

Unas palabras sobre homología (y Análisis Topológico de Datos)

Miguel A. Maldonado

mmaldonado@uaz.edu.mx

7ma Jornada estatal de ciencia y tecnología 2025

Unidad Académica de Matemáticas,
Universidad Autónoma de Zacatecas



Motivación

¿Qué es la Topología?

¿Qué es homología?

Homología simplicial

Homología persistente

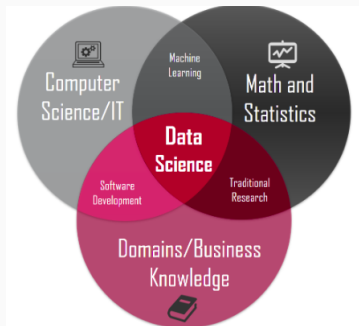
Comentarios finales

Referencias



Motivación

La **ciencia de datos** es la disciplina y la técnica de convertir datos en conocimiento, con el fin de tomar decisiones y realizar predicciones de manera informada.



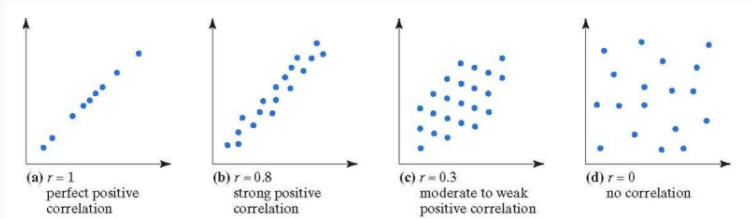
En particular, se busca obtener, tratar, analizar, depurar y manipular información (datos).

En esta área multidisciplinaria, las matemáticas y las ciencias computacionales tienen una gran presencia.



Visualización de datos

Una **nube de datos** es la representación gráfica (p.e. $\mathbb{R}^2, \mathbb{R}^3$) de los datos. Es de gran relevancia buscar patrones que reflejen relaciones entre ellos:



La manera en que se distribuyen los datos determina el modelo aplicado para analizarlos.

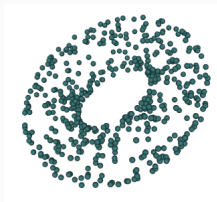


Ejemplo

Inferencia: *dada una muestra discreta (de un objeto) determinar características del objeto subyacente*



Si la muestra es suficientemente densa entonces refleja de mejor forma información del objeto.



Ciertos métodos topológicos se usan para reconocer dicha información; en particular, para detectar huecos en el objeto.



¿Qué es la Topología?

Lo que no es Topología



Topología

Es un tipo de geometría que estudia la *forma* de los objetos permitiendo que sean estirados y contraídos pero no rotos.

- Área **joven** de las matemáticas:
el concepto de espacio
topológico se debe a J.B.
Listing (1847).
- Ofrece estudio **cualitativo** de los
objetos (ángulos y distancias
son irrelevantes)
- Estudia propiedades **robustas**; es decir, aquellas que no
cambian con pequeñas variaciones.



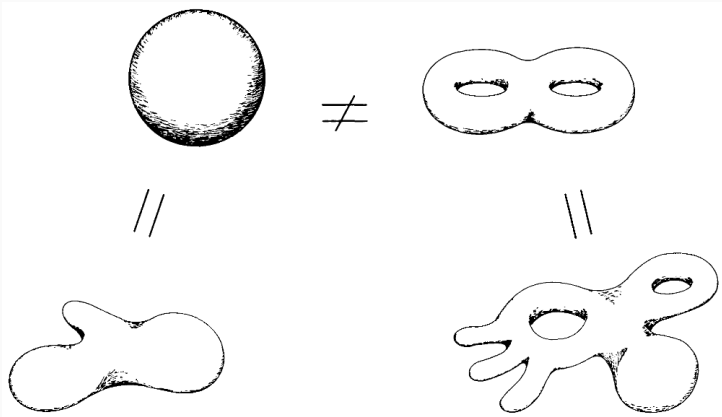


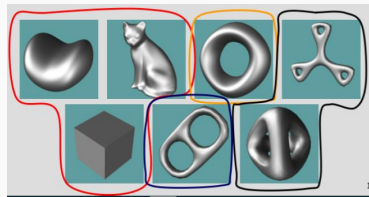
Figure 1: Objetos no equivalentes y sus posibles deformaciones



Las **propiedades topológicas** son aquellas que no cambian bajo deformaciones continuas; en particular, objetos equivalentes tienen las mismas propiedades topológicas.

