

Seminario de Persistencia en Geometría 2024-1

5.2. Complejos de Čech y Rips, y Análisis Topológico de Datos.

Miguel Evangelista

23 de noviembre de 2023

Contenidos

Consideraciones

Complejo de Rips

Complejo de Čech

Consideraciones

Sea M una variedad Riemmaniana y sea $X \subseteq M$ una muestra finita de puntos de M .

Consideraciones

Sea M una variedad Riemmaniana y sea $X \subseteq M$ una muestra finita de puntos de M .

Dado lo anterior, surgen las siguientes preguntas.

- ▶ Teniendo sólo X , ¿es posible reconstruir M ?,
- ▶ ¿Qué tan bien podemos reconstruir M ?

Consideraciones

Denotemos por d la distancia de Riemann en M (y por consecuencia la distancia inducida en X). Decimos que el espacio métrico (X, d) modela una nube de datos.

Consideraciones

Denotemos por d la distancia de Riemann en M (y por consecuencia la distancia inducida en X). Decimos que el espacio métrico (X, d) modela una nube de datos.

Para responder la pregunta anterior, se tiene que que reconstruir la topología de M utilizando la geometría “local” del conjunto X

Consideraciones

Denotemos por d la distancia de Riemann en M (y por consecuencia la distancia inducida en X). Decimos que el espacio métrico (X, d) modela una nube de datos.

Para responder la pregunta anterior, se tiene que que reconstruir la topología de M utilizando la geometría “local” del conjunto X

Dado que uno de los principios del análisis de datos topológicos es “no confiar en grandes distancias”. (Ver Michael W. Mahoney, et al, *Algorithmic and statistical challenges in modern large-scale data analysis are the focus of MMDS 2008*, 2008)

Contenidos

Consideraciones

Complejo de Rips

Complejo de Čech

Complejo de Rips

Consideramos un espacio métrico finito (X, d) .

Se define el complejo simplicial $R_\alpha(X)$ de la siguiente manera:

- ▶ los vértices de $R_\alpha(X)$ son los puntos del conjunto X , y

Complejo de Rips

Consideramos un espacio métrico finito (X, d) .

Se define el complejo simplicial $R_\alpha(X)$ de la siguiente manera:

- ▶ los vértices de $R_\alpha(X)$ son los puntos del conjunto X , y
- ▶ los $k + 1$ puntos en X determinan un k -simplejo

$$\sigma = [x_0, \dots, x_k] \text{ si } d(x_i, x_j) < \alpha \quad \forall i, j$$

para $0 < \alpha \in \mathbf{R}$

Obs.

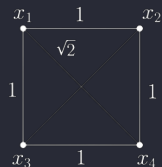
- ▶ El complejo de Rips está completamente determinado por su 1—esqueleto, es de hecho, un complejo de bandera.

Obs.

- ▶ El complejo de Rips está completamente determinado por su 1-esqueleto, es de hecho, un complejo de bandera.
- ▶ Para $0 < \alpha \leq \min_{x,y \in X, x \neq y} d(x,y)$ el complejo $R_\alpha(X)$ es una colección finita de puntos.

Complejo de Rips

Esta construcción se ilustra en la siguiente figura



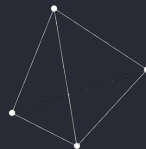
(X, d)



$R(X, \alpha)$
for $0 < \alpha \leq 1$



$R(X, \alpha)$
for $1 < \alpha \leq \sqrt{2}$



$R(X, \alpha)$
for $\alpha > \sqrt{2}$

Para $\alpha \neq \beta$ existe una función simplicial *natural*

$$i_{\alpha,\beta} : R_{\alpha}(X) \rightarrow R_{\beta}(X).$$

Para $\alpha \neq \beta$ existe una función simplicial *natural*

$$i_{\alpha,\beta} : R_{\alpha}(X) \rightarrow R_{\beta}(X).$$

Tomando $V_{\alpha}(X) = H_*(R_{\alpha}(X))$ y $\pi_{\alpha,\beta} = (i_{\alpha,\beta})_*$, obtenemos un módulo de persistencia, que denominamos módulo de Rips.

