguir entre dos definiciones equivalentes: los complejos simpliciales *abstractos* y los complejos simpliciales *geométricos*.

Siguiendo el enfoque combinatorio, comenzamos definiendo los complejos simpliciales abstractos y algunas nociones relacionadas.

Definición 1. Un *complejo simplicial abstracto* es una colección, \mathcal{V} , de subconjuntos no vacíos de un conjunto, \mathcal{V}_0 , que verifica las siguientes propiedades:

1. Si $v \in \mathcal{V}_0$, entonces $\{v\} \in \mathcal{V}$
2. Si $\sigma \in \mathcal{V}$ y $\tau \subset \sigma$, entonces $\tau \in \mathcal{V}$

A los elementos de \mathcal{V} los llamaremos *símplices*, más concretamente: dado $\sigma \in \mathcal{V}$, diremos que σ tiene *dimensión* p , y que σ es un *p-símplice*, si $|\sigma| = p + 1$. Asimismo, definimos la *dimensión de* \mathcal{V} como el máximo de las dimensiones de sus símplices y denotaremos por \mathcal{V}_p a la colección de los p -símplices de \mathcal{V} .

En relación con el concepto de símplice y de dimensión surge la siguiente noción:

Definición 2. Sean σ y τ dos símplices de \mathcal{V} tales que $\tau \subset \sigma$. Entonces diremos que τ es una *cara* de σ , y si las dimensiones de σ y τ difieren por a , diremos que τ es una cara de σ de *codimensión* a .

Ahora que hemos definido los complejos simpliciales abstractos veamos un pequeño ejemplo para fijar ideas.

Ejemplo 1. Supongamos el siguiente complejo simplicial abstracto:

$$\mathcal{V} = \{\{a\}, \{b\}, \{c\}, \{d\}, \{a, b\}, \{a, c\}, \{a, d\}, \{b, c\}, \{b, d\}, \{c, d\}, \{a, b, c\}, \{a, b, d\}, \\ \{a, c, d\}, \{a, b, c, d\}\}$$

Así, tenemos que la dimensión de \mathcal{V} es 3. También observamos que el 3-símplice $\{a, b, c, d\}$ tiene por caras de codimensión 1 a los 2-símplices $\{a, b, c\}$, $\{a, b, d\}$ y $\{a, c, d\}$. Veamos su representación geométrica.

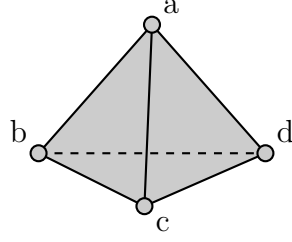


Figura 1: Representación geométrica del complejo simplicial \mathcal{V} .

Esta representación es única salvo homeomorfismo. Observamos que interpretando \mathcal{V} como un subconjunto de \mathbb{R}^n obtenemos un tetraedro. Esta idea motiva el otro enfoque de los complejos simpliciales: el enfoque geométrico. ◀

Siguiendo el enfoque geométrico es necesario que, antes de llegar a la definición de complejo simplicial geométrico, veamos unos conceptos previos relacionados con la propia definición.

Definición 3. Sean $\{u_0, u_1, \dots, u_k\} \subset \mathbb{R}^n$. Diremos que los $k+1$ puntos son *afínmente independientes* si los k vectores $u_1 - u_0, u_2 - u_0, \dots, u_k - u_0$ son linealmente independientes.

Sea $x \in \mathbb{R}^n$. Diremos que x es una *combinación afín* de los u_i si $\exists \lambda_0, \dots, \lambda_k$ tales que $x = \sum_{i=0}^k \lambda_i u_i$ y $\sum_{i=0}^k \lambda_i = 1$

Definición 4. Sean $\{u_0, u_1, \dots, u_k\} \subset \mathbb{R}^n$ $k+1$ puntos afínmente independientes y $x = \sum_{i=0}^k \lambda_i u_i$ una combinación afín. Diremos que x es una *combinación convexa* de los u_i si todos los λ_i son no negativos.

Definimos la *clausura convexa* de los u_i como el conjunto de todas sus posibles combinaciones convexas.

Ahora que ya contamos con estas nociones previas pasamos a definir la pieza clave en la definición de complejo simplicial geométrico: el *símplice*.

Definición 5. Definimos un *k-símplice* como la clausura convexa de $k+1$ puntos afínmente independientes. Lo denotaremos por $\sigma = \text{conv}\{u_0, u_1, \dots, u_k\}$, y diremos que la *dimensión de σ* es k .

Llamamos *cara* de σ a cualquier combinación convexa de un subconjunto no vacío de los u_i . Lo denotaremos por \leq .

Para los casos $k = 0, 1, 2, 3$ diremos que σ es un vértice, arista, triángulo, tetraedro respectivamente.

Habiendo definido todos los conceptos previos necesarios pasamos a definir *complejo simplicial geométrico*.

Definición 6. Llamamos *complejo simplicial geométrico* a la colección finita de símplices \mathcal{V} verificando las siguientes propiedades:

1. Si $\sigma \in \mathcal{V}$ y $\tau \leq \sigma \implies \tau \in \mathcal{V}$
2. Si $\sigma_1, \sigma_2 \in \mathcal{V} \implies \sigma_1 \cap \sigma_2 = \emptyset$ o $\sigma_1 \cap \sigma_2$ es una cara común a ambos.

De aquí en adelante emplearemos la definición de complejo simplicial abstracto, pues es la más adecuada para el presente trabajo.

Ahora que ya hemos definido los objetos con los que vamos a trabajar, procedemos a definir las aplicaciones entre ellos.

Definición 7. Una *aplicación entre complejos simpliciales*, $f : \mathcal{V} \rightarrow \mathcal{V}'$, es una aplicación $f : \mathcal{V}_0 \rightarrow \mathcal{V}'_0$ tal que $f(\sigma) \in \mathcal{V}' \ \forall \sigma \in \mathcal{V}$.

Teniendo definidas las aplicaciones entre complejos simpliciales, vamos a dotar a los complejos simpliciales de una cierta estructura que será de lo más útil para los propósitos del presente trabajo.