Ejemplos: el número de partes, tener orilla/frontera, ser *abierto*, ser cerrado y acotado (**compacto**), cardinalidad, etc.

Esto permite clasificar objetos:
si no tienen las mismas propiedades,
no pueden ser equivalentes



Los hoyos o cavidades en un objeto son propiedades topológicas de suma importancia...

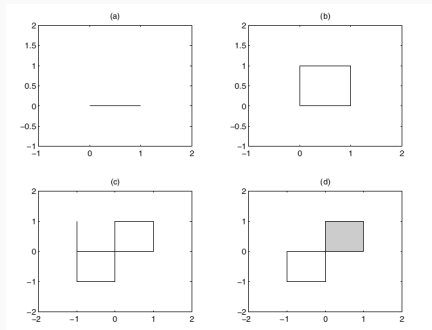


Figure 2: Objetos con y sin hoyos

Detectarlos y clasificarlos es primordial para la Topología.

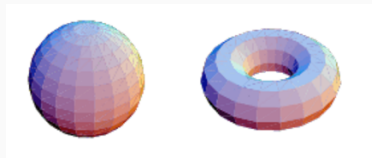


Figure 3: Existen distintos tipos de cavidades





¿Qué es homología?

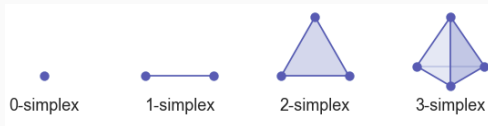
Homología = *algebrizar* la geometría



1er paso: definir objetos

Un q -simplejo es un conjunto de puntos distintos, ordenados, no colineales:

$$\sigma_q = (v_0, v_1, \dots, v_q) \in \mathbb{R}^{q+1}$$

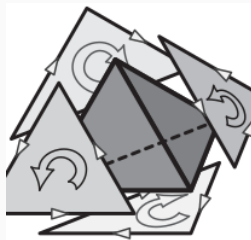
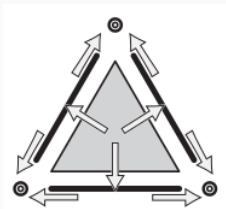


Observemos que

- **0-simplejos** = vértices (v)
- **1-simplejos** = aristas (v_1, v_2)
- **2-simplejos** = caras triangulares (v_1, v_2, v_3)
- **3-simplejos** = tetraedros rellenos (v_1, v_2, v_3, v_4)



Las **caras** de un q -simplejo σ_q son aquellos simplejos de dimensión menor contenidos en él



Así, un q -simplejo tiene caras de dimensión $q - 1, q - 2, \dots, 1, 0$.



Un **complejo simplicial** K es una colección de simplejos que se intersectan bien:

1. Si $\sigma \in K$, también están sus caras.
2. Si $\sigma, \tau \in K$, $\sigma \cap \tau = \emptyset$ ó $\sigma \cap \tau$ es una cara en común.

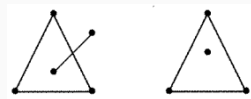
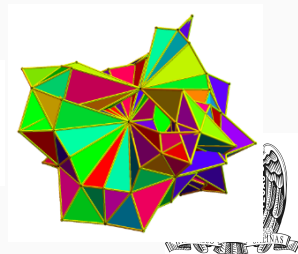
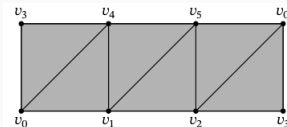
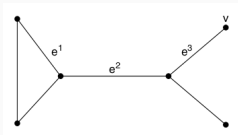


Figure 4: Patologías

Ejemplos:

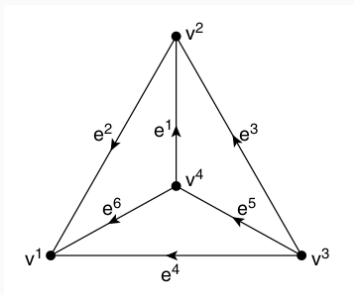


2do paso: algebrizamos

$C_q(X)$ es el espacio vectorial generado por los q -simplejos de X

$$c \in C_q(X), \quad c = \alpha_1 \sigma_q^1 + \alpha_2 \sigma_q^2 + \cdots + \alpha_m \sigma_q^m, \quad \text{con } \alpha_i \in \mathbb{R}$$

son llamadas **q -cadenas**.