Para $\alpha \neq \beta$ existe una función simplicial *natural*

$$i_{\alpha,\beta} : R_{\alpha}(X) \rightarrow R_{\beta}(X).$$

Tomando $V_{\alpha}(X) = H_*(R_{\alpha}(X))$ y $\pi_{\alpha,\beta} = (i_{\alpha,\beta})_*$, obtenemos un módulo de persistencia, que denominamos módulo de Rips.

Obs.

Los complejos de Rips fueron introducidos por primera vez por Vietoris.

Contenidos

Consideraciones

Complejo de Rips

Complejo de Čech

Para $t > 0$, escribimos $B_t(x)$ que denota la bola abierta con centro en x de radio t con respecto a la métrica d .

Complejo de Čech

Para $t > 0$, escribimos $B_t(x)$ que denota la bola abierta con centro en x de radio t con respecto a la métrica d .

Definición 5.2.1

Sea $U = \{U_i\}$ una colección finita de subconjuntos de un conjunto A . Definimos el complejo Čech $\check{C}(U)$ asociado a U como:

- ▶ Los vértices son los conjuntos U_i .
- ▶ Una colección ordenada $\sigma = [U_0, \dots, U_k]$ es un k -simplejo si $\bigcap U_j \neq \emptyset$.
- ▶ Los operadores de frontera se definen de la manera estándar, lo que permite a uno considerar la homología correspondiente $H_*\check{C}(U)$.

Remark 5.2.2

Un caso especial que será de particular interés para nosotros es el siguiente: Sea (X, d) y fije $t > 0$.

¹Esto es para que la notación sea similar a la del complejo Rips

Remark 5.2.2

Un caso especial que será de particular interés para nosotros es el siguiente: Sea (X, d) y fije $t > 0$.

Considere la colección de bolas abiertas de radio $\frac{t}{2}$ centradas en el punto $x_i \in X$ denotado por $U_i = B_{\frac{t}{2}}(x_i)$.

¹Esto es para que la notación sea similar a la del complejo Rips

Remark 5.2.2

Un caso especial que será de particular interés para nosotros es el siguiente: Sea (X, d) y fije $t > 0$.

Considere la colección de bolas abiertas de radio $\frac{t}{2}$ centradas en el punto $x_i \in X$ denotado por $U_i = B_{\frac{t}{2}}(x_i)$.

Examinaremos la homología asociada al complejo Čech de esta colección $U_t = \{U_i\}$, y la denotaremos por $\check{C}_t(X)^1$.

¹Esto es para que la notación sea similar a la del complejo Rips

Complejo de Čech

Variando $t > 0$, de manera similar al caso del complejo Rips, podemos considerar el módulo de persistencia $H_*\check{C}(U)$ con los mapeos de persistencia inducidos por los mapeos de inclusión $i_{s,t} : \check{C}_s(X) \rightarrow \check{C}_t(X)$

Complejo de Čech

Variando $t > 0$, de manera similar al caso del complejo Rips, podemos considerar el módulo de persistencia $H_*\check{C}(U)$ con los mapeos de persistencia inducidos por los mapeos de inclusión $i_{s,t} : \check{C}_s(X) \rightarrow \check{C}_t(X)$

Notemos que los los módulos de persistencia provenientes de los complejos Čech y Rips correspondientes a un espacio métrico finito (X, d) son 1–entrelazados (after passing to a *logarithmic scale*).

Complejos de Čech y Rips

Lema 5.2.3

Sea (X, d) un espacio métrico finito. Tome $V_a = H_*(R_{2^a}(X))$ y $W_a = H_*(\check{C}_{2^a}(X))$, con morfismos inducidos a partir de los complejos Rips y Čech respectivamente. Entonces V y W son 1–entrelazados.

Complejos de Čech y Rips

Dem.

- ▶ Si $[y_0, \dots, y_k]$ es un simplejo en $R_t(X)$ se tiene que $d(y_i, y_j) < t$, entonces $y_i \in B_t(y_j)$ para todo i, j .
En particular, y_0 es un punto común para todo $B_t(y_j)$, por lo que $[y_0, \dots, y_k]$ determina un k -simplejo en \check{C}_{2t} . Por tanto, $R_t \subset \check{C}_{2t}$.

