

Universidad  
Rey Juan Carlos

Escuela Técnica Superior  
de Ingeniería Informática

Grado en Matemáticas

Curso 2023-2024

Trabajo Fin de Grado

**ESTUDIO DE LA CENTRALIDAD EN  
HIPERGRAFOS A TRAVÉS DE SUS GRAFOS  
LINEALES ASOCIADOS**

**Autor: Sergio Hernández Sandoval**

**Tutor: Regino Criado Herrero**

©2024 <Sergio Hernández Sandoval>  
Algunos derechos reservados  
Este documento se distribuye bajo la licencia “Atribución-CompartirIgual 4.0 Internacional”  
de Creative Commons, disponible en:  
<https://creativecommons.org/licenses/by-sa/4.0/deed.es>

# Agradecimientos

Deseo expresar mi profundo agradecimiento a mi tutor del TFG, Regino Criado Herrero, por su valiosa guía y apoyo a lo largo de este trabajo. Su compromiso y disposición constante para brindarme la orientación necesaria han sido fundamentales. Siempre estuvo disponible para responder a mis preguntas y proporcionar valiosas sugerencias que han enriquecido significativamente mi investigación.



# Resumen

Este Trabajo de Fin de Grado aborda el estudio de la centralidad en hipergrafos con el objetivo de poder determinar la influencia de los elementos y, consecuentemente, identificar aquellos de mayor relevancia en redes complejas caracterizadas por relaciones que involucran dos o más elementos.

Tras una primera y segunda sección donde se introduce el trabajo y se plantean los objetivos, en la tercera sección se proporcionará al lector una introducción a los conceptos fundamentales sobre grafos e hipergrafos que son esenciales para la comprensión de este trabajo. La cuarta sección se centra en un análisis preliminar de la centralidad en grafos, con un enfoque específico en la centralidad de vector propio, para posteriormente llevar a cabo la extensión a la centralidad en hipergrafos. Se destaca la importancia crucial del Teorema de Perron-Frobenius, ya que resulta esencial al examinar la centralidad de los grafos lineales asociados a los hipergrafos. Esto, a su vez, sienta las bases para la extensión al final de la sección para el cálculo de la centralidad en hipergrafos. La quinta sección se dedicará a la resolución de una aplicación práctica tomando como base un problema contextualizado en la vida cotidiana. Los resultados obtenidos revelan la importancia crucial que tiene el Teorema de Perron-Frobenius, siendo clave para el cálculo de la centralidad en los grafos lineales asociados a los hipergrafos.

## Palabras clave:

- Grafo
- Grafo lineal
- Hipergrafo
- Centralidad
- Vector propio
- Teorema de Perron-Frobenius



# Índice de contenidos

Índice de tablas	IX
Índice de figuras	XI
<b>1. Introducción</b>	<b>1</b>
<b>2. Objetivos</b>	<b>3</b>
<b>3. Fundamentos teóricos</b>	<b>5</b>
3.1. Teoría de Grafos . . . . .	5
3.2. Teoría de Hipergrafos . . . . .	9
<b>4. Contenidos principales</b>	<b>15</b>
4.1. Centralidad en grafos . . . . .	15
4.1.1. Centralidad de vector propio . . . . .	17
4.2. Centralidad en hipergrafos . . . . .	30
<b>5. Aplicación</b>	<b>39</b>
5.1. Trabajador más influyente . . . . .	39
<b>6. Conclusiones y trabajos futuros</b>	<b>57</b>
6.1. Conclusiones . . . . .	57
6.2. Trabajos futuros . . . . .	57
<b>Bibliografía y referencias</b>	<b>58</b>
<b>Anexos</b>	<b>61</b>
<b>A. Tabla de tiempos</b>	<b>63</b>



# Índice de tablas

5.1. Registro de trabajadores. . . . .	40
5.2. Registro de conferencias. . . . .	40
5.3. Resultados obtenidos del problema del trabajador más influyente.	54
5.4. Ranking de influencia tras los resultados. . . . .	54



# Índice de figuras

3.1. Ilustración de un grafo simple no dirigido (izquierda) y un grafo simple dirigido (derecha). . . . .	7
3.2. Grafo original (izquierda) y su grafo lineal (derecha). . . . .	9
3.3. Ilustración de un ejemplo de hipergrafo. . . . .	10
3.4. Hipergrafo original (izquierda) y su grafo lineal asociado (derecha). . . . .	13
4.1. Esquema de los pasos a seguir para el cálculo de la centralidad en hipergrafos. . . . .	31
5.1. Representación gráfica del hipergrafo creado a partir de la información del problema. Los nodos representan los trabajadores y las hiperaristas las conferencias. . . . .	41
5.2. Esquema extraído a partir del hipergrafo graficado del problema en el que se muestran los colores de las hiperaristas que inciden en cada uno de los nodos. . . . .	42
5.3. Representación gráfica del grafo lineal creado a partir del hipergrafo. . . . .	43
A.1. Parte 1 de la tabla de tiempos. . . . .	63
A.2. Parte 2 de la tabla de tiempos. . . . .	64
A.3. Parte 3 de la tabla de tiempos. . . . .	65
A.4. Parte 4 de la tabla de tiempos. . . . .	66



# 1

## Introducción

En el ámbito de la rama de las matemáticas encargada del estudio de las redes complejas, los llamados hipergrafos han captado de forma inmediata la atención de los investigadores a causa de su capacidad para representar relaciones de alta cardinalidad entre elementos, las cuales serían difíciles de modelar a través de simples grafos. A diferencia de estos que solo permiten uniones entre dos elementos, los hipergrafos involucran a un conjunto de elementos en una sola conexión. Esto hace que adquieran una importancia vital en áreas donde hay relaciones mucho más complejas.

Debido a ello, en este trabajo nos enfocaremos en el concepto de hipergrafo y, concretamente, en el estudio de la centralidad en ellos. La centralidad es una propiedad que asocia una importancia cuantitativa a cada nodo de la red, permitiendo distinguir cuales de ellos presentan una mayor trascendencia e influencia en determinados contextos. Un papel fundamental en este panorama lo juegan los grafos lineales, a partir de los cuales, construidos en base a los hipergrafos, y mediante la aplicación de un teorema crucial, conseguiremos lograr nuestro objetivo. Para ello, abordaremos todos los conceptos necesarios para comprender en profundidad este importante aspecto perteneciente a la Teoría de Hipergrafos.

El trabajo se llevará a cabo a través de la revisión de libros y documentos existentes sobre los grafos, los hipergrafos y la centralidad en ambos, revisando y tratando todos los conceptos teóricos que permitan la comprensión del mismo. Este enfoque nos permitirá abordar de manera efectiva la centralidad en hipergrafos y su importancia en la resolución de problemas tanto en la vida cotidiana como en una inmensa variedad de campos como en inteligencia artificial, bioinformática o teoría de redes.



# 2

## Objetivos

El objetivo de este Trabajo de Fin de Grado es estudiar y analizar la centralidad en hipergrafos a partir del análisis y estudio de la centralidad en sus correspondientes grafos lineales. La centralidad es una medida que determina la importancia e influencia de los diferentes nodos en una red, por lo que se convierte en una propiedad clave para el estudio y análisis de numerosos problemas en los que se requiere identificar los elementos cruciales. El hecho de que la centralidad solo se pueda calcular con relaciones binarias, parece indicar que imposibilita su cálculo en hipergrafos. No obstante, al poseer las mismas características los hipergrafos y sus grafos lineales, mediante el estudio de la centralidad del grafo lineal asociado a un hipergrafo, va a ser posible estudiar y determinar la centralidad de dicho hipergrafo.

Se comenzará definiendo un marco histórico, teórico y contextual de los grafos e hipergrafos para situar al lector y sentar los conocimientos necesarios para la comprensión del tema central sobre el que gira el documento. Se continuará con el estudio de la centralidad en grafos, tratando el Teorema de Perron-Frobenius, resultado que nos dará las herramientas imprescindibles para posteriormente llevar a cabo la extensión al estudio de la centralidad en hipergrafos.

Una posible aplicación práctica del objeto de estudio que se atenderá al final del documento es estudiar la importancia e influencia que tienen los trabajadores del departamento de una empresa en una semana concreta tras guardar todas las conferencias realizadas dicha semana junto con los participantes. El objetivo sería determinar el trabajador más influyente de la empresa durante esa semana. Este escenario se modelaría con un hipergrafo, y habría que estudiar su centralidad.



# 3

## Fundamentos teóricos

### 3.1. Teoría de Grafos

El origen de los grafos se remonta al siglo XVIII, más concretamente al año 1736, cuando Leonhard Euler, un matemático suizo, consiguió encontrar solución al famoso problema matemático de los siete puentes de la ciudad de Königsberg [1]. Este consistía en encontrar un camino que permitiera pasar por los siete puentes situados sobre el río Pregel de forma que se cruzase solo una vez por cada uno de ellos.

Para resolverlo, Euler ideó una forma de representar dicho problema mediante un diagrama en el que sustituía cada zona de la ciudad con un punto y cada puente con una línea entre pares de los puntos anteriores. De esta forma, consiguió reducirlo a decidir si existía un camino que comenzase en un punto y que pasase por cada línea una sola vez. Dicho diagrama creado por Euler fue la primera noción de lo que actualmente conocemos como *grafo*, y con ello pudo demostrar que era imposible recorrer todos los puentes sin repetir ninguno sin necesidad de comprobar todos los caminos posibles.

La resolución de este problema supuso el primer resultado de la conocida *Teoría de Grafos*, que hace referencia al área de las matemáticas que se encarga de estudiar las propiedades de los grafos.

En el año 1847, el físico prusiano Gustav Kirchoff fue la primera persona en aplicar la Teoría de Grafos a un problema de la vida real en el ámbito de la ingeniería. Llevó a cabo un análisis de redes eléctricas publicando las leyes que

reciben su propio nombre, las *Leyes de Kirchoff*, las cuales rigen el comportamiento de las corrientes eléctricas en un circuito cerrado.

Unos años más tarde, en 1852, el matemático sudafricano Francis Guthrie presentó el famoso problema de los cuatro colores, el cual afirma que es posible colorear cualquier mapa de países usando solo cuatro colores, de modo que dos países vecinos nunca compartan el mismo color. Este problema consiguió resolverse un siglo después de la mano de Kenneth Appel y Wolfgang Haken en el año 1976. Su resolución marcó un antes y un después en el desarrollo de la Teoría de Grafos, ya que los matemáticos tuvieron que introducir y definir términos y conceptos teóricos esenciales en este campo.

En el año 1857, el matemático Arthur Cayley aplicó la Teoría de Grafos en el ámbito de la química para estudiar y resolver el problema de enumeración de los isómeros, que son compuestos químicos con la misma fórmula molecular pero con distinta forma estructural y propiedades. Para ello, representó cada compuesto químico como un grafo en el cual los átomos eran los puntos y los enlaces químicos las líneas que unen pares de dichos puntos.

No fue hasta el año 1936 cuando se publicó el primer libro de la Teoría de Grafos. Fue escrito por el matemático húngaro Dénes Kónig y titulado *Teoría de Grafos finitos e infinitos* [2].

Desde mediados del siglo XX y a lo largo de todo el siglo XXI hasta la fecha actual, la Teoría de Grafos ha experimentado un significativo avance en todos los aspectos. Se han realizado numerosos descubrimientos en el estudio y aplicación de esta rama de las matemáticas, las cuales pueden aplicarse en prácticamente cualquier ámbito de la vida real. Los grafos se han convertido en una herramienta esencial que ayudan a resolver problemas reales como analizar redes sociales, entender las redes de información, construir modelos de redes de ordenadores, diagnosticar enfermedades neurodegenerativas en neurociencia, optimizar rutas en logística o diseñar circuitos electrónicos. A medida que avanzamos hacia el futuro, se espera que los estudios en grafos sigan adquiriendo mayor importancia, proporcionando nuevas soluciones a problemas complejos y sentando bases sólidas para desarrollar algoritmos más eficientes y tecnologías innovadoras.

Una vez situado el contexto, comenzamos dando la definición básica de grafo:

**Definición 3.1.1.** Sea  $V = \{1, 2, \dots, N\}$  un conjunto de puntos llamados *vértices* o *nodos*, y  $E = \{\{u, v\} \mid u, v \in V \wedge u \neq v\}$  un conjunto de pares no ordenados de vértices pertenecientes a  $V$  llamados *aristas*. Entonces el par  $G = (V, E)$  se denomina *grafo*.

Consideraremos que las aristas no pueden unir un vértice a sí mismo, lo que se define como lazo, y además cada arista solo puede relacionar a dos vértices, llamados cada uno de ellos *extremos* de la arista.

Si el grafo presenta un número finito de vértices y aristas, será un *grafo finito*, mientras que si el número de vértices y aristas es infinito, será un *grafo infinito*. No obstante, en este documento siempre se va a trabajar con grafos finitos.

Hablaremos de un *grafo no dirigido* si las aristas del grafo no tienen una orientación definida de forma que se considera sentido bidireccional, es decir, si la arista  $\{u, v\}$  pertenece a  $E$ , entonces la arista  $\{v, u\}$  también. Sin embargo, estaremos hablando de un *grafo dirigido* o *digrafo* si las aristas presentan una orientación definida de forma que cada flecha se considera como sentido único, es decir, la existencia de la arista  $\{u, v\}$  no implica la existencia de la arista  $\{v, u\}$ .

Si dado un grafo, el conjunto de aristas  $E$  puede contener elementos repetidos, es decir, la arista  $\{u, v\} \in E$  puede aparecer varias veces, entonces estaremos hablando de un *multigrafo*. En caso contrario, hablaremos de *grafo simple*.

En La Figura 3.1 se puede apreciar la representación gráfica de dos grafos simples. La ilustración de la izquierda, como las aristas no tienen orientación ni se repite ninguna, representa un grafo simple no dirigido. La ilustración de la derecha, como las aristas tienen orientación y no se repite ninguna, representa un grafo simple dirigido.

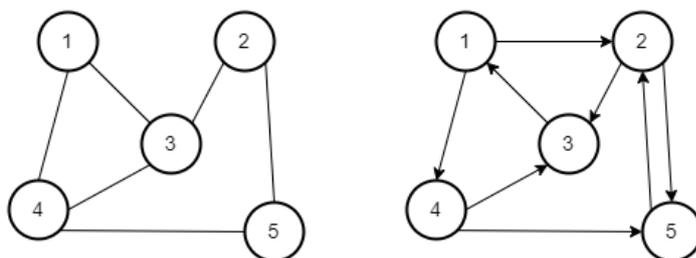


Figura 3.1: Ilustración de un grafo simple no dirigido (izquierda) y un grafo simple dirigido (derecha).

Dados dos nodos  $u, v \in V$  de un grafo no dirigido, decimos que estos dos nodos son adyacentes si existe una arista que une a ambos, es decir, si existe  $\{u, v\} \in E$ . Por otro lado, dados dos nodos  $u, v \in V$  de un grafo dirigido, decimos que el nodo  $v$  es adyacente al nodo  $u$  si existe una arista desde  $u$  hasta  $v$ , es decir, si existe  $\{u, v\} \in E$ .

Una de las principales formas de representar un grafo  $G$ , ya sea simple o multigrafo, es mediante una matriz de adyacencia, la cual nos permite deducir todas las propiedades que posee el grafo. La *matriz de adyacencia* de un grafo  $G = (V, E)$  de  $n$  vértices es una matriz  $n \times n$  en la cual en la diagonal habrá ceros, ya que las aristas no unen un nodo consigo mismo al considerar la no existencia de lazos, y en cada posición  $a_{ij}$  se colocará el número total de aristas que unen al nodo  $i$  con el  $j$ . Si el grafo es no dirigido, la matriz será simétrica puesto que la

existencia de la arista  $\{u, v\}$  implica la existencia de la arista  $\{v, u\}$ . Si el grafo es simple, la matriz solo estará formada por ceros y unos ya que un nodo solo puede unirse a otro mediante una única arista, mientras que si es un multigrafo pueden tomar cualquier valor finito.

