

Universidad
Rey Juan Carlos

Escuela Técnica Superior
de Ingeniería Informática

Grado en Matemáticas

Curso 2023-2024

Trabajo Fin de Grado

**FUNDAMENTOS DEL TEOREMA DE
PERRON-FROBENIUS Y SU UTILIDAD EN EL
ALGORITMO PAGERANK**

Autor: Guillermo Grande Santi

Tutor: Regino Criado Herrero

© 2024 **Guillermo Grande Santi**

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

Quiero expresar mi profundo agradecimiento a mi familia, cuyo apoyo incondicional ha sido la fuerza impulsora detrás de este logro y me han acompañado y apoyado a lo largo de toda mi carrera. Las lecciones, la confianza y el amor que mi familia me ha ofrecido resultan fundamentales e insustituibles. Su presencia constante ha sido mi mayor motivación, inspirándome a esforzarme más allá de mis límites y a perseverar en la búsqueda del conocimiento.

También agradezco en especial a mi pareja, Rebeca, cuya presencia ha sido una fuente inigualable de apoyo a lo largo de mi carrera académica. Ha demostrado comprenderme mejor que nadie, ofreciendo un respaldo incondicional que ha sido esencial para superar obstáculos y convirtiéndola en un pilar fundamental en el éxito de este logro.

A mis amigos y compañeros de clase, gracias por compartir este viaje académico conmigo y por las experiencias inolvidables que hemos vivido juntos. Ha sido todo un placer aprender a vuestro lado.

Agradezco a mi tutor, Regino Criado, por su sabia orientación y por introducirme en un tema que ha despertado un profundo interés en mí.

Resumen

El **Teorema de Perron-Frobenius**, una herramienta fundamental en la teoría de matrices, ha encontrado aplicaciones significativas en diversos campos, y una de las áreas más destacadas es su relación con el algoritmo PageRank. Este teorema, desarrollado por Oskar Perron y Ferdinand Frobenius, establece propiedades clave de las matrices no negativas e irreducibles, un tipo especial de matrices que surge en una variedad de contextos matemáticos y científicos.

En el ámbito de la informática y la web, **el algoritmo PageRank**, creado por Larry Page y Sergey Brin en el contexto de Google, utiliza el Teorema de Perron-Frobenius de manera crucial. PageRank es un algoritmo de clasificación de páginas web que asigna a cada página un valor numérico que representa su importancia relativa en la red.

En este trabajo de fin de grado, se explorará en detalle la utilidad del Teorema de Perron-Frobenius en el contexto del algoritmo PageRank. Tras haber expuesto unos fundamentos de Álgebra Lineal y Teoría de Grafos, investigaré cómo la teoría de Perron-Frobenius afecta a las clasificaciones y, en concreto, garantiza la existencia y unicidad de la solución del PageRank.

Índice de contenidos

Índice de tablas	IX
Índice de figuras	XI
Índice de métodos	XIII
1. Introducción	1
2. Objetivos	3
2.1. Fundamentos de la Teoría de Perron-Frobenius	3
2.2. Utilidad en el Algoritmo PageRank	4
3. Contenidos principales	5
3.1. Fundamentos de Álgebra Lineal	5
3.2. Fundamentos de Teoría de Grafos	9
3.2.1. Conceptos Básicos	9
3.2.2. Relación entre Conectividad y Matrices Irreducibles	12
3.3. Teoría de Perron-Frobenius	19
3.3.1. Teorema de Perron-Frobenius	19
3.3.2. Aplicación a un Problema Real	27
4. Utilidad en el Algoritmo PageRank	33
4.1. El Modelo de la Red	34
4.2. El Surfista Aleatorio	36
4.3. Cadenas de Márkov y Matrices de transición	38
4.3.1. Conceptos Básicos	38
4.3.2. Ejemplo	42
4.4. PageRank	46
4.4.1. Problema con los Nodos Dangling y su Ajuste Correspondiente	46
4.4.2. Teorema de Perron-Frobenius y PageRank: Una Conexión Fundamental	48
4.4.3. Definición Formal de PageRank	48
4.5. Método de las potencias	50