Consideremos \mathbb{Z}_2 el cuerpo de dos elementos. Dado un complejo simplicial \mathcal{V} , denotaremos por $C_p(\mathcal{V})$ al \mathbb{Z}_2 -espacio vectorial libre cuya base viene dada por los p -símplices de \mathcal{V} . Ahora, para cualquier $p \in \{1, 2, \dots\}$ definimos la siguiente aplicación:

$$\begin{aligned} \partial_p : C_p(\mathcal{V}) &\rightarrow C_{p-1}(\mathcal{V}) \\ c &\mapsto \sum_{d \subset c, d \in \mathcal{V}_{p-1}} d \end{aligned} \quad (2.1)$$

Si $p = 0$ definimos $\partial_0 = 0$. Intuitivamente, ∂_p le asigna a cada p -símplice su borde, esto es, la suma de sus caras de codimensión 1. Esta aplicación tiene una propiedad muy importante, que motivará la siguiente subsección:

Proposición 1. Sea ∂_p definida como en 2.1. Entonces para todo $p \in \{0, 1, 2, \dots\}$ $\partial_p \circ \partial_{p+1} = 0$. Coloquialmente, «el borde del borde es vacío».

Demostración. Sea $c \in C_{p+1}(\mathcal{V})$ y $v \in \mathcal{V}$ el símplice representado por c . Veamos que $\partial_p(\partial_{p+1}(c)) = 0$.

En efecto, notemos que v posee $\binom{p+2}{p}$ caras distintas de codimensión 2. Sea τ una de ellas, es decir, τ es un p -símplice y $\tau \subset v$.

Si probamos que τ aparece en 2 caras de codimensión 1 de v habremos terminado, pues aparecerá 2 veces como vector al hacer $\partial_p(\partial_{p+1}(c))$ y como estamos en \mathbb{Z}_2 se anulará. Esto implica lo que queremos probar.

Observemos que τ tiene dimensión p mientras que v tiene dimensión $p+2$. Por lo tanto, supongamos, sin pérdida de generalidad, que τ viene dado por los p últimos elementos de v . Así, tenemos dos elementos libres en v , y al calcular las caras de codimensión 1 de v , con los p últimos elementos fijos, tendremos únicamente 2 caras que contienen a τ . \square

Nota. La elección del cuerpo sobre el que se toman los espacios vectoriales es muy significativa. De hecho, si escogemos otro cuerpo, los resultados serán muy distintos, y los cálculos para llegar a ellos, serán más engorrosos. Veremos este hecho en los siguientes ejemplos.

Veamos un ejemplo que ilustre la proposición anterior, es decir, que «el borde del borde es vacío».

Ejemplo 2. Supongamos el complejo simplicial \mathcal{V} del ejemplo anterior y $\sigma = \{a, b, c\} \in \mathcal{V}$.

Así pues, tendremos $c \in C_2(\mathcal{V})$, con $c = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}$ la representación de σ en $C_2(\mathcal{V})$. Ahora expresamos las aplicaciones ∂_2 y ∂_1 en forma matricial:

$$\partial_2 = \begin{pmatrix} 1 & 1 & 0 \\ 1 & 0 & 1 \\ 0 & 1 & 1 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \quad \text{y} \quad \partial_1 = \begin{pmatrix} 1 & 1 & 1 & 0 & 0 & 0 \\ 1 & 0 & 0 & 1 & 1 & 0 \\ 0 & 1 & 0 & 1 & 0 & 1 \\ 0 & 0 & 1 & 0 & 1 & 1 \end{pmatrix}$$

Ahora, teniendo en cuenta que estamos operando en un cuerpo de característica 2, hacemos $\partial_1(\partial_2(c))$:

$$\begin{aligned} \partial_1(\partial_2(c)) &= \begin{pmatrix} 1 & 1 & 1 & 0 & 0 & 0 \\ 1 & 0 & 0 & 1 & 1 & 0 \\ 0 & 1 & 0 & 1 & 0 & 1 \\ 0 & 0 & 1 & 0 & 1 & 1 \end{pmatrix} \cdot \left(\begin{pmatrix} 1 & 1 & 0 \\ 1 & 0 & 1 \\ 0 & 1 & 1 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \cdot \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} \right) = \\ &= \begin{pmatrix} 1 & 1 & 1 & 0 & 0 & 0 \\ 1 & 0 & 0 & 1 & 1 & 0 \\ 0 & 1 & 0 & 1 & 0 & 1 \\ 0 & 0 & 1 & 0 & 1 & 1 \end{pmatrix} \cdot \begin{pmatrix} 1 \\ 1 \\ 0 \\ 1 \\ 0 \\ 0 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \\ 0 \end{pmatrix} = \vec{0} \end{aligned}$$

Hemos comprobado que, en efecto, «el borde del borde» de σ es 0. Para comprobarlo para cualquier vector bastará observar que:

$$\begin{pmatrix} 1 & 1 & 1 & 0 & 0 & 0 \\ 1 & 0 & 0 & 1 & 1 & 0 \\ 0 & 1 & 0 & 1 & 0 & 1 \\ 0 & 0 & 1 & 0 & 1 & 1 \end{pmatrix} \cdot \begin{pmatrix} 1 & 1 & 0 \\ 1 & 0 & 1 \\ 0 & 1 & 1 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}$$

En este ejemplo apreciamos lo significativo de elegir el cuerpo \mathbb{Z}_2 , pues en otro caso, los productos matriciales son más difíciles de calcular y podrían no anularse. ◀

De la proposición anterior se desprende que $Im(\partial_{p+1}) \subset Ker(\partial_p)$. Este hecho motiva la siguiente noción importante del presente trabajo: los *grupos de homología*.

2.2. Homología. Homología persistente

En la subsección anterior, más concretamente en 2.1, hemos introducido la aplicación «borde». A lo largo de esta subsección vamos a profundizar más en ella, y en los espacios sobre los que está definida, llegando de una manera natural a la definición de *grupo de homología*.

En primer lugar, vamos a centrarnos en $C_p(\mathcal{V})$ y sus elementos.

Tal y como hemos mencionado anteriormente, podemos ver $C_p(\mathcal{V})$ como un \mathbb{Z}_2 -espacio vectorial libre cuya base viene dada por los p -símplices de \mathcal{V} . Así, si $c \in C_p(\mathcal{V})$, entonces c es un vector que representa a un p -símplice v . De esta manera, podemos ver v como suma de los p -símplices de las componentes no nulas de c . Más formalmente:

Definición 8. Sea \mathcal{V} un complejo simplicial y $p \in \mathbb{N} \cup \{0\}$ tal que $p \leq \dim \mathcal{V}$. Una *p-cadena* es una suma formal de p -símplices de \mathcal{V} .

Con la noción de p -cadena, ya podemos formalizar la definición de $C_p(\mathcal{V})$.

