

Análisis topológico de datos: complejo de Morse-Smale discreto

DANIEL AMAYA ZABALA*
UNIVERSIDAD DE ANTIOQUIA

DR. CARLOS SEGOVIA**
UNIVERSIDAD NACIONAL AUTÓNOMA DE MÉXICO

Junio 2024

Descripción:

El Análisis topológico de datos (TDA) ha experimentado un auge en la comunidad científica en las últimas décadas. Los topólogos algebraicos han tenido un papel fundamental al aplicar estas técnicas al análisis de datos, haciendo uso de las nuevas capacidades computacionales. La combinación de la teoría topológica con las herramientas computacionales nos permiten explorar y comprender estructuras intrínsecas en los datos, lo que resulta valioso en áreas como: Visualización y procesamiento de imágenes, la clasificación y detección de patrones. En este proyecto emplearemos la herramienta Topology ToolKit (TTK), la cual utilizaremos para estudiar el complejo de Morse-Smale discreto para aproximar visualizaciones de variedades de 3 dimensiones tomando como referencia los artículos [3][5].

*Daniel.amayaz@udea.edu.co

**csegovia@im.unam.mx

1. Introducción

En las matemáticas aplicadas el análisis topológico de datos (TDA) es un enfoque al análisis de conjuntos de datos utilizando técnicas de la topología. La extracción de información de los conjuntos de datos que son de grandes dimensiones, incompletos y ruidosos usualmente representan un gran reto. TDA provee herramientas de trabajo para analizar tales datos de una forma más independiente de una métrica escogida y proporciona una reducción de dimensión. La motivación inicial es el estudio de la forma de los datos. TDA ha sido un puente entre la topología algebraica y otras ramas de las matemáticas puras, un ejemplo de esto es el uso de la homología de persistencia, una adaptación de la homología a las nubes de datos.

2. Funciones de Morse

Considere el disco abierto de puntos a una distancia menor que uno del origen, $D = \{x \in \mathbb{R}^2 \mid \|x\| < 1\}$. Es homeomorfo a \mathbb{R}^2 . Esto es fácil de ver, considere la función $f : D \rightarrow \mathbb{R}^2$ definida por $f(x) = \frac{x}{1-\|x\|}$. Llamaremos a cualquier subconjunto de un espacio topológico que es homeomorfo a D un disco abierto.

Definición. una 2 variedad sin frontera es un espacio topológico \mathbb{K} cuyos puntos pertenecen en discos abiertos. Intuitivamente esto significa que \mathbb{K} localmente parece un plano.

Definición (Compacidad). Decimos que \mathbb{K} es compacto si para toda cobertura de \mathbb{K} por conjuntos abiertos, llamados cubiertas abiertas, podemos encontrar un número finito de conjuntos que cubren a \mathbb{K} . Decimos que la cubierta finita siempre tiene subcobertura finita.

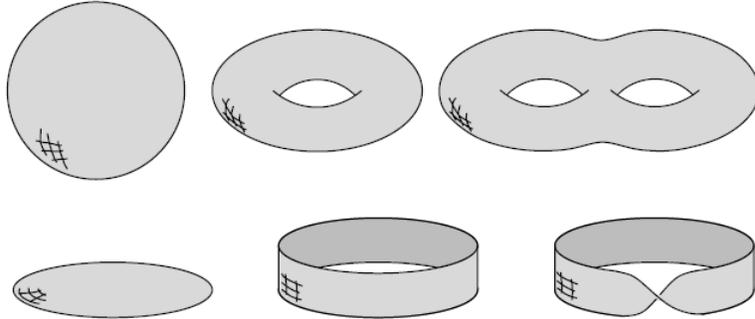


Figura 1: Desde arriba a la izq. Tenemos la esfera, el toro, y el doble toro. De abajo hacía la derecha, el disco, un cilindro y la banda de Möbius. Imagen tomada de [4]

Ahora haremos una breve introducción a la teoría de Morse, la cual nos permite de forma elegante estudiar la topología de las superficies a través del estudio de funciones en éstas, más aún, estudiar su homotopía.