Tomando como referencia el grafo no dirigido de La figura 3.1, su correspondiente matriz de adyacencia sería

$$A = \begin{pmatrix} 0 & 0 & 1 & 1 & 0 \\ 0 & 0 & 1 & 0 & 1 \\ 1 & 1 & 0 & 1 & 0 \\ 1 & 0 & 1 & 0 & 1 \\ 0 & 1 & 0 & 1 & 0 \end{pmatrix}$$

En un grafo no dirigido, se define el *grado* de un nodo como el número total de aristas que salen del nodo. En un grafo dirigido, se define el *grado de entrada* de un nodo como el número total de aristas que entran al nodo, mientras que el *grado de salida* de un nodo como el número total de aristas que salen del nodo. Si el grado de entrada de un nodo es cero y el de salida positivo, se denomina *nodo fuente*, mientras que si el grado de salida de un nodo es cero y el de entrada positivo, se llama *nodo sumidero*.

Dado un grafo  $G = (V, E)$ , un *camino* es una sucesión de aristas de forma que el extremo derecho de una arista coincida con el izquierdo de la consecutiva. Un posible camino entre los nodos  $v_j$  y  $v_m$  sería  $\{ \{v_j, v_k\}, \{v_k, v_l\}, \{v_l, \dots\}, \dots, \{\dots, v_r\}, \{v_r, v_m\} \}$ . El número de aristas que conforman el camino se denomina *longitud*. Si el camino empieza y termina en el mismo vértice, estaríamos hablando de un *camino cerrado*.

Dados dos vértices de un grafo, la *distancia* entre ambos nodos es la longitud del camino más corto entre ambos nodos.

Un grafo se dice que es *conexo* si existe siempre un camino entre dos vértices distintos cualesquiera del grafo.

A continuación, presentamos el concepto de grafo lineal:

**Definición 3.1.2.** Sea  $G = (V, E)$  un grafo, se define su grafo lineal  $L(G)$  como aquel grafo que cumple las siguientes condiciones:

- Cada nodo de  $L(G)$  se corresponde con una arista de  $G$ .
- Dos nodos en  $L(G)$  son adyacentes si y solo si ambas aristas en  $G$  comparten al menos un nodo.

Dicho de otra forma, los grafos lineales corresponden con el grafo de intersección de las aristas del grafo original, es decir, permiten representar las relaciones

entre las diferentes aristas. Por ello, su uso se expande a numerosas aplicaciones en las cuales las aristas adquieren una mayor atención como puede ser la planificación y optimización de rutas y comunicaciones, además de la identificación de propiedades del grafo original.

En La Figura 3.2 se puede observar un grafo original a la izquierda y su grafo lineal asociado a la derecha.

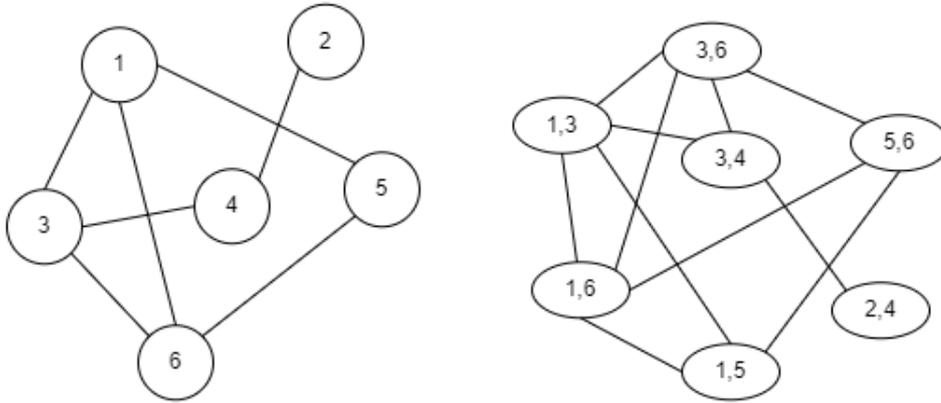


Figura 3.2: Grafo original (izquierda) y su grafo lineal (derecha).

## 3.2. Teoría de Hipergrafos

Desde sus inicios y hasta el día de hoy, la Teoría de Grafos ha demostrado ser una herramienta con un valor incalculable en todo tipo de campos y con numerosas aplicaciones para la resolución de problemas en la vida real. Sin embargo, existen una inmensidad de escenarios para los cuales los grafos tradicionales se quedan cortos en su capacidad para representar esas relaciones de manera efectiva. Es en este contexto donde surge el concepto de hipergrafo. Los hipergrafos son una extensión de los grafos que nos permiten representar y analizar relaciones de mayor cardinalidad que involucran conjuntos de elementos. Esta nueva forma de representar relaciones más complejas se desarrolló incrementalmente a finales del siglo XX y principios del siglo XXI con una gran cantidad de resultados obtenidos tras el esfuerzo y trabajo de numerosos investigadores. No obstante, no fue hasta 1960, de la mano del matemático francés Claude Berge, cuando este campo de las matemáticas en auge se convirtió en una teoría independiente, la Teoría de Hipergrafos.

Introducimos la definición básica de hipergrafo:

**Definición 3.2.1.** Sea  $X = \{1, 2, \dots, N\}$  un conjunto finito de nodos y sea  $\varepsilon = \{h_1, h_2, \dots, h_n\}$  una familia de subconjuntos no vacíos formados con elementos de

$X$ , llamados *hiperaristas* o *hiperbordes*. Entonces el par  $\mathcal{H} = (X, \mathcal{E})$  se denomina *hipergrafo*.

A partir de este concepto podemos presentar las características más importantes de los hipergrafos, algunas de las cuales son análogas a la de los grafos mientras que otras son exclusivas, y nuevas definiciones aún no introducidas, tomando como base [3] y [4].

Una hiperarista  $h \in \mathcal{E}$  que solo relaciona a un único nodo, es decir,  $|h| = 1$ , se denomina *bucle*. No obstante, debido a que los bucles no aportan información adicional en términos de relaciones y conexiones entre nodos, solo se trabajará con hipergrafos sin bucles, es decir, con hiperaristas que incluyan al menos dos nodos. De esta forma, cada hiperarista perteneciente a  $\mathcal{E}$  contendrá un número de vértices comprendido entre dos y el cardinal de  $X$ , ambos límites incluidos. La *dimensión* de una hiperarista  $h \in \mathcal{E}$  siempre será  $|h| - 1$ .

Un hipergrafo sin bucles y sin hiperaristas repetidas se denomina *hipergrafo simple*. Un hipergrafo que es simple y que además presenta únicamente pares de hiperaristas que comparten como máximo un nodo se denomina *hipergrafo lineal*.

Como se aprecia en La Figura 3.3, el hipergrafo de la ilustración presenta cinco vértices y tres hiperaristas. La hiperarista  $h_1$  conecta los vértices 1, 2 y 3, la hiperarista  $h_2$  conecta los vértices 3, 4 y 5, y por último, la hiperarista  $h_3$  conecta los vértices 2 y 5.

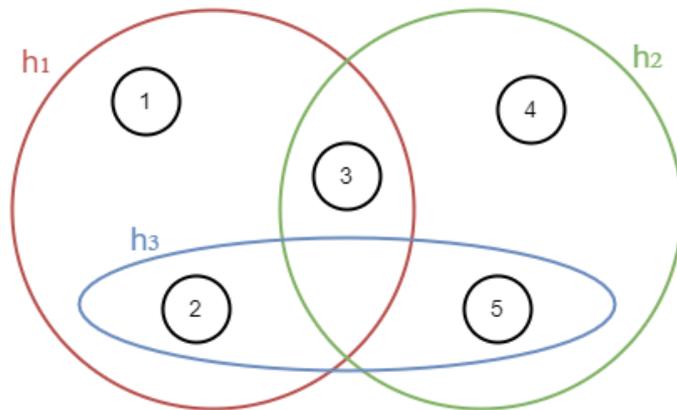


Figura 3.3: Ilustración de un ejemplo de hipergrafo.

El *orden* de un hipergrafo corresponde con el número de elementos de  $X$ , es decir, el número total de nodos que tiene el hipergrafo, mientras que el *tamaño* de un hipergrafo corresponde con el número de elementos de  $\mathcal{E}$ , es decir, el número total de hiperaristas que tiene el hipergrafo.

Dos hiperaristas  $h_i$  y  $h_j$  se dice que son *incidentes* si tienen al menos un nodo

en común, es decir,  $h_i \cap h_j \neq \emptyset$ .

Dos nodos son *adyacentes* si existe al menos una hiperarista que contiene a ambos.

Dado un nodo  $i$ , se define la *estrella* de  $i$ ,  $E(i)$ , como el conjunto de todas las hiperaristas que contienen al nodo  $i$ , es decir,  $E(i) = \{h \mid i \in h\}$ . Si dados dos nodos  $i$  y  $j$ , sus respectivas estrellas son iguales,  $E(i) = E(j)$ , se dice que los nodos son *iguales*, puesto que participan en las mismas hiperaristas.

El *grado* de un nodo se define como el número total de hiperaristas en las que participa, siendo igual a la cardinalidad de su estrella, denotándose  $gr(i) = |E(i)|$ . Si todos los nodos de un hipergrafo tienen el mismo grado  $k$ , entonces podemos hablar de *hipergrafo  $k$ -regular* o simplemente *hipergrafo regular*. El máximo grado de un hipergrafo  $\mathcal{H}$  corresponde con el mayor grado de sus nodos, y se denota como  $\Delta(\mathcal{H})$ .

El *rango* de un hipergrafo  $\mathcal{H}$  se define como la máxima cardinalidad de todas sus hiperaristas, mientras que el *co-rango* como la mínima cardinalidad de sus hiperaristas. Formalmente,  $rg(\mathcal{H}) = \max_{h \in \mathcal{H}} \{|h|\}$  y  $corg(\mathcal{H}) = \min_{h \in \mathcal{H}} \{|h|\}$ . Si el rango y co-rango toman el mismo valor  $k$  en el hipergrafo, entonces podemos hablar de *hipergrafo  $k$ -uniforme* o simplemente *hipergrafo uniforme*.

Sean  $\mathcal{H} = (X, \varepsilon)$  y  $\mathcal{H}' = (X', \varepsilon')$  dos hipergrafos, se dice que  $\mathcal{H}'$  es un *subhipergrafo* de  $\mathcal{H}$  si  $X'$  está contenido en  $X$  y si para toda hiperarista  $h'$  perteneciente a  $\varepsilon'$ , solo existe una única hiperarista perteneciente a  $\varepsilon$  tal que  $h' \subset h$ . Es decir, si los nodos de  $X'$  también se hallan en  $X$  y si toda hiperarista de  $\varepsilon'$  solo está contenida en una hiperarista de  $\varepsilon$ .

Si tenemos el hipergrafo  $\mathcal{H} = (X, \varepsilon)$ , dado un conjunto  $Y$  tal que  $Y \subset X$ , el *subhipergrafo inducido por  $H$  en  $Y$*  es  $\mathcal{H}_Y = (Y, \varepsilon_Y)$ , donde  $\varepsilon_Y$  es la familia de hiperaristas que conectan únicamente nodos de  $Y$ . Dicho de otra forma,  $\varepsilon_Y$  tiene las mismas hiperaristas que  $\varepsilon$ , pero quitando aquellas que contienen por lo menos a un nodo que está en  $X$  y no en  $Y$ .

Dado un hipergrafo  $\mathcal{H}$ , se define un *camino* desde el nodo  $u$  hasta el nodo  $v$  como una secuencia en la que se alternan nodos e hiperaristas de forma que  $u = n_1, h_1, n_2, h_2, \dots, n_s, h_s, n_{s+1} = v$ , donde dos nodos consecutivos deben ser distintos y pertenecer a la misma hiperarista. Si dicho camino no repite nodos pero pudiendo repetir aristas, se llama *recorrido*, mientras que si dicho camino no tiene elementos repetidos, es decir, todas los nodos e hiperaristas por los que pasa son distintos, se denomina *ruta*.

Un hipergrafo se dice que es *conexo* si existe siempre un camino entre cualquier par de nodos diferentes.

Dado un camino de un hipergrafo, se define *longitud* del camino como el número de hiperaristas que atraviesa. Por ello, la *distancia* entre dos nodos de

un hipergrafo es la longitud mínima del camino que los une. Se denota  $d(i, j)$ , siendo  $i$  y  $j$  dos nodos del hipergrafo. Si entre los nodos  $i$  y  $j$  no existiera un camino en el hipergrafo, la distancia entre ambos sería infinito. El *diámetro* de un hipergrafo es la mayor distancia entre dos nodos cualesquiera.

Una de las formas más sencillas de representar un hipergrafo es mediante su matriz de incidencia. Sea  $\mathcal{H}$  un hipergrafo con  $n$  vértices y  $m$  hiperaristas, la matriz de incidencia será una matriz  $n \times m$ . En cada posición  $a_{nm}$ , se colocará un 1 si la hiperarista  $m$  incide en el nodo  $n$ , mientras que se colocará un 0 si no incide.

Tomando como referencia el hipergrafo de La Figura 3.3, su correspondiente matriz de incidencia sería

$$I(\mathcal{H}) = \begin{pmatrix} 1 & 0 & 0 \\ 1 & 0 & 1 \\ 1 & 1 & 0 \\ 0 & 1 & 0 \\ 0 & 1 & 1 \end{pmatrix}$$

Otra forma de representar los hipergrafos es mediante una matriz de adyacencia. Sea  $\mathcal{H}$  un hipergrafo con  $n$  vértices y un número finito de hiperaristas, la matriz de adyacencia será una matriz  $n \times n$ . En cada posición  $a_{ij}$ , se colocará el número de hiperaristas que comparte el nodo  $i$  con el nodo  $j$ . Será una matriz simétrica con diagonal de ceros.

Tomando como referencia el hipergrafo de La Figura 3.3, su correspondiente matriz de adyacencia sería

$$A(\mathcal{H}) = \begin{pmatrix} 0 & 1 & 1 & 0 & 0 \\ 1 & 0 & 1 & 0 & 1 \\ 1 & 1 & 0 & 1 & 1 \\ 0 & 0 & 1 & 0 & 1 \\ 0 & 1 & 1 & 1 & 0 \end{pmatrix}$$

Otra forma útil de representar un hipergrafo es mediante una matriz de frecuencia de relaciones. Sea  $\mathcal{H}$  un hipergrafo con  $n$  vértices y un número finito de hiperaristas, la matriz de frecuencia de relaciones será una matriz  $n \times n$ . En cada posición  $a_{ij}$ , si  $i = j$  se colocará el grado del nodo  $i$ , es decir, el número total de hiperaristas que inciden en él, mientras que si  $i \neq j$  se colocará el número total de hiperaristas que inciden simultáneamente en los nodos  $i$  y  $j$ . Será una matriz simétrica.

Tomando como referencia el hipergrafo de La Figura 3.4, su correspondiente matriz de frecuencia de relaciones sería

$$F(\mathcal{H}) = \begin{pmatrix} 1 & 0 & 1 & 1 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 1 \\ 1 & 0 & 2 & 2 & 1 & 0 \\ 1 & 0 & 2 & 3 & 2 & 1 \\ 0 & 0 & 1 & 2 & 2 & 1 \\ 0 & 1 & 0 & 1 & 1 & 2 \end{pmatrix}$$

A continuación, presentamos el concepto de grafo lineal asociado a un hipergrafo, muy similar al visto en la subsección anterior con grafos:

**Definición 3.2.2.** *Sea  $\mathcal{H}$  un hipergrafo, se define su grafo lineal  $L(\mathcal{H})$  asociado como aquel que cumple las siguientes condiciones:*

- *Cada nodo de  $L(\mathcal{H})$  se corresponde con una hiperarista de  $\mathcal{H}$ .*
- *Dos nodos en  $L(\mathcal{H})$  son adyacentes si y solo si ambas hiperaristas en  $\mathcal{H}$  comparten al menos un nodo.*

En La Figura 3.4 se puede observar un grafo original a la izquierda y su grafo lineal asociado a la derecha. El hipergrafo original presenta 6 nodos y 4 hiperaristas, por lo que cada nodo del grafo lineal reproduce cada una de las hiperaristas del hipergrafo. En el grafo lineal los nodos  $h_1$  y  $h_2$  son adyacentes porque ambas hiperaristas en el hipergrafo inciden en el nodo 3 y 4, los nodos  $h_1$  y  $h_4$  son adyacentes porque ambas hiperaristas inciden en el nodo 4, los nodos  $h_2$  y  $h_4$  son adyacentes porque ambas hiperaristas inciden en el nodo 4 y 5, y por último, los nodos  $h_3$  y  $h_4$  son adyacentes porque ambas hiperaristas inciden en el nodo 6.