Definición 9. Sea \mathcal{V} un complejo simplicial abstracto. Definimos el *grupo de p-cadenas* de \mathcal{V} como el conjunto de todas las p -cadenas de \mathcal{V} , con la operación suma componente a componente con coeficientes en \mathbb{Z}_2 . Lo denotaremos por $(C_p(\mathcal{V}), +)$ o simplemente $C_p(\mathcal{V})$.

De la definición anterior se desprende el siguiente resultado:

Proposición 2. Sea \mathcal{V} un complejo simplicial abstracto. Para cada $p \in \mathbb{N} \cup \{0\}$ tal que $p \leq \dim \mathcal{V}$, entonces $(C_p(\mathcal{V}), +)$ es grupo abeliano.

Demostración. La asociatividad se tiene por herencia de la suma en \mathbb{Z}_2 . La existencia de elemento neutro es clara, pues bastará considerar el vector nulo. La existencia de opuesto también es inmediata, ya que todo elemento es opuesto de sí mismo. Finalmente, como la suma en \mathbb{Z}_2 es conmutativa, se sigue que $(C_p(\mathcal{V}), +)$ es abeliano. \square

En segundo lugar, y habiendo definido los grupos de p -cadenas, pasamos a hacer un estudio más detallado de la aplicación «borde» definida en 2.1. Vamos con su definición:

Definición 10. Sea \mathcal{V} un complejo simplicial abstracto, $\sigma \in \mathcal{V}$ y $\sigma = \{u_1, \dots, u_p\}$. Definimos el *homomorfismo borde para un símplice* como:

$$B_p(\sigma) = \sum_{j=0}^p \{u_1, \dots, \widehat{u_j}, \dots, u_p\}$$

Donde $\widehat{u_j}$ indica que omitimos u_j . Notemos que este homomorfismo puede extenderse para cadenas, en concreto:

$$\begin{aligned} \partial_p : C_p(\mathcal{V}) &\rightarrow C_{p-1}(\mathcal{V}) \\ c = \sum \sigma_i &\mapsto \partial_p(c) = \sum B_p(\sigma_i) \end{aligned} \tag{2.2}$$

A ∂_p lo llamaremos *homomorfismo borde para cadenas*, aunque por simplicidad nos referiremos a él como *homomorfismo borde*.

A los elementos de $\text{Im}(\partial_p)$ los llamaremos *(p-1)-bordes*, y a los elementos de $\text{Ker}(\partial_p)$, *p-ciclos*.

Demostración. Vamos a probar que ∂_p es, en efecto, un homomorfismo.

Sean $\sigma, \tau \in \mathcal{V}_p$ tales que $\sigma = \{u_1, \dots, u_p\}$ y $\tau = \{w_1, \dots, w_p\}$.

$$B_p(\sigma) + B_p(\tau) = \sum_{j=0}^p \{u_1, \dots, \widehat{u_j}, \dots, u_p\} + \sum_{j=0}^p \{w_1, \dots, \widehat{w_j}, \dots, w_p\} = \sum_{i=0}^p \sum_{j=0}^p \{u_1 + w_1, \dots, \widehat{u_i} + \widehat{w_j}, \dots, u_p + w_p\} = B_p(\sigma + \tau).$$

Esto prueba que B_p conmuta con la suma para símlices. Se sigue que ∂_p conmuta para cadenas.

Ahora, $0 = B_p(\sigma) + B_p(\sigma) = B_p(\sigma + \sigma) = B_p(0)$. Esto implica que B_p deja fijo el símlice neutro. Al igual que antes, esta propiedad se extiende para cadenas.

Finalmente, $0 = B_p(0) = B_p(\sigma - \sigma) = B_p(\sigma) + B_p(-\sigma)$ y sumando $-B_p(\sigma)$ a ambos lados de la ecuación se sigue que B_p conmuta con el opuesto. Una vez más, este hecho se extiende a cadenas. \square

Este homomorfismo tiene propiedades muy interesantes, entre ellas:

Proposición 3. *Sea ∂_p el homomorfismo borde definido en 2.2. Entonces $Im(\partial_p)$ es subgrupo de $C_{p-1}(\mathcal{V})$ y $Ker(\partial_p)$ es subgrupo de $C_p(\mathcal{V})$. Más aún, ambos son subgrupos normales de $C_{p-1}(\mathcal{V})$, $C_p(\mathcal{V})$ respectivamente.*

Demostración. Vamos a probar ambos hechos por separado:

- Probaremos primero que $Im(\partial_p)$ es subgrupo:

Sean $a, b \in Im(\partial_p)$ tales que $a = \partial_p(c)$ y $b = \partial_p(d)$, $c, d \in C_p(\mathcal{V})$.

Entonces $a + b = \partial_p(c) + \partial_p(d) = \partial_p(c + d)$. Esto implica que $Im(\partial_p)$ es cerrado para la suma.

Ahora, $0 = \partial_p(0) = \partial_p(c + c) = \partial_p(c) + \partial_p(c) = a + a$. Esto implica que $Im(\partial_p)$ es cerrado para opuestos.

Se sigue que $Im(\partial_p) \leq C_{p-1}(\mathcal{V})$.

- Probemos ahora que $Ker(\partial_p)$ es subgrupo:

Al igual que antes, sean $a, b \in Ker(\partial_p)$.

Entonces $0 = \partial_p(a) + \partial_p(b) = \partial_p(a + b)$. Esto prueba que $Ker(\partial_p)$ es cerrado para la suma.

Ahora, $0 = \partial_p(0) = \partial_p(a - a) = \partial_p(a) + \partial_p(-a) = \partial_p(-a)$. Esto prueba que $Ker(\partial_p)$ es cerrado para opuestos.

Se sigue que $Ker(\partial_p) \leq C_p(\mathcal{V})$.

En virtud de la proposición 2 y lo ya probado, se concluye la normalidad. \square

Además de la propiedad anterior, y tal y como vimos en la subsección anterior, el homomorfismo borde tiene la propiedad fundamental de que $\partial_p \circ \partial_{p+1} = 0$. Este hecho implica que $Im(\partial_{p+1}) \subset Ker(\partial_p)$. Por lo tanto, podemos considerar el cociente entre ambos grupos: $Ker(\partial_p)/Im(\partial_{p+1})$. Este cociente constituye una de las piezas clave del presente trabajo:

Definición 11. Sea $p \in \mathbb{N} \cup \{0\}$ y \mathcal{V} un complejo simplicial. Definimos el p -ésimo grupo de homología de \mathcal{V} como el grupo cociente $Ker(\partial_p)/Im(\partial_{p+1})$, donde ∂_p está definida como en 2.2. Lo denotaremos por $H_p(\mathcal{V})$.

