

MEMORIAS

COLOQUIO DEL VERANO CIENTÍFICO UNPA 2025

Complejos simpliciales y una aplicación introductoria en ciencia de datos

López Ortega Víctor Jared¹, Kantún Montiel Aura Lucina², Méndez Salinas Víctor Manuel², Rojas Carrasco Anahí^{2*}

¹Licenciatura en Matemáticas Aplicadas, Universidad del Papaloapan.

²Instituto de Agroingeniería, Universidad del Papaloapan.

*email: arojas@unpa.edu.mx

INTRODUCCIÓN

En la actualidad, el análisis de datos se ha consolidado como una herramienta esencial en la generación de conocimiento en diversas áreas, desde la biología y la ingeniería hasta las ciencias sociales. Su propósito es descubrir patrones, relaciones y estructuras ocultas en grandes volúmenes de información. No obstante, muchos conjuntos de datos reales presentan formas no lineales o estructuras complejas –como agrupamientos, ciclos o vacío– que no pueden capturarse adecuadamente mediante métodos estadísticos tradicionales.

La topología algebraica ofrece un marco adecuado para estudiar estas propiedades globales a través de invariantes topológicos¹. En particular, los complejos simpliciales constituyen una herramienta fundamental, pues permiten representar relaciones geométricas y de proximidad entre puntos de un conjunto de datos mediante objetos que son computacionalmente manejables^{2,3}. Este trabajo tiene como objetivo presentar una introducción accesible a los complejos simpliciales, ilustrando su construcción y aplicación básica en el contexto de la ciencia de datos.

MATERIALES Y MÉTODOS

Con el propósito de establecer las bases formales del estudio, en esta sección se presentan las definiciones fundamentales que permitieron construir y comprender los complejos simpliciales. Estas nociones constituyen el marco metodológico desde el cual se exploró su aplicación al análisis de datos.

En primer lugar, se introdujeron los conceptos de combinación convexa y envoltura convexa, que permitieron formalizar la idea de regiones generadas por puntos en un espacio euclidiano.

Definición 1: Sean u_0, u_1, \dots, u_k puntos en \mathbb{R}^n . Un punto $x = \sum_{i=0}^k \lambda_i u_i$, con cada $\lambda_i \in \mathbb{R}$, es una combinación afín de u_0, u_1, \dots, u_k si $\sum_{i=0}^k \lambda_i = 1$. La envoltura afín es el conjunto de todas las combinaciones afines.

Definición 2: Sean u_0, u_1, \dots, u_k puntos en \mathbb{R}^n . Una combinación afín $x = \sum_{i=0}^k \lambda_i u_i$, se dice que es una combinación convexa si $\lambda_i \geq 0$ para cada $1 \leq i \leq k$. Al conjunto de todas las combinaciones convexas se le llama envoltura convexa.

Posteriormente, se definió el concepto de puntos afinmente independientes, condición esencial para construir símplices bien definidos.

Definición 3: Se dice que los puntos $u_0, u_1, \dots, u_k \in \mathbb{R}^n$ son afinmente independientes si se cumple que: dos combinaciones afines $x = \sum_{i=0}^k \lambda_i u_i$ y $y = \sum_{i=0}^k \mu_i u_i$ son iguales si y sólo si $\lambda_i = \mu_i$, para cada $1 \leq i \leq k$.

A partir de estos elementos, se establecieron las nociones de k -símplice, cara y cara propia, las cuales describen las unidades geométricas básicas que conforman un complejo simplicial.

Definición 4: Un k -símplice es la envoltura convexa de $k + 1$ puntos afinmente independientes u_0, u_1, \dots, u_k de \mathbb{R}^n . Denotaremos a este conjunto como

$$\sigma = \text{conv}\{u_0, u_1, \dots, u_k\}.$$

Se emplean nombres particulares para los primeros casos: vértice (0-símplice), arista (1-símplice), triángulo (2-símplice) y tetraedro (3-símplice).

Definición 5: Sea $\sigma = \text{conv}\{u_0, u_1, \dots, u_k\}$ un k -símplice. Una cara de σ es la envoltura convexa de un subconjunto no vacío, A de $\{u_0, u_1, \dots, u_k\}$, y ésta es propia si el subconjunto A es un subconjunto propio de $\{u_0, u_1, \dots, u_k\}$.

Si τ es una cara de σ , escribiremos $\tau \leq \sigma$, y $\tau < \sigma$, si τ es una cara propia de σ .

Finalmente, se presentó la definición formal de complejo simplicial, entendido como una colección finita de símplices que satisface condiciones de cierre bajo inclusión de caras y de intersección.

Definición 6: Un complejo simplicial es una colección finita de símplices, K , que satisface las siguientes propiedades:

1. si $\sigma \in K$ y $\tau \leq \sigma$, entonces $\tau \in K$,
2. si $\sigma, \sigma_0 \in K$, entonces $\sigma \cap \sigma_0 = \phi$ o $\sigma \cap \sigma_0 \leq \sigma$ y $\sigma \cap \sigma_0 \leq \sigma_0$.

Este marco teórico-metodológico proporcionó los elementos necesarios para representar estructuras discretas a partir de conjuntos finitos de puntos, lo que permitió estudiar sus propiedades geométricas y topológicas en el contexto del análisis de datos.