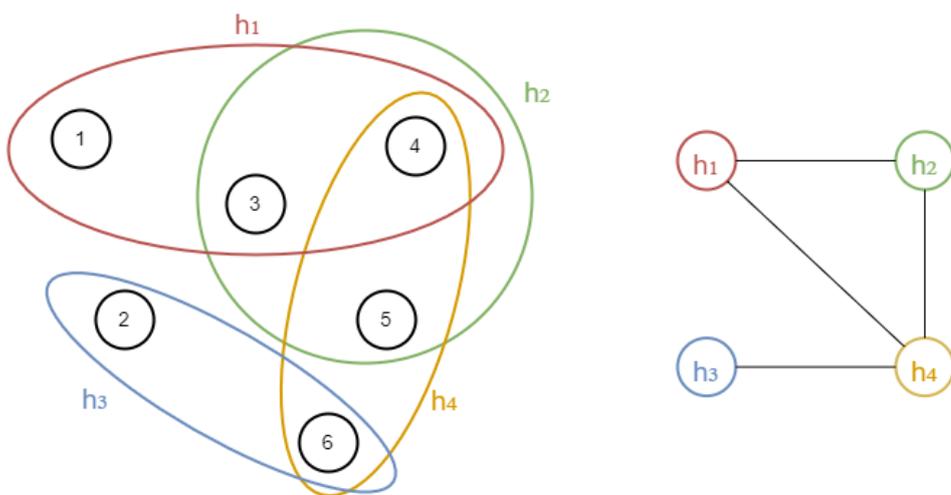


Figura 3.4: Hipergrafo original (izquierda) y su grafo lineal asociado (derecha).



# 4

## Contenidos principales

### 4.1. Centralidad en grafos

Como hemos visto en la sección anterior, los grafos que estamos tratando en este documento están formados por una cantidad finita de nodos. No obstante, como en cualquier ámbito de la vida real en el que tengamos una red conectada, no todos los elementos presentan la misma importancia. En un callejero encontramos calles con más afluencia de gente que en otras, en una empresa hay trabajadores con mayor responsabilidad y presencia, o en las redes sociales siempre hay famosos con millones de seguidores con mayor repercusión sobre el resto. De esta forma, cada nodo en un grafo va a tener su correspondiente importancia, la cual va a depender de la serie de factores que se tengan en cuenta. Esto es lo que se conoce como *medida de centralidad de un grafo*.

Existen numerosas definiciones sobre la centralidad de un grafo en función de los aspectos a tener en cuenta en dicha estructura. Tomando como referencia [5], encontramos tres tipos de centralidades.

Por un lado, tenemos la *centralidad de grado*, la cual asocia la importancia cuantitativa de un nodo a su número total de conexiones. Esto hace que el nodo más importante sea aquel que presente la mayor cantidad de enlaces, mientras que el menos importante será aquel que presente menos enlaces. No obstante, esta centralidad ignora totalmente la calidad de las vecindades, es decir, no tiene en cuenta si los vecinos son a su vez importantes, además de la estructura de la red más allá de un salto. Su uso suele estar relacionado con el estudio de la

propagación de información en una red o de un virus. Supongamos que estamos analizando la propagación de un virus informático a través de una red social que es usada por millones de personas. Cada una de estas personas en la red se corresponde con un nodo, y las amistades entre ellas se identifican con los enlaces. Utilizando esta centralidad de grado, se puede determinar quiénes son los individuos con más amigos, es decir, los nodos con una alta centralidad de grado. Debido a que tienen muchas conexiones, la probabilidad de que entren en contacto con una cantidad mayor de personas aumenta considerablemente. Si uno de estos nodos con un alto grado está infectado con el virus, es más probable que lo transmita a muchas otras personas con las que está conectado. Por lo tanto, estos nodos con alto grado pueden jugar un papel importante en el inicio de la propagación del virus a través de la red, y es vital detectarlos y localizarlos.

Por otro lado, tenemos la *centralidad de intermediación*, la cual asocia la importancia cuantitativa de un nodo a su posición en la red para facilitar la comunicación entre otros nodos, es decir, a su papel como intermediario. A diferencia de la centralidad de grado, esta centralidad no se basa en la cantidad de conexiones sino en cómo un nodo es capaz de actuar como un puente o intermediario para que otros nodos se comuniquen. Supongamos que tenemos delante nuestra un mapa del metro urbano de una ciudad, donde las estaciones son los nodos y los enlaces son los caminos que unen una estación con otra. Algunas estaciones pueden tener muchas conexiones, pero lo que realmente importa es cuántas rutas pasan a través de ellas. Esto es lo que mide la centralidad de intermediación. Si un nodo tiene una alta centralidad de intermediación en la red de metro, significa que muchas rutas pasan por esa estación en concreto, por lo que actúa como un punto crucial para que las personas se trasladen de un punto a otro en la ciudad. Si se produjera alguna avería o incidencia en alguna estación o parte de línea de metro que supusiera algún cierre de estaciones, identificar las estaciones con alta centralidad de intermediación sería vital para frenar de forma efectiva el flujo de personas en la red de metro.

Por último, tenemos la *centralidad de vector propio*, la cual asocia la importancia cuantitativa de un nodo tanto a la cantidad como a la calidad de sus conexiones en la red. Supongamos que estamos analizando una red social que es usada por millones de personas. Cada una de estas personas en la red se corresponde con un nodo, y las amistades entre ellas se identifican con los enlaces. En este contexto, la centralidad de vector propio permite determinar quiénes son los usuarios más influyentes en dicha red. Un nodo tendrá una alta centralidad de vector propio si está conectado a otros nodos que también son a la vez influyentes, es decir, no solo importa la cantidad de amigos que tengas en la red social, sino que también importa la influencia en la red de tus amigos y la de los amigos de tus amigos. De esta manera, se crea un efecto cascada de influencia que se propaga por la red.

La centralidad que nos va a permitir resolver nuestra aplicación práctica descrita en el apartado de objetivos dándonos las herramientas necesarias para posteriormente el cálculo de la centralidad en hipergrafos es la centralidad de vector propio, ya que nos interesa tanto la cantidad de conexiones como la calidad de las mismas. Por ello, en el siguiente apartado nos enfocaremos en su estudio, análisis y método para calcularla. Esta centralidad valora tanto la cantidad como la calidad de las conexiones de un nodo, lo que la hace especialmente relevante en situaciones donde no solo importa cuántas conexiones tienes, sino también quiénes son tus conexiones y cuánta influencia tienen en la totalidad de la red.

### 4.1.1. Centralidad de vector propio

La centralidad de vector propio, como su mismo nombre sugiere, tiene una estrecha y directa relación con los llamados vectores propios o autovectores, que corresponden con conceptos fundamentales de álgebra lineal. La obtención de estos vectores revela información crucial sobre la importancia y la estructura de los nodos en la red. Debido a esto y en este contexto, la centralidad de vector propio se convierte en una herramienta muy valiosa para entender la dinámica y las interconexiones dentro de sistemas complejos.

Antes de adentrarnos en los detalles de esta centralidad, es esencial repasar algunos conceptos clave de álgebra lineal. Estos fundamentos proporcionan el marco teórico necesario para comprender cómo los vectores propios y sus autovalores asociados son cruciales para calcular la centralidad de vector propio.

En álgebra lineal, los vectores propios de un operador lineal son todos aquellos vectores, distintos del vector nulo, que cuando son transformados por ese operador lineal, generan un múltiplo escalar de sí mismos sin cambiar su dirección y proporcionando información sobre dicho operador. Dicho múltiplo escalar se define como autovalor asociado al autovector, y se suele denotar con  $\lambda$ .

Un operador lineal  $F$  es una función que, dados dos vectores  $u$  y  $v$ , y un escalar  $c$ , cumple

- $F(u + v) = F(u) + F(v)$
- $F(c \cdot u) = c \cdot F(u)$

En el contexto de este documento, un claro operador lineal que cumple todas estas condiciones son las matrices cuadradas. Dada una matriz cuadrada  $A$ , dos vectores  $u$  y  $v$ , y un escalar  $c$ , se cumple que

- $A(u + v) = Au + Av$
- $A(c \cdot u) = c \cdot (Au)$

Volviendo a la definición de vector propio, sea  $A$  una matriz cuadrada y  $v$  un vector propio de esta, al aplicar la transformación lineal multiplicando ambos, se obtiene un nuevo vector que conforma una versión escalada del autovector  $v$ , preservando la dirección y orientación de la matriz.

Escribiéndolo formalmente, sea  $\lambda \in \mathbb{R}$  un escalar,  $A$  una matriz cuadrada y  $v$  un autovector de  $A$  no nulo, se tiene que

$$A \cdot v = \lambda \cdot v \quad (4.1)$$

Llamaremos a  $\lambda$  autovalor de la matriz  $A$  y a  $v$  autovector de  $A$  asociado al autovalor  $\lambda$ .

Cabe resaltar que, por todo lo que se acaba de explicar, solo tiene sentido hablar de autovalores y autovectores cuando tratamos matrices cuadradas. Los autovectores siempre están asociados a su respectivo autovalor y viceversa. Cada autovector solo puede tener asociado un único autovalor; no obstante, un autovalor si que puede tener asociados más de un autovector distinto, y es lo que se conoce como multiplicidad de un autovalor, el número de autovectores que tiene asociados un autovalor. El conjunto de todos los autovalores de una matriz  $A$  cuadrada dada se denota con  $\sigma(A)$ . El máximo de estos autovalores pertenecientes a  $\sigma(A)$ , todos ellos en valor absoluto, recibe el nombre de radio espectral de  $A$  y se denota con  $\rho(A)$ .

Una vez introducido y explicado el concepto de autovector en el área de álgebra lineal, estamos preparados para entrar en materia de la centralidad en cuestión empleando como guía [5], [6] y [7].

La centralidad de vector propio consiste en obtener un vector, llamémoslo  $c$ , que cuantifique la importancia de todos los nodos de un grafo. Este vector tendrá tantas componentes como nodos tenga el grafo, asociando la componente  $c_i$  a la importancia del nodo  $i$ . Además,  $c = (c_1, c_2, \dots, c_n) \in \mathbb{R}^n$ , siendo  $n$  el número total de nodos del grafo, debe satisfacer:

- Toda componente  $c_i$  debe ser mayor o igual a 0, para todo  $1 \leq i \leq n$ . Esto se debe a que la importancia no se puede cuantificar de forma negativa, debe ser un valor positivo o igual a 0.
- La norma 1 del vector, que corresponde con la suma del valor absoluto de todas sus componentes, debe ser igual a 1. No obstante, por el punto

anterior, al ser todas las componentes mayores que cero basta únicamente en este caso con sumar todas las componentes, es decir,  $\|c\|_1 = \|c_1\| + \dots + \|c_n\| = c_1 + \dots + c_n = 1$ . Esto se debe a que los valores normalizados son más fáciles de interpretar y analizar, ya que están en una escala que varía de 0 a 1, permitiendo comparar la importancia cuantitativa de los nodos de una forma más efectiva.

Como en esta centralidad se tiene en cuenta además del número de conexiones la calidad de estas, entonces la centralidad de cada nodo va a depender también de la centralidad de sus vecinos. Por ejemplo, una página web puede ser considerada importante si otras páginas webs de alta afluencia contienen enlaces a dicha página, un cantante si es elegido por otro de gran éxito para participar en una colaboración en su disco, una persona en Facebook si gente famosa le sigue o un artículo si aparece como referencia en otros de alta consideración. En general, estas relaciones se modelan con grafos dirigidos, ya que cada conexión tiene una dirección que indica la dirección del flujo de información. Es por ello que en los siguientes párrafos nos referiremos a grafos dirigidos.

De esta forma, podríamos definir en primera instancia la centralidad del nodo  $i$  como la suma de las centralidades de todos los nodos  $j$  que se relacionan con el  $i$ , resultando la expresión

$$c_i = \sum_{j \rightarrow i} c_j$$

No obstante, esta igualdad definida no es del todo correcta. En realidad, la centralidad de un nodo es proporcional a la centralidad de sus vecinos. Esto significa que cuanto más central sea un nodo, más probable es que sus vecinos también lo sean. Y viceversa, cuanto más central sean los vecinos de un nodo, más central será él mismo. Por ello, a la anterior igualdad es necesario añadirle un parámetro, al que llamaremos  $\lambda$ , obteniendo

$$c_i = \frac{1}{\lambda} \cdot \sum_{j \rightarrow i} c_j$$

Como debemos tener en cuenta tanto el número de conexiones como la calidad de estas, necesitamos una representación del grafo que nos proporcione la información que deseamos y que a la vez nos permita trabajar con ella. La que encaja a la perfección es la matriz de adyacencia. A través de ella, de una forma fácil y sencilla se puede obtener con qué elementos está relacionado cada elemento del grafo únicamente recorriendo las filas y las columnas. Si queremos conocer los

nodos conectados con el nodo  $i$ , habrá que recorrer la columna  $i$  y cada posición  $j$  que tenga coeficiente 1, significará que existe conexión de  $j$  con  $i$ . Esta matriz de adyacencia la definiremos como  $A = (a_{ij}) \in M_{n \times n}$ , siendo  $n$  el número total de nodos del grafo. Esta matriz vendrá dada por

$$a_{ij} = \begin{cases} 1 & \text{si existe la arista } i \rightarrow j \text{ en el grafo} \\ 0 & \text{en cualquier otro caso} \end{cases}$$

Como nos interesa lo contrario, es decir, la existencia de las aristas  $j \rightarrow i$ , entonces trabajaremos con la traspuesta de esta matriz de adyacencia,  $A = (a_{ji})$ .

En el sumatorio estamos incluyendo la centralidad de todos los nodos que conectan con el  $i$ . Por ello, debemos mantener constante este nodo fijándolo, e ir recorriendo todos los nodos  $j$  restantes para conocer quiénes conectan con él. Si el nodo es adyacente,  $a_{ji}$  será igual a 1, por lo que la centralidad de ese nodo se añadirá al sumatorio. Por otro lado, si no es adyacente,  $a_{ji}$  será igual a 0 y su centralidad no se incluirá. Por lo consiguiente, obtenemos la expresión

$$c_i = \frac{1}{\lambda} \cdot \sum_{j=1}^n a_{ji} \cdot c_j$$

Expresado matricialmente, se obtiene la siguiente igualdad en la que ya conseguimos el vector  $c$  que buscábamos:

$$c = \frac{1}{\lambda} \cdot A^T \cdot c,$$

o lo que es lo mismo, multiplicando a ambos lados por  $\lambda$ ,

$$c \cdot \lambda = A^T \cdot c \tag{4.2}$$

Si nos fijamos en el resultado obtenido, La Ecuación 4.2 tiene exactamente el mismo aspecto que La Ecuación 4.1. En ambas encontramos un vector, un escalar y una matriz cuadrada.

En el caso en el que la matriz de adyacencia sea simétrica, que será cuando las conexiones del grafo sean bidireccionales implicando a la vez a ambos nodos, podremos sustituir la traspuesta de  $A$  por la propia matriz  $A$ .

Por tanto, tras la deducción anterior llegamos a la conclusión de que el vector de centralidad  $c$  que buscamos coincide exactamente con un autovector de la traspuesta de la matriz de adyacencia del grafo en cuestión,  $A^T$ , asociado al autovalor  $\lambda$ .

En este punto aparece un teorema que se vuelve esencial para la tarea en cuestión. Como acabamos de ver, para hallar el vector de centralidad es necesario obtener cierto autovector de la traspuesta de la matriz de adyacencia del grafo. Por ello, este teorema, que enunciaremos y demostraremos a continuación, garantiza la existencia y unicidad de dicho autovector para matrices con ciertos requisitos.

El nombre de este teorema, Teorema de Perron-Frobenius, proviene de los dos matemáticos que hicieron contribuciones significativas a su desarrollo y comprensión: Oskar Perron y Ferdinand Frobenius. Es un resultado clave que establece propiedades especiales para ciertos tipos de matrices. Este teorema asegura que en matrices cuadradas, no negativas e irreducibles, existe un autovalor dominante real positivo y su correspondiente autovector, el cual va a ser el autovector de la traspuesta de la matriz de adyacencia que buscamos y que nos dará la centralidad del grafo. El autovalor dominante y su correspondiente autovector son únicos y definen las características de la red. Con este teorema se asegura su existencia y unicidad.