A su dimensión, $dim H_p(\mathcal{V}) = dim Ker(\partial_p) - dim Im(\partial_{p+1})$, la llamaremos p -ésimo número de Betti, y la denotaremos por $\beta_p(\mathcal{V})$.

Intuitivamente, los p -ciclos que no son p -bordes representan agujeros p -dimensionales. Por lo tanto, $\beta_p(\mathcal{V})$ representa el número de p -agujeros de \mathcal{V} . Además, notemos que si $dim \mathcal{V} = n$, entonces $\forall p > n \ H_p(\mathcal{V}) = \emptyset$, pues $\mathcal{V} = \emptyset$.

Ahora veamos un ejemplo en el que calculamos los números de Betti dado un complejo simplicial abstracto.

Ejemplo 3. Supongamos el siguiente complejo simplicial abstracto:

$$\mathcal{V} = \{\{a\}, \{b\}, \{c\}, \{d\}, \{e\}, \{a, b\}, \{a, c\}, \{a, d\}, \{b, c\}, \{c, d\}, \{a, b, c\}\}$$

Construimos la secuencia de grupos de cadenas asociados:

$$\emptyset \longrightarrow C_2(\mathcal{V}) \xrightarrow{\partial_2} C_1(\mathcal{V}) \xrightarrow{\partial_1} C_0(\mathcal{V}) \xrightarrow{\partial_0} \emptyset$$

Calculamos ∂_2 y ∂_1 , y los expresamos de manera matricial:

$$\partial_2 = \begin{pmatrix} 1 \\ 1 \\ 0 \\ 1 \\ 0 \end{pmatrix} \quad \partial_1 = \begin{pmatrix} 1 & 1 & 1 & 0 & 0 \\ 1 & 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 1 & 1 \\ 0 & 0 & 1 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix}$$

Ahora para calcular $\beta_0(\mathcal{V})$ y $\beta_1(\mathcal{V})$ bastará calcular el rango de las anteriores matrices. Observamos que:

$$\partial_1 = \begin{pmatrix} 1 & 1 & 1 & 0 & 0 \\ \boxed{1} & 0 & 0 & 1 & 0 \\ 0 & \boxed{1} & 0 & 1 & 1 \\ 0 & 0 & \boxed{1} & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix}$$

Con lo que, $dim Im(\partial_1) = 3$ y $dim Ker(\partial_1) = 2$, y por lo tanto, $\beta_1(\mathcal{V}) = 1$ y $\beta_0(\mathcal{V}) = 2$. La representación gráfica de \mathcal{V} en \mathbb{R}^2 nos queda:

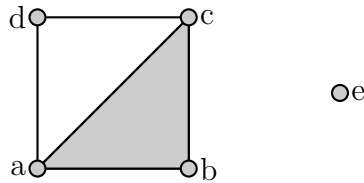


Figura 2: Representación geométrica del complejo simplicial \mathcal{V} .

Tal y como ya hemos comentado, los números de Betti nos cuentan los agujeros p -dimensionales. En este caso, si observamos la representación anterior, vemos que tenemos

un agujero 1-dimensional y dos componentes conexas, que se corresponde con los números de Betti que hemos calculado.

Notemos que, al igual que en el ejemplo anterior, si trabajamos sobre otro cuerpo que no sea \mathbb{Z}_2 , los rangos de las matrices podrían ser distintos. En este caso, si trabajamos sobre \mathbb{Q} , el rango de ∂_1 es 4. Esto afecta a los números de Betti y perjudica la eficiencia computacional de los algoritmos empleados para su cálculo. ◀

Al igual que en la subsección anterior, veamos como podemos definir morfismos, de manera más general, entre los objetos que estamos manejando.

Consideremos una aplicación entre complejos simpliciales, $f : \mathcal{V} \rightarrow \mathcal{V}'$. Aplicando el mismo razonamiento que para la definición del homomorfismo borde, tenemos que f induce un homomorfismo entre grupos de cadenas :

$$\begin{aligned} \overline{f_p} : C_p(\mathcal{V}) &\rightarrow C_p(\mathcal{V}') \\ c = \sum_{\sigma \in \mathcal{V}_p} \sigma &\mapsto \overline{f_p}(c) = \sum_{f(\sigma) \in \mathcal{V}'_p} f(\sigma) \end{aligned}$$

Además, tal f nos permite construir la secuencia:

$$\begin{array}{ccccccccccc} \emptyset & \rightarrow & C_p(\mathcal{V}) & \xrightarrow{\partial_p} & C_{p-1}(\mathcal{V}) & \xrightarrow{\partial_{p-1}} & \cdots & \xrightarrow{\partial_2} & C_1(\mathcal{V}) & \xrightarrow{\partial_1} & C_0(\mathcal{V}) & \xrightarrow{\partial_0} & \emptyset \\ & & \downarrow \overline{f_p} & & \downarrow \overline{f_{p-1}} & & & & \downarrow \overline{f_1} & & \downarrow \overline{f_0} & & \\ \emptyset & \rightarrow & C_p(\mathcal{V}') & \xrightarrow{\partial'_p} & C_{p-1}(\mathcal{V}') & \xrightarrow{\partial'_{p-1}} & \cdots & \xrightarrow{\partial'_2} & C_1(\mathcal{V}') & \xrightarrow{\partial'_1} & C_0(\mathcal{V}') & \xrightarrow{\partial'_0} & \emptyset \end{array}$$

De esta secuencia observamos que:

$$\overline{f_{p-1}} \circ \partial_p = \partial'_p \circ \overline{f_p}$$

En consecuencia, $\overline{f_p}$ induce un homomorfismo entre grupos de homología:

$$\begin{aligned} f_p : H_p(\mathcal{V}) &\rightarrow H_p(\mathcal{V}') \\ [c] &\mapsto [\overline{f_p}(c)] \end{aligned}$$

Concluimos que, dada una aplicación f entre complejos simpliciales, siempre es posible asociarle una aplicación f_p entre grupos de homología.

Nota. Esta propiedad es muy importante, de hecho, se conoce como *functorialidad* y pertenece al ámbito de la teoría de categorías que queda fuera del alcance del presente trabajo. Sin embargo, observemos que esta propiedad es la responsable de que podamos ver $C_p(\mathcal{V})$ como grupo o como espacio vectorial según nos convenga.