Ejemplos de cadenas:

$$2e^1 + e^2 - e^6 - e^3$$

$$v^2 - v^4 - v^3$$



Suma de q -cadenas:

$$(\sum \alpha_i \sigma_q^i) + (\sum \beta_i \sigma_q^i) := \sum (\alpha_i + \beta_i) \sigma_q^i$$

Inversos: $c = \sum \alpha_i \sigma_q^i$ es $-c = \sum (-\alpha_i) \sigma_q^i$.

Multiplicación por escalar:

$$\alpha c = \alpha \alpha_1 \sigma_q^1 + \alpha \alpha_2 \sigma_q^2 + \cdots + \alpha \alpha_m \sigma_q^m, \quad \alpha \in \mathbb{R}$$



De lo anterior se obtiene una colección de espacios vectoriales $\{C_q(X)\}_{q \geq 0}$ relacionados entre sí mediante el **operador frontera...**

- Para 0-simplejos:
 $\partial(v) = 0$
- Para dimension 1:
 $\partial(v_1, v_2) := (v_2) - (v_1)$
- Para dimension 2:
 $\partial(v_1, v_2, v_3) := (v_2, v_3) - (v_1, v_3) + (v_1, v_2)$
- Para 3-simplejos:
 $\partial(v_1, v_2, v_3, v_4) := (v_2, v_3, v_4) - (v_1, v_3, v_4) + (v_1, v_2, v_4) - (v_1, v_2, v_3)$

Para una cadena $c = \sum \alpha_i \sigma_q^i \in C_q(X)$ definimos

$$\partial_q(c) := \sum \alpha_i \partial_q \sigma_q^i$$

Notemos que $\partial_q(c) \in C_{q-1}(X)$.



Ejemplo

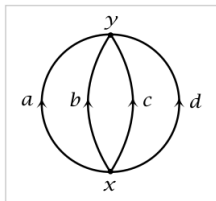


Figure 5: Tres 1-ciclos

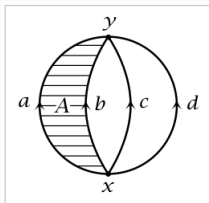


Figure 6: Dos 1-ciclos

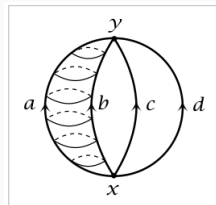


Figure 7: Dos 1-ciclos, un 2-ciclo



Ejemplo

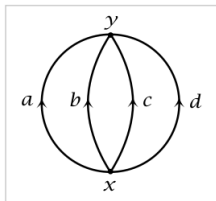


Figure 5: Tres 1-ciclos

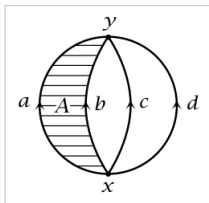


Figure 6: Dos 1-ciclos

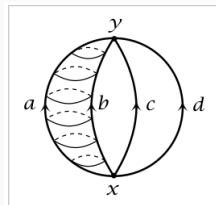


Figure 7: Dos 1-ciclos, un 2-ciclo



Notemos que los ciclos que corresponden a huecos o cavidades en el objeto son aquellos que no son frontera de una simplejo de dimensión mayor.



Por el resultado sobre el operador frontera

$$B_k(X) \subset Z_k(X), \quad \forall k \geq 0$$

Nos interesa determinar aquellos elementos de $Z_k(X)$ que **no** están en $B_k(X)$; es decir, determinar las distintas clases laterales del cociente $Z_k(X)/B_k(X)$.

q -ésima **homología** de X :

$$H_k(X) := Z_k(X)/B_k(X)$$



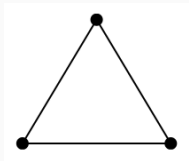
Los **números de Betti** de X

$$\beta_k(X) := \dim H_k(X)$$

indican cuántos generadores tiene $H_k(X)$.



Ejemplo: triángulo hueco



Sólo hay información en dimensiones 0, 1:
 $C_0(X)$, $C_1(X)$, tienen 3 generadores cada uno.