Por ello, a continuación introducimos el enunciado del Teorema de Perron-Frobenius:

**Teorema 4.1.1.1. Perron-Frobenius.** *Si  $A = (a_{ij})$  es una matriz cuadrada, no negativa e irreducible, entonces se cumplen las siguientes afirmaciones:*

- (I) *El radio espectral de  $A$  es un autovalor de  $A$ , es decir,  $\rho(A) \in \sigma(A)$ , y además la multiplicidad de este autovalor  $\rho(A)$  es igual a 1.*
- (II) *Existe un vector  $u \in \mathbb{R}^n$  con todas sus componentes mayores que 0, con norma igual a 1 y que resulta ser un autovector de la traspuesta de  $A$  (también de  $A$ ) asociado al radio espectral.*
- (III) *Este vector  $u$  no negativo y con norma 1 es el único que cumple las anteriores condiciones, y si existiese otro autovector  $v$  que satisface las mismas condiciones, entonces necesariamente  $u = v$ .*

Debido a que vamos a demostrar dicho teorema de forma geométrica, vamos a enunciarlo seguidamente de otra manera que se adecue a su posterior demostración, tomando como base [8].

**Teorema 4.1.1.2. Perron-Frobenius (Versión 2).** *Sea  $A$  una matriz no negativa e irreducible de tamaño  $n \times n$ . Entonces, existe un autovalor simple y positivo  $\lambda$  de  $A$  que tiene asociado un autovector positivo (es decir, todas sus coordenadas son positivas), y que tiene el valor absoluto más alto entre los módulos de los otros autovalores de  $A$ .*

***Demostración***

Es necesario introducir en primera instancia la definición de rayo, la cual va a usarse en toda la demostración. Un rayo en  $\mathbb{R}^n$  será el conjunto  $r[v] = \{\mu v \mid \mu > 0\}$ , es decir, el conjunto de todos los múltiplos positivos de un vector dado, los cuales presentan la misma dirección al ser proporcionales.

En la demostración consideramos  $\mathbb{R}^+$  como el conjunto de rayos no negativos, es decir, definiremos  $\mathbb{R}^+$  como el conjunto de todas las semirrectas que tienen su origen en el origen de coordenadas y que están contenidas en el cuadrante que se corresponde geoméricamente con el que todas las coordenadas son no negativas. Este cuadrante se define como

$$C^+ = \{(x_1, \dots, x_n) \in \mathbb{R}^n / x_i \geq 0, i = 1, \dots, n\}.$$

(i) Debido a que los coeficientes de la matriz  $A$  y las coordenadas de los rayos que tratamos son no negativas, entonces al actuar  $A$  sobre uno de ellos, devuelve otro rayo no negativo. Esto hace que sea imposible devolver coordenadas negativas. Por ello, podemos definir la matriz  $A$  como una aplicación cerrada en el conjunto  $\mathbb{R}^+$  de rayos no negativos, es decir,  $A(\mathbb{R}^+) \subseteq \mathbb{R}^+$ . No obstante, aún se puede restringir más, demostrando a continuación que ningún rayo en  $\mathbb{R}^+$  se transforma en el elemento nulo por la acción de  $A$ .

Para demostrar esto, vamos a utilizar un método de razonamiento por reducción al absurdo, suponiendo unas condiciones iniciales y llegando a contradicciones.

Supongamos que al actuar la matriz  $A$  sobre  $v = (0, \dots, v_i, \dots, 0)$  se obtiene el elemento nulo, es decir,  $A \cdot v^T = (a_{1i}v_i, a_{2i}v_i, \dots, a_{ni}v_i)^T = (0, \dots, 0)^T$ . Esto impone necesariamente que la columna  $i$ -ésima de la matriz  $A$ , constituida por  $a_{1i}, a_{2i}, \dots, a_{ni}$ , esté llena de ceros, ya que  $v_i$  es distinto de cero. Al existir en  $A$  una columna nula, obliga a que esta sea reducible.

Ahora, en vez de suponer que existe una sola coordenada distinta de cero, supongamos que existen  $k$  coordenadas distintas de cero. Como todos los elementos tanto de  $A$  como de  $v$  son no negativos, el resultado de la multiplicación generaría coordenadas resultantes de  $k$  sumandos no negativos. Pero como el vector que se obtiene es  $(0, \dots, 0)^T$ , necesariamente todos los sumandos deben ser iguales a cero, obligando a que sean nulos los coeficientes de la matriz  $A$  que participan en estos sumandos. Esto hace que haya  $k$  columnas de ceros en  $A$ . Al existir en

A columnas nulas, obliga a que esta sea reducible.

Como por hipótesis sabemos que  $A$  es irreducible, hemos llegado a una contradicción. Por tanto, ningún rayo en  $\mathbb{R}^+$  se transforma en el elemento nulo por la acción de  $A$ .

(ii) Lo siguiente a demostrar es que ningún rayo en  $\partial\mathbb{R}^+$  queda invariante bajo la acción de  $A$ . Al igual que en (i), utilizaremos una reducción al absurdo.

Supongamos que en  $\partial C^+$ , el borde de dicho cuadrante, existe un rayo, definámoslo  $r = r[v]$  ( $v$  es un vector perteneciente a dicho rayo), para el cual  $A(r) = r$ . Es decir, al aplicar  $A$  sobre el rayo, se obtiene el mismo rayo.

Como  $r$  yace en  $\partial C^+$ , los vectores de este rayo tendrán al menos una coordenada igual a cero. Por ello, y al ser no nulo,  $v$  contendrá  $k$  coordenadas iguales a cero, siendo  $0 < k < n$ . Estos ceros pueden colocarse en las  $k$  primeras coordenadas de  $v$  tras una reordenación.

Por hipótesis tenemos que  $A(r) = r$ , y sabiendo que las primeras  $k$  coordenadas de  $v$  son cero, entonces la matriz  $A$  debe ser de la forma

$$\begin{pmatrix} M & 0 \\ P & N \end{pmatrix}$$

donde  $M$  es una matriz cuadrada  $k \times k$  llena de ceros y  $N$  es una matriz  $(n - k) \times (n - k)$ . Esto entraría en contradicción con la irreducibilidad de  $A$ .

Veamos esto con un ejemplo. Si estamos en  $\mathbb{R}^7$ , y sea  $v = (0, 0, 0, 0, p, q, r)$ . Como  $r[v] = \{\mu v \mid \mu > 0\}$ , y por hipótesis  $A(r) = r$ , entonces obligatoriamente  $A$  debe ser de la forma:

$$\begin{pmatrix} a_{11} & a_{12} & a_{13} & a_{14} & 0 & 0 & 0 \\ a_{21} & a_{22} & a_{23} & a_{24} & 0 & 0 & 0 \\ a_{31} & a_{32} & a_{33} & a_{34} & 0 & 0 & 0 \\ a_{41} & a_{42} & a_{43} & a_{44} & 0 & 0 & 0 \\ a_{51} & a_{52} & a_{53} & a_{54} & a_{55} & a_{56} & a_{57} \\ a_{61} & a_{62} & a_{63} & a_{64} & a_{65} & a_{66} & a_{67} \\ a_{71} & a_{72} & a_{73} & a_{74} & a_{75} & a_{76} & a_{77} \end{pmatrix} \cdot \begin{pmatrix} 0 \\ 0 \\ 0 \\ 0 \\ p \\ q \\ r \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \\ 0 \\ \mu p \\ \mu q \\ \mu r \end{pmatrix}$$

(iii) Es necesario introducir en este apartado el Teorema del Punto Fijo de Brouwer. Dicho teorema afirma que en un espacio euclídeo, toda aplicación continua de un conjunto no vacío, compacto y convexo en sí mismo tiene un punto fijo.

Para poder aplicarlo a nuestro conjunto  $\mathbb{R}^+$ , necesitamos verlo como un conjunto que cumpla las condiciones del teorema. La forma correcta de hacerlo es

identificar cada rayo de  $\mathbb{R}^+$  con su punto perteneciente a  $M = \{(x_1, \dots, x_n)/x_1 + \dots + x_n = 1, x_i \geq 0, i = 1, \dots, n\}$ .

Por tanto, al identificar  $\mathbb{R}^+$  con  $M$ , tenemos que se cumple que es no nulo, compacto y convexo. Y como por (i) sabemos que  $A(\mathbb{R}^+) \subseteq \mathbb{R}^+$ , entonces por el Teorema del Punto Fijo de Brouwer se asegura que existe un punto fijo en  $M$ , o lo que es lo mismo, un rayo invariante  $r$  en  $\mathbb{R}^+$ . Como por (ii) sabemos que ningún rayo en el borde del cuadrante queda invariante bajo la acción de  $A$ , entonces necesariamente, si existe un rayo fijo, este debe ser positivo. Por tanto, al tomar un vector  $v$  perteneciente al rayo (misma dirección), este vector tiene todas sus componentes positivas. Si  $v$  es un autovector, entonces su autovalor asociado,  $\lambda$ , también será positivo ya que tanto  $v$  como su imagen pertenecen a  $C^+$ .

Existe una correspondencia entre los rayos de  $\mathbb{R}^+$  y la esfera  $S^{n-1}$ . Lo vamos a estudiar para  $\mathbb{R}^3$ , y se haría de forma análoga para  $\mathbb{R}^n$ .

En  $\mathbb{R}^3$ , dado un plano  $p$ , la intersección de dicho plano con la esfera  $S^2$  es la circunferencia  $S_p^1$ . Partiendo de este concepto, supongamos que tenemos un plano, llamémoslo  $\pi$ , que es invariante y que contiene al rayo fijo  $r$ . La intersección entre  $\mathbb{R}^+$  y  $S_\pi^1$  es un arco, llamémosle  $L$ . Como  $L$  pertenece al plano invariante y por (i), entonces  $A(L) \subset L$ .

La aplicación  $A$  restringida a este plano  $\pi$  es una aplicación lineal y podemos buscar su forma de Jordan. Como sabemos que  $\lambda$  es uno de sus autovalores, tenemos tres posibilidades distintas:

- a)  $\text{mult}_{\text{algebraica}}(\lambda) = \text{mult}_{\text{geométrica}}(\lambda) = 2$ . En este caso, su forma de Jordan sería

$$A|_\pi = \begin{pmatrix} \lambda & 0 \\ 0 & \lambda \end{pmatrix}$$

- b)  $\text{mult}_{\text{algebraica}}(\lambda) = 2$  y  $\text{mult}_{\text{geométrica}}(\lambda) = 1$ . En este caso, su forma de Jordan sería

$$A|_\pi = \begin{pmatrix} \lambda & 0 \\ 1 & \lambda \end{pmatrix} \quad \text{ó} \quad A|_\pi = \begin{pmatrix} \lambda & 1 \\ 0 & \lambda \end{pmatrix}$$

- c)  $\text{mult}_{\text{algebraica}}(\lambda) = \text{mult}_{\text{geométrica}}(\lambda) = 1$ . En este caso, su forma de Jordan sería

$$A|_\pi = \begin{pmatrix} \lambda & 0 \\ 0 & \mu \end{pmatrix}$$

(1) Por reducción al absurdo, para el caso a), supongamos que existen dos vectores independientes asociados al autovalor  $\lambda$ ,  $v_1$  y  $v_2$ . Tomándolos como base

del plano, cualquier otro vector del plano podría expresarse como una combinación de ambos, es decir,  $w = a \cdot v_1 + b \cdot v_2$ . Como la matriz  $A$  es una aplicación lineal, entonces  $A \cdot w = \lambda \cdot w$ . Esto implica que  $A$  deja invariante a todos los rayos del plano, incluidos los del borde, lo que contradice lo que hemos demostrado en (ii). Por tanto, el caso a) no es posible.

(2) Por reducción al absurdo, para el caso b), supongamos que la matriz  $A|_{\pi}$  es semejante a la matriz cuya forma de Jordan es

$$\begin{pmatrix} \lambda & 0 \\ 1 & \lambda \end{pmatrix}$$

Al ser semejantes, pueden considerarse dos formas de una misma transformación lineal pero respecto a bases distintas, por lo que fijan otros vectores. Tenemos que

$$\begin{pmatrix} \lambda & 0 \\ 1 & \lambda \end{pmatrix} \cdot \begin{pmatrix} a \\ b \end{pmatrix} = \begin{pmatrix} a\lambda \\ a + b\lambda \end{pmatrix}$$

La aplicación lineal dada por  $A$  fija los rayos que pasan por  $(0,1)$  y  $(0,-1)$ , y todos los demás giran en sentido anti-horario. No obstante, cuanto más cerca está un vector de los vectores fijos, menor es el giro, al no poder sobrepasarlo.

Si la matriz  $A|_{\pi}$  es semejante, va a realizar la misma transformación, dejando dos vectores fijos y el resto girando en sentido anti-horario con amplitud menor cuanto más cerca esté de los fijos.

No obstante, los rayos fijos a través de  $A|_{\pi}$  no pueden estar en  $\partial\mathbb{R}^+$  (contradiría (ii)) y no pueden ser distintos del eje vertical (porque el vector que pasa por  $(1,0)$  saldría de  $\mathbb{R}^+$  contradiciendo (i)). Por tanto,  $A|_{\pi}$  no es semejante a la de opción b (razonamiento análogo para la otra forma de Jordan de la opción b))

De (1) se deduce que la multiplicidad geométrica es 1, y por (2) que la multiplicidad algebraica es 1, dando lugar a que el único caso posible es la opción c).

Por tanto, se acaba de demostrar que el autovalor  $\lambda$  correspondiente al rayo  $r$  fijo es simple, positivo y tiene asociado un autovector positivo.

**(iv)** Lo único que falta por demostrar es que  $|\lambda| \geq |\mu|$  para cualquier otro valor propio  $\mu$ . Para ello, vamos a utilizar al igual que en (i) y en (ii) un razonamiento por reducción al absurdo.

Supongamos que existe un  $\mu$  tal que  $|\mu| > |\lambda|$ . Encontramos dos casos:

- $\mu$  es un valor propio real:

Imaginemos que tenemos un autovector  $v_\mu$  asociado al valor propio real  $\mu$ . Este autovector define un rayo con su dirección, llamémoslo  $r_\mu$ . Asumamos que este  $r_\mu$  no pertenece a  $\mathbb{R}^+$ , puesto que si perteneciera ya estaría demostrado el teorema. Entonces  $A^n$  actúa en  $S^1$ , que representa todos los rayos del plano generado por  $v_\mu$  y  $v_\lambda$ , fijando el conjunto  $\{\pm r_\mu, \pm r_\lambda\}$ . Como por hipótesis hemos asumido que  $|\mu| > |\lambda|$ , la acción de  $A^n$  en  $S^1$  presenta dos puntos atractores,  $\{\pm r_\mu\}$  y dos puntos repulsores,  $\{\pm r_\lambda\}$ . Este escenario nos lleva a una contradicción, puesto que el rayo  $r_\mu$  atraería uno de los dos puntos de la intersección de  $\partial\mathbb{R}^+$  con  $S^1$  fuera de  $\mathbb{R}^+$ , y esto no podría ocurrir.

- $\mu$  es un valor propio complejo:

La matriz  $A$  actúa de  $C^n$  en  $C^n$ . El espacio propio generado por un autovalor complejo corresponde con el conjunto de autovectores asociados a él, y está contenido en  $C^n$ . No obstante, solo queremos su corte con  $\mathbb{R}^n$ . Por ello, dado un autovector complejo  $v = x + iy$ , la intersección de  $\mathbb{R}^n$  con el espacio generado por  $v$  origina un plano de  $\mathbb{R}^n$  que contiene a los vectores  $x$  e  $y$ . Este plano, llamémoslo  $P$ , puede definirse como  $P = \langle \text{real}(v), \text{imaginaria}(v) \rangle$ . La transformación que sufre dicho espacio propio generado por  $v$  tras actuar  $A$  es una rotación y una dilatación. Como este espacio propio es invariante, tanto él como su imagen por  $A$  presentan la misma parte real, correspondiente al plano  $P$ . Por ello, se llega a que dicho plano  $P$  es invariante bajo la acción de  $A$ .