RESULTADOS Y DISCUSIÓN

Para ejemplificar la aplicación de los complejos simpliciales en el análisis de datos, se consideró un conjunto finito de puntos en el plano \mathbb{R}^2 , v_0, v_1, v_2 . A cada punto se le asoció una bola de radio fijo, y las intersecciones entre ellas determinaron las relaciones de proximidad que dieron origen al complejo simplicial correspondiente.

- En un primer paso, cada punto constituyó un 0-símplice.
- Al unirlos dos a dos se generaron los 1-símplices.
- Finalmente, los tres puntos en conjunto formaron un 2-símplice, es decir un triángulo.

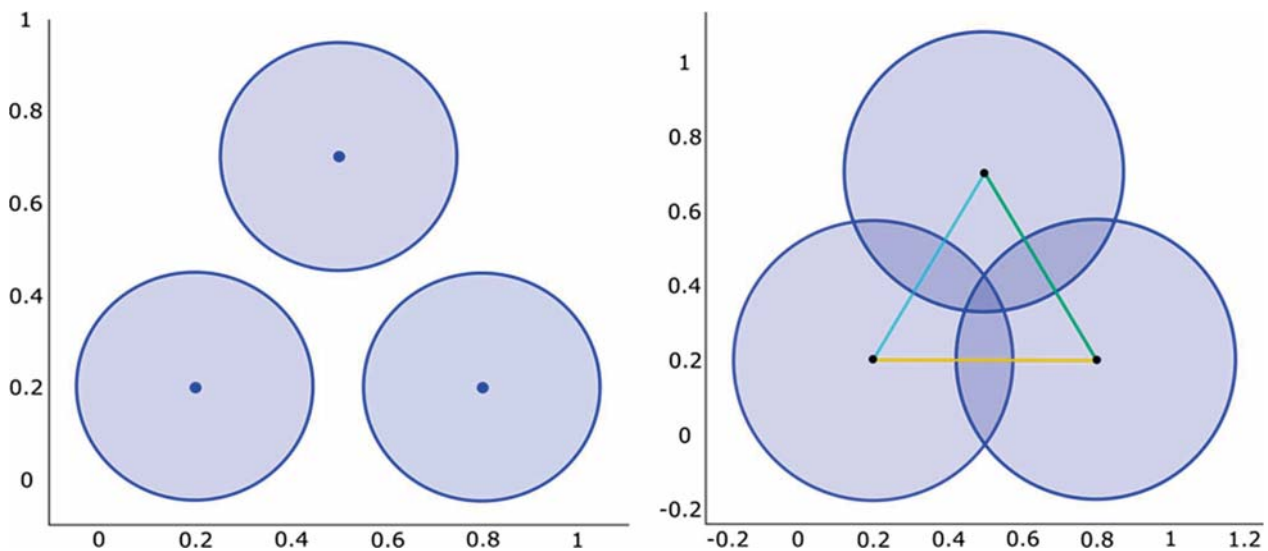
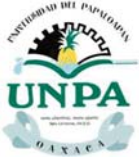


Figura 1. Construcción de complejos simpliciales a partir de una nube de puntos. A la izquierda, con un radio pequeño, las bolas alrededor de cada punto no se intersecan y sólo aparecen vértices aislados. A la derecha, con un radio mayor, las intersecciones generan aristas y triángulos, formando un complejo simplicial que refleja la estructura de los datos.



MEMORIAS

COLOQUIO DEL VERANO CIENTÍFICO UNPA 2025

De esta forma, el conjunto v_0, v_1, v_2 dio lugar a un complejo simplicial que consta de tres vértices, tres aristas y un triángulo.

De este modo, una nube de puntos pudo transformarse en un complejo simplicial, y su análisis permitió descubrir patrones en los datos. La ventaja de este enfoque es que estudiar datos se convirtió en estudiar propiedades topológicas, como la conectividad o la persistencia de ciclos, lo que brindó un marco teórico claro y a la vez herramientas prácticas para la ciencia de datos.

En casos más generales, cuando los datos consisten en colecciones grandes de puntos en espacios métricos de mayor dimensión, el mismo procedimiento se extiende para obtener representaciones discretas que sirven como base para cálculos topológicos más avanzados, tales como la homología simplicial y la homología persistente.

CONCLUSIÓN

Este trabajo presentó una introducción formal y visual al uso de los complejos simpliciales en el análisis de datos. A partir de definiciones fundamentales y un ejemplo ilustrativo, se mostró cómo estas construcciones permiten representar relaciones geométricas y de proximidad entre los puntos de un conjunto de datos, ofreciendo una forma discreta de analizar su estructura. El ejemplo desarrollado cumplió con el propósito del trabajo, al mostrar cómo los complejos simpliciales actúan como puente entre la topología algebraica y la ciencia de datos. De esta manera, se sentaron las bases conceptuales para herramientas más avanzadas, como la homología persistente, y se destacó el valor de la abstracción matemática como medio para comprender patrones y formas presentes en los datos.

El enfoque buscó ser accesible a estudiantes de licenciatura y fomentar la exploración de un campo actual y en expansión, donde la teoría matemática se conecta directamente con problemas prácticos de análisis e interpretación de datos.

REFERENCIAS

- [1] Hatcher, A. (2002). Algebraic Topology. England: Cambridge University Press.
- [2] Herbert, E.; Harer, J.L. (2009). Computational Topology: An introduction. United States of America: American Mathematical Society.
- [3] Chazal, F.; Michel, B. (2017). An introduction to Topological Data Analysis: fundamental and practical aspects. Foundations of Computational Mathematics 17 (16), 1505-1566.