Complejos de Čech y Rips

Dem.

- ▶ Si $[y_0, \dots, y_k]$ es un simplejo en $R_t(X)$ se tiene que $d(y_i, y_j) < t$, entonces $y_i \in B_t(y_j)$ para todo i, j .
En particular, y_0 es un punto común para todo $B_t(y_j)$, por lo que $[y_0, \dots, y_k]$ determina un k -simplejo en \check{C}_{2t} . Por tanto, $R_t \subset \check{C}_{2t}$.
- ▶ Si $[y_0, \dots, y_k]$ es un simplejo $\check{C}_t(X)$, es decir, $\bigcap_j B_{\frac{t}{2}}(y_j) \neq \emptyset$, en particular para cada y_i, y_j las bolas $B_{\frac{t}{2}}(y_i)$ y $B_{\frac{t}{2}}(y_j)$ se intersectan, dado que $d(y_i, y_j) < t$.
Entonces $[y_0, \dots, y_k]$ es un simplejo en $R_t(X) \subseteq R_{\frac{t}{2}}(X)$. Por tanto, Por tanto, $\check{C}_{2t} \subset R_t$.

Complejos de Čech y Rips

Dem.

Pasando a la “escala logarítmica” y considerando los módulos de persistencia como antes, vemos que V y W son 1—entrelazados, con mapas entrelazados inducidos por la identidad. □

Remark 5.2.4

Mencionemos que los “módulos de persistencia” V y W reescalados logarítmicamente discutidos aquí no cumplen completamente con la definición 1.1.1 de módulos de persistencia.

Remark 5.2.4

Mencionemos que los “módulos de persistencia” V y W reescalados logarítmicamente discutidos aquí no cumplen completamente con la definición 1.1.1 de módulos de persistencia.

Es decir, la propiedad 4, que hace que los módulos de persistencia desaparezcan por la izquierda a partir de algún momento, no se cumple. Sin embargo, V y W son módulos de persistencia de tipo localmente finito, como se define en la sección 2,4.

Ejemplo 5.2.5

Considere un hexágono regular con una longitud de lado 1 (ver la siguiente figura).

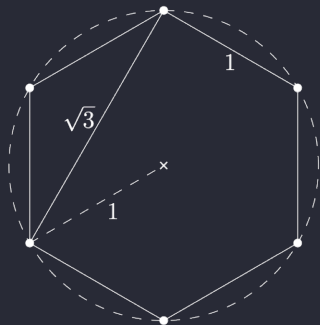
Calculemos los códigos de barras correspondientes a sus complejos Rips y Čech.

Ejemplo

Ejemplo 5.2.5

Considere un hexágono regular con una longitud de lado 1 (ver la siguiente figura).

Calculemos los códigos de barras correspondientes a sus complejos Rips y Čech.



Ejemplo

Complejo de Rips

Tenemos los siguientes grupos de homología a medida que t varía.

Ejemplo

Complejo de Rips

Tenemos los siguientes grupos de homología a medida que t varía.

- ▶ Para $0 < t \leq 1$, tenemos seis puntos distintos, es decir,
 $H_0 = \mathbf{R}^6$.

Ejemplo

Complejo de Rips

Tenemos los siguientes grupos de homología a medida que t varía.

- ▶ Para $0 < t \leq 1$, tenemos seis puntos distintos, es decir, $H_0 = \mathbf{R}^6$.
- ▶ Para $1 < t \leq \sqrt{3}$, tenemos un \mathbf{S}^1 , entonces $H_0 = \mathbf{R}$ y $H_1 = \mathbf{R}$.

Ejemplo

Complejo de Rips

Tenemos los siguientes grupos de homología a medida que t varía.