Tomando dicho plano invariante  $P$  correspondiente al autovalor  $\mu$ ,  $P$  no contiene ningún rayo de  $\mathbb{R}^+$ . Esto se debe a que habría tres posibilidades para la intersección de  $P$  con  $\mathbb{R}^+$ :

- a) Se cortan en un único rayo. Para ello sería necesario que el rayo perteneciera a  $\partial\mathbb{R}^+$ . Entonces su imagen sería el propio rayo, ya que  $A(\mathbb{R}^+) \subseteq \mathbb{R}^+$ ,  $P$  es invariante y su único elemento perteneciente a  $\mathbb{R}^+$  es este rayo. Esto entra en contradicción con (ii).
- b) Parte de  $P$  está contenido en  $\mathbb{R}^+$ . Como cada punto de  $S^1$  representa a cada rayo de  $P$ , toda la colección de rayos de  $\mathbb{R}^+$  contenida en  $P$  se corresponde con un arco  $L$  de  $S^1$ . Al ser  $A$  una rotación,  $A(L)$  no estaría contenida en  $L$ , lo que contradice  $A(\mathbb{R}^+) \subseteq \mathbb{R}^+$ .
- c)  $P$  no corta a  $\mathbb{R}^+$ . Al demostrar que tanto a) como b) son imposibles, esta es la única opción.

Consideremos ahora el espacio tridimensional  $E^3$  generado por el plano invariante  $P$  y el autovector  $v_\lambda$  asociado al autovalor propio  $\lambda$ . Como por

hipótesis  $|\mu| > |\lambda|$ , la dinámica de la acción de  $A$  en  $E^3$  genera una línea repulsora generada por  $v_\lambda$  y una atracción al plano  $P$ . Esto hace que los rayos de  $\mathbb{R}^+$  en  $E^3$ , excluyendo  $r_\lambda$ , se acerquen al plano atractor  $P$  tanto como queramos por la acción repetida de  $A$ . De esta forma, se deduce que todos los rayos abandonan  $\mathbb{R}^+$  tras un número de aplicaciones de  $A$ , entrando en contradicción con  $A(\mathbb{R}^+) \subseteq \mathbb{R}^+$ .

Esto último se puede demostrar formalmente así: sea  $r[x]$  un rayo de  $\mathbb{R}^+ \setminus v_\lambda$  con  $x = v_\lambda + w$ , con  $w \neq 0$  y  $w \in P$ . Entonces se tiene que

$$A^n \cdot \left( \frac{x}{|\mu|^n} \right) = \frac{A^n(v_\lambda) + A^n(w)}{|\mu|^n} = \frac{\lambda^n}{|\mu|^n} \cdot v_\lambda + \frac{\mu^n}{|\mu|^n} \cdot w$$

donde  $\frac{\lambda^n}{|\mu|^n} \rightarrow 0$  cuando  $n \rightarrow \infty$ . Esto implica que el rayo  $A^n(r[x])$  puede acercarse al plano atractor  $P$  tanto como queramos, ya que el módulo del otro sumando,  $\frac{\mu^n}{|\mu|^n} \cdot w$ , es constante.

Por tanto,  $|\lambda| \geq |\mu|$  para cualquier otro valor propio  $\mu$ .

■

Volviendo al primer enunciado del teorema, cabe destacar que las matrices originales y sus traspuestas tienen los mismos autovalores en términos de valor absoluto, de forma que si  $|\lambda|$  es un autovalor de  $A$  con un autovector asociado  $v$ , entonces  $|\lambda|$  también es un autovalor de  $A^T$  con el mismo autovector  $v$ . Como estamos cogiendo el autovalor dominante, que es el mayor en valor absoluto, entonces este será el dominante para ambas matrices, la original y su traspuesta, y su autovector asociado, que es el que queremos, también será el mismo. Por ello, podemos utilizar cualquiera en función del contexto del problema que tengamos. No obstante, tras el razonamiento empleado para obtener La Ecuación 4.2, empleamos en el enunciado del teorema el autovector de la traspuesta.

Como indica el enunciado del teorema, para poder aplicarlo, la matriz debe de cumplir una serie de condiciones. Veámoslas en profundidad.

En primer lugar, las matrices deben ser cuadradas, lo que significa que tienen tantas filas como columnas. Nuestra matriz de adyacencia es cuadrada, de forma que lo cumple. Esta condición es imprescindible para garantizar el cálculo de sus autovalores y autovectores al ser un operador lineal, lo que es esencial para nuestro objetivo.

En segundo lugar, estas matrices deben ser no negativas, es decir, todos los coeficientes deben ser mayores o iguales a 0. Nuestra matriz de adyacencia cumple con este requisito, puesto que los grafos que estamos tratando no pueden tener conexiones negativas y además los nodos no se relacionan consigo mismos. Esto es

importante porque, en el mundo real, no podemos tener valores negativos cuando estamos hablando de ciertas cantidades y asegura que no haya cancelaciones entre los elementos al calcular los autovectores.

Finalmente, las matrices deben ser irreducibles, lo que implica que no se pueden dividir en partes más pequeñas sin perder su esencia. Esto hace que todos los elementos estén interconectados y tengan influencia directa entre sí, sin permitir la existencia de nodos aislados. Todos los nodos tienen algún camino hacia otros nodos en la red. Esto es esencial para identificar los nodos centrales más importantes que tienen influencia directa sobre otros nodos en la red. Una forma de comprobar si la matriz es irreducible es revisando que su grafo dirigido asociado es fuertemente conexo. Las anteriores condiciones siempre se cumplían por ser una matriz de adyacencia; no obstante, este requerimiento si que es necesario comprobarlo porque no siempre implica su cumplimiento. En el caso en el que la matriz de adyacencia no fuera irreducible, no se podría aplicar el Teorema de Perron-Frobenius directamente al grafo completo. Pero si el grafo se puede dividir en componentes irreducibles, el teorema se podría aplicar a cada uno de estos componentes por separado.

Todas estas condiciones nos permiten garantizar la existencia y unicidad del autovector de la matriz, el cual recibe el nombre de *vector de Perron*.

La existencia del vector propio positivo se logra cuando la matriz cumple con las condiciones del teorema de Perron-Frobenius, que incluyen ser una matriz no negativa y que sea irreducible, lo que garantiza que siempre haya al menos un valor propio real no negativo, es decir, un autovalor dominante.

La unicidad del vector propio positivo se asegura cuando la multiplicidad del autovalor dominante (es decir, el número de veces que el autovalor dominante se repite como raíz característica de la matriz) es igual a 1. Si la multiplicidad del autovalor dominante fuera mayor que 1, entonces habría varios vectores propios asociados al mismo autovalor, y no habría un criterio claro para elegir cuál de ellos representa mejor las características del sistema.

Aunque el teorema nos garantiza la existencia de este autovector, es necesario calcular el autovalor dominante y su autovector asociado mediante métodos clásicos de obtención de estos, los cuales veremos posteriormente en una aplicación práctica.

Por todo esto, el vector de Perron, que tendrá tantas componentes como nodos el grafo, corresponde con el vector de centralidad deseado del grafo en cuestión y nos proporciona la importancia cuantitativa de cada nodo. De esta forma, hemos obtenido lo que buscábamos.

Una vez explicado el cálculo de la centralidad de vector propio en grafos teóricamente, para su mejor comprensión se considera necesario ejemplificarlo.

Tomando como referencia el grafo dirigido de La Figura 3.1, vamos a llevar a cabo el cálculo de su centralidad de vector propio.

El primer paso es obtener su matriz de adyacencia, la cual tiene la forma

$$A = \begin{pmatrix} 0 & 1 & 0 & 1 & 0 \\ 0 & 0 & 1 & 0 & 1 \\ 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 1 \\ 0 & 1 & 0 & 0 & 0 \end{pmatrix}$$

A continuación, tenemos que comprobar que esta matriz cumple las condiciones del Teorema de Perron-Frobenius, es decir, que sea cuadrada, no negativa e irreducible:

- Es una matriz cuadrada por tener el mismo número de filas que de columnas, concretamente una matriz  $5 \times 5$ .
- Es una matriz no negativa, puesto que todos sus coeficientes son mayores o igual a cero.
- Verifiquemos que es irreducible. Para ello, debemos comprobar que su grafo dirigido asociado, que corresponde con nuestro inicial de La Figura 3.1, es fuertemente conexo, es decir, que existe un camino entre cualquier par de nodos distintos. Se puede apreciar que lo cumple fácilmente al tratarse de un grafo muy pequeño. Si fuese de mayores dimensiones, se puede verificar si cumple esta condición de una manera más fácil utilizando librerías o funciones en algunos lenguajes de programación tardando apenas segundos.

Como nuestra matriz  $A$  cumple todas las condiciones, podemos aplicar el Teorema de Perron-Frobenius, asegurando que existe un autovector único asociado al autovalor dominante de la matriz que nos dará la centralidad.

Utilizando cualquier método de obtención de valores propios de una matriz, obtenemos los siguientes:

$$\begin{aligned} \lambda_1 &= 1,5214 \\ \lambda_2 &= -0,7607 + 0,8579i \\ \lambda_3 &= -0,7607 - 0,8579i \\ \lambda_4 &= 0 \\ \lambda_5 &= 0 \end{aligned}$$

Nuestro autovalor dominante es  $\lambda = 1,5214$ , con multiplicidad igual a 1. El último paso que nos queda es obtener el autovector asociado a dicho autovalor.

Utilizando cualquier método de obtención de autovectores de una matriz a partir de un autovalor y normalizándolo en norma 1, obtenemos el autovector de Perron

$$c = (0'2718, 0'2068, 0'1787, 0'2068, 0'1359),$$

el cual nos proporciona el vector de centralidad del grafo que estamos estudiando. La componente  $c_i$  del vector se identifica con la centralidad de vector propio del nodo  $i$  del grafo. Se verifica que la suma de todas las componentes es igual a 1.

## 4.2. Centralidad en hipergrafos

Hasta el momento hemos explorado cómo la centralidad en grafos proporciona una visión crucial sobre la importancia de los nodos en redes con conexiones binarias. Sin embargo, cuando se trata de modelar relaciones más complejas de mayor cardinalidad, las limitaciones de los grafos se vuelven notables. Aquí es donde entran en juego los hipergrafos y su respectiva medida de centralidad, el objeto de estudio de este apartado.

Determinar la importancia de los diferentes trabajadores de una empresa a partir de su participación en conferencias o la importancia de autores en función de su colaboración en papers constituyen diferentes tipos de aplicaciones prácticas de la vida cotidiana en los cuales se requiere el cálculo de la centralidad en hipergrafos.

En el apartado anterior hemos visto la manera de calcular la centralidad en grafos simples. El primer pensamiento que se puede tener es que es posible aplicar el mismo método a hipergrafos puesto que también presentan una matriz de adyacencia para su representación. Pero esta suposición es totalmente falsa. El Teorema de Perron-Frobenius, el cual usamos en grafos simples para el cálculo del vector de Perron, no se puede aplicar directamente a hipergrafos debido a las diferencias estructurales que poseen ambos tipos de redes, aunque los dos presenten nodos y relaciones entre estos. Las hiperaristas al conectar dos o más nodos, añade una complejidad adicional en la estructura de su matriz de adyacencia.

Es en este momento cuando entra en acción el concepto de grafo lineal. Hemos visto en la sección de fundamentos teóricos que a partir de un hipergrafo podemos crear su grafo lineal asociado, pasando las hiperaristas a ser los nodos y dos de

estas serán adyacentes si presentan al menos un nodo en común. Un grafo lineal es un grafo simple con relaciones binarias, y con él si que se puede aplicar el Teorema de Perron-Frobenius para calcular su centralidad de vector propio. De esta manera, obtendríamos la centralidad del grafo lineal asociado al hipergrafo, lo que equivale a tener la importancia cuantitativa de las hiperaristas. Pero a nosotros nos interesa la de los nodos. Solo nos faltaría extender esa centralidad de hiperaristas a los nodos en el hipergrafo, y es justo lo que se desarrollará a continuación. Véase este razonamiento con una representación gráfica en La Figura 4.1.

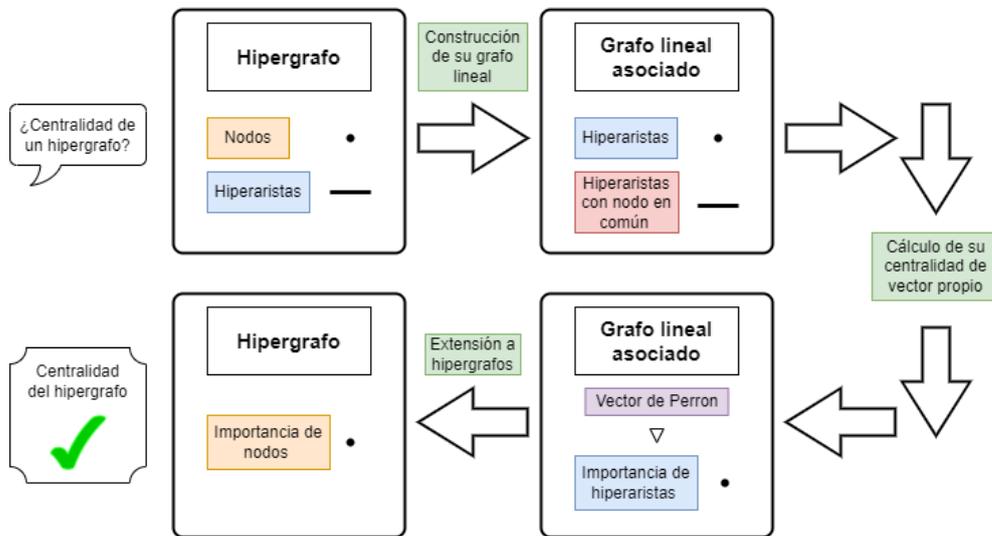


Figura 4.1: Esquema de los pasos a seguir para el cálculo de la centralidad en hipergrafos.

Para tratar toda la extensión de la centralidad en grafos a hipergrafos, nos apoyaremos en [9].

En este punto, supondremos que tenemos construido el grafo lineal asociado y hemos hallado su vector de Perron siguiendo el apartado anterior, que nos proporciona la importancia de las hiperaristas del hipergrafo.

Llamemos al vector de Perron obtenido del grafo lineal asociado  $c(h)$ , por dar la importancia de todas las hiperaristas, cada una de ellas una componente. Es decir, la hiperarista 1 tiene asociada la importancia de la componente 1 y la hiperarista  $n$  tiene asociada la importancia de la componente  $n$ .

Gracias a este vector  $c(h)$  conocemos la importancia de las hiperaristas, pero lo que queremos es averiguar la importancia de los nodos del hipergrafo. Pero para ello, este vector es crucial, puesto que gracias a él vamos a poder ir calculando uno por uno la centralidad de cada nodo para lograr el vector de centralidad

general del hipergrafo.

El vector de centralidad del hipergrafo, llamémoslo  $c$ , será un vector de vectores no negativo  $c = (c_1, c_2, \dots, c_n) \in (\mathbb{R}^{D-1})^n$ , siendo  $n$  el número de nodos total del hipergrafo y  $D$  el tamaño de la hiperarista que más nodos conecta. Cada componente de este vector corresponde con la centralidad vectorial de cada nodo  $i \in V$ , siendo  $V$  el conjunto de todos los nodos del hipergrafo. Cada  $c_i$  es un vector no negativo de la forma

$$c_i = (c_{i2}, \dots, c_{iD}) \in \mathbb{R}^{D-1}.$$

Cada una de las componentes de este  $c_i$  se hallará empleando la siguiente fórmula, para cada  $2 \leq k \leq D$ ,

$$c_{ik} = \frac{1}{k} \cdot \sum_{\substack{i \in h \in \mathcal{E} \\ |h|=k}} c(h), \quad (4.3)$$

siendo  $k = |h|$ , es decir, el tamaño de la hiperarista  $h$ .