5. Conclusiones y trabajos futuros	53
Bibliografía	57

Índice de tablas

3.1. Tabla de clasificación del campeonato por victorias en los partidos	28
3.2. Tabla de victorias de cada equipo	29

Índice de figuras

3.1. Ejemplo de un grafo no dirigido y su matriz de adyacencia	10
3.2. Ejemplo de un grafo dirigido y su matriz de adyacencia	10
3.3. Grafo fuertemente conexo - Matriz de adyacencia irreducible	18
3.4. Grafo no fuertemente conexo - Matriz de adyacencia reducible	18
3.5. Ejemplo de un grafo dirigido y su matriz de adyacencia	24
3.6. Ejemplo de un grafo dirigido y su matriz de adyacencia	26
3.7. Ejemplo Partidos NBA Fuertemente Conexo	30
4.1. Ejemplo 1 - Importancia Páginas Web	34
4.2. Ejemplo 2 - Importancia Páginas Web	35
4.3. Ejemplo Paseo Aleatorio	37
4.4. Ejemplo Grafo Web - Matrices de transición	42

Índice de métodos

3.1. Métodos de Perron-Frobenius.py	31
3.2. Datos devueltos por Perron-Frobenius.py	31
3.3. Más Métodos de Perron-Frobenius.py	32
3.4. Resultado Perron-Frobenius.py	32
4.1. Resultado de transition-matrix.py	44

1

Introducción

El campo de las matemáticas alberga una diversidad de teoremas cuyas aplicaciones trascienden la percepción general, a pesar de su relevancia. Este estudio se enfocará en uno de dichos teoremas, específicamente el **Teorema de Perron-Frobenius**. Aunque no goza de una notoriedad generalizada, este teorema se erige como un fundamento crucial para una amplia gama de investigaciones vinculadas a disciplinas como la economía, las ciencias sociales, la biología, la ecología y la informática.

El propósito de este trabajo consiste en desarrollar la Teoría de Perron-Frobenius, la cual, basada en los principios de la Teoría de Grafos y el Álgebra Lineal, arroja resultados que posteriormente se analizarán en términos de su utilidad principal: la capacidad para establecer 'rankings' o clasificaciones. En otras palabras, la habilidad de **ordenar** nodos en una red de acuerdo con su **relevancia**. Esta capacidad encuentra utilidad en diversas áreas, desde cuantificar la influencia de individuos en redes sociales hasta la creación de algoritmos para la jerarquización de páginas web.

Un claro ejemplo es el algoritmo **PageRank**, utilizado por el motor de búsqueda Google, el cual se impuso sobre otros buscadores en 1998, fundamentándose en este teorema matemático. Este algoritmo, desarrollado por Larry Page y Sergey Brin, ha revolucionado la forma en que evaluamos y jerarquizamos la información en la era digital. Por este motivo, se explorará la interconexión entre el Teorema de Perron-Frobenius y el algoritmo PageRank, con el objetivo de detallar cómo Brin y Page desarrollaron un modelo integral para la World Wide Web, permitiendo una utilización efectiva de dicho teorema.

2

Objetivos

En este trabajo de fin de grado existen dos objetivos principales que guiarán la investigación. El primero se centra en la exploración y desarrollo de los fundamentos de la teoría de Perron-Frobenius. Por otro lado, el segundo objetivo se enfoca en desentrañar el desarrollo del algoritmo PageRank, con especial énfasis en cómo se modeló la red para poder aplicar el teorema anterior. Ambos objetivos, en conjunto, constituyen la base esencial para el desarrollo de un marco teórico sólido y la comprensión profunda de la teoría de Perron-Frobenius, el algoritmo PageRank y su intersección.

2.1. Fundamentos de la Teoría de Perron-Frobenius

En primer lugar, se tiene como objetivo establecer los fundamentos teóricos esenciales para comprender el Teorema de Perron-Frobenius. Se abordarán conceptos clave de la **Teoría de Grafos** y del **Álgebra Lineal** que sirven como cimientos para el desarrollo y la comprensión de la Teoría de Perron-Frobenius y que se utilizarán a lo largo de todo el trabajo. En concreto, se busca explorar las propiedades de las **matrices irreducibles** y su conexión con la conectividad de grafos dirigidos, explicar cómo el teorema de Perron-Frobenius nos ayuda en la **identificación de un vector de ordenaciones**, incluir una **demostración del teorema**, apoyar lo explicado con **ejemplos prácticos** y, en general, proporcionar una base sólida para la posterior utilización del teorema en el ámbito del algoritmo PageRank.

