

El laplaciano en la teoría espectral de grafos

Jesús Antonio Ramírez Aliphat
Departamento Académico de Matemáticas
ITAM
jesus.ramirez.ilp@gmail.com

1. Introducción

La teoría espectral de grafos es un puente entre el álgebra lineal, la combinatoria y múltiples aplicaciones prácticas que modelan fenómenos a través de redes, desde sistemas de transporte y circuitos eléctricos hasta redes sociales o neuronales.

Por medio de la representación matricial de un grafo, podemos recuperar ideas geométricas y aprovechar la estructura del espacio vectorial de funciones definidas en los vértices para obtener descripciones útiles de fenómenos de interés en la red que representa. Podemos preguntarnos qué tan bien conectado está el grafo, es decir, qué tan fácil es llegar de un vértice a otro, si el flujo es bueno a largo plazo o si se forman pequeñas comunidades en él. Para esto, nuestra herramienta principal será un operador que está relacionado con la difusión de *algo* en el grafo. Según el contexto, ese *algo* puede representar calor, corriente eléctrica, información, o incluso personas.

Por otro lado, la estructura que se obtiene a partir de la representación matricial es útil para encontrar buenas representaciones visuales de un grafo, fáciles de interpretar y que revelen información relevante, es decir, un buen dibujo. Veremos que se pueden utilizar las soluciones de sistemas de ecuaciones que involucran esta matriz como coordenadas para los vértices.

Antes de entrar de lleno en el contenido del artículo, comienzo con algunos conceptos preliminares para que quien lee se familiarice con el vocabulario utilizado y algunos resultados que se dan por hecho. Si

así se lo desea, se puede consultar [6] para refrescar algunos antecedentes de álgebra lineal y [7] para una introducción a la teoría de grafos convencional y sus aplicaciones.

Un grafo (o gráfica) se define como una pareja $G = (V, E)$ formada por un conjunto de vértices (nodos) y un conjunto de aristas que los unen. Decimos que dos vértices son adyacentes, denotado $i \sim j$, si están conectados por una arista, y llamamos la vecindad de un vértice al conjunto de vértices adyacentes a él, denotada $N(v) = \{u : u \sim v\}$ y definimos el grado de un vértice como $deg(v) = |N(v)|$.

Un grafo es simple si sus aristas no son dirigidas y no hay más de una arista entre un par de vértices, ni aristas de un vértice a sí mismo (lazos). Por otro lado, un grafo es dirigido si sus aristas tienen una orientación. Si $(i, j) \in E$, decimos que hay un arco de i a j , denotado $i \rightarrow j$. Además, es posible definir una función $w : E \rightarrow \mathbb{R}^+$ que le asigna un peso a las aristas del grafo (sea o no dirigido), en este caso se extiende la definición de grado a $deg(v) = \sum_{v \in e} w(e)$. Este peso puede ser interpretado, según el contexto, como una distancia, una probabilidad, o un costo por tomar una arista, etc.

En un grafo simple con n vértices, la matriz de adyacencia se define como la matriz binaria $A(G) \in M_{n \times n}(\mathbb{R})$ donde la entrada (i, j) es uno si $i \sim j$ y cero en otro caso. Sin embargo, para algunas partes de este artículo es necesario trabajar con una definición más general, en la que $A_{ij} = w(e_{ij})$, donde $w(e_{ij})$ denota el peso de arco de i a j (similarmente $A_{ij} = 0$ si $i \not\sim j$).

La teoría espectral se ocupa de estudiar propiedades de interés de los grafos a través de matrices asociadas a ellos y su espectro. Una de las matrices protagonistas en esta área de estudio es la matriz laplaciana, a veces llamada simplemente el laplaciano del grafo.

Definición 1.1. Para un grafo no dirigido, el laplaciano $L(G)$ se define:

$$L(G)_{ij} = \begin{cases} deg(i) & \text{si } i = j, \\ -w(e_{ij}) & \text{si } i \neq j, \end{cases}$$

Equivalentemente, podemos definir $L(G) = D(G) - A(G)$, donde $D(G)$ es la matriz diagonal de grados. Cuando no exista ambigüedad respecto al grafo del que se está hablando, se escribirá simplemente L para simplificar la notación.

La matriz L es real y simétrica, por lo que el *teorema espectral* [6] garantiza que tiene eigenvalores reales y sus eigenvectores forman una base ortonormal de \mathbb{R}^n . Además, se tiene que $\forall x \in \mathbb{R}^n : x^T L x \geq 0$, es decir, L es positiva semidefinida, lo que implica que todos los eigenvalores son no negativos. Más aún, si $\lambda_1 \leq \lambda_2 \leq \dots \leq \lambda_n$ son

los eigenvalores de L , se tiene que $\lambda_1 = 0$ con eigenvector constante $\mathbf{1}$, como se puede comprobar fácilmente:

$$(L\mathbf{1})_i = \text{deg}(i) - \sum_{j \sim i} w(e_{ij}) = 0.$$

La multiplicidad del eigenvalor cero es igual al número de componentes conexas del grafo. Por ejemplo, si los vértices de G pueden partitionarse como $V = V_1 \cup V_2$ de modo que no existen aristas entre ellos, entonces la matriz laplaciana será de la forma:

$$L = \begin{pmatrix} L_1 & 0 \\ 0 & L_2 \end{pmatrix},$$

y un vector indicador de V_i será un eigenvector del cero para $i \in \{1, 2\}$. Inductivamente se concluye que si existen k componentes conexas la matriz laplaciana tendrá k eigenvectores linealmente independientes correspondientes al eigenvalor cero.

En este artículo solo nos ocuparemos de grafos conexos, es decir, aquellos en los que existe un camino entre cualquier par de vértices, de modo que λ_2 es estrictamente positivo.

2. Construcción del laplaciano

Una definición más natural de la matriz laplaciana, que nos permitirá entender mejor el origen de su nombre y su relación con el operador diferencial, requiere de la matriz de incidencia, que se define a continuación.

Definición 2.1. Sea G un grafo dirigido con n vértices y m aristas, la matriz de incidencia $B \in M_{n \times m}(\mathbb{R})$ está definida como:

$$B_{ij} = \begin{cases} -1 & \text{si } e_j \text{ sale de } i, \\ 1 & \text{si } e_j \text{ termina en } i, \\ 0 & \text{en otro caso.} \end{cases}$$

A partir de esta matriz, podemos generalizar la definición del laplaciano para grafos dirigidos como $L = BB^T$. Además, es posible recuperar la matriz laplaciana de un grafo no dirigido mediante este método asignando una orientación arbitraria a las aristas.

El operador laplaciano es útil para medir la difusión de un cierto flujo (de energía, información, etc.) en la red que el grafo representa. Para ilustrar este hecho, consideremos el siguiente ejemplo:

Consideremos $x : V \rightarrow \mathbb{R}$ una función definida en los vértices del grafo (calor, por ejemplo) y analicemos el resultado de aplicar el operador.

