Herramientas de análisis topológico de datos en el estudio de sistemas dinámicos discretos

Horacio Araiza, Brayan Ramírez, Dra. Rosalía Hernández

Abstract

Se presenta una propuesta de visualización para el análisis de la dinámica de un sistema dinámico discreto, basada en herramientas de análisis topológico de datos. En particular, ésta se muestra en el contexto del modelo de Vicsek (C.M. Topaz, et. al.) en un espacio fase con condiciones de periodicidad en la frontera, el cual describe el movimiento colectivo de grupos de organismos. Se presenta también una comparación con otras variables de orden usualmente utilizadas, así como con la función de entropía persistente.

Agregados biológicos

Son grupos de organismos tales como parvadas, enjambres, manadas, etc. en donde las interacciones sociales entre los miembros pueden jugar un rol crucial en la formación de éstos. Bastan comportamientos como atracción, repulsión o alineamiento para producir una gran variedad de morfologías.

Complejos Simpliciales

Para obtener un objeto geométrico global a partir de una nube discreta de N puntos es necesario construir un *complejo simplicial*, el cual es un conjunto finito de *k-simplejos*, donde un 0-simplejo es un vértice, un 1-simplejo es una arista, un 2-simplejo un triángulo, etc.

Complejo de Vietoris-Rips

Dada una nube de puntos en $\mathbb{R}^n$, el Complejo de Vietoris-Rips, con parámetro de proximidad ε, es el complejo simplicial abstracto formado por todos los k-simplejos $\{v_0, v_1, ..., v_k\}$, donde la distancia entre cada par de puntos es menor o igual que ϵ.

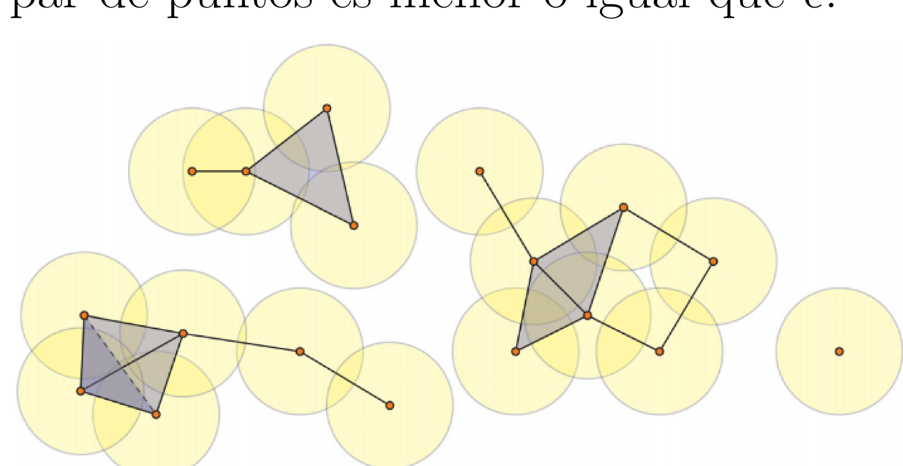

Fig. 1:Complejo de Vietoris-Rips para un conjunto de puntos

Homología

La homología es una manera de develar los "agujeros" k-dimensionales en un complejo simplicial. Esto requiere imponer una estructura algebraica en el complejo simplicial S: para cada $k \in \{0, 1, ..., \dim S\}$ se crea un espacio vectorial C_k cuya base es el conjunto de k-simplejos de S, con coeficientes en $\mathbb{Z}_2$. Los elementos de C_k son llamados k-cadenas.
Para calcular la homología, se requiere expresar algebraicamente la frontera ∂_k de un k-simplejo σ, la cual es la unión de los $(k-1)$-simplejos $\tau \subseteq \sigma$.
Los operadores frontera conectan los espacios vectoriales C_k en un *complejo de cadenas*:

$$\cdots \longrightarrow C_{k+1} \xrightarrow{\partial_{k+1}} C_k \xrightarrow{\partial_k} C_{k-1} \longrightarrow \cdots \longrightarrow C_2 \xrightarrow{\partial_2} C_1 \xrightarrow{\partial_1} C_0 \xrightarrow{\partial_0} 0$$