Basta estudiar el operador frontera $\partial : C_1(X) \rightarrow C_0(X)$ para obtener:

$$\begin{bmatrix} -1 & 0 & -1 & 0 \\ 1 & 0 & 0 & -1 \\ 0 & -1 & 1 & 0 \\ 0 & 1 & 0 & 1 \end{bmatrix} \begin{bmatrix} \alpha_1 \\ \alpha_2 \\ \alpha_3 \\ \alpha_4 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \end{bmatrix}.$$

solving

$$\begin{bmatrix} -\alpha_1 - \alpha_3 \\ \alpha_1 - \alpha_4 \\ -\alpha_2 + \alpha_3 \\ \alpha_2 + \alpha_4 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \end{bmatrix},$$

$$\alpha_1 = -\alpha_2 = -\alpha_3 = \alpha_4.$$

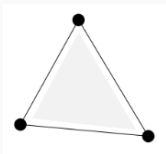
$$\beta_0(X) = 1$$

$$\beta_1(X) = 1$$

$$\beta_k(X) = 0, \forall k \geq 2$$



Ejemplo: triángulo relleno



Hay información en dimensiones
0, 1, 2: $C_0(X)$, $C_1(X)$, $C_2(X)$, tienen
3, 3 y 1 generadores,
respectivamente.

Basta estudiar el operador frontera $\partial : C_1(X) \rightarrow C_0(X)$ para
obtener:

$$\beta_0(X) = 1, \quad \beta_k(X) = 0, \quad \forall k \geq 1$$



Propiedades

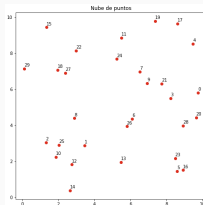
Componentes conexas: son subconjuntos conexos maximales de un espacio, los *pedazos* de los que está compuesto.

Ejemplo.

_____ X _____ Y

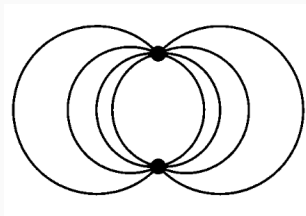
tienen una y dos componentes, respectivamente.

Una nube de datos $\mathcal{P}$ tiene $|\mathcal{P}|$ componentes



Un 1-ciclo no trivial es una curva cerrada simple:

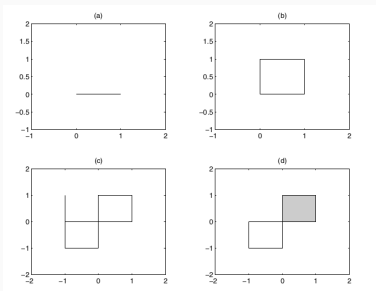
En el plano, los $\beta_1(X)$ corresponde al número de fronteras de regiones cerradas y acotadas.



Teorema Los generadores de $H_1(X)$ corresponden a los 1-ciclos que no son frontera de triángulos; es decir, $\beta_1(X)$ refleja cuántos huecos tiene X .



Ejemplo. Para las regiones



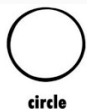
$$\beta_1(X_a) = 0, \quad \beta_1(X_b) = 1$$

$$\beta_1(X_c) = 2, \quad \beta_1(X_d) = 1$$

Observemos también que $\beta_0 = 1$ para todos los casos anteriores. ◀



Homología de esferas



$$\beta_1(S^1) = 1, \quad \beta_k(S^1) = 0, k \neq 1$$

$$\beta_2(S^2) = 1, \quad \beta_k(S^2) = 0, k \neq 2$$

En general
$$\beta_k(S^n) = \begin{cases} 1, & k = 0, n \\ 0, & k \geq 1 \end{cases}$$

Más aún:

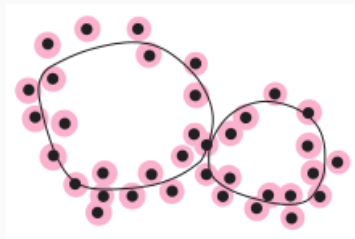
- Si $n \neq m$, $S^n \not\cong S^m$
- La homología distingue los distintos tipos de huecos que tienen.
- Si $\beta_k(X) \neq 0$, entonces X contiene una k -esfera: $S^k \subset X$