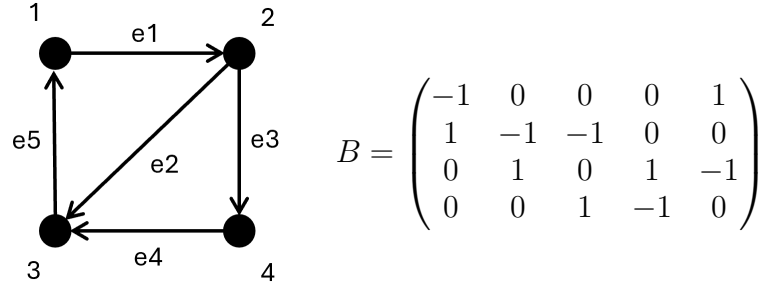


Figura 1. Un grafo dirigido y su matriz de incidencia.

Primero, consideramos la aplicación de B^T a nuestra función x .

$$B^T x = \begin{pmatrix} -1 & 1 & 0 & 0 \\ 0 & -1 & 1 & 0 \\ 0 & -1 & 0 & 1 \\ 0 & 0 & 1 & -1 \\ 1 & 0 & -1 & 0 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{pmatrix} = \begin{pmatrix} x_2 - x_1 \\ x_3 - x_2 \\ x_4 - x_2 \\ x_3 - x_4 \\ x_1 - x_3 \end{pmatrix}$$

El resultado se puede interpretar como el flujo que lleva cada arista en una unidad de tiempo, por ejemplo, la primera entrada nos dice que la arista e_1 es la diferencia entre los valores de la función en los vértices 1 y 2, una especie de derivadas parciales respecto a las aristas. En nuestra interpretación B^T juega el papel del gradiente de la función.

Por otro lado, al multiplicarse por un vector de aristas, el operador B nos dice cuánta energía entra o sale de cada nodo.

$$B e = \begin{pmatrix} -1 & 0 & 0 & 0 & 1 \\ 1 & -1 & -1 & 0 & 0 \\ 0 & 1 & 0 & 1 & -1 \\ 0 & 0 & 1 & -1 & 0 \end{pmatrix} \begin{pmatrix} e_1 \\ e_2 \\ e_3 \\ e_4 \\ e_5 \end{pmatrix} = \begin{pmatrix} e_5 - e_1 \\ e_1 - e_2 - e_3 \\ e_2 + e_4 - e_5 \\ e_3 - e_4 \end{pmatrix}$$

En nuestro ejemplo, el vértice 2 está recibiendo el flujo de la arista 1 y dejando salir el de las aristas 2 y 3, así que la matriz B funciona como la divergencia. De modo que el operador L es análogo al operador diferencial laplaciano $\Delta = \nabla \cdot \nabla$, es decir, la divergencia del gradiente.

$$BB^T x = \begin{pmatrix} 2 & -1 & -1 & 0 \\ -1 & 3 & -1 & -1 \\ -1 & -1 & 3 & -1 \\ 0 & -1 & -1 & 2 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{pmatrix} = \begin{pmatrix} 2x_1 - x_2 - x_3 \\ -x_1 + 3x_2 - x_3 - x_4 \\ -x_1 - x_2 + 3x_3 - x_4 \\ -x_2 - x_3 + 2x_4 \end{pmatrix}$$

Así, en una unidad de tiempo, el vértice 1 tendrá dos veces su calor original menos los valores iniciales de los vértices 2 y 3.

Esta conexión con la física puede ser aprovechada para ganar cierta intuición sobre el comportamiento de una red. En la siguiente sección se discute cómo podemos explotar estas ideas para entender mejor la estructura del grafo.

3. Conexidad y conductancia

Para estudiar el flujo de información en una red es necesario preguntarnos qué tan bien conectada está. La teoría espectral utiliza herramientas del álgebra lineal para cuantificar la conexidad de un grafo en términos de sus eigenvalores y encontrar particiones y comunidades en él.

Comencemos por estudiar la forma cuadrática del laplaciano:

$$\begin{aligned} x^T Lx &= \sum_{i=1}^n x_i \left(\deg(i)x_i - \sum_{j \sim i} w(e_{ij})x_j \right) = \sum_{i=1}^n x_i \left(\sum_{j \sim i} w(e_{ij})(x_i - x_j) \right) \\ &= \sum_{i=1}^n \sum_{j \sim i} w(e_{ij})(x_i^2 - x_i x_j) = \sum_{i \sim j} w(e_{ij})(x_i - x_j)^2. \end{aligned} \quad (1)$$

Esta es una medida de las diferencias de los valores de x entre vértices adyacentes, ponderada por el peso de las aristas entre ellos. En particular, en un grafo simple se tiene $x^T Lx = \sum_{i \sim j} (x_i - x_j)^2$.

Como se mencionó antes, la matriz laplaciana es singular, simétrica y semidefinida positiva, por lo que sus eigenvalores cumplen el *teorema de Courant-Fischer-Weyl* [6].

Teorema 3.1. Sean $0 = \lambda_1 \leq \lambda_2 \leq \dots \leq \lambda_n$ los eigenvalores del laplaciano, entonces:

$$\lambda_2 = \min_{x^T \mathbf{1} = 0} \frac{x^T Lx}{x^T x}; \quad v_2 = \arg \min_{x^T \mathbf{1} = 0} \frac{x^T Lx}{x^T x}. \quad (2)$$

El resultado anterior tiene como consecuencia que los primeros eigenvectores de L tienen valores similares en las entradas correspondientes a vértices adyacentes, por lo que el segundo eigenvalor más pequeño resulta ser muy importante, tanto que recibe un nombre propio.

Definición 3.2. La conexidad algebraica de un grafo conexo se define como el mínimo eigenvalor positivo λ_2 de su matriz laplaciana.

En algunos textos [10] se refiere a la conexidad algebraica como el valor de Fiedler, y a un respectivo eigenvector v_2 como vector de Fiedler. ¿Pero qué tiene que ver este eigenvalor con qué tan bien conectado está un grafo? Para contestar esta pregunta, veremos su relación con la frontera de un subconjunto de vértices, definida a continuación.

Definición 3.3. Dado $S \subset V$, definimos la frontera ∂S como el conjunto de aristas que van de un vértice en S a uno en su complemento $\bar{S} = V \setminus S$.

Sea χ_S el vector característico del conjunto S , es decir:

$$\chi_S(i) = \begin{cases} 1 & \text{si } i \in S, \\ 0 & \text{si } i \notin S, \end{cases}$$

entonces, la forma cuadrática del laplaciano de un grafo da como resultado la suma de los pesos de las aristas en su frontera;

$$\chi_S^T L \chi_S = \sum_{i \sim j} w(e_{ij}) (\chi_S(i) - \chi_S(j))^2 = w(\partial S), \quad (3)$$

o simplemente el número de aristas $|\partial S|$ en un grafo simple. Notaremos que esta forma cuadrática no cambia si se suma un vector constante, en particular, el vector x , definido como $x(i) = \chi_S(i) - |S|/|V|$, cumple que $x^T L x = \chi_S^T L \chi_S$, pero $\sum_{i=1}^n x_i = 0$, así que es ortogonal al vector $\mathbf{1}$ y el teorema 3.1 nos permite relacionar el eigenvalor λ_2 con las fronteras.

Esta cantidad es importante, pues dividir el grafo en dos grupos de vértices con pocas conexiones entre los vértices de distintos conjuntos se puede reducir a minimizar el número (o peso) de aristas entre un conjunto de vértices y su complemento.