En esta ecuación, por cada nodo  $i$  del hipergrafo, tenemos que ir recorriendo todos los tamaños de las hiperaristas en las cuales el nodo participa desde 2 hasta el rango del hipergrafo (máximo tamaño de todas las hiperaristas) e ir sumando todas sus centralidades. Dicho de otra forma, la componente  $c_{ik}$  será la suma de todas las centralidades de las hiperaristas de tamaño  $k$  que incidan en el nodo  $i$ . El valor de  $c(h)$  corresponde con la importancia obtenida en el vector de Perron para la hiperarista correspondiente. Si un nodo no participa en ninguna hiperarista de cierto tamaño  $k$ , el valor de centralidad  $c_{ik}$  para dicho tamaño será cero. Por ejemplo, si tuviésemos un hipergrafo de 10 elementos y la hiperarista de mayor tamaño engloba 6 nodos, entonces La Fórmula 4.3 deberá aplicarse 5 veces por cada nodo  $i$ ,  $1 \leq i \leq 10$ . Cada una de las cinco veces es por cada tamaño de las hiperaristas,  $k = 2, 3, 4, 5, 6$ . Se empieza en el dos debido a que no consideramos hiperaristas que solo englobe un elemento al no proporcionar ningún tipo de información valiosa. En este ejemplo, el vector de centralidad del nodo 2 se calcularía así:

$$c_2 = (c_{22}, c_{23}, c_{24}, c_{25}, c_{26}) = \left( \frac{1}{2} \sum_{\substack{2 \in h \in \varepsilon \\ |h|=2}} c(h), \frac{1}{3} \sum_{\substack{2 \in h \in \varepsilon \\ |h|=3}} c(h), \frac{1}{4} \sum_{\substack{2 \in h \in \varepsilon \\ |h|=4}} c(h), \frac{1}{5} \sum_{\substack{2 \in h \in \varepsilon \\ |h|=5}} c(h), \frac{1}{6} \sum_{\substack{2 \in h \in \varepsilon \\ |h|=6}} c(h) \right)$$

Por tanto, empleando esto somos capaces de obtener nuestro vector de centralidad del hipergrafo.

No obstante, nos falta por comprobar que nuestra medida de centralidad del hipergrafo está correctamente normalizada. Veámoslo a continuación.

Por un lado, tenemos la suma de las centralidades de cada uno de los nodos pertenecientes al hipergrafo, todas ellas normalizadas en norma 1, es decir,

$$\sum_{i \in V} \|c_i\|_1 \tag{4.4}$$

Este sumatorio puede descomponerse en

$$\sum_{i \in V} \sum_{k=2}^D c_{ik} \tag{4.5}$$

correspondiente a la suma de las centralidades de todos los nodos  $i$  pertenecientes al hipergrafo en función de su participación en hiperaristas de diferentes tamaños.

Sustituyendo La Ecuación 4.3 en La Ecuación 4.5, se obtiene que

$$\sum_{i \in V} \sum_{k=2}^D \sum_{\substack{i \in h \in \varepsilon \\ |h|=k}} \frac{c(h)}{k} \tag{4.6}$$

Las diferentes partes que componen La Ecuación 4.6 son las siguientes:

$$\underbrace{\sum_{\substack{i \in h \in \varepsilon \\ |h|=k}} \frac{c(h)}{k}}$$

Suma de todas las centralidades de las hiperaristas de tamaño  $k$  que inciden en el nodo  $i$

$$\underbrace{\sum_{k=2}^D \sum_{\substack{i \in h \in \varepsilon \\ |h|=k}} \frac{c(h)}{k}}$$

Suma de todas las centralidades de las hiperaristas de todos los tamaños  $2 \leq k \leq D$  que inciden en el nodo  $i$

$$\underbrace{\sum_{i \in V} \sum_{k=2}^D \sum_{\substack{i \in h \in \varepsilon \\ |h|=k}} \frac{c(h)}{k}}$$

Suma de todas las centralidades de las hiperaristas de todos los tamaños  $2 \leq k \leq D$  que inciden en cada uno de los nodos  $i$

Ahora, volviendo a La Ecuación 4.6, se tiene que para cada nodo  $i$  del hipergrafo,

$$\sum_{k=2}^D \sum_{\substack{i \in h \in \varepsilon \\ |h|=k}} \frac{c(h)}{k} = \sum_{i \in h \in \varepsilon} \frac{c(h)}{|h|}, \quad (4.7)$$

debido a que ambos lados de la ecuación están sumando sobre los mismos conjuntos  $h$  en  $\varepsilon$ . En el lado izquierdo, estamos sumando sobre todos los conjuntos  $h$  de diferentes tamaños (desde  $k$  igual a 2 hasta  $k$  igual a  $D$ ), y para cada conjunto de tamaño  $k$ , estamos dividiendo  $c(h)$  por  $k$ . En el lado derecho, estamos sumando sobre los mismos conjuntos  $h$ , pero esta vez, estamos dividiendo  $c(h)$  por el tamaño real de cada conjunto, que es  $|h|$ . Esto asegura que cada  $c(h)$  se cuente exactamente una vez sin importar el tamaño de  $h$ . Es decir, dado un nodo  $i$ , se obtiene el mismo resultado sumando todas las centralidades de las hiperaristas yendo por tamaños de menor a mayor dividiendo la suma de centralidades de tamaño  $k$  por  $k$  (expresión izquierda de La Ecuación 4.7), que recorriendo todas las hiperaristas en las que participa el nodo sin llevar un orden de tamaños y

dividir cada centralidad de hiperarista por su tamaño  $|h|$  (expresión derecha de La Ecuación 4.7). Por ello, con ambas expresiones se llega a lo mismo.

Sustituyendo La Ecuación 4.7 en La Ecuación 4.6, se tiene que

$$\sum_{i \in V} \sum_{k=2}^D \sum_{\substack{i \in h \in \varepsilon \\ |h|=k}} \frac{c(h)}{k} = \sum_{i \in V} \sum_{i \in h \in \varepsilon} \frac{c(h)}{|h|} = \sum_{h \in \varepsilon} \sum_{i \in h} \frac{c(h)}{|h|} = \sum_{h \in \varepsilon} c(h) \quad (4.8)$$

Cambiando el orden de los sumatorios y considerando que cada término  $\frac{c(h)}{|h|}$  aparece justo  $|h|$  veces, se puede tener la última igualdad en La Ecuación 4.8.

Retomando el otro lado de la cadena de igualdades que acabamos de hacer e igualando ambos extremos, se tiene que

$$\sum_{i \in V} \|c_i\|_1 = \sum_{h \in \varepsilon} c(h) \quad (4.9)$$

Pero sabemos por el Teorema de Perron-Frobenius que la suma de todas las componentes de nuestro  $c(h)$ , coincidente con el vector de Perron, es 1, por lo que

$$\sum_{i \in V} \|c_i\|_1 = \sum_{h \in \varepsilon} c(h) = 1, \quad (4.10)$$

o lo que es lo mismo,

$$\sum_{i \in V} \|c_i\|_1 = 1, \quad (4.11)$$

justo lo que queríamos.

Con esto finaliza la explicación teórica de la extensión de la centralidad en hipergrafos.

Una vez explicado el cálculo de la centralidad de hipergrafos a partir de su grafo lineal asociado teóricamente, para su mejor comprensión se considera necesario ejemplificarlo.

Tomando como referencia el hipergrafo con su correspondiente grafo lineal de La Figura 3.4, vamos a llevar a cabo el cálculo de su centralidad.

Como ya tenemos su grafo lineal construido, llevaremos a cabo el cálculo de su centralidad de vector propio. Este grafo tiene la siguiente matriz de adyacencia:

$$A = \begin{pmatrix} 0 & 1 & 0 & 1 \\ 1 & 0 & 0 & 1 \\ 0 & 0 & 0 & 1 \\ 1 & 1 & 1 & 0 \end{pmatrix}$$

Esta matriz cumple todas las condiciones del Teorema de Perron-Frobenius, puesto que es cuadrada por tener el mismo número de filas que de columnas, es no negativa porque sus coeficientes son todos mayores o igual a cero, y es irreducible por ser su grafo dirigido asociado fuertemente conexo (puede comprobarse fácilmente al tener una cantidad muy pequeña de nodos). Además, otra característica de la matriz de adyacencia de los grafos lineales asociados a hipergrafos es que es simétrica, luego no importa trabajar con la original o la traspuesta para el teorema. Por ello, nos aseguramos la existencia y unicidad del vector de Perron.

Los autovalores obtenidos tras su cálculo son:

$$\begin{aligned} \lambda_1 &= -1,4812 \\ \lambda_2 &= -1 \\ \lambda_3 &= 0,3111 \\ \lambda_4 &= 2,1701 \end{aligned}$$

Nos quedamos con el autovalor mayor en valor absoluto, que corresponde con  $\lambda = 2,1701$ . A partir de él, buscamos su autovector asociado, que una vez obtenido y normalizado en norma 1, se obtiene que es

$$c = (0'2696, 0'2696, 0'1454, 0'3154).$$

Por tanto, este es nuestro vector de Perron, cuyas componentes suman 1, el cual define la importancia de las hiperaristas. El siguiente paso es extenderlo usando lo visto en este apartado.

A continuación tenemos que aplicar La Fórmula 4.3 para cada nodo y así ir calculando la centralidad de cada uno. Cada centralidad va a constituir una respectiva componente en el vector de centralidad del hipergrafo que buscamos. Los cálculos con sus correspondientes resultados son los siguientes:

- Centralidad del nodo 1:

$$c_{12} = \frac{1}{2} \cdot (0) = 0$$
$$c_{13} = \frac{1}{3} \cdot (0,2696) = 0,0898$$

El vector de centralidad del nodo 1 es  $c_1 = (0,0'0898)$ . Sumando todas las componentes, obtenemos una centralidad del nodo 1 de 0,0898.

- Centralidad del nodo 2:

$$c_{22} = \frac{1}{2} \cdot (0'1454) = 0,0727$$
$$c_{23} = \frac{1}{3} \cdot (0) = 0$$

El vector de centralidad del nodo 2 es  $c_2 = (0'0727, 0)$ . Sumando todas las componentes, obtenemos una centralidad del nodo 2 de 0,0727.

- Centralidad del nodo 3:

$$c_{32} = \frac{1}{2} \cdot (0) = 0$$
$$c_{33} = \frac{1}{3} \cdot (0,2696 + 0,2696) = 0,1797$$

El vector de centralidad del nodo 3 es  $c_3 = (0,0'1797)$ . Sumando todas las componentes, obtenemos una centralidad del nodo 3 de 0,1797.

- Centralidad del nodo 4:

$$c_{42} = \frac{1}{2} \cdot (0) = 0$$
$$c_{43} = \frac{1}{3} \cdot (0,2696 + 0,2696 + 0,3154) = 0,2848$$

El vector de centralidad del nodo 4 es  $c_4 = (0, 0'2848)$ . Sumando todas las componentes, obtenemos una centralidad del nodo 4 de 0,2848.

- Centralidad del nodo 5:

$$c_{52} = \frac{1}{2} \cdot (0) = 0$$
$$c_{53} = \frac{1}{3} \cdot (0,2696 + 0,3154) = 0,195$$

El vector de centralidad del nodo 5 es  $c_5 = (0, 0'195)$ . Sumando todas las componentes, obtenemos una centralidad del nodo 5 de 0,195.

- Centralidad del nodo 6:

$$c_{62} = \frac{1}{2} \cdot (0,1454) = 0,0727$$
$$c_{63} = \frac{1}{3} \cdot (0,3154) = 0,1051$$

El vector de centralidad del nodo 6 es  $c_6 = (0'0727, 0'1051)$ . Sumando todas las componentes, obtenemos una centralidad del nodo 6 de 0,1778.

Una vez calculada la centralidad de cada uno de los nodos, construimos nuestro vector de centralidad general con cada una de las componentes, el cual resulta de la forma

$$c = (0'0898, 0'0727, 0'1797, 0'2848, 0'195, 0'1778),$$

dando la suma de todos ellos igual a 1.

Finalmente, este vector  $c$  obtenido corresponde con la medida de centralidad de nuestro hipergrafo en cuestión, y cada componente  $c_i$  define la importancia cuantitativa de cada nodo  $i$ .

# 5

## Aplicación

Una vez tratados y abordados todos los fundamentos teóricos y los contenidos sobre el estudio de la centralidad en hipergrafos, es necesario introducirse en su aplicación en la vida cotidiana para la resolución de problemas.

A continuación, mostraremos un ejemplo sobre la resolución de este tipo de problemas ya planteados en el trabajo.

### 5.1. Trabajador más influyente

El departamento de marketing de una empresa está formado por 15 trabajadores, y durante todas las semanas, se llevan a cabo numerosas conferencias. Durante la última semana, se han registrado todas las conferencias realizadas junto con los trabajadores que fueron partícipes en ellas, y se desea determinar quién ha sido el trabajador más influyente y con más importancia en esa semana. La información necesaria para calcularlo se proporciona a continuación.

El registro de los trabajadores del departamento se muestra en La Tabla [5.1](#), mientras que el registro de las conferencias se muestra en La Tabla [5.2](#).

Tabla 5.1: Registro de trabajadores.

	Nombre	Abreviatura		Nombre	Abreviatura
Trabajador 1	Ana	A	Trabajador 9	Irene	I
Trabajador 2	Blanca	B	Trabajador 10	Jimena	J
Trabajador 3	Carlos	C	Trabajador 11	Laura	L
Trabajador 4	Daniel	D	Trabajador 12	Marcos	M
Trabajador 5	Esther	E	Trabajador 13	Nadia	N
Trabajador 6	Fiona	F	Trabajador 14	Óscar	O
Trabajador 7	Gerardo	G	Trabajador 15	Paula	P
Trabajador 8	Hugo	H			

Tabla 5.2: Registro de conferencias.

	Día de la semana	Nº de participantes	Participantes
Conferencia 1	Lunes	2	A,F
Conferencia 2	Lunes	4	B,C,F,G
Conferencia 3	Martes	3	G,L,N
Conferencia 4	Martes	5	H,I,L,N,O
Conferencia 5	Martes	5	C,D,J,M,L
Conferencia 6	Miércoles	7	A,D,E,J,N,O,P
Conferencia 7	Miércoles	4	B,F,M,I
Conferencia 8	Jueves	5	E,F,L,M,O
Conferencia 9	Jueves	2	G,P
Conferencia 10	Jueves	5	A,C,D,H,P
Conferencia 11	Jueves	3	B,E,H
Conferencia 12	Viernes	3	A,B,C
Conferencia 13	Viernes	3	G,H,O

En primer lugar, tenemos que determinar la estructura en la cual organizar el problema planteado. Como tenemos una serie de elementos los cuales se relacionan de forma múltiple, la estructura ideal consiste en un hipergrafo.

El hipergrafo que modela esta situación presenta un vértice por cada trabajador del departamento y una hiperarista por cada conferencia realizada. De esta manera, los trabajadores que han estado presentes en una misma conferencia estarán en una misma hiperarista. Su representación queda reflejada en La Figura 5.1.

La resolución de este problema se puede definir como el cálculo de la centralidad del hipergrafo, puesto que queremos obtener que importancia tiene cada trabajador de forma cuantitativa. El trabajador más influyente y más importante durante esa semana en la cual se han ido registrando los datos será aquel que tenga un mayor valor en su respectiva componente en el vector de centralidad que tendremos que hallar.

Para conseguir nuestro objetivo de forma efectiva, los pasos clave que debemos seguir una vez se tiene el hipergrafo de estudio representado, vistos teóricamente en secciones previas, son construir el grafo lineal asociado al hipergrafo, calcular la centralidad de este para obtener la importancia de sus nodos (equivalentemente, la importancia de las hiperaristas del hipergrafo, que son las conferencias) y posteriormente extenderlo para calcular la importancia de los nodos del hipergrafo (equivalentemente la importancia de los trabajadores).

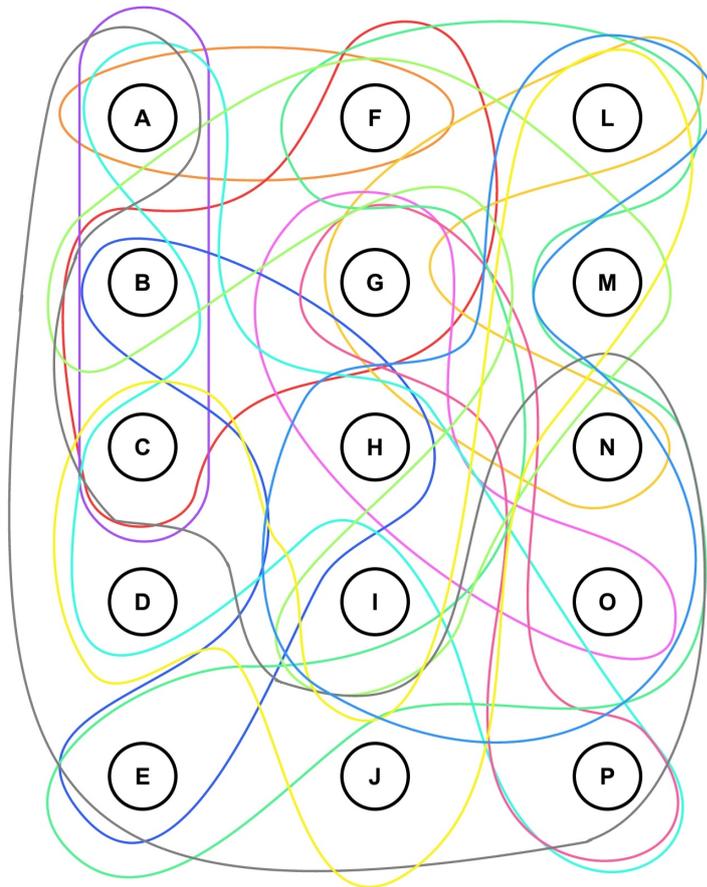


Figura 5.1: Representación gráfica del hipergrafo creado a partir de la información del problema. Los nodos representan los trabajadores y las hiperaristas las conferencias.