Los grupos de homología de S se definen como

$$H_k = \frac{Ker(\partial_k)}{Im(\partial_{k+1})}$$

y los números de Betti, definidos por $b_k = dim H_k$, dan información sobre los huecos de dimensión k del complejo simplicial. Por ejemplo, b_0 representa el número de componentes conexas, b_1 el número de círculos topológicos, b_2 los volúmenes atrapados, etc.

Homología Persistente

Las características topológicas de una nube de puntos dependen fuertemente del parámetro de proximidad ε, por tanto, es preferible visualizar los números de Betti con respecto a un intervalo de valores de ε, y estudiar entonces la respectiva filtración de subcomplejos. Así, las características que persisten en un amplio rango de valores de ε son consideradas señales de la topología subyacente global.

Modelo de Vicsek

El modelo de Vicsek es un sistema dinámico en tiempo discreto y espacio continuo, que describe el movimiento de partículas puntuales interactuantes en un cuadrado con condiciones de periodicidad en la frontera.

Las ecuaciones que lo describen son:

$$\begin{aligned}
\theta_i(t+\Delta t) &= \frac{1}{N}\left(\sum_{|\mathbf{x}_i - \mathbf{x}_j| \le R} \theta_j(t)\right) + U(-\eta/2, \eta/2), \\
\mathbf{v}_i(t+\Delta t) &= v_0(\cos\theta_i(t+\Delta t), \sin\theta_i(t+\Delta t)), \\
\mathbf{x}_i(t+\Delta t) &= \mathbf{x}_i(t) + \mathbf{v}_i(t+\Delta t)\,\Delta t.
\end{aligned}$$

Fig. 2:Modelo de Vicsek

Modelo de D'Orsogna

El modelo de D'Orsogna se basa en interacciones de atracción y repulsión entre partículas. Las ecuaciones modelan un sistema dinámico (continuo en el tiempo) que describe el movimiento en un plano de partículas puntuales interactuantes. Las ecuaciones son:

$$\begin{aligned}
\dot{\mathbf{x}}_i &= \mathbf{v}_i, \\
m\dot{\mathbf{v}}_i &= (\alpha - \beta|\mathbf{v}_i|^2)\mathbf{v}_i - \nabla_i Q_i, \\
Q_i &= \sum_{j \neq i} C_r e^{-|\mathbf{x}_i - \mathbf{x}_j|/L_r} - C_a e^{-|\mathbf{x}_i - \mathbf{x}_j|/L_a}.
\end{aligned}$$

Fig. 3:Modelo de D'Orsogna

Entropía Persistente

La entropía persistente se define como la entropía de Shannon del código de barras de una filtración de un conjunto finito de puntos, y es un estadístico topológico relacionado con la cantidad de información y la incertidumbre de los datos. Se calcula como:

$$S = -\sum_i p_i \log p_i$$

donde $p_i = l_i/L$ es el "peso" de una barra, $l_i = b_i - a_i$ su longitud y $L = \sum_i l_i$ es la suma de las longitudes de todas las barras.