Homología persistente

Análisis Topológico de Datos

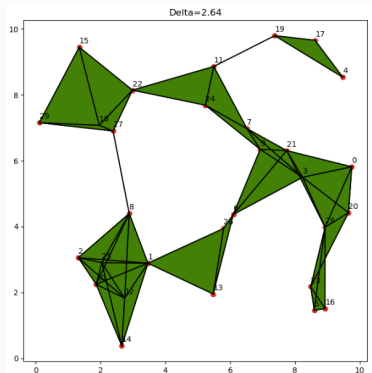
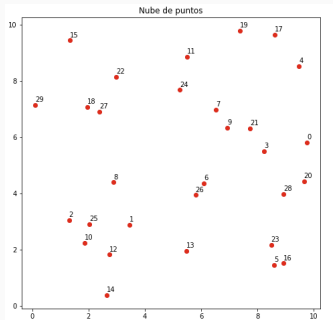
ATD es el uso de herramientas topológicas en la ciencia de datos para, entre otras cosas, detectar patrones en su representación gráfica.



Pregunta: si los cálculos de homología permiten detectar huecos, ¿cómo hacemos para detectar los huecos de la figura?



Una nube de datos no tiene información relevante pues sólo es una colección de puntos...

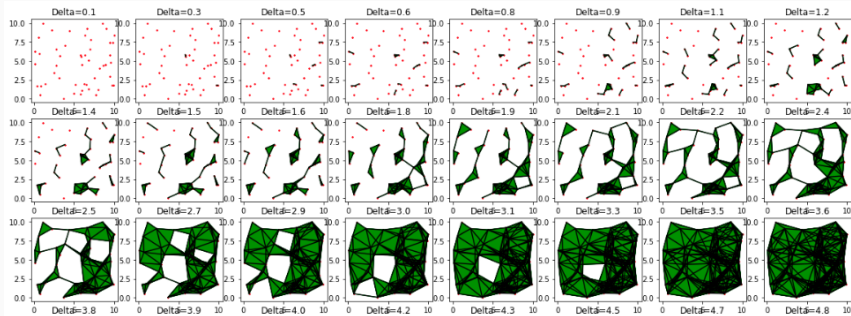


A partir de una nube se construyen objetos geométricos para identificar cómo se relacionan entre si



Ejemplo

Nube de datos y una colección de complejos con parámetro $\delta > 0$

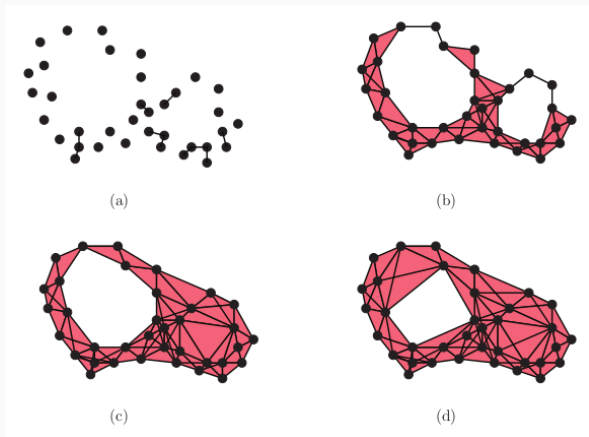


Tomando a δ como un parámetro de tiempo resulta importante determinar cuánto tiempo están presentes las características detectadas.



Ejemplo

La nube de datos muestra cierta disposición entre ellos...



De los huecos que se forman, uno sobrevive más que otro; tiene una **vida** más larga. ◀



La colección de objetos forma una **filtración**:

$$X_0 \longrightarrow X_1 \longrightarrow \cdots \longrightarrow X_{n-1} \longrightarrow X_n,$$

Consideramos la homología para detectar los huecos:

$$H_q(X_0) \longrightarrow H_q(X_1) \longrightarrow \cdots \longrightarrow H_q(X_{n-1}) \longrightarrow H_q(X_n)$$

Este sistema es llamado el **módulo de persistencia**.



Estudiar cómo cambia $H_q(X_i)$ a lo largo de la filtración nos dirá qué tan predominante es la estructura detectada.