- ▶ Para $0 < t \leq 1$, tenemos seis puntos distintos, es decir, $H_0 = \mathbf{R}^6$.
- ▶ Para $1 < t \leq \sqrt{3}$, tenemos un \mathbf{S}^1 , entonces $H_0 = \mathbf{R}$ y $H_1 = \mathbf{R}$.
- ▶ Para $\sqrt{3} < t \leq 2$, tenemos una esfera \mathbf{S}^2 , que se obtiene pegando dos discos a lo largo de su frontera. Estos discos están creados por los triángulos que se muestran en la siguiente figura.

Por tanto, $H_0 = \mathbf{R}$, $H_1 = 0$ y $H_2 = \mathbf{R}$

Ejemplo

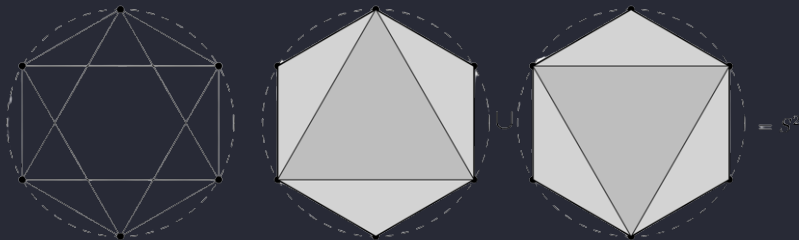
Complejo de Rips

- Para $t > 2$, obtenemos un 5-simplejo creado por los vértices del hexágono, por lo que solo nos queda $H_0 = \mathbf{R}$.

Ejemplo

Complejo de Rips

- Para $t > 2$, obtenemos un 5-simplejo creado por los vértices del hexágono, por lo que solo nos queda $H_0 = \mathbf{R}$.

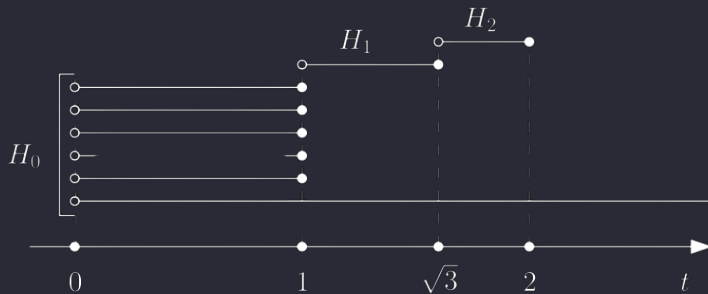


Ejemplo

Por último, el código de barras queda de la siguiente manera.

Ejemplo

Por último, el código de barras queda de la siguiente manera.



Ejemplo

Complejo de Čech

Tenemos los siguientes grupos de homología a medida que t varía.

Ejemplo

Complejo de Čech

Tenemos los siguientes grupos de homología a medida que t varía.

- ▶ Para $0 < t \leq 1$, tenemos, $H_0 = \mathbf{R}^6$.

Ejemplo

Complejo de Čech

Tenemos los siguientes grupos de homología a medida que t varía.

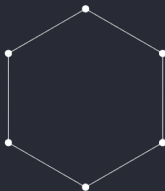
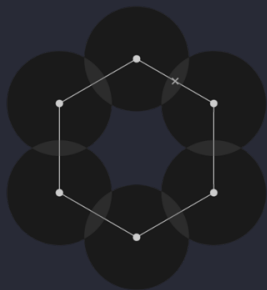
- ▶ Para $0 < t \leq 1$, tenemos, $H_0 = \mathbf{R}^6$.
- ▶ Para $1 < t \leq \sqrt{3}$, tenemos un \mathbf{S}^1 , entonces $H_0 = \mathbf{R}$ y $H_1 = \mathbf{R}$. Ver la siguiente figura.

Ejemplo

Complejo de Čech

Tenemos los siguientes grupos de homología a medida que t varía.

- ▶ Para $0 < t \leq 1$, tenemos, $H_0 = \mathbf{R}^6$.
- ▶ Para $1 < t \leq \sqrt{3}$, tenemos un \mathbf{S}^1 , entonces $H_0 = \mathbf{R}$ y $H_1 = \mathbf{R}$. Ver la siguiente figura.