Resultados

- Evolución de los huecos 0-dimensionales para el modelo de Vicsek

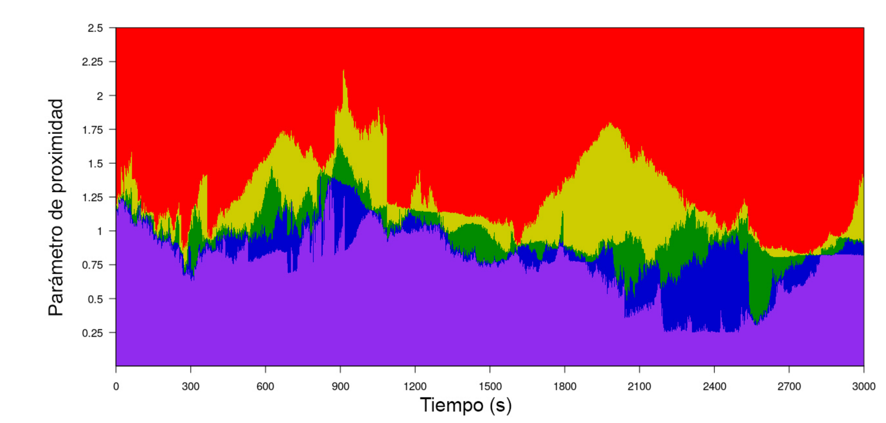

Fig. 4:Simulación con N=300, l=25, $\eta = 0.1$

- Evolución de los huecos 1-dimensionales para el modelo de D'Orsogna

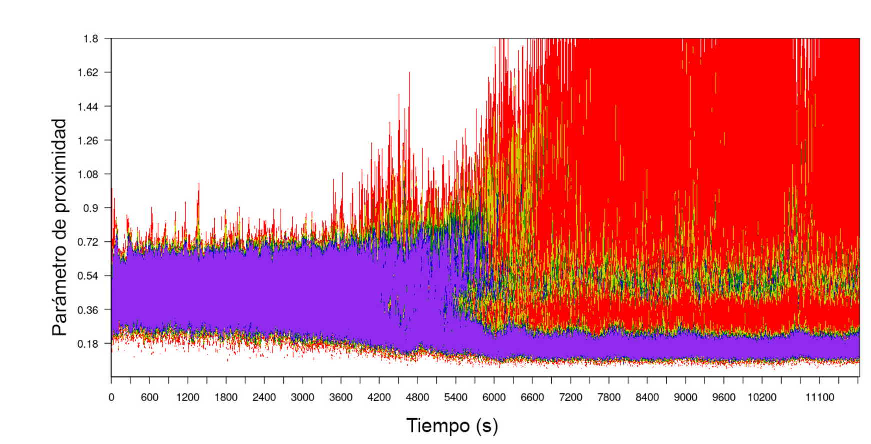

Fig. 5:Simulación con N=500, $\alpha = 1.5, \beta = 0.5, C_r = 1, L_r = 0.5, C_a = 0.5, L_a = 2$

- Entropía persistente 0-dimensional para el modelo de Vicsek

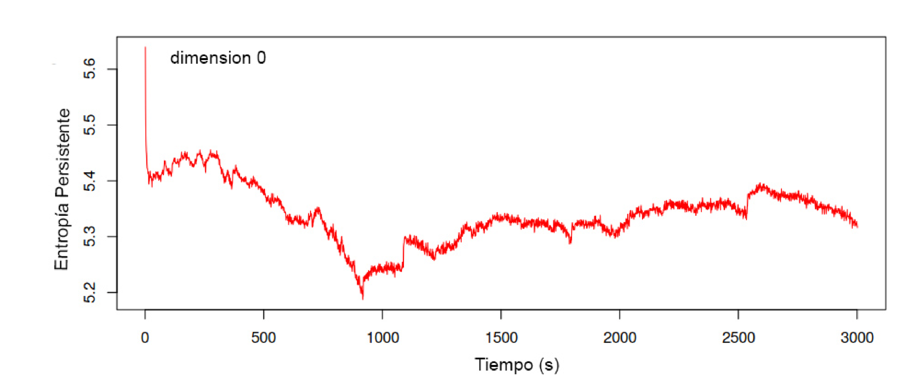

Fig. 6:Simulación con N=300, l=25, $\eta = 0.1$

References

[1] Topaz CM Ziegelmeier L, Halverson T.
Topological data analysis of biological aggregation models.
PLoS ONE 10(5), 2015.