2.3. Homología persistente

Tras todos los conceptos previos, ya estamos en una buena posición para definir el concepto central del trabajo: **la homología persistente**. Sin embargo, aún vamos a necesitar una definición más.

Definición 12. Sea \mathcal{V} un complejo simplicial finito y $\mathcal{K}_1 \subset \mathcal{K}_2 \subset \dots \subset \mathcal{K}_{n-1} \subset \mathcal{K}_n = \mathcal{K}$ una cadena de subcomplejos simpliciales de \mathcal{V} . Al complejo simplicial \mathcal{V} y a su cadena asociada los llamaremos *complejo simplicial filtrado*.

Ahora, y sin más dilación, introducimos el concepto que da nombre a este trabajo.

Definición 13. Sea $\mathcal{K}_1 \subset \mathcal{K}_2 \subset \dots \subset \mathcal{K}_{n-1} \subset \mathcal{K}_n = \mathcal{K}$ un complejo simplicial filtrado. La p -ésima *homología persistente* de \mathcal{V} es el par:

$$(\{H_p(\mathcal{K}_i)\}_{1 \leq i \leq r}, \{f_{i,j}\}_{1 \leq i \leq j \leq r})$$

Donde $\forall i, j \in \{1, 2, \dots, r\}$ con $i \leq j$, las aplicaciones lineales $f_{i,j} : H_p(\mathcal{K}_i) \rightarrow H_p(\mathcal{K}_j)$ son las inducidas por las inclusiones $\mathcal{K}_i \hookrightarrow \mathcal{K}_j$.

De la definición anterior, observamos que el concepto de homología persistente es más "rico" que el de homología, ya que, la homología persistente nos ofrece más información acerca de un complejo simplicial filtrado que la consideración de la homología de los subcomplejos simpliciales. Para visualizar la homología persistente nos serán de mucha utilidad los *diagramas de persistencia*, así como los *códigos de barras*.

Veamos un ejemplo que ilustre lo que hemos definido.

Ejemplo 4. Ejemplo adecuado de complejo simplicial filtrado + código de barras + diagrama de persistencia. (En construcción). ◀

3. Homología persistente en redes neuronales

En la sección anterior hemos discutido todas las cuestiones acerca de la homología persistente en complejos simpliciales. Pero, ¿Cómo aplica toda esta teoría a las redes neuronales?

Para responder a esta cuestión debemos pensar en las redes neuronales como si fueran grafos, cuyos nodos serán las neuronas de la red, y las aristas, las relaciones entre las neuronas. De este modo podremos construir complejos simpliciales asociados al grafo y aplicarles toda la teoría que ya conocemos.

A lo largo de esta sección veremos con detalle como hacer esta correspondencia, así como algunos ejemplos ilustrativos de los conceptos que irán apareciendo.

(En el futuro habrá que añadir alguna disquisición previa sobre redes neuronales a la sección anterior).

3.1. Construcción de un complejo simplicial a partir de una red neuronal

Consideraremos el conjunto de las neuronas como el conjunto de vértices, es decir, $\nu_0 = \{k_0, \dots, k_n\}$ con $n+1$ el número de neuronas. Pensando de esta manera, vemos las

redes neuronales como grafos dirigidos con pesos ω_{ij} , con ω_{ij} el peso entre k_i y k_j . Notemos que $\omega_{ij} = 0$ si y sólo si k_i y k_j no están conectadas. Con esto en mente, definimos la importancia de k_i (salida) para k_j (llegada) como:

$$R_{ij} = \begin{cases} 1 & \text{si } i = j \\ \omega_{ij}^+ / \sum_{k, k \neq j} \omega_{kj}^+ & \text{si } i \neq j \end{cases} \quad (3.1)$$

Donde ω_{ij}^+ es la parte positiva del peso ω_{ij} , esto es, $\omega_{ij}^+ := \max\{0, \omega_{ij}\}$.

Observamos que la importancia de una neurona para sí misma es de 1, y la importancia entre neuronas distintas es la proporción del peso entre ellas con respecto al resto de pesos de la neurona de llegada.

(La elección de ω_{ij}^+ se toma motivada por la *regla- z^+* definida en la “*descomposición profunda de Taylor*”. Preguntar).

Para poder definir los complejos simpliciales tenemos que extender la definición de la importancia entre neuronas, para aquellas que no estén directamente conectadas. Consideremos las neuronas k_0 y k_2 conectadas por el camino: $k_2 \rightarrow k_1 \rightarrow k_0$, la importancia de k_2 para k_0 es, según el camino entre ellas, $R_{21} \cdot R_{10}$. Por lo tanto, definimos:

$$\overline{R_{ij}} = \max\{R_{k_i k_{m_1}} \cdots R_{k_{m_n} k_j} \mid (k_i, k_{m_1}, \dots, k_{m_n}, k_j) \in C_{ij}\} \quad (3.2)$$

Donde C_{ij} denota el conjunto de todos los posibles caminos de k_i a k_j .

(Se puede definir $\overline{R_{ij}}$ considerando varios caminos en C_{ij} . Elegimos el máximo por eficiencia computacional. Preguntar).

De aquí en adelante, numeraremos las neuronas de una red neuronal en orden ascendente, desde las neuronas de llegada hasta las de salida.

Veamos un ejemplo sencillo para interiorizar estas definiciones que serán clave a lo largo de esta sección.

Ejemplo 5. Supongamos la siguiente representación de una red neuronal con sus correspondientes pesos:

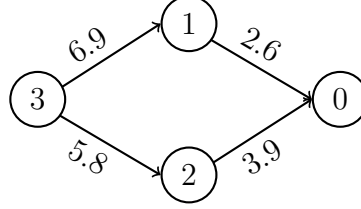


Figura 3: Representación de una red neuronal de 4 neuronas y 1 capa.