Para esto traeremos un ejemplo clásico. Consideremos el toro.

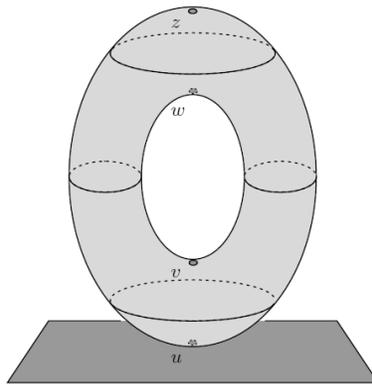


Figura 2: Toro 3D con puntos críticos medidos por la función de altura. Imagen obtenida de [4]

Sea \mathbb{M} un toro 2-dimensional y $f(x)$ la altura del punto $x \in \mathbb{M}$ sobre el plano horizontal en el cual reposa el toro. Como en la figura 2 llamamos $f : \mathbb{M} \rightarrow \mathbb{R}$ la función altura, cada número real a tiene una preimagen,

$f^{-1}(a)$, el cual se refiere a un conjunto de nivel.

$$\mathbb{M}^a = f^{-1}(-\infty, a] = \{x \in \mathbb{M} \mid f(x) \leq a\}$$

Nosotros estamos interesados en la evolución del conjunto de subnivel a medida que incrementamos el nivel a . En particular los cambios más importantes se dan cuando a pasa a través de los puntos críticos del toro u, v, w, z

Esto lo podemos ver en la figura. Para $a < f(u)$ el conjunto de subnivel es vacío. Para $f(u) < a < f(w)$ el conjunto de subnivel es un cilindro, el cual tiene la homotopía de un círculo. Así continuamos hasta tener $f(z) < a$ donde tenemos el toro completo.

Las palabras suave y diferenciable serán usadas intercambiabilmente para significar diferenciable de clase C^∞ . El espacio tangente de una variedad suave M en un punto r será denotado por TM_r . Si $g : \mathbb{M} \rightarrow \mathbb{N}$ es una función suave con $g(r) = s$ entonces la función lineal inducida de espacios tangentes será denotada por: $g_* : TM_r \rightarrow TM_s$.

Ahora, sea f una función suave real valuada en una variedad M . Decimos que $r \in M$ es llamado un **punto crítico** de f si la función inducida $f_* : TM_r \rightarrow T\mathbb{R}_{f(r)}$ es cero. Si escogemos un sistema coordenado (x^1, \dots, x^n) en una vecindad de U de r esto significa que:

$$\frac{\partial f}{\partial x^1}(r) = \dots = \frac{\partial f}{\partial x^n}(r) = 0$$

El valor real de $f(r)$ es llamado el **valor crítico** de f . Un punto crítico r se dice que es no degenerado si y sólo si la matriz:

$$\left(\frac{\partial^2 f}{\partial x^i \partial x^j}(r) \right)$$

Es no singular.

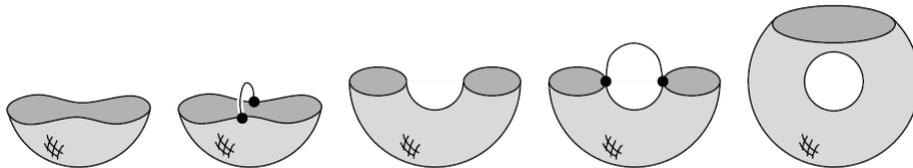


Figura 3: Homotopías del toro. Imagen obtenida de [4]

Definición. Una función de Morse es una función suave en una variedad M , $f : M \rightarrow \mathbb{R}$ tal que:

1. Todos los puntos críticos son no degenerados, y
2. Los puntos críticos tienen distintos valores en la función.

Lema (Morse 1). Sea r un punto crítico no degenerado para f . Entonces existe un sistema coordenado local (y^1, \dots, y^n) en una vecindad U de r con $y^i(r) = 0$ para todos los i tales que la identidad $f = f(r) - (y^1)^2 - \dots - (y^\lambda)^2 + (y^{\lambda+1})^2 + \dots + (y^n)^2$ es válida en todo U , donde λ es el índice de f en r .