2.2. Utilidad en el Algoritmo PageRank

En segundo lugar, se pretende examinar la implementación práctica del Teorema de Perron-Frobenius en la concepción del ampliamente reconocido algoritmo PageRank. En este sentido, se desglosarán los pasos seguidos por Brin y Page para modelar la World Wide Web de manera que el teorema fuera aplicable y asegurara la obtención del vector de ordenaciones deseado para las páginas web. Para ello, se buscará explicar el proceso de **modelado de la red** como un **extenso grafo dirigido**, destacando conceptos clave como la **hipótesis del surfista aleatorio** o conceptos básicos sobre las **cadenas de Márkov**, que adquieren relevancia para la generación de **matrices de transición estocásticas**. Estas matrices son esenciales para derivar el **vector PageRank**, que a su vez proporciona la **ordenación** deseada de las **páginas web**. Además, se abordará un método para abordar la magnitud de las matrices, conocido como el **método de las potencias**, y todo ello con **ejemplos** ilustrativos.

3

Contenidos principales

En la secciones [3.1](#) y [3.2](#) se establecerán las bases teóricas que se utilizarán posteriormente para desarrollar la teoría de Perron-Frobenius. Para establecer estos conceptos se han utilizado artículos como [\[1\]](#) y [\[2\]](#).

3.1. Fundamentos de Álgebra Lineal

En principio, revisaremos algunos principios matriciales fundamentales que serán empleados en el desarrollo de la Teoría de Perron-Frobenius. Se destacará especialmente la consideración de matrices irreducibles, matrices no negativas, y ciertos aspectos relacionados con los autovalores de una matriz.

Definición 3.1.1. *Dada una matriz cuadrada A , un número complejo λ se llama autovalor de A si existe un vector no nulo \mathbf{v} tal que*

$$A\mathbf{v} = \lambda\mathbf{v}$$

El vector \mathbf{v} se llama autovector correspondiente al autovalor λ . Además, los autovalores de una matriz A son las raíces del polinomio característico $p(\lambda) = \det(A - \lambda I)$, donde \det denota el determinante e I es la matriz identidad.

Definición 3.1.2. *Dada una matriz cuadrada A , el conjunto de autovalores de A se denota por $\sigma(A)$, es llamado espectro de A y viene dado por:*

$$\sigma(A) = \{\lambda : \det(A - \lambda I) = 0\}$$

Por otro lado, el máximo de los autovalores, también conocido como el **radio espectral** de A , se denota por $\rho(A)$ y está dado por:

$$\rho(A) = \max_{\lambda_i \in \sigma(A)} \{|\lambda_i|\}$$

Proposición 3.1.1. *Dada una matriz cuadrada A cualquiera y su traspuesta A^T , el espectro de A y A^T es el mismo, es decir, la matriz A y su traspuesta A^T tienen los mismos autovalores, y por tanto, también el mismo radio espectral.*

Demostración:

Para demostrar que los autovalores de A y A^T son los mismos, consideremos el polinomio característico de A . Recordando que la propiedad de transposición del determinante nos dice que $\det(B) = \det(B^T)$ para cualquier matriz cuadrada B , se tiene entonces que:

$$p^T(\lambda) = \det(A^T - \lambda I) = \det((A - \lambda I)^T) = \det(A - \lambda I) = p(\lambda).$$

Esto demuestra que el polinomio característico de A es igual al polinomio característico de A^T , lo que implica que A y A^T tienen los mismos autovalores.

Finalmente, el radio espectral de una matriz se define como el máximo valor absoluto de sus autovalores. Dado que A y A^T tienen los mismos autovalores, concluimos que también tienen el mismo radio espectral. ■

Definición 3.1.3. *La multiplicidad algebraica de λ es el número de veces que λ aparece como raíz del polinomio característico de A , definido como:*

$$\det(A - \lambda I) = 0$$

donde I es la matriz identidad. La multiplicidad algebraica de λ es el mayor exponente al cual λ aparece como raíz en la factorización del polinomio característico.

Por otro lado, la multiplicidad geométrica de un valor propio es la dimensión del espacio propio asociado.