Tal y como vemos, las neuronas ya han sido ordenadas de manera correcta. Además en este caso todos los pesos son positivos, por lo que no nos tenemos que preocupar escoger la parte positiva. Vamos a calcular algunas importancias entre neuronas:

$$\begin{aligned}
 R_{31} &= \frac{6.9}{6.9} = 1 & R_{32} &= \frac{5.8}{5.8} = 1 \\
 R_{10} &= \frac{2.6}{6.5} = 0.4 & R_{20} &= \frac{3.9}{6.5} = 0.6 \\
 \overline{R_{30}} &= \max\{R_{31} \cdot R_{10}, R_{32} \cdot R_{20}\} = R_{32} \cdot R_{20} = 0.6
 \end{aligned}$$

Este ejemplo pone de manifiesto la intuición detrás de la definición de la importancia entre neuronas. Lo que hace es medir la aportación de la neurona emisora, k_i , con respecto al resto de neuronas emisoras de k_j . ◀

Teniendo en cuenta la definición de $\overline{R_{ij}}$, y el orden en una red, ya podemos construir un complejo simplicial filtrado. En primer lugar, definimos los p -símplices a partir de \mathcal{V}_0 como sigue:

$$\mathcal{K}_p^t = \begin{cases} \nu_0 & \text{si } p = 0 \\ \{(k_{a_0}, \dots, k_{a_p}) \mid k_{a_s} \in \nu_0, \overline{R_{a_s a_r}} \geq t, \forall a_s > a_r\} & \text{si } p \geq 1 \end{cases} \quad (3.3)$$

Donde $0 \leq t \leq 1$ es un parámetro.

Notemos que de la construcción que hemos hecho de los complejos simpliciales caben dos interpretaciones: la primera consiste en tomar un símplice si la importancia entre dos vértices calculada *localmente* pasa el filtro t ; la segunda se distingue de la primera en que la importancia entre dos vértices se calcula *globalmente*.

Veamos la diferencia entre ambas interpretaciones con el siguiente ejemplo.

Ejemplo 6. Supongamos la siguiente representación de una red neuronal con las importancias entre neuronas ya calculadas:

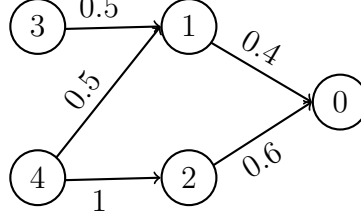


Figura 4: Representación de una red neuronal de 5 neuronas y 1 capa.

Vamos calcular $\mathcal{K}_2^{0.4}$ para ver las diferencias entre las interpretaciones. En primer lugar, listamos los posibles 2-símplices:

$$\{4, 2, 0\}, \{4, 1, 0\}, \{3, 1, 0\}$$

■ Interpretación local

Como la importancia entre 4-2, 2-0 y 4-0 es mayor que 0.4 $\Rightarrow \{4, 2, 0\} \in \mathcal{K}_2^{0.4}$, donde importancia entre 4-0 viene dada por:

$$\overline{R_{40}} = \max\{R_{42} \cdot R_{20}\} = R_{42} \cdot R_{20} = 0.6$$

Siguiendo la misma regla tenemos que $\{4, 1, 0\}, \{3, 1, 0\} \notin \mathcal{K}_2^{0.4}$. Por lo tanto, $\mathcal{K}_2^{0.4} = \{\{4, 2, 0\}\}$.

■ Interpretación global

Razonando igual que antes, $\{4, 2, 0\} \in \mathcal{K}_2^{0.4}$. Ahora bien, como la importancia entre 4-1 y 1-0 es mayor o igual que 0.4, y

$$\overline{R_{40}} = \max\{R_{42} \cdot R_{20}, R_{41} \cdot R_{10}\} = R_{42} \cdot R_{20} = 0.6$$

Entonces $\{4, 1, 0\} \in \mathcal{K}_2^{0.4}$. Sin embargo, $\{3, 2, 0\} \notin \mathcal{K}_2^{0.4}$ y así, $\mathcal{K}_2^{0.4} = \{\{4, 2, 0\}, \{4, 1, 0\}\}$

El ejemplo pone de manifiesto la principal diferencia entre ambas interpretaciones: en la primera, el máximo se calcula sobre los caminos que aparecen en el p-símplice; en la segunda, el máximo se calcula sobre los caminos que aparecen en todos los p-símplices.

Nota. Mientras que la interpretación global es consistente para los 1-símplices, la interpretación local no lo es. Para el cálculo de los 1-símplices en la interpretación local se toma el máximo entre todos los posibles caminos.



Ahora que ya tenemos definidos los p-símplices, vamos con la construcción del complejo simplicial abstracto. Para ello damos el siguiente resultado:

Proposición 4. Sea $\nu_0 = \{k_0, \dots, k_n\}$ un conjunto finito, y $\{\omega_{ij}\}_{0 \leq j \leq i \leq n}$ un conjunto de números reales. Sea $\overline{R_{ij}}$ la importancia entre neuronas definida en (3.2), y \mathcal{K}_p^t los p-símplices definidos en (3.3) con t parámetro real entre 0 y 1. Entonces, $\mathcal{K}^t = \bigcup_{s=0}^{s=n} \mathcal{K}_s^t$ es un complejo simplicial.

Demostración. Supongamos hipótesis generales.

Para probar que \mathcal{K}^t es un complejo simplicial debemos ver:

1. $v \in \nu_0 \implies \{v\} \in \mathcal{K}^t$
2. $\sigma \in \mathcal{K}^t \wedge \tau \subset \sigma \implies \tau \in \mathcal{K}^t$

Notemos que la primera propiedad se deduce inmediatamente de (3.3) y de la definición de \mathcal{K}^t .

Así pues, vamos a probar la segunda.

$\sigma = (k_{m_0}, \dots, k_{m_p}) \in \mathcal{K}^t \implies \overline{R_{m_i m_j}} \geq t \ \forall m_i \geq m_j$. Ahora sea $\tau \subset \sigma$, entonces,

$\tau = (k_{n_0}, \dots, k_{n_q})$, y como $\{n_0, \dots, n_q\} \subset \{m_0, \dots, m_p\}$, se tendrá que $\overline{R_{n_i n_j}} \geq t$,

$\forall n_i \geq n_j$. El resultado se sigue inmediatamente. \square

Ahora que ya tenemos construido nuestro complejo simplicial vamos a proceder a su filtración. Para ello será necesario el siguiente resultado:

Proposición 5. Sea $(t_i)_{i=1}^n$ una sucesión, monótona decreciente, de números reales entre 1 y 0. Entonces, $\nu_0 = \emptyset$ y $\mathcal{K}_i = \mathcal{K}^{t_i}$ con $1 \leq i \leq n$, es un complejo simplicial filtrado.

Demostración. Supongamos hipótesis generales.

Por la proposición anterior, sabemos que \mathcal{K}^{t_n} es un complejo simplicial. Ahora bien,

$t_i > t_j \implies \mathcal{K}_p^{t_i} \subset \mathcal{K}_p^{t_j}$ por la definición (3.3). Entonces, $\emptyset = \nu_0 \subset \mathcal{K}_1 \subset \dots \subset \mathcal{K}_n = \mathcal{K}^{t_n}$.