Sin embargo, optimizar esto con vectores discretos con entradas en $\{0, 1\}$ es muy complicado, por lo que se propone relajar esta restricción y permitir un vector con entradas en los reales. Esta versión simplificada del problema da lugar a una heurística para particionar el grafo utilizando un vector de Fiedler. La idea es tomar conjuntos de nivel en las entradas de v_2 , es decir, establecer un umbral c (usualmente cero), y definir un vector característico χ con:

$$\chi(i) = \begin{cases} 1 & \text{si } v_2(i) \geq c, \\ 0 & \text{si } v_2(i) < c. \end{cases}$$

Esta técnica es muy útil para encontrar *bottlenecks* o cuellos de botella. Formalmente, buscamos encontrar un conjunto de corte, es decir, la menor cantidad de aristas que, al ser removidas, desconectarían el grafo en dos partes con aproximadamente la misma cantidad de vértices. Por ejemplo, en la figura 2, las aristas punteadas forman un cuello de botella entre S y T . Los cuellos de botella nos ayudan a separar una red en dos y son relevantes en el análisis de redes de transporte y comunicación, así como en la optimización de flujos y la robustez de algunos sistemas de información. Además, pueden ser útiles computacionalmente para procesar la información de un grafo por partes, ya que una permutación de los vértices (ordenarlos según los valores de

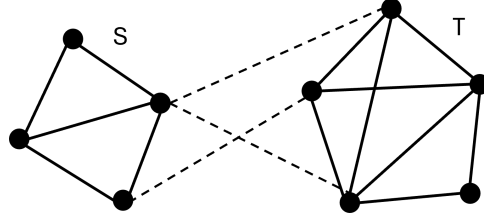


Figura 2. Un cuello de botella.

v_2) nos permite representar su matriz de adyacencia como una matriz por bloques:

$$A = \begin{pmatrix} A_S & \Theta_1 \\ \Theta_2 & A_T \end{pmatrix},$$

donde $\Theta_2 = \Theta_1^T$ son matrices ralas, i.e. compuestas en su mayoría por ceros, pues existen muy pocas aristas entre S y T , lo que permite estudiar las componentes mejor conectadas por separado, minimizando la pérdida de información.

Para grafos sin pesos, podemos pensar en la propensión de un grafo a tener cuellos de botella en términos de un invariante que relaciona la proporción entre las aristas en la frontera de un conjunto de vértices y su tamaño.

Definición 3.4. Sea G un grafo sin pesos, definimos el número isoperimétrico $\theta(G)$ como

$$\theta(G) = \min_{|S| \leq n/2} \frac{|\partial S|}{|S|}.$$

Lema 3.5. Para un grafo conexo G , su conectividad algebraica λ_2 satisface $\lambda_2/2 \leq \theta(G)$.

Demostración. Sea $|V| = n$ y $S \subset V$ con $|S| \leq n/2$ y definamos $s = |S|/n$. Consideramos su vector característico χ_S como arriba, y para aprovechar el teorema de Courant-Fischer (2) tomamos $x = \chi_S - s\mathbf{1}$ de modo que $x^T \mathbf{1} = 0$. Entonces

$$\lambda_2 \leq \frac{x^T Lx}{x^T x},$$

pero la ecuación (3) implica que

$$x^T Lx = \sum_{j \sim i} (x_i - x_j)^2 = \sum_{j \sim i} (\chi_S(i) - \chi_S(j))^2 = |\partial S|.$$

Por otro lado;

$$x^T x = \sum_{i=1}^n (\chi_S(i) - s)^2 = \sum_{i \in S} (1 - s)^2 + \sum_{j \notin S} s^2$$

$$= ns(1-s)^2 + (n-ns)s^2 = n(s-s^2).$$

Sustituyendo $s = |S|/n$, se obtiene $x^T x = \frac{1}{n}|S||\bar{S}|$ y recordando que $|\bar{S}| \geq n/2$, tenemos

$$\lambda_2 \leq \frac{n|\partial S|}{|S||\bar{S}|} \leq \frac{2|\partial S|}{|S|}. \quad \square$$

Es fácil notar que grafos con un bajo número isoperimétrico pueden ser bisectados con facilidad, mientras que si $\theta(G)$ es relativamente grande, no existen cuellos de botella.

Para extender esta noción a grafos pesados, que surgen en múltiples aplicaciones, es necesario trabajar con el volumen de un subconjunto de vértices, definido como $Vol(S) = \sum_{i \in S} deg(i)$. A partir de esto, podemos generalizar el número isoperimétrico a través de la conductancia.

Definición 3.6. La conductancia de un subconjunto de vértices S se define como:

$$\Phi(S) = \frac{w(\partial S)}{\min\{Vol(S), Vol(V \setminus S)\}}.$$

Es inmediato que la conductancia de S es igual a la de su complemento $\bar{S} = V \setminus S$. La conductancia del grafo $\Phi(G)$ se define como la mínima conductancia de todos sus subconjuntos propios de vértices, es decir $\Phi(G) = \min_{S \subset V} \Phi(S)$.

El nombre de conductancia proviene de tratar el grafo como una red de resistores [10], siendo el equivalente a la conductancia eléctrica, es decir, el recíproco de la resistencia. Intuitivamente, la conductancia es una medida de qué tan fácil es salir de un conjunto de vértices, o qué tan bueno es el flujo en un grafo, por lo que guarda relación con el comportamiento de caminatas aleatorias.

Es posible obtener una desigualdad parecida a la del lema 3.2 para la conductancia de un grafo en términos del espectro de una versión regularizada de su laplaciano.

Definición 3.7. Sea L la matriz laplaciana de un grafo y D su matriz diagonal de grados. El laplaciano normalizado (simétrico) está definido como $\mathcal{L} = D^{-1/2}LD^{-1/2}$.

Para entender la relación entre el espectro de la matriz laplaciana y su versión normalizada y porqué nos interesa, hacemos notar que

$$\sum_{i=1}^n deg(i)\chi_S(i)^2 = Vol(S),$$

y tenemos la siguiente identidad:

$$\frac{\chi_S^T L \chi_S}{\chi_S^T D \chi_S} = \frac{w(\partial S)}{Vol(S)}.$$

Ahora, si denotamos $y = D^{1/2}x$ se puede expresar $(x^T Lx)/(x^T Dx)$ como el cociente de Rayleigh [6] de la matriz \mathcal{L} como se muestra a continuación:

$$\frac{x^T Lx}{x^T Dx} = \frac{y^T D^{-1/2} L D^{-1/2} y}{y^T y} = \frac{y^T \mathcal{L} y}{y^T y},$$

pero el teorema 3.1 nos dice que los primeros eigenvectores de \mathcal{L} minimizan este cociente. Si y es un eigenvector de \mathcal{L} correspondiente al valor ν , se satisface la relación:

$$\mathcal{L}y = \nu y \iff LD^{-1/2}y = \nu D^{1/2}y \iff Lx = \nu Dx, \quad (4)$$

de modo que $0 = \nu_1 \leq \dots \leq \nu_n \leq 2$ son soluciones al problema de eigenvalores generalizado [6, 10] dado por las matrices simétricas L y D .