Lema (Morse 2). Sea f una función suave real-valuada en una variedad M . Sea $a \leq b$ y suponga que el conjunto $f^{-1}[a, b]$, consiste de todos los $r \in M$ con $a \leq f(r) \leq b$, es compacta, y no contiene ningún punto crítico de f . Entonces M^a es difeomorfo a M^b . Además, M^a es una deformación retraible de M^b , así la función inclusión $M^a \rightarrow M^b$ es una equivalencia de homotopía.

Teorema (Morse 3). Si f es una función diferenciable en una variedad M sin puntos críticos degenerados, y si cada M^a es compacto, entonces M tiene el tipo de homotopía de un complejo CW, con una célula de dimensión λ por cada punto crítico de índice λ .

3. Funciones de Morse discretas

En este apartado estudiaremos la teoría de Morse discreta y su necesidad para el estudio de datos discretos. Sea M un complejo simplicial finito, K el conjunto de simplices de M , y K_p los simplices de dimensión p . Una función discreta de Morse en M será de hecho una función de K . Esto es, asigna un número real a cada simplejo en M . Escribimos $\sigma^{(p)}$ si σ tiene dimensión p , y $\tau > \sigma$ si σ pertenece a la frontera de τ . Decimos que una función: $f : K \rightarrow \mathbb{R}$ es una función de Morse discreta si para todo $\sigma^{(p)} \in K_p$:

1.
$$\#\{\tau^{(p+1)} > \sigma \mid f(\tau) \leq f(\sigma)\} \leq 1$$
2.
$$\#\{\nu^{(p-1)} < \sigma \mid f(\nu) \geq f(\sigma)\} \leq 1$$

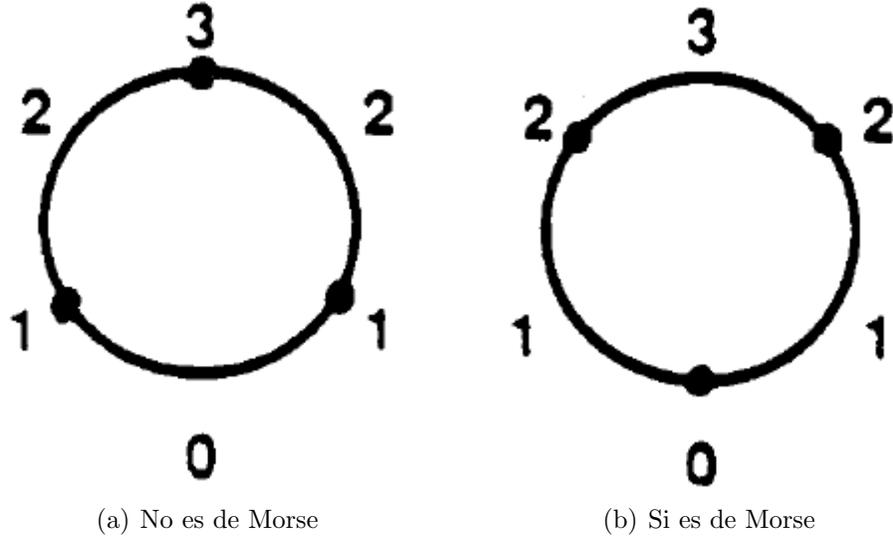


Figura 4: Ejemplos de funciones que son y no de Morse. Imagen obtenida de [2]

La función **a** no es una función de Morse discreta ya que el borde $f^{-1}(0)$ falla en la regla (2) y el vértice $f^{-1}(3)$ falla en la condición (1).

Definición. Decimos que $\sigma^{(p)}$ es crítico (con índice p) si:

1. $\# \{ \tau^{(p+1)} > \sigma \mid f(\tau) \leq f(\sigma) \} = 0$
2. $\# \{ \nu^{(p-1)} < \sigma \mid f(\nu) \geq f(\sigma) \} = 0$

La función **b** en la figura 4 es una función de Morse, además, $f^{-1}(0)$ es un punto crítico de índice 0, $f^{-1}(3)$ es un simplejo crítico de índice 1, y no hay más simplejos críticos.