Se sigue inmediatamente el resultado. \square

Ahora que ya contamos con todas las herramientas necesarias, veamos un ejemplo completo en el que calculemos los números de Betti y los diagramas correspondientes. Haremos el desarrollo del ejemplo siguiendo las dos interpretaciones vistas en el ejemplo 6.

Ejemplo 7. Supongamos la siguiente representación de una red neuronal con las importancias entre neuronas ya calculadas:

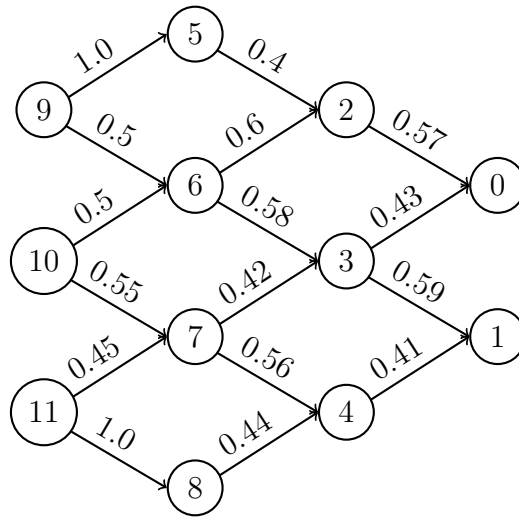
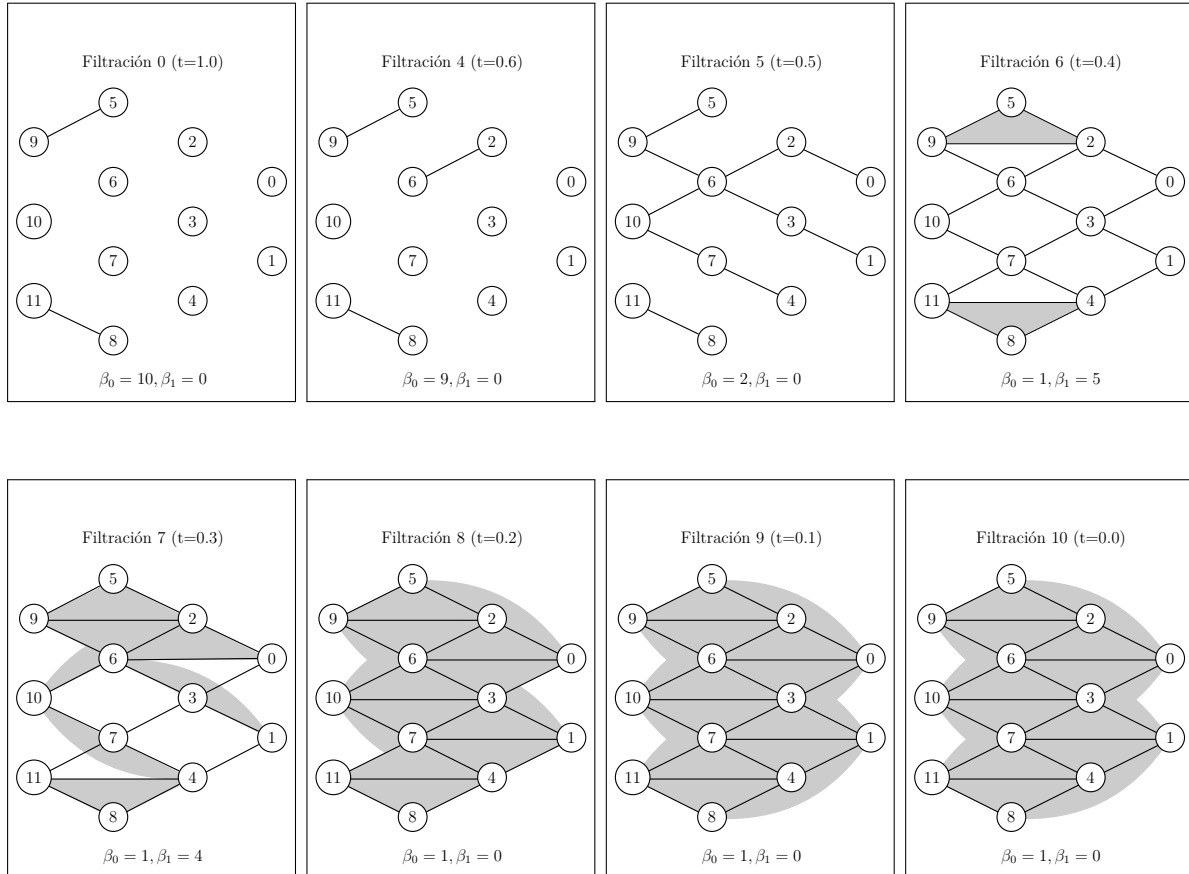
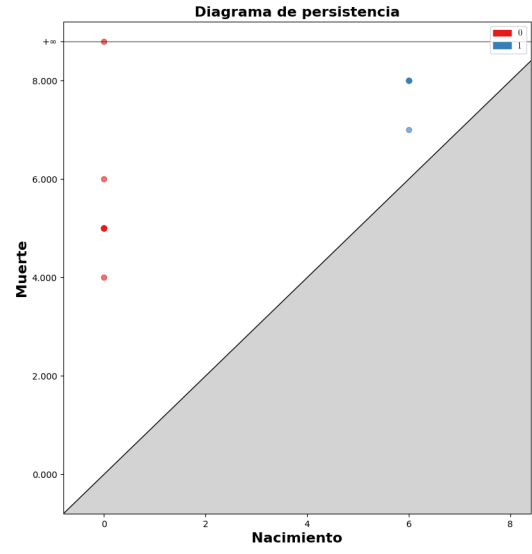
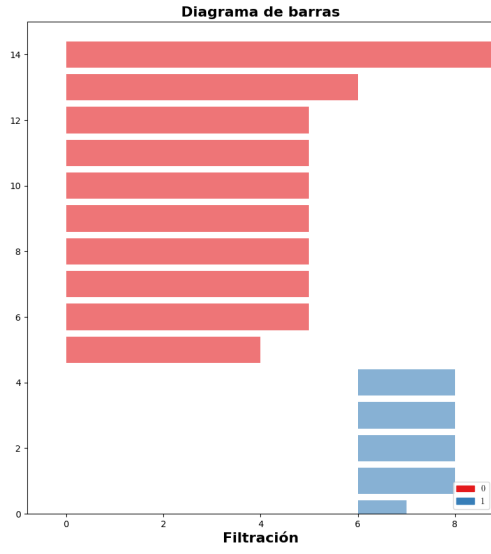


Figura 5: Representación de una red neuronal de 12 neuronas y 2 capas.