Aunque en general no es posible expresar los eigenvalores de \mathcal{L} en términos de los λ_i , la relación dada por (4) nos permite acotarlos. Para ν_i eigenvalor de \mathcal{L} se cumplen las desigualdades:

$$\lambda_i/d_{max} \leq \nu_i \leq \lambda_i/d_{min}, \quad (5)$$

donde d_{max} y d_{min} representan, respectivamente, el grado máximo y mínimo del grafo.

Podemos relacionar la conductancia de un grafo y el segundo eigenvalor de su laplaciano normalizado por medio del siguiente resultado.

Teorema 3.8. *Sea G un grafo conexo y $0 = \nu_1 < \nu_2 \leq \dots \leq \nu_n$ los eigenvalores de su matriz laplaciana normalizada. Entonces se cumple la **desigualdad de Cheeger**:*

$$\nu_2/2 \leq \Phi(G) \leq \sqrt{2\nu_2}.$$

La demostración del lado izquierdo de la desigualdad es similar a la del lema anterior. La del lado derecho es innecesariamente larga para este artículo, pero puede consultarse en [2, 10].

A continuación se presentan algunos ejemplos de instancias en donde se aproximan ambas cotas de la desigualdad. Cabe mencionar que el espectro del laplaciano de la definición 1.1 está mucho mejor estudiado que el de su versión normalizada, por lo que facilitamos algunos cálculos asumiendo que las aristas tienen peso unitario y considerando grafos regulares, es decir, aquellos donde todos los vértices tienen el mismo grado. Digamos que $deg(v) = d$ para todo vértice. Entonces, d_{max} es igual a d_{min} y la ecuación (5) implica que $\nu_i = \lambda_i/d$.

Sea Q_k un hipercubo de dimensión k , este grafo tiene $n = 2^k$ nodos y es regular de grado k . Para todo $k \geq 1$, la conexidad algebraica de Q_k es $\lambda_2(Q_k) = 2$ [5], así que $\nu_2(Q_k) = 2/k$.

Además, el conjunto que realiza la conductancia son exactamente 2^{k-1} aristas que dividen el grafo en dos (aquellas punteadas en la figura 3). En cada mitad hay 2^{k-1} vértices de grado k , por lo que $\Phi(Q_k) = 1/k$. Este resultado muestra la existencia de toda una familia de grafos que alcanza la igualdad del lado izquierdo del teorema 3.3.

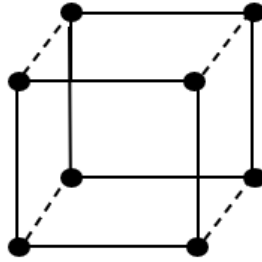


Figura 3. El hipercubo Q_3 .

Consideremos ahora C_n un ciclo con n vértices, por facilidad, asumimos que n es un número par. Entonces, si removemos dos aristas opuestas, como en la figura 4, podemos partir el grafo en conjuntos de $n/2$ vértices, todos de grado dos, por lo que $\Phi(C_n) = 2/n$.

Por otro lado, es un resultado conocido en teoría espectral de grafos que la conectividad algebraica de un ciclo es $\lambda_2(C_n) = 2 - 2\cos(\pi/n)$ [9, 10]. Para n grande, la expansión de Taylor del coseno nos permite aproximar $\lambda_2(C_n) \approx \pi^2/n^2$, y aprovechando la 2-regularidad del ciclo, obtenemos que $\nu_2(C_n) = \lambda_2(C_n)/2$ y, en consecuencia, $\sqrt{2\nu_2} \approx \pi/n$, aproximando la cota derecha.

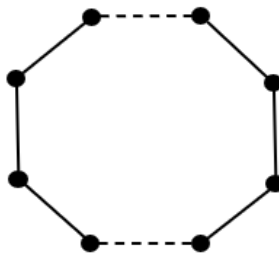


Figura 4. El ciclo C_8 cortado a la mitad.

Ambos lados de la desigualdad son útiles, ya que, si la conductancia del grafo es alta, significa que no hay una manera obvia de separar los vértices de la red, pero que es fácil que la información llegue de un nodo

a otro. En particular, el laplaciano normalizado \mathcal{L} es similar a $I - P$, donde I es la matriz identidad y P es la matriz de transición de una caminata aleatoria en el grafo [8]. Intuitivamente, si uno se encuentra en un vértice i , a cada unidad de tiempo se mueve hacia el vértice j con probabilidad P_{ij} , dada por:

$$P_{ij} = \begin{cases} \frac{1}{\deg(i)} & \text{si } i \sim j, \\ 0 & \text{si } i \not\sim j. \end{cases}$$

Así, si al tiempo cero se asigna una distribución de probabilidades $q \in \mathbb{R}^{1 \times n}$ sobre los vértices, después de m unidades de tiempo, las probabilidades de encontrarse en cada vértice están dadas por qP^m , y a largo plazo, la proporción del tiempo que se pasará en el vértice i será de $q_*(i) = \deg(i) / \sum_{j \in V} \deg(j)$ [8]. Este vector renglón q_* satisface la ecuación $q_*P = q_*$ y se conoce como la distribución estacionaria. Una alta conductancia quiere decir que el flujo en el grafo es bueno, así el tiempo de mezcla de una caminata aleatoria es bajo y esta converge rápidamente a la distribución estacionaria [2, 1].

En contraste, si la conductancia es baja, es más fácil separar el grafo e identificar comunidades de vértices bien conectados entre sí, lo que permite implementar técnicas como el *spectral clustering*, así como encontrar buenas inmersiones del grafo en el espacio euclidiano mediante las técnicas descritas a continuación.

4. Dibujo espectral y clustering

Uno de los retos que se enfrentan al trabajar con grafos es encontrar una inmersión de los vértices en el plano o, de manera informal, un dibujo que sea fácil de interpretar y estético. Algunas propiedades deseables son que los vértices adyacentes queden cerca entre sí, que las aristas sean líneas rectas y que haya el menor número posible de cruces entre ellas.

De acuerdo con las ideas desarrolladas en la sección anterior, podemos aprovechar que los primeros eigenvectores no triviales v_2 y v_3 tienen valores similares en las entradas correspondientes a nodos adyacentes y utilizarlos como coordenadas en el plano, o, en general, utilizar $\{v_2, \dots, v_{d+1}\}$ como coordenadas en el espacio \mathbb{R}^d . A esta técnica se le conoce como el dibujo espectral, o encaje espectral (spectral embedding en inglés).

Según la intuición física que hemos construido, así como la matriz L es un análogo discreto del operador Δ , la ecuación (1) muestra que su forma cuadrática $\sum_{i \sim j} w(e_{ij})(x_i - x_j)^2$ mide la variación de una función.

Esta juega el papel de la energía de Dirichlet [3] del sistema, por lo que los primeros eigenvectores del operador minimizan esta energía.

Consideremos algunos ejemplos sencillos y bien estudiados: un ciclo de $2n$ vértices C_{2n} y una trayectoria P_n . El caso del ciclo resulta más simple, pues $L(C_{2n})$ es una matriz circulante [6]. Recordamos que una matriz $A = (a_{ij})$ es circulante si $a_{i,j} = a_{i+1,j+1}$, es decir, si cada renglón se obtiene desplazando a la derecha las entradas del anterior.