Una vez graficado el hipergrafo que modela el problema en La Figura 5.1 y con la ayuda del esquema extraído en La Figura 5.2, el siguiente paso es construir su grafo lineal asociado, en el cual los nodos serán las hiperaristas, es decir, las diferentes conferencias, y dos de estas serán adyacentes si ambas inciden en el hipergrafo en al menos un nodo en común, es decir, si ambas conferencias contienen al menos a un trabajador en común. En La Figura 5.3 podemos ver la representación gráfica del grafo lineal asociado al hipergrafo del problema. Presenta 13 nodos, coincidente con las 13 conferencias o 13 hiperaristas del hipergrafo, los cuales han sido numerados con el orden de las conferencias y con los mismos colores de las hiperaristas del hipergrafo para facilitar posteriormente los cálculos y hacerlo más visual y menos lioso.

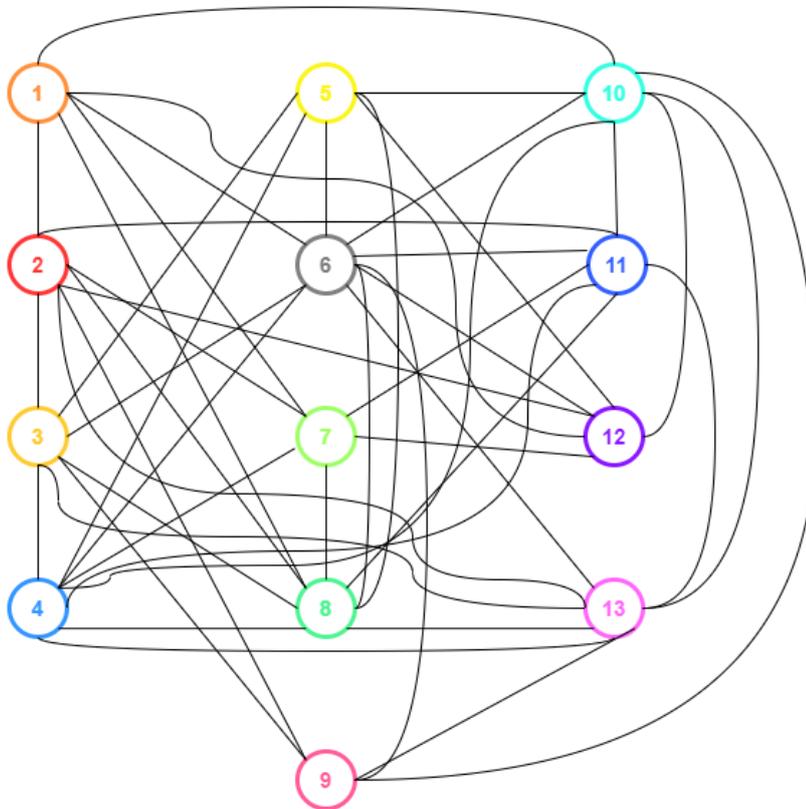


Figura 5.3: Representación gráfica del grafo lineal creado a partir del hipergrafo.

Una vez obtenido el grafo lineal asociado al hipergrafo, lo siguiente es construir su matriz de adyacencia. Como el grafo tiene 13 nodos, será una matriz  $13 \times 13$ , como es un grafo no dirigido, será simétrica, y como un nodo no puede relacionarse consigo mismo ya que no tiene sentido hacer conferencias en las que participe solo una persona, tendrá ceros en la diagonal. Por tanto, su matriz de adyacencia será

$$A(L(G)) = \begin{pmatrix} 0 & 1 & 0 & 0 & 0 & 1 & 1 & 1 & 0 & 1 & 0 & 1 & 0 \\ 1 & 0 & 1 & 0 & 1 & 0 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \\ 0 & 1 & 0 & 1 & 1 & 1 & 0 & 1 & 1 & 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 & 1 & 1 & 1 & 1 & 0 & 1 & 1 & 0 & 1 \\ 0 & 1 & 1 & 1 & 0 & 1 & 1 & 1 & 0 & 1 & 0 & 1 & 0 \\ 1 & 0 & 1 & 1 & 1 & 0 & 0 & 1 & 1 & 1 & 1 & 1 & 1 \\ 1 & 1 & 0 & 1 & 1 & 0 & 0 & 1 & 0 & 0 & 1 & 1 & 0 \\ 1 & 1 & 1 & 1 & 1 & 1 & 1 & 0 & 0 & 0 & 1 & 0 & 1 \\ 0 & 1 & 1 & 0 & 0 & 1 & 0 & 0 & 0 & 1 & 0 & 0 & 1 \\ 1 & 1 & 0 & 1 & 1 & 1 & 0 & 0 & 1 & 0 & 1 & 1 & 1 \\ 0 & 1 & 0 & 1 & 0 & 1 & 1 & 1 & 0 & 1 & 0 & 1 & 1 \\ 1 & 1 & 0 & 0 & 1 & 1 & 1 & 0 & 0 & 1 & 1 & 0 & 0 \\ 0 & 1 & 1 & 1 & 0 & 1 & 0 & 1 & 1 & 1 & 1 & 0 & 0 \end{pmatrix}$$

Se puede apreciar que la matriz de adyacencia obtenida del grafo lineal es cuadrada, simétrica, no negativa y además irreducible. Cuadrada porque tiene el mismo número de columnas que de filas, simétrica porque es igual a su traspuesta debido a que el grafo lineal obtenido es bidireccional, no negativa porque todos los coeficientes son mayores o igual a cero, e irreducible porque si construimos el grafo dirigido asociado a la matriz de adyacencia obtenemos un grafo fuertemente conexo, es decir, existe un camino entre cualquier par de nodos. Al ser un grafo no muy grande, se puede comprobar nodo a nodo que se cumple considerando en el grafo lineal construido anteriormente que cada arista implica dos aristas, una en cada sentido (si la cantidad de nodos fuese muy elevada, esta condición se puede verificar mediante cualquier librería en lenguajes de programación en apenas segundos). Por tanto, tenemos una matriz que cumple todas las premisas del Teorema de Perron-Frobenius. De esta manera, este teorema nos garantiza la existencia de un autovalor cuyo valor es el máximo en valor absoluto de todos los autovalores de la matriz y que tiene multiplicidad 1, llamado autovalor dominante, y la existencia y unicidad de un vector asociado a dicho autovalor tal que todas sus componentes son positivas y además su norma en 1 es igual a 1. Al ser simétrica, da igual coger esta matriz de adyacencia o su traspuesta para emplear el teorema.

Para obtener los autovalores de la matriz se puede utilizar cualquier método estándar de obtención de autovalores de una matriz. Por ejemplo, se puede seguir la resolución de la ecuación característica  $\det(A(L(G)) - \lambda I) = 0$ , siendo  $A(L(G))$  la matriz de adyacencia del grafo lineal,  $I$  la matriz identidad del mismo tamaño que  $A(L(G))$  y  $\lambda$  el autovalor que se busca. Los valores de  $\lambda$  coinciden con las raíces del polinomio característico, y tendrá tantas como indique su grado. Una vez se tiene el polinomio, las raíces pueden hallarse mediante el método de Ruffini.

Tras aplicar este método, los autovalores obtenidos son:

$$\begin{array}{ll}
 \lambda_1 = -3,2085 & \lambda_8 = 0,1000 \\
 \lambda_2 = -2,6173 & \lambda_9 = 0,5102 \\
 \lambda_3 = -2,4219 & \lambda_{10} = 0,7569 \\
 \lambda_4 = -1,7411 & \lambda_{11} = 1,3886 \\
 \lambda_5 = -1,4517 & \lambda_{12} = 2,2169 \\
 \lambda_6 = -0,9332 & \lambda_{13} = 8,0541 \\
 \lambda_7 = -0,6530 &
 \end{array}$$

De todos estos autovalores obtenidos, nos debemos quedar con el autovalor mayor en valor absoluto, que corresponde con  $\lambda = 8,0541$ . A partir de este autovalor dominante, vamos a calcular su vector asociado, que es el que nos interesa.

Para obtener el autovector asociado al autovalor de una matriz se puede utilizar cualquier método estándar de obtención de autovectores a partir de un autovalor de una matriz. Por ejemplo, se puede seguir la resolución del sistema de ecuaciones lineal  $(A - \lambda I)\mathbf{x} = \mathbf{0}$ , siendo  $A(L(G))$  la matriz de adyacencia del grafo lineal,  $I$  la matriz identidad del mismo tamaño que  $A(L(G))$ ,  $\lambda$  el autovalor previamente calculado y  $\mathbf{x}$  el autovector que queremos. Una vez hallado, lo normalizamos en norma 1 para que todas las componentes sumen 1, dividiendo cada componente entre el total de la suma de todas ellas.

Tras aplicar este método, se obtiene que el vector normalizado es

$$x = (0'0615, 0'0913, 0'07, 0'0801, 0'0802, 0'0928, 0'0685, 0'0876, \\
 0'052, 0'0855, 0'0813, 0'0697, 0'0795).$$

La suma de todas sus componentes es igual a 1, y cumple todos los requisitos.

Este vector obtenido, llamado vector de Perron, nos proporciona la importancia de cada nodo del grafo lineal, lo que implica que en este momento conocemos la importancia de cada hiperarista en el hipergrafo, puesto que los nodos en el grafo lineal equivalen a las hiperaristas en el hipergrafo. Es decir, conocemos la importancia de cada conferencia. Ahora solo nos falta calcular la de cada trabajador a partir de la de las conferencias.

Llegados a este momento, tenemos que aplicar la fórmula fundamental de extensión de centralidad a hipergrafos para cada nodo y así ir calculando la centralidad de cada uno. Cada centralidad va a constituir una respectiva componente en el vector de centralidad del hipergrafo que buscamos. Los cálculos con sus co-

rrespondientes resultados los llevamos a cabo a continuación:

- Centralidad del nodo A:

$$c_{A2} = \frac{1}{2} \cdot (0,0615) = 0,0307$$

$$c_{A3} = \frac{1}{3} \cdot (0,0697) = 0,0232$$

$$c_{A4} = \frac{1}{4} \cdot (0) = 0$$

$$c_{A5} = \frac{1}{5} \cdot (0,0855) = 0,0171$$

$$c_{A6} = \frac{1}{6} \cdot (0) = 0$$

$$c_{A7} = \frac{1}{7} \cdot (0,0928) = 0,0132$$

El vector de centralidad del nodo A es  $c_A = (0'0307, 0'0232, 0, 0'0171, 0, 0'0132)$ . Sumando todas las componentes, obtenemos una centralidad del nodo A de 0,0842.

- Centralidad del nodo B:

$$c_{B2} = \frac{1}{2} \cdot (0) = 0$$

$$c_{B3} = \frac{1}{3} \cdot (0,0813 + 0,0697) = 0,0503$$

$$c_{B4} = \frac{1}{4} \cdot (0,0913 + 0,0685) = 0,0399$$

$$c_{B5} = \frac{1}{5} \cdot (0) = 0$$

$$c_{B6} = \frac{1}{6} \cdot (0) = 0$$

$$c_{B7} = \frac{1}{7} \cdot (0) = 0$$

El vector de centralidad del nodo B es  $c_B = (0, 0'0503, 0'0399, 0, 0, 0)$ . Sumando todas las componentes, obtenemos una centralidad del nodo B de 0,0902.

- Centralidad del nodo C:

$$c_{C2} = \frac{1}{2} \cdot (0) = 0$$

$$c_{C3} = \frac{1}{3} \cdot (0,0697) = 0,0232$$

$$c_{C4} = \frac{1}{4} \cdot (0,0913) = 0,0228$$

$$c_{C5} = \frac{1}{5} \cdot (0,0802 + 0,0855) = 0,0331$$

$$c_{C6} = \frac{1}{6} \cdot (0) = 0$$

$$c_{C7} = \frac{1}{7} \cdot (0) = 0$$

El vector de centralidad del nodo C es  $c_C = (0, 0'0232, 0'0228, 0'0331, 0, 0)$ . Sumando todas las componentes, obtenemos una centralidad del nodo C de 0,0791.

- Centralidad del nodo D:

$$c_{D2} = \frac{1}{2} \cdot (0) = 0$$

$$c_{D3} = \frac{1}{3} \cdot (0) = 0$$

$$c_{D4} = \frac{1}{4} \cdot (0) = 0$$

$$c_{D5} = \frac{1}{5} \cdot (0,0802 + 0,0855) = 0,0331$$

$$c_{D6} = \frac{1}{6} \cdot (0) = 0$$

$$c_{D7} = \frac{1}{7} \cdot (0,928) = 0,0132$$

El vector de centralidad del nodo D es  $c_D = (0, 0, 0, 0'0331, 0, 0'0132)$ . Sumando todas las componentes, obtenemos una centralidad del nodo D de 0,0463.

- Centralidad del nodo E:

$$c_{E2} = \frac{1}{2} \cdot (0) = 0$$

$$c_{E3} = \frac{1}{3} \cdot (0,0813) = 0,0271$$

$$c_{E4} = \frac{1}{4} \cdot (0) = 0$$

$$c_{E5} = \frac{1}{5} \cdot (0,0876) = 0,0175$$

$$c_{E6} = \frac{1}{6} \cdot (0) = 0$$

$$c_{E7} = \frac{1}{7} \cdot (0,928) = 0,0132$$

El vector de centralidad del nodo E es  $c_E = (0, 0'0271, 0, 0'0175, 0, 0'0132)$ . Sumando todas las componentes, obtenemos una centralidad del nodo E de 0,0578.

- Centralidad del nodo F:

$$c_{F2} = \frac{1}{2} \cdot (0,0615) = 0,0307$$

$$c_{F3} = \frac{1}{3} \cdot (0) = 0$$

$$c_{F4} = \frac{1}{4} \cdot (0,0913 + 0,0685) = 0,0399$$

$$c_{F5} = \frac{1}{5} \cdot (0,0876) = 0,0175$$

$$c_{F6} = \frac{1}{6} \cdot (0) = 0$$

$$c_{F7} = \frac{1}{7} \cdot (0) = 0$$

El vector de centralidad del nodo F es  $c_F = (0'0307, 0, 0'0399, 0'0175, 0, 0)$ . Sumando todas las componentes, obtenemos una centralidad del nodo F de 0,0881.

- Centralidad del nodo G:

$$c_{G2} = \frac{1}{2} \cdot (0,052) = 0,026$$

$$c_{G3} = \frac{1}{3} \cdot (0,07 + 0'0795) = 0,0498$$

$$c_{G4} = \frac{1}{4} \cdot (0,0913) = 0,0228$$

$$c_{G5} = \frac{1}{5} \cdot (0) = 0$$

$$c_{G6} = \frac{1}{6} \cdot (0) = 0$$

$$c_{G7} = \frac{1}{7} \cdot (0) = 0$$

El vector de centralidad del nodo G es  $c_G = (0'026, 0'0498, 0'0228, 0, 0, 0, 0)$ . Sumando todas las componentes, obtenemos una centralidad del nodo G de 0,0986.

- Centralidad del nodo H:

$$c_{H2} = \frac{1}{2} \cdot (0) = 0$$

$$c_{H3} = \frac{1}{3} \cdot (0,0813 + 0,0795) = 0,0536$$

$$c_{H4} = \frac{1}{4} \cdot (0) = 0$$

$$c_{H5} = \frac{1}{5} \cdot (0,0801 + 0,0855) = 0,0331$$

$$c_{H6} = \frac{1}{6} \cdot (0) = 0$$

$$c_{H7} = \frac{1}{7} \cdot (0) = 0$$

El vector de centralidad del nodo H es  $c_H = (0, 0'0536, 0, 0'0331, 0, 0, 0)$ . Sumando todas las componentes, obtenemos una centralidad del nodo H de 0,0867.