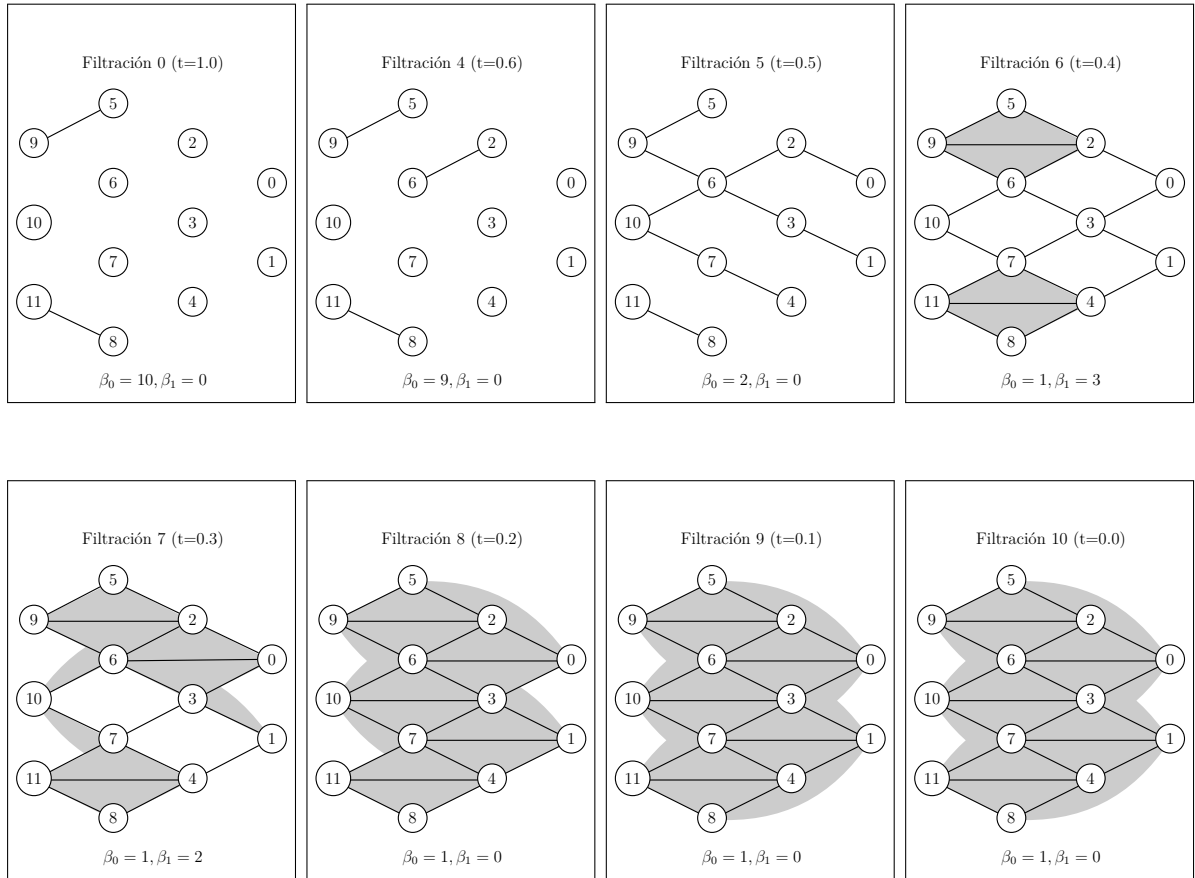
Vamos a calcularle el complejo simplicial filtrado asociado. Para ello vamos a ilustrar unos cuantos pasos en la filtración con los correspondientes números de Betti asociados. También añadimos los correspondientes diagramas de persistencia y de barras realizados con *GUDHI* y *Dionysus*.

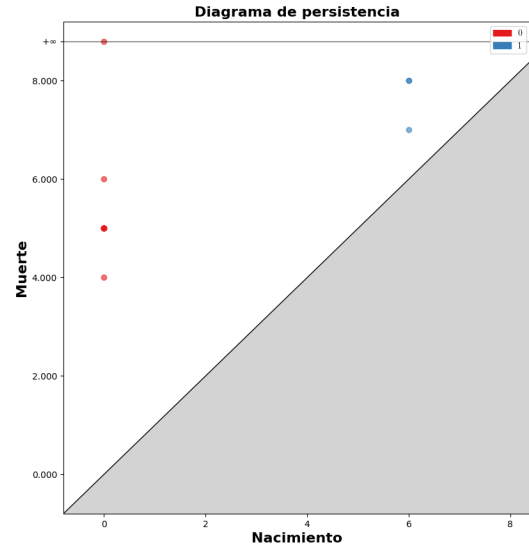
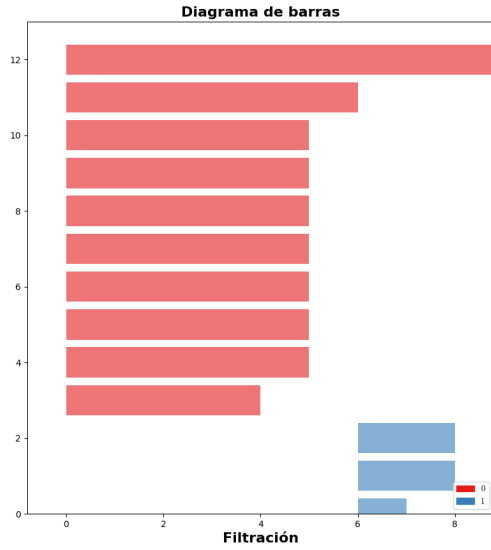
■ Interpretación local





■ Interpretación global





Nota. Para una mayor claridad, en los dibujos de las filtraciones se han omitido algunas aristas.



Del ejemplo anterior podemos extraer unas conclusiones muy importantes. Por una parte, observamos que si las neuronas de entrada se conectan directamente a las de salida, el conocimiento de la red será “pobre” ya que será equivalente a la detección de patrones. Por otra parte, el incremento del número de Betti β_1 indica que la red determina la neurona de llegada por combinación de las neuronas de salida. De este modo, podemos suponer que el aumento de β_1 releja la complejidad del conocimiento adquirido por la red. Por lo tanto, mediante el uso de la homología persistente seremos capaces de medir la complejidad del conocimiento adquirido por la red.

4. Experimentos

(En construcción.)