La matriz laplaciana del ciclo es de la forma

$$L(C_{2n}) = \begin{pmatrix} 2 & -1 & \cdots & 0 & -1 \\ -1 & 2 & -1 & \cdots & 0 \\ \vdots & -1 & 2 & \cdots & \vdots \\ 0 & & \ddots & \ddots & -1 \\ -1 & 0 & \cdots & -1 & 2 \end{pmatrix}.$$

Las matrices de este tipo tienen propiedades útiles; en particular, el espectro de una matriz circulante de tamaño m puede ser expresado en términos de la raíz m -ésima de la unidad $\omega = e^{2\pi i/m}$. Por ejemplo, sus eigenvalores son combinaciones lineales de potencias de ω , pero como $L(C_{2n})$ es una matriz simétrica sus eigenvalores son reales, en efecto:

$$\lambda_k = 2 - 2 \cos\left(\frac{2\pi k}{2n}\right) = 2 - 2 \cos\left(\frac{\pi k}{n}\right); \quad 0 \leq k \leq n.$$

Para $1 \leq k < n$, la multiplicidad (algebraica y geométrica) de λ_k es igual a dos y su eigenspacio está generado por los siguientes vectores:

$$x_k(j) = \cos\left(\frac{\pi k j}{n}\right), \quad y_k(j) = \sin\left(\frac{\pi k j}{n}\right), \quad (6)$$

mientras que tanto y_0 como y_n son ignorados y x_0 es un vector constante correspondiente al eigenvalor cero. Geométricamente, parece natural que los eigenvectores del ciclo sean coordenadas del círculo unitario, como se observa a la izquierda de la figura 5.

Vale la pena remarcar que la notación utilizada en el ejemplo anterior no es consistente con el resto del artículo, pues llamamos $\lambda_0 = 0$ al primer eigenvalor y cada λ_k se repite dos veces, excepto cuando $k = n$. Esto se debe a que la notación elegida simplifica algunas expresiones, pero se espera que sea suficientemente clara.

Por otro lado, el laplaciano de una trayectoria de n vértices no es una matriz circulante, de hecho

$$L(P_n) = \begin{pmatrix} 1 & -1 & \cdots & 0 & 0 \\ -1 & 2 & -1 & \cdots & 0 \\ \vdots & -1 & 2 & \cdots & \vdots \\ 0 & & \ddots & \ddots & -1 \\ 0 & 0 & \cdots & -1 & 1 \end{pmatrix}.$$

Sin embargo, una construcción inteligente nos permite pensar en el grafo de la trayectoria P_n como un cociente del ciclo de $2n$ vértices. Para esto es necesario pedir ciertas condiciones de simetría en los eigenvectores de $L(C_{2n})$, a saber, requerimos que $v_k(j) = v_k(2n - j + 1)$. Los eigenvectores obtenidos en (6) no satisfacen esta identidad, sin embargo, si definimos nuevos vectores ψ_k como:

$$\psi_k(j) := \cos\left(\frac{\pi k(j - \frac{1}{2})}{n}\right) = \cos\left(\frac{k\pi j}{n} - \frac{k\pi}{2n}\right), \quad (7)$$

estos nuevos vectores cumplen la simetría deseada y no es difícil comprobar que se satisface

$$2\psi_k(j) - \psi_k(j - 1) - \psi_k(j + 1) = \psi_k(j) \left(2 - 2 \cos\left(\frac{k\pi}{n}\right)\right), \quad (8)$$

es decir, $L(C_{2n})\psi_k(j) = \lambda_k\psi_k(j)$, y los ψ_k son eigenvectores del laplaciano del ciclo. Argumentamos que un vector con las primeras n entradas de ψ_k descritas en (7) es un eigenvector de $L(P_n)$ correspondiente al eigenvalor λ_k .

Para $1 < j < n$ es inmediato que se satisface la ecuación (8), y se puede verificar haciendo uso de identidades trigonométricas que

$$\psi_k(1) - \psi_k(2) = \lambda_k\psi_k(1), \quad \text{y} \quad \psi_k(n) - \psi_k(n - 1) = \lambda_k\psi_k(n),$$

de modo que $L(P_n)\psi_k(j) = \lambda_k\psi_k(j)$ para $i \in \{1, \dots, n\}$.

Tomando $v_k \in \mathbb{R}^n$ con las primeras n entradas de ψ_k , tenemos que el espectro de la trayectoria está dado por:

$$\lambda_k(P_n) = 2 - 2 \cos\left(\frac{k\pi}{n}\right) \quad ; \quad v_k(j) = \cos\left(\frac{k\pi j}{n} - \frac{k\pi}{2n}\right).$$

Intuitivamente, lo que ocurre es que estamos «doblando» el ciclo sobre el eje que atraviesa las aristas $\{1, 2n\}$ y $\{n, n + 1\}$ e identificamos un vértice con su imagen bajo esta reflexión (véase la figura 5). Por ejemplo, si $n = 5$, consideramos las clases de equivalencia $\{1, 10\}$, $\{2, 9\}$, \dots $\{5, 6\}$. La condición de simetría asegura que vértices identificados tengan los mismos valores en ψ , se refiere a quien lee a [9] para los detalles.

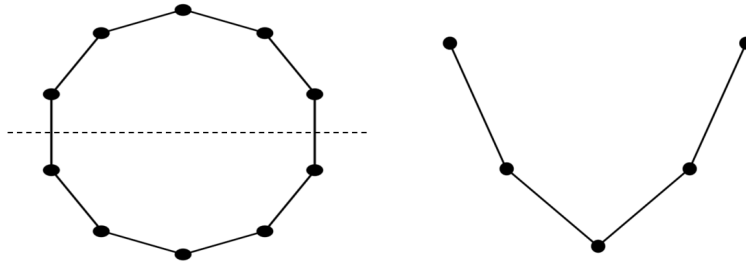


Figura 5. Dibujo espectral de P_5 obtenido como cociente de C_{10} .

Una de las ventajas de esta técnica es que la distancia euclidiana entre vértices adyacentes es menor, por lo que se obtiene una representación visual fiel a la estructura del grafo. En el mejor de los casos, el dibujo espectral es útil para identificar comunidades en una red. Una excelente manera de ilustrar esto es simular un *pequeño mundo*, siguiendo las ideas de Watts y Strogatz [13].

Un grafo de Watts-Strogatz con parámetros (n, k, p) se construye colocando n vértices en un anillo y conectando a cada uno con sus k vecinos más cercanos, obteniendo un grafo k -regular. Posteriormente, para cada vértice $u \in V$, cada una de sus aristas es *recableada* con probabilidad p , es decir, se elimina la arista y se conecta con un vértice elegido uniformemente al azar de $V \setminus (N(u) \cup \{u\})$, o sea, excluyendo al nodo mismo y sus vecinos.

Esta es una herramienta que permite modelar fenómenos como conexiones neuronales o redes sociales, pues se obtiene un grafo donde la distancia entre cualquier par de vértices es relativamente corta, pero que mantiene una estructura de comunidades. En la figura 6 podemos observar grafos de Watts-Strogatz con 20 y 25 nodos dibujados con un algoritmo dirigido por fuerzas (arriba) y usando el dibujo espectral.