La definición anterior nos da un análogo discreto de la noción de puntos críticos de índice P para funciones suaves. Antes de dar los principales teoremas, presentaremos una definición. Suponga que f es una función de Morse discreta en un complejo simplicial M. Para cualquier $c \in \mathbb{R}$ definimos el subcomplejo de nivel $M(c)$ por:

$$M(c) = \bigcup_{f(\tau) \leq c} \bigcup_{c \leq f(\sigma)} \sigma$$

Esto es, $M(c)$ es el subcomplejo de M consistiendo de todos los simplejos τ con $f(\tau) \leq c$, al igual que todas sus caras.

Teorema. Suponga que el intervalo $[a, b]$ no contiene valores críticos de f . Entonces $M(a)$ es una deformación retraible de $M(b)$. Por otra parte, $M(b)$ colapsa simplicialmente sobre $M(a)$

Teorema. Suponga $\sigma^{(p)}$ es un simplejo crítico con $f(\sigma) \in [a, b]$, y no hay otros simplejos críticos con valores en $[a, b]$. Entonces $M(b)$ es homotopicamente equivalente a:

$$M(a) \bigcup_{\dot{e}^{(p)}} e^{(p)}$$

Donde $e^{(p)}$ es una p -célula, y está pegada a $M(a)$ por su frontera $\dot{e}^{(p)}$

Corolario. Suponga que M es un complejo simplicial con una función de Morse discreta. Entonces M es homotopicamente equivalente a un complejo CW con exactamente una célula de dimensión p por cada simplejo crítico de índice p .

Presentamos un ejemplo simple. Vea la Figura 5. Aquí $f^{-1}(0)$ es un simplejo crítico de índice 0, $f^{-1}(8)$ es un simplejo crítico de índice 1, y no hay más simplejos críticos. Así, por el corolario anterior se tiene que M es homotopicamente equivalente a un círculo.

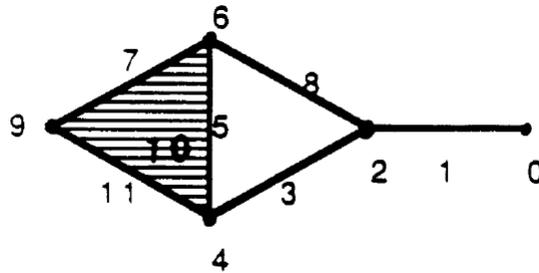


Figura 5: Complejo simplicial

Es útil indicar pictóricamente el colapso simplicial referido en el teorema anterior. Suponga $\sigma^{(p)}$ es un simplejo no crítico con $\tau^{(p+1)} > \sigma$ satisfaciendo $f(\tau) \leq f(\sigma)$. Entonces dibujamos una flecha que sale de σ a τ . El diagrama resultante puede verse en la figura 6. Estas flechas pueden ser vistas como un análogo al gradiente del campo vectorial de una función de Morse.

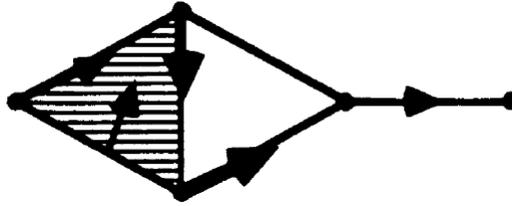


Figura 6: Colapso simplicial. Imagen obtenida de [2]

4. Ejemplo de Morse-Smale

El complejo de Morse-Smale ha probado ser una herramienta muy útil para extraer y visualizar características de datos de valores escalares.

El siguiente ejemplo es tomado de la base de datos de TTK.