Las propiedades/características detectadas se representan por elementos $z_i \in H_q(X_i)$. Al determinar

- su **nacimiento** $b(z_i)$ (cuándo aparecen)
- su **muerte** $d(z_i)$ (cuándo desaparecen)

se calcula su **persistencia**

$$lf(z_i) := d(z_i) - b(z_i)$$

Elementos con $lf(z_i)$ grande reflejan características predominantes

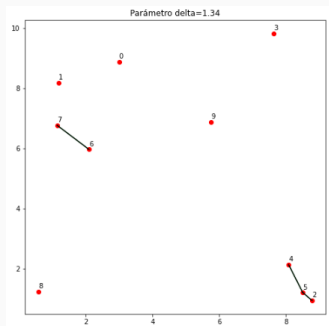
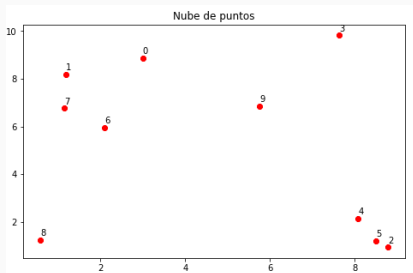


Ejemplo

El módulo

$$H_0(X_0) \longrightarrow H_0(X_1) \longrightarrow \cdots \longrightarrow H_0(X_{n-1}) \longrightarrow H_0(X_n)$$

permite conocer cómo varían las componentes conexas.



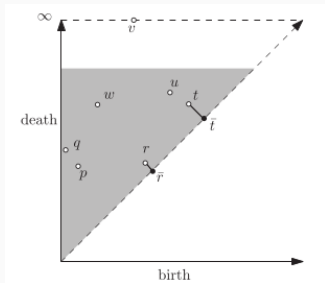
H_0 cambia al conectar puntos. ◀



Diagrama de persistencia

Para los elementos z_i de la homología consideramos el conjunto de parejas

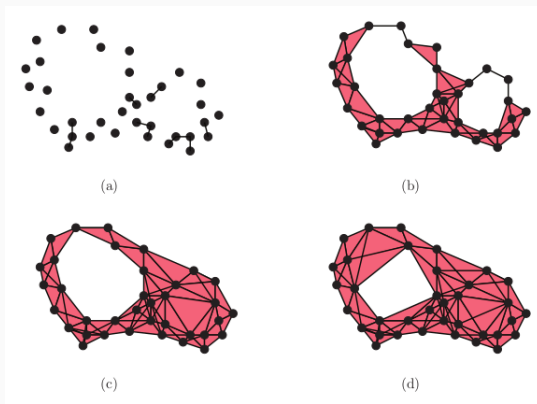
$$\{(b(z_0), d(z_0)), (b(z_1), d(z_1)), \dots\}$$



El **diagrama de persistencia** es la representación de los puntos en el plano $\overline{\mathbb{R}^2}$



Ejemplo



El análisis homológico muestra la existencia de dos clases z , w correspondientes a los huecos que se forman; una de ellas tiene mayor persistencia que la otra. ◀



Comentarios finales

- Hay una fuerte componente computacional en el área:
programación, visualización, procesamiento de imágenes.



- Hay una fuerte componente computacional en el área:
programación, visualización, procesamiento de imágenes.
- Hay una fuerte componente teórica en el ATD: *álgebra lineal, álgebra conmutativa, teoría de representaciones, topología*



- Hay una fuerte componente computacional en el área: *programación, visualización, procesamiento de imágenes.*
- Hay una fuerte componente teórica en el ATD: *álgebra lineal, álgebra conmutativa, teoría de representaciones, topología*
- El estudio del diagrama de persistencia se encuentra con la estadística y el análisis funcional cuando se recurre a ciertas estructuras métricas.



- Hay una fuerte componente computacional en el área:
programación, visualización, procesamiento de imágenes.
- Hay una fuerte componente teórica en el ATD: *álgebra lineal, álgebra conmutativa, teoría de representaciones, topología*
- El estudio del diagrama de persistencia se encuentra con la estadística y el análisis funcional cuando se recurre a ciertas estructuras métricas.
- Las aplicaciones de estos métodos son muchas:
detección/clasificación de cáncer, estudio de actividad cerebral, estudios de lingüística, estadísticas deportivas,...



¡Gracias!

Miguel A. Maldonado

mmaldonado@uaz.edu.mx

Unidad Académica de Matemáticas,
Universidad Autónoma de Zacatecas



Referencias

Aplicaciones de la topología

- Adams, Franzosa, **Intro. to Topology, Pure and Applied**, Pearson, 2008
- Edelsbrunner, Harer, **Computational Topology**, AMS, 2010
- Ghrist, R., **Elementary Algebraic Topology**, 2014.
- Fraleigh, **A First Course in Abstract Algebra**, Pearson, 2002.
- Wildberger, *An introduction to homology*, Insights into Mathematics, video de YouTube.
- Nanda, V., **Computational Algebraic Topology**.