En ambos casos, los clusters o comunidades son mucho más fáciles de identificar en el dibujo espectral (abajo), aunque algunos vértices resultan demasiado cercanos entre sí.

Una de las principales aplicaciones es el *spectral clustering* [12], donde, en lugar de utilizar directamente algoritmos de clustering (como k -means) sobre una nube de datos, se obtiene una triangulación de los puntos o se utiliza alguna técnica que permita recuperar un grafo de los datos espaciales y se encaja el grafo resultante en \mathbb{R}^d utilizando los primeros d eigenvectores no triviales y solo entonces se implementa algún método tradicional para clustering.

Este procesamiento de los datos permite encontrar clusters que los métodos más tradicionales no recuperan, como en la figura 7.

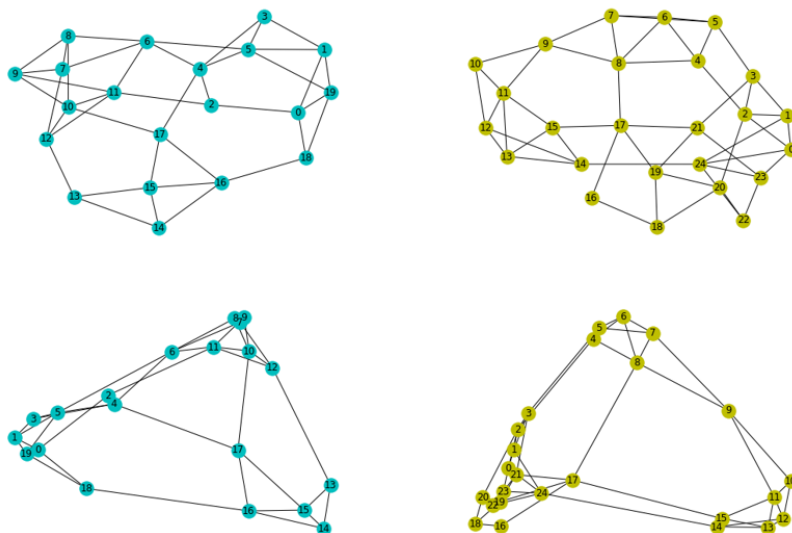


Figura 6. Grafos de Watts-Strogatz con $k = 4$ y $p = 0.2$.

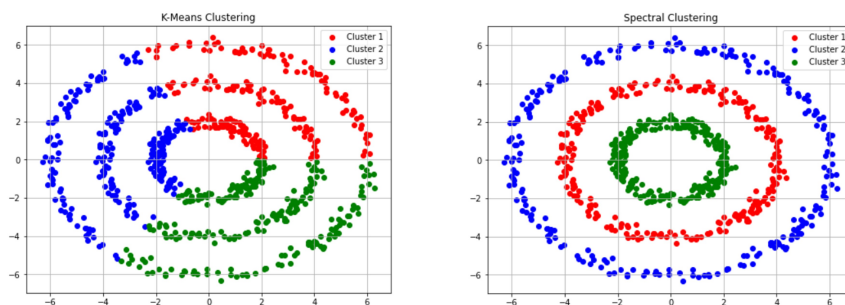


Figura 7. kMeans (izquierda) vs Spectral clustering (derecha).

Sin embargo, es importante mencionar que el dibujo espectral tiene limitaciones y no siempre produce resultados estéticos. Los problemas surgen principalmente para grafos muy densos (con muchas aristas). Para ilustrar este hecho, consideremos un grafo aleatorio de Erdős-Renyi [4], que toma cada arista del grafo completo de n vértices K_n con probabilidad p . En la figura 8 se contrasta el dibujo espectral de un grafo de este tipo con un dibujo circular del mismo.

Como se puede apreciar, en el dibujo espectral hay subconjuntos de vértices que terminan muy juntos entre sí (en ocasiones incluso en la misma posición), mientras que algunos vértices quedan muy aislados, resultando en un dibujo encimado y confuso.

De manera general, un buen dibujo espectral depende de qué tan denso es el grafo, esto es, la proporción de aristas que tiene respecto a K_n .

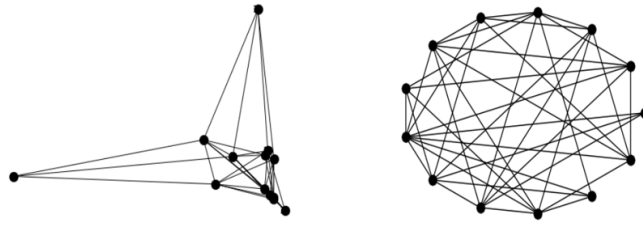


Figura 8. Dibujo espectral y circular de un grafo ER con $n = 13$ y $p = 0.5$.

Un mejor dibujo espectral es posible con grafos de baja conductancia, y en consecuencia, con valores pequeños de λ_2 .

Aunque esta técnica es útil para representar características relevantes en muchos grafos, en ocasiones es necesario recurrir a otras técnicas para hacer evidentes otras propiedades de interés del grafo. Para ilustrarlo, comparemos dos dibujos distintos del dodecaedro en la figura 9.

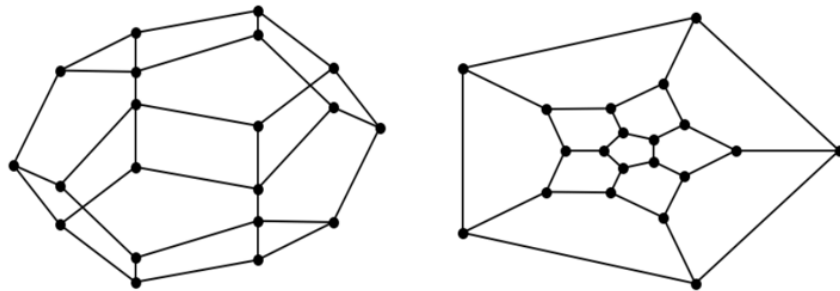


Figura 9. Dibujo espectral (izquierda) vs. dibujo de Tutte (derecha) del dodecaedro.

Aunque es cierto que el dibujo espectral es interpretable y refleja la geometría del dodecaedro, este dibujo pasa por alto una característica importante, pues este grafo puede dibujarse en el plano sin que haya cruces de aristas. En contraste, la técnica con la que se obtuvo el dibujo de la derecha sí refleja la planaridad del grafo. Este es un ejemplo de un dibujo de Tutte, que se discute en la siguiente sección.

5. Dibujo de Tutte y planaridad

Algunos algoritmos populares utilizan analogías con sistemas físicos para obtener dibujos que minimicen algún tipo de energía. Los llamados

algoritmos dirigidos por fuerzas modelan las aristas del grafo como resortes, con alguna constante elástica y longitud ideal fijas, mientras que los nodos son tratados como imanes con la misma polaridad, de modo que los vértices se repelen entre ellos, mientras que las aristas (resortes) mantienen cerca a los vértices adyacentes, resultando en dibujos donde los vértices están moderadamente alejados, evitando que se concentren demasiado y que las aristas sean demasiado largas.