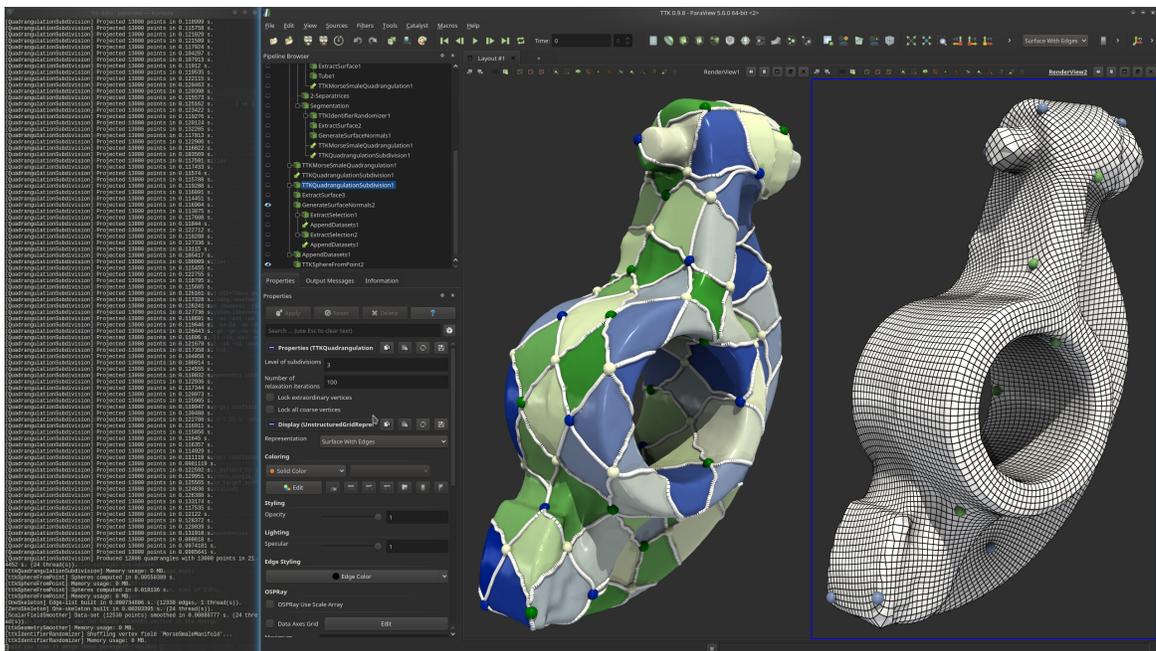


Figura 7: Cuadrangulación del complejo de Morse-Smale. Imagen obtenida de [6]

Este ejemplo carga primero un modelo mecánico como una malla triangular 2D desde el disco. Este modelo mecánico incorpora una colección de

campos escalares que corresponde a la salida del módulo EigenField. Este módulo generó una familia de funciones que están acopladas a la forma del conjunto de datos (básicamente, son funciones propias de la matriz laplaciana de la triangulación, ordenadas por magnitud de valor propio decreciente).

Referencias

- [1] J.W. Milnor. *Morse Theory*. Annals of mathematics studies. Princeton University Press, 1963. ISBN: 9780691080086. URL: https://books.google.com.co/books?id=A9QZZ3S_QxwC.
- [2] Robin Forman. «Morse Theory for Cell Complexes». En: *Advances in Mathematics* 134.1 (1998), págs. 90-145. ISSN: 0001-8708. DOI: 10.1006/aima.1997.1650. URL: <https://www.sciencedirect.com/science/article/pii/S0001870897916509>.
- [3] Attila Gyulassy et al. «A Practical Approach to Morse-Smale Complex Computation: Scalability and Generality». En: *IEEE Transactions on Visualization and Computer Graphics* 14.6 (2008), págs. 1619-1626. DOI: 10.1109/TVCG.2008.110.
- [4] H. Edelsbrunner y J. Harer. *Computational Topology: An Introduction*. Applied Mathematics. American Mathematical Society, 2010. ISBN: 9780821849255. URL: <https://books.google.com.co/books?id=MDXa6gFRZuIC>.
- [5] Julien Tierny et al. «The Topology ToolKit». En: *CoRR* abs/1805.09110 (2018). arXiv: 1805.09110. URL: <http://arxiv.org/abs/1805.09110>.
- [6] The TTK Development Team. *Morse-Smale Segmentation Example*. Accessed: 2024-05-16. 2024. URL: https://topology-tool-kit.github.io/examples/morseSmaleSegmentation_at/.