- Centralidad del nodo I:

$$c_{I2} = \frac{1}{2} \cdot (0) = 0$$

$$c_{I3} = \frac{1}{3} \cdot (0) = 0$$

$$c_{I4} = \frac{1}{4} \cdot (0,0685) = 0,0171$$

$$c_{I5} = \frac{1}{5} \cdot (0,0801) = 0,016$$

$$c_{I6} = \frac{1}{6} \cdot (0) = 0$$

$$c_{I7} = \frac{1}{7} \cdot (0) = 0$$

El vector de centralidad del nodo I es  $c_I = (0, 0, 0,0171, 0,016, 0, 0)$ . Sumando todas las componentes, obtenemos una centralidad del nodo I de 0,0331.

- Centralidad del nodo J:

$$c_{J2} = \frac{1}{2} \cdot (0) = 0$$

$$c_{J3} = \frac{1}{3} \cdot (0) = 0$$

$$c_{J4} = \frac{1}{4} \cdot (0) = 0$$

$$c_{J5} = \frac{1}{5} \cdot (0,0802) = 0,016$$

$$c_{J6} = \frac{1}{6} \cdot (0) = 0$$

$$c_{J7} = \frac{1}{7} \cdot (0,928) = 0,0132$$

El vector de centralidad del nodo J es  $c_J = (0, 0, 0, 0,016, 0, 0,0132)$ . Sumando todas las componentes, obtenemos una centralidad del nodo J de 0,0292.

- Centralidad del nodo L:

$$c_{L2} = \frac{1}{2} \cdot (0) = 0$$

$$c_{L3} = \frac{1}{3} \cdot (0,07) = 0,0233$$

$$c_{L4} = \frac{1}{4} \cdot (0) = 0$$

$$c_{L5} = \frac{1}{5} \cdot (0,0801 + 0,0802 + 0,0876) = 0,0496$$

$$c_{L6} = \frac{1}{6} \cdot (0) = 0$$

$$c_{L7} = \frac{1}{7} \cdot (0) = 0$$

El vector de centralidad del nodo L es  $c_L = (0, 0,0233, 0, 0,0496, 0, 0)$ . Sumando todas las componentes, obtenemos una centralidad del nodo L de 0,0729.

- Centralidad del nodo M:

$$c_{M2} = \frac{1}{2} \cdot (0) = 0$$

$$c_{M3} = \frac{1}{3} \cdot (0) = 0$$

$$c_{M4} = \frac{1}{4} \cdot (0,0685) = 0,0171$$

$$c_{M5} = \frac{1}{5} \cdot (0,0802 + 0,0876) = 0,0336$$

$$c_{M6} = \frac{1}{6} \cdot (0) = 0$$

$$c_{M7} = \frac{1}{7} \cdot (0) = 0$$

El vector de centralidad del nodo M es  $c_M = (0, 0, 0,0171, 0,0336, 0, 0)$ . Sumando todas las componentes, obtenemos una centralidad del nodo M de 0,0507.

- Centralidad del nodo N:

$$c_{N2} = \frac{1}{2} \cdot (0) = 0$$

$$c_{N3} = \frac{1}{3} \cdot (0,07) = 0,0233$$

$$c_{N4} = \frac{1}{4} \cdot (0) = 0$$

$$c_{N5} = \frac{1}{5} \cdot (0,0801) = 0,016$$

$$c_{N6} = \frac{1}{6} \cdot (0) = 0$$

$$c_{N7} = \frac{1}{7} \cdot (0,928) = 0,0132$$

El vector de centralidad del nodo N es  $c_N = (0, 0'0233, 0, 0'016, 0, 0'0132)$ . Sumando todas las componentes, obtenemos una centralidad del nodo N de 0,0525.

- Centralidad del nodo O:

$$c_{O2} = \frac{1}{2} \cdot (0) = 0$$

$$c_{O3} = \frac{1}{3} \cdot (0,0795) = 0,0265$$

$$c_{O4} = \frac{1}{4} \cdot (0) = 0$$

$$c_{O5} = \frac{1}{5} \cdot (0,0801 + 0,0876) = 0,0335$$

$$c_{O6} = \frac{1}{6} \cdot (0) = 0$$

$$c_{O7} = \frac{1}{7} \cdot (0,928) = 0,0132$$

El vector de centralidad del nodo O es  $c_O = (0, 0'0265, 0, 0'0335, 0, 0'0132)$ . Sumando todas las componentes, obtenemos una centralidad del nodo O de 0,0732.

- Centralidad del nodo P:

$$c_{P2} = \frac{1}{2} \cdot (0,052) = 0,026$$

$$c_{P3} = \frac{1}{3} \cdot (0) = 0$$

$$c_{P4} = \frac{1}{4} \cdot (0) = 0$$

$$c_{P5} = \frac{1}{5} \cdot (0,0855) = 0,0171$$

$$c_{P6} = \frac{1}{6} \cdot (0) = 0$$

$$c_{P7} = \frac{1}{7} \cdot (0,928) = 0,132$$

El vector de centralidad del nodo P es  $c_P = (0'026, 0, 0, 0'0171, 0, 0'0132)$ . Sumando todas las componentes, obtenemos una centralidad del nodo P de 0,0563.

Una vez calculada la centralidad de cada uno de los nodos, construimos nuestro vector de centralidad general con cada una de las componentes, el cual resulta de la forma

$$c = (0'0842, 0'0902, 0'0791, 0'0463, 0'0578, 0'0881, 0'0986, 0'0867, 0'0331, 0'0292, 0'0729, 0'0507, 0'0525, 0'0732, 0'0563),$$

dando la suma de todos ellos igual a 1.

Obsérvense los resultados gráficamente en La Tabla 5.3, en la que hay cuatro columnas que relacionan la componente del vector de centralidad con su correspondiente nodo, nombre del trabajador e importancia calculada.

Tabla 5.3: Resultados obtenidos del problema del trabajador más influyente.

Componente del vector	Nodo	Trabajador asociado	Importancia[0-1]
1	A	Ana	0.0842
2	B	Blanca	0.0902
3	C	Carlos	0.0791
4	D	Daniel	0.0463
5	E	Esther	0.0578
6	F	Fiona	0.0881
<b>7</b>	<b>G</b>	<b>Gerardo</b>	<b>0.0986</b>
8	H	Hugo	0.0867
9	I	Irene	0.0331
10	J	Jimena	0.0292
11	L	Laura	0.0729
12	M	Marcos	0.0507
13	N	Nadia	0.0525
14	O	Oscar	0.0732
15	P	Paula	0.0563

Suma = 1

Tabla 5.4: Ranking de influencia tras los resultados.

Posición	Nombre	Importancia
1	Gerardo	0.0986
2	Blanca	0.0902
3	Fiona	0.0881
4	Hugo	0.0867
5	Ana	0.0842
6	Carlos	0.0791
7	Óscar	0.0732
8	Laura	0.0729

Posición	Nombre	Importancia
9	Esther	0.0578
10	Paula	0.0563
11	Nadia	0.0525
12	Marcos	0.0507
13	Daniel	0.0463
14	Irene	0.0331
15	Jimena	0.0292

Es cierto que solo nos solicitaban el trabajador más influyente, pero resulta interesante poder apreciar de un solo vistazo la posición que ocupa el resto de trabajadores. Por ello, en La Tabla 5.4 se proporciona un ranking de importancia de todos los trabajadores durante esa semana en las conferencias, desde el más influyente en la posición 1 hasta el menos influyente en la posición 15.

Tras esto, se puede apreciar como el nodo G, asociado a la componente 7 del vector, es el que presenta el máximo valor, 0'0986. Por ello, este nodo, que hace referencia a Gerardo, es el más influyente.

Por tanto, podemos concluir que tras la semana de conferencias en el departamento de marketing de la empresa, Gerardo es el trabajador que ha tenido una mayor influencia e importancia.



# 6

## Conclusiones y trabajos futuros

### 6.1. Conclusiones

A lo largo de los cinco años de mi formación en el grado de matemáticas, he notado que el Teorema de Perron-Frobenius ha pasado prácticamente desapercibido. A pesar de su potencial y relevancia, este teorema ha sido escasamente abordado durante mi trayectoria académica, únicamente siendo nombrado un par de veces. Llama la atención cómo una herramienta tan valiosa, especialmente en el cálculo de la centralidad en grafos e hipergrafos y con una inmensa variedad de aplicaciones en numerosas áreas, ha recibido una atención limitada en algunas asignaturas. No obstante, al profundizar en su análisis para llevar a cabo la realización de este trabajo, he llegado a apreciar su versatilidad y aplicabilidad, llegando a ser vital para el cálculo de la centralidad en hipergrafos.

### 6.2. Trabajos futuros

En el ámbito de posibles futuras investigaciones, considerando mi trabajo actual basado en la centralidad de hipergrafos en cuanto a vectores propios, surge como un rumbo propicio la exploración de la centralidad basada en la intermediación, también conocida como *betweenness*. Mientras que la investigación de este trabajo se ha centrado en la importancia de los nodos según sus vectores propios, el estudio de la centralidad de intermediación podría proporcionar una perspectiva diferente y enriquecedora. Investigar cómo se distribuye la importan-

cia en función de la participación de los nodos en las rutas más cortas entre otros nodos podría manifestar patrones interesantes en la estructura de hipergrafos. Este enfoque no solo ampliaría el alcance de mi trabajo actual, sino que también contribuiría a una comprensión más completa de las dinámicas y relaciones en hipergrafos, ofreciendo así una perspectiva más integral sobre la centralidad en estas estructuras complejas.

# Bibliografía y referencias

- [1] B. Hopkins and R. Wilson, *The Truth About Konigsberg*. The College Mathematics Journal, 2005, vol. 35.
- [2] D. Konig, *Theorie der endlichen und unendlichen Graphen*. American Mathematical Soc., 2001.
- [3] S. Boccaletti, P. De Lellis, C. del Genio, K. Alfaro-Bittner, R. Criado, S. Jalan, and M. Romance, *The structure and dynamics of networks with higher order interactions*. Physics Reports 1018 (2023) 1–64.
- [4] A. Bretto, *Hypergraph Theory: An Introduction*. Springer International Publishing Switzerland, 2013.
- [5] K. Engø-Monsen and G. S. Canright, *Introducing Network Analysis*, 2008.
- [6] R. Criado, M. Romance, and L. E. Sola, *Teoría de Perron-Frobenius: importancia, poder y centralidad*. La Gaceta de la RSME, vol. 17, no. 3, pp. 485–514, 2014.
- [7] F. A. Rodrigues, *Network centrality: an introduction*. p. 3–12, 2019.
- [8] A. Borobia and U. R. Trías, *A geometric proof of the Perron-Frobenius theorem*. Rev. Mat. Univ. Complut. Madrid, vol. 5, no. 1, pp. 57–63, 1992.
- [9] K. Kovalenko, M. Romance, E. Vasilyeva, D. Aleja, R. Criado, D. Musatov, A. Raigorodskii, J. Flores, I. Samoylenko, K. Alfaro-Bittner, M. Perc, and S. Boccaletti, *Vector centrality in hypergraphs*. Chaos, Solitons and Fractals 162 (2022) 112397.



# Anexos





## Tabla de tiempos

Tarea	Fecha	Horas
Reuniones con tutor	Junio	15
Lectura de documentos para la selección del tema	Junio	10
Búsqueda plantilla e intalación Latex	Martes 13 Junio	3
Lectura de documentación	Miércoles 14 Junio	3
Lectura de documentación	Jueves 8 Junio	2
Lectura de documentación	Viernes 9 Junio	3
Lectura de documentación	Sábado 10 Junio	3
Redacción de la sección 1 y 2	Domingo 11 Junio	5
Lectura de documentación	Martes 13 Junio	2
Redacción de la sección 3.1	Jueves 15 Junio	4
Redacción de la sección 3.1	Viernes 16 Junio	3
Redacción de la sección 3.2	Domingo 18 Junio	3
Redacción de la sección 3.2	Martes 20 Junio	2
Revisión y corrección de fallos de la sección 3	Miércoles 21 Junio	4

Figura A.1: Parte 1 de la tabla de tiempos.

Lectura de documentación	Martes 22 Ago	2
Lectura de documentación	Miércoles 23 Ago	3
Lectura de documentación	Jueves 24 Ago	3
Realización de ejercicios prácticos a papel para comprobar entendimiento	Lunes 28 Ago	3
Realización de ejercicios prácticos a papel para comprobar entendimiento	Martes 29 Ago	2
Realización de ejercicios prácticos a papel para comprobar entendimiento	Miércoles 31 Ago	2
Revisión de lo que llevo actualmente	Viernes 1 Sept	4
Reunión con el tutor	Jueves 7 Sept	3
Repaso, reestructuración y corrección de errores	Lunes 11 Sept	2
Redacción 5.1	Martes 12 Sept	3
Redacción 5.1	Miércoles 13 Sept	3
Redacción 5.1	Jueves 14 Sept	3
Redacción 5.1	Viernes 15 Sept	3
Lectura de documentación	Lunes 18 Sept	3
Lectura de documentación	Martes 19 Sept	3
Redacción 4.1	Miércoles 20 Sept	2
Redacción 4.1	Jueves 21 Sept	3
Redacción 4.1	Viernes 22 Sept	1
Revisión de lo escrito en 4.1	Domingo 24 Sept	2
Lectura de documentación	Lunes 25 Sept	3
Lectura de documentación	Martes 26 Sept	3
Lectura de documentación	Jueves 28 Sept	2
Redacción 4.1.1	Sábado 30 Sept	2

Figura A.2: Parte 2 de la tabla de tiempos.

Capítulo A. Tabla de tiempos

---

Redacción 4.1.1	Lunes 2 Oct	2
Revisión de lo escrito en 4.1.1	Miércoles 4 Oct	3
Reunión con tutor	Jueves 5 Oct	3
Redacción 4.1.1	Sábado 7 Oct	4
Redacción 4.1.1	Lunes 9 Oct	3
Redacción 4.1.1	Martes 10 Oct	3
Revisión de lo escrito en 4.1.1	Jueves 12 Oct	3
Lectura de documentación	Viernes 13 Oct	2
Lectura de documentación	Sábado 14 Oct	1
Lectura de documentación	Domingo 15 Oct	3
Redacción 4.2	Lunes 16 Oct	3
Redacción 4.2	Miércoles 18 Oct	3
Redacción 4.2	Jueves 19 Oct	2
Revisión de lo escrito en 4.2	Viernes 20 Oct	3
Redacción 4.2	Domingo 22 Oct	4
Redacción 4.2	Martes 24 Oct	3
Revisión de lo escrito en 4.2	Miércoles 25 Oct	3
Redacción 4.2	Jueves 26 Oct	2
Lectura detenida y corrección de errores puntos 1-2	Martes 7 Nov	3
Lectura detenida y corrección de errores punto 3	Viernes 10 Nov	3
Lectura detenida y corrección de errores punto 4	Sábado 11 Nov	4
Lectura detenida y corrección de errores punto 5	Domingo 12 Nov	2
Reunión con el tutor	Miércoles 15 Nov	2

Figura A.3: Parte 3 de la tabla de tiempos.

Reunión con el tutor	Miércoles 15 Nov	2
Arreglos sección 5	Miércoles 15 Nov	3
Redacción de conclusiones y trabajos futuros	Jueves 16 Nov	2
Redacción de agradecimientos y resumen	Viernes 17 Nov	2
Lectura de la demostración del Teorema de Perron-Frobenius y arreglos generales	Sábado 18 Nov	5
Lectura de la demostración del Teorema de Perron-Frobenius	Domingo 19 Nov	5
Redacción de la demostración del Teorema de Perron-Frobenius	Lunes 20 Nov	3
Redacción de la demostración del Teorema de Perron-Frobenius	Martes 21 Nov	2
Redacción de la demostración del Teorema de Perron-Frobenius	Miércoles 22 Nov	3
Revisión de todo el documento	Jueves 23 Nov	3
Correcciones tras revisión del tutor	Diciembre	15

<b>TOTAL HORAS</b>
<b>229</b>

Figura A.4: Parte 4 de la tabla de tiempos.