Un ejemplo particularmente simple de este tipo de algoritmos fue propuesto por Tutte en 1963 en su artículo *How to Draw a Graph* [11]. Este método no considera fuerzas de repulsión, sencillamente modela a las aristas del grafo como resortes y se fija estratégicamente la posición de algunos vértices.

Sea $\vec{x}_i = (x_i, y_i)^T$ la posición del vértice i en el plano. Entonces, la ley de Hooke [3] nos dice que la fuerza ejercida sobre un vértice es:

$$F_i = \sum_{j \sim i} -k(\vec{x}_j - \vec{x}_i),$$

donde k es la constante elástica, pero podemos asumir $k = 1$ sin perder generalidad. Entonces, tenemos que el sistema está en equilibrio cuando los vértices se encuentran en el centro de masa de sus vecinos, es decir, cuando se cumple:

$$(L\vec{x})_i = \sum_{j \sim i} (\vec{x}_i - \vec{x}_j) = 0, \quad (9)$$

por lo que buscamos resolver el sistema de ecuaciones $L\vec{x} = 0$ (en lugar de un vector de incógnitas, tenemos una matriz de $n \times 2$, por lo que en realidad se resuelven dos sistemas).

La ecuación $Lz = 0$ siempre tiene como solución a un vector constante $c\vec{1}$, es decir, enviar todos los vértices a una misma posición. Además, esta solución es única, pues en un grafo conexo la multiplicidad de $\lambda_1 = 0$ es igual a uno, y el espacio nulo de la matriz está dado por $\alpha\mathbf{1}$, para $\alpha \in \mathbb{R}$, pero la solución constante no sirve para nuestro propósito de dibujar el grafo.

Para solucionar este problema, vamos a fijar los vértices de un ciclo antes de resolver para las posiciones de los vértices libres. De este modo, la ecuación (9) se convierte en:

$$\begin{pmatrix} L_{fijo} & M^T \\ M & L_{libre} \end{pmatrix} \begin{pmatrix} \vec{x}_{fijo} \\ \vec{x}_{libre} \end{pmatrix} = \begin{pmatrix} E \\ 0 \end{pmatrix},$$

donde E no es relevante, pues representa la energía de los vértices fijos. Resolviendo, se obtiene:

$$M\vec{x}_{fijo} + L_{libre}\vec{x}_{libre} = 0 \iff \vec{x}_{libre} = -L_{libre}^{-1}M\vec{x}_{fijo}. \quad (10)$$

Sabemos que L_{libre} es invertible, pues es una submatriz principal del laplaciano, por lo que es una matriz simétrica y diagonalmente dominante (SDD) [6], es decir, para cada renglón se cumple que el valor absoluto de la entrada diagonal es mayor a la suma de los valores absolutos del resto de las entradas, lo que implica que su determinante es distinto de cero, por lo que el sistema en (10) siempre tiene solución.

Como se ha dicho antes, un buen dibujo debería buscar disminuir los cruces de aristas, evitándolos por completo cuando sea posible, lo que nos lleva a la idea de planaridad.

Definición 5.1. Un grafo se dice aplanable si puede ser inmerso en el plano sin que haya cruces de aristas; se le denomina grafo plano una vez dibujado sin cruces.

Definición 5.2. En un grafo plano, decimos que los vértices de un ciclo C forman una cara si todas las aristas del grafo y todos los vértices de $V \setminus C$ están completamente dentro o completamente fuera de la región que C delimita en el plano.

En general, puede haber más de un dibujo plano para un grafo. Por ejemplo, en la figura 10, los vértices $\{1, 2, 3, 7, 8, 5\}$ forman una cara en el dibujo de la izquierda, pero no en el de la derecha. Notemos algunas

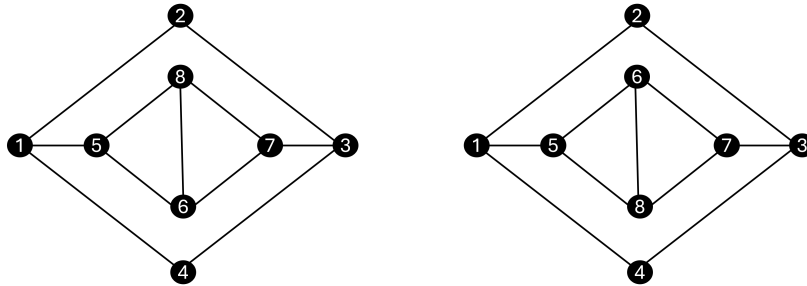


Figura 10. Distintos dibujos planos del mismo grafo.

características del grafo de la figura, por ejemplo, si uno removiera los vértices 5 y 7, el grafo se separaría en dos piezas (componentes conexas); el cuadrado exterior y una arista en medio. En este caso, se dice que $\{5, 7\}$ es un *conjunto de corte*.

Definición 5.3. Un grafo es 3-conexo si cualesquiera 2 vértices **no** forman un conjunto de corte, es decir, si se puede eliminar cualquier par de vértices sin desconectar el grafo.

Recordemos que en un grafo conexo siempre existe una trayectoria entre cualquier par de vértices. Es de esperarse que grafos mejor conectados tengan rutas alternativas. Decimos que dos trayectorias son

internamente disjuntas si solo coinciden en sus extremos, i.e. los vértices al principio y al final. Regresando al ejemplo de la figura 10, tenemos que 1,2,3 y 1,5,6,7,3 son trayectorias internamente disjuntas entre los vértices 1 y 3.

Teorema 5.4. *Sea G un grafo conexo y sean $u, v \in V$ dos vértices no adyacentes. Entonces, el máximo número de trayectorias internamente disjuntas entre u y v es igual al tamaño del mínimo conjunto de corte que los desconecta.*

Esta es una versión del famoso *teorema de Menger* [7], que nos permite trabajar con una definición alternativa. Un grafo es 3 conexo si entre cualquier par de vértices no adyacentes existen al menos 3 trayectorias internamente disjuntas.

Los grafos aplanables y 3-conexos son de nuestro interés debido a que tienen un único dibujo plano, en el sentido de que si los vértices de un ciclo forman una cara en un dibujo plano, también serán una cara en cualquier otro. Más aún, Tutte [11] demostró el siguiente resultado.

Teorema 5.5. *Sea G un grafo aplanable y 3-conexo. Entonces, si se fijan los vértices de una de sus caras en aquellos de un polígono convexo y se dejan a las aristas comportarse como resortes, cada vértice interior termina en el centro de masa (baricentro) de sus vecinos y cada cara resulta ser un polígono convexo. Además, el dibujo plano obtenido es único.*

En la práctica, el llamado dibujo de Tutte se obtiene de resolver el sistema de ecuaciones descrito en (10). Si bien este sistema tiene solución para cualquier grafo conexo, solo se obtendrá un dibujo sin cruces de aristas si el grafo en cuestión es aplanable, mientras que la unicidad del dibujo obtenido depende de la 3-conexidad.

En particular, cualquier grafo obtenido a partir de un poliedro convexo cumple con estas características, por ejemplo, los sólidos platónicos, como el icosaedro en la figura 11.