Ejemplo

Complejo de Čech

- ▶ Para $\sqrt{3} < t \leq 2$, tenemos un \mathbf{S}^1 , que se obtiene de la unión de triángulos que es homotópico a \mathbf{S}^1 , es decir, todavía tenemos $H_0 = \mathbf{R}$ y $H_1 = \mathbf{R}$. Ver la siguiente figura.

Ejemplo

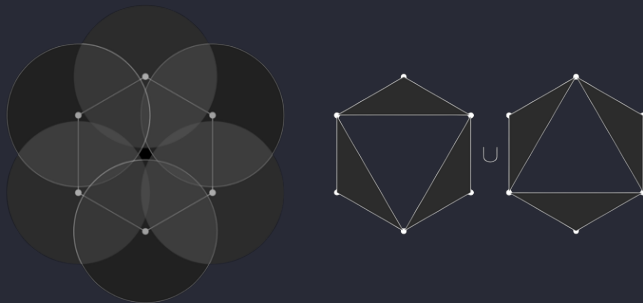
Complejo de Čech

- ▶ Para $\sqrt{3} < t \leq 2$, tenemos un \mathbf{S}^1 , que se obtiene de la unión de triángulos que es homotópico a \mathbf{S}^1 , es decir, todavía tenemos $H_0 = \mathbf{R}$ y $H_1 = \mathbf{R}$. Ver la siguiente figura. Hay que tener en cuenta que los grandes triángulos equiláteros todavía no están presentes.

Ejemplo

Complejo de Čech

- ▶ Para $\sqrt{3} < t \leq 2$, tenemos un \mathbf{S}^1 , que se obtiene de la unión de triángulos que es homotópico a \mathbf{S}^1 , es decir, todavía tenemos $H_0 = \mathbf{R}$ y $H_1 = \mathbf{R}$. Ver la siguiente figura. Hay que tener en cuenta que los grandes triángulos equiláteros todavía no están presentes.



Ejemplo

Complejo de Čech

- ▶ Para $t > 2$, obtenemos un 5—simplejo, por lo que solo nos queda $H_0 = \mathbf{R}$

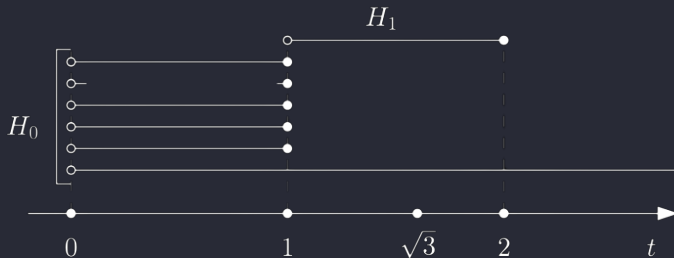
Por último, el código de barras queda de la siguiente manera.

Ejemplo

Complejo de Čech

- ▶ Para $t > 2$, obtenemos un 5-simplejo, por lo que solo nos queda $H_0 = \mathbf{R}$

Por último, el código de barras queda de la siguiente manera.



Ejemplo

Comparando los dos códigos de barras, notamos que el complejo Rips captura una celda bidimensional “redundante” (en el sentido de la topología del hexágono), mientras que el complejo Čech no.

Ejemplo

Comparando los dos códigos de barras, notamos que el complejo Rips captura una celda bidimensional “redundante” (en el sentido de la topología del hexágono), mientras que el complejo Čech no.

Sin embargo, observemos que el complejo Čech tiene la desventaja de ser más difícil de calcular y manejar (ver [76], Capítulo 5), ya que necesitamos conocer (y almacenar) la información sobre todos los posibles símlices (es decir, intersecciones) de cualquier número de bolas alrededor de los puntos muestreados).

Ejemplo

Al mismo tiempo, para el complejo Rips, como mencionamos en el Ejemplo 1.1.5, la información requerida es solo sobre 1-símples, es decir, sobre las distancias entre cada par de puntos del conjunto de muestra

Ejemplo

Ejercicio 5.2.6

Verifique que los dos códigos de barras que encontramos coincidan después de pasar a la *escala logarítmica* descrita en el Lema 5.2.3.