Cabe mencionar que es importante hacer una buena elección de la cara que se fija como externa. Por ejemplo, los siguientes dos son dibujos planos de un grafo propuesto por el mismo Tutte, pero en el de la izquierda, los vértices interiores terminan muy juntos entre sí.

Finalmente, para que aplique el teorema de Tutte y podamos garantizar que este método obtiene un dibujo plano, es necesario que el ciclo que se fija en la cara exterior sea, en efecto, una cara, como se aprecia en la figura 13, donde se muestran dos dibujos de un clásico balón de fútbol (icosaedro truncado). A la izquierda, podemos observar un dibujo de Tutte adecuado, donde se toma como cara externa a uno de los hexágonos del poliedro, mientras que a la derecha se fijan en el exterior a los vértices del ciclo dibujado en azul.

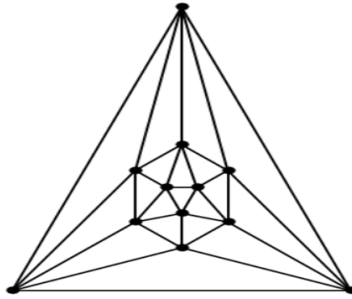


Figura 11. El dibujo de Tutte del Icosaedro.

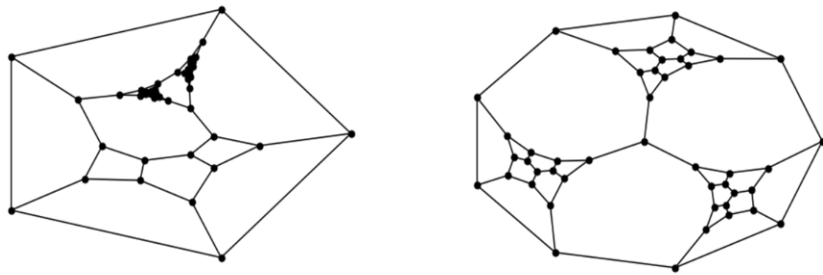


Figura 12. El grafo de Tutte dibujado con distintas caras externas.

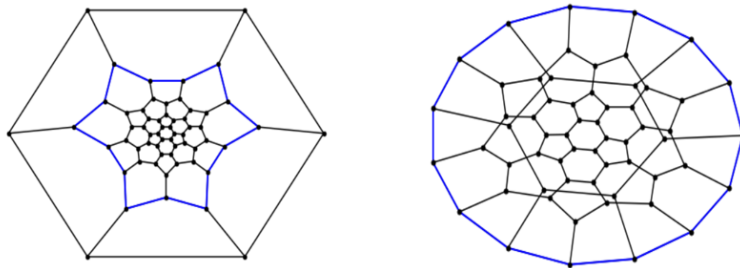


Figura 13. Balón de fútbol con distintas caras externas.

Como podemos darnos cuenta, el segundo dibujo presenta cruces de aristas. Esto se debe a que este ciclo en realidad no forma una verdadera cara del grafo, en el sentido de la definición 5.1, pues existen vértices tanto en el interior como en el exterior de la región que delimita.

6. Conclusiones

Hemos visto como la representación matricial de los grafos permite aprovechar las herramientas del álgebra lineal para analizar las propiedades de una red. En particular, la difusión de energía en un sistema, estudiada mediante el operador laplaciano, nos da información acerca de lo bien conectado que está un grafo y cómo fluye la información en él, permitiendo identificar comunidades.

Por otro lado, comparar un grafo con un sistema físico permite implementar ideas bien estudiadas en otras disciplinas para describirlo desde nuevas perspectivas y recuperar información relevante sobre su estructura. Además, es posible obtener métodos para la inmersión del grafo en un espacio euclidiano por medio de la minimización de cierta energía, obteniendo dibujos útiles e interpretables.

Estas son solo algunas de las aplicaciones de una poderosa teoría que busca cambiar la manera en que se estudian los grafos y, en consecuencia, un sinnúmero de fenómenos son modelados por ellos en áreas como la física, probabilidad, combinatoria y ciencias de la computación.

Agradecimientos

Quiero agradecer sinceramente a Ernesto Pérez Chavela por alentarme a escribir este artículo, así como a Pablo Castañeda, quien organiza los coloquios de matemáticas del ITAM donde estudiantes como yo tienen la oportunidad de exponer sobre los temas que les apasionan.

De igual manera, me siento profundamente agradecido con las doctoras Gloria López Chávez y Ana Paulina Figueroa, quienes despertaron mi interés por los grafos y me refirieron a buenos libros y artículos, además del apoyo y enseñanza de su parte.

Por último, agradezco a Rubén Martínez Avendaño su asesoramiento en algunos temas que se tratan en el artículo, y sus valiosos consejos para su redacción.

Bibliografía

- [1] F. Chung, *Four proofs for the Cheeger inequality and graph partitioning algorithms*, vol. II, International Congress of Chinese Mathematicians (ICCM), 2007.
- [2] F. R. K. Chung, *Spectral graph theory*, CBMS Regional Conference Series in Mathematics, American Mathematical Society, Providence, Rhode Island, 1997, <https://doi.org/10.1090/cbms/092>.
- [3] R. Courant y D. Hilbert, *Methods of mathematical physics*, paperback.^a ed., vol. 1, Wiley-VCH, New York, 1989, <https://doi.org/10.1002/9783527617210>.

- [4] P. Erdős y A. Rényi, «On random graphs», *Publicationes Mathematicae Debrecen*, vol. 6, 1959, 290–297.
- [5] S. J. Florkowski, «Spectral graph theory of the hypercube», tesis de maestría, Naval Postgraduate School, Monterey, California, USA, 2008.
- [6] R. A. Horn y C. R. Johnson, *Matrix analysis*, 2.^a ed., Cambridge University Press, New York, 2013, <https://doi.org/10.1017/CBO9781139020411>.
- [7] D. Jungnickel, *Graphs, networks and algorithms*, 4.^a ed., Algorithms and Computation in Mathematics, Springer, London, 2013, <https://doi.org/10.1007/978-3-642-32278-5>.
- [8] D. A. Levin, Y. Peres y E. L. Wilmer, *Markov chains and mixing times*, 2.^a ed., American Mathematical Society, Providence, Rhode Island, 2017.
- [9] D. A. Spielman, «Spectral graph theory and its applications», Lecture Notes, Lecture 2. Yale University, <https://www.cs.yale.edu/homes/spielman/561/2004/>, 2004.
- [10] ———, «Spectral and algebraic graph theory», 2025, Preprint, <http://cs-www.cs.yale.edu/homes/spielman/sagt>.
- [11] W. T. Tutte, «How to draw a graph?», *Proceedings of the London Mathematical Society*, vol. s3-13, núm. 1, 1963, 743–768, <https://doi.org/10.1112/plms/s3-13.1.743>.
- [12] U. von Luxburg, «A tutorial on spectral clustering», *Statistics and Computing*, vol. 17, núm. 4, 2007, 395–416, <https://doi.org/10.1007/s11222-007-9033-z>.
- [13] D. J. Watts y S. H. Strogatz, «Collective dynamics of ‘small-world’ networks», *Nature*, vol. 393, núm. 6684, 1998, 440–442, <https://doi.org/10.1038/30918>.