Definición 3.1.4. *Dado un vector \mathbf{v} en \mathbb{R}^n , se dice que \mathbf{v} es un vector positivo si todos sus componentes son mayores a cero. En otras palabras, se tiene que para cada $i \in \{1, \dots, n\}$ entonces $v_i > 0$.*

Extendiendo la definición a matrices, se tiene que una matriz es positiva si

$$\forall i, j \in \{1, \dots, n\} \text{ se tiene } a_{ij} > 0.$$

Análogamente, una matriz es negativa si para cada $i, j \in \{1, \dots, n\}$, se tiene $a_{ij} < 0$. Por último, se dice que una matriz es no negativa si

$$\forall i, j \in \{1, \dots, n\} \text{ se tiene } a_{ij} \geq 0.$$

Observación. Para expresar que una matriz es positiva escribiremos $A > 0$. Para expresar que una matriz es no negativa escribiremos $A \geq 0$.

Definición 3.1.5. Una matriz cuadrada A de orden $n \times n$ se dice que es primitiva si para alguna de las potencias positivas de A , la matriz resultante es positiva, es decir, la matriz A es primitiva si y solo si existe un número natural p tal que todos los coeficientes de la matriz A^p son positivas.

$$(A^p)_{ij} > 0 \quad \text{para todo } i, j = 1, 2, \dots, n$$

donde $(A^p)_{ij}$ representa la entrada en la fila i -ésima y la columna j -ésima de la matriz A^p .

Definición 3.1.6. Una matriz cuadrada $P \in \mathbb{R}^{n \times n}$ es una matriz de permutación si en cada fila y columna de P hay un elemento igual a 1 y el resto de los elementos son cero. La suma de cada fila y cada columna es igual a 1.

Proposición 3.1.2. Sea $P \in \mathbb{R}^{n \times n}$ una matriz de permutación. Entonces se cumplen las siguientes propiedades:

- (i) $P^{-1} = P^T$.
- (ii) P^T también es una matriz de permutación.

Demostración:

- (i) Para demostrar que $P^{-1} = P^T$, consideremos que una matriz de permutación P se obtiene a partir de la matriz identidad I mediante una serie de permutaciones de filas. La operación de transposición P^T implica realizar las mismas permutaciones pero en las columnas. Dado que permutar las filas de I de la misma manera que permutamos las columnas para obtener P^T nos devuelve la matriz identidad, esto implica que P^T es la inversa de P , es decir, $P^{-1} = P^T$.
- (ii) Trivial, ya que la transposición de una matriz no cambia el valor de los elementos por fila y columna, por lo que si P tiene en cada fila y columna

de P un elemento igual a 1 y el resto de los elementos son cero, P^T tiene exactamente un elemento igual a 1 en cada fila y en cada columna y el resto de elementos son cero, por lo que es una matriz de permutación. ■

Definición 3.1.7. Una matriz cuadrada A de orden $n \times n$ se dice que es irreducible si no existe ninguna matriz de permutación P tal que $P^T A P$ sea una matriz por bloques, es decir,

$$P^T A P = \begin{bmatrix} B & C \\ 0 & D \end{bmatrix},$$

donde B y D son matrices cuadradas no nulas cuyo orden es menor que n .

Observación. Toda matriz primitiva es irreducible. No obstante, es conocido [3] que sólo las matrices irreducibles y no negativas con algún elemento diagonal no nulo son primitivas.

Definición 3.1.8. Una matriz estocástica es aquella matriz cuyas entradas son todas no negativas, y la suma de las entradas de cada columna es igual a 1, es decir:

$$A = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{bmatrix}$$

es estocástica si $a_{ij} \geq 0$ y $\sum_i a_{ij} = 1$ para cada j fijo.

Además, A^T cumple que $a_{ij} \geq 0$ y $\sum_j a_{ij} = 1$ para cada i fijo.

Teorema 3.1.1. Sea $A = (a_{ij})_{1 \leq i, j \leq n} \in \mathbb{R}^{n \times n}$ una matriz estocástica. Entonces $\lambda = 1$ es un valor propio de A y además, el radio espectral de A es 1.

Demostración:

Sea $v = (1, 1, \dots, 1)^T$. Veamos que $Av = v$. Para ello, se toma primero A^T tal que cada fila de la matriz suma 1 (estocástica por filas):

$$A^T v = \begin{pmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & \ddots & \vdots \\ a_{n1} & \cdots & a_{nn} \end{pmatrix} \begin{pmatrix} 1 \\ \vdots \\ 1 \end{pmatrix} = \begin{pmatrix} \sum_{j=1}^n a_{1j} \\ \vdots \\ \sum_{j=1}^n a_{nj} \end{pmatrix} = \begin{pmatrix} 1 \\ \vdots \\ 1 \end{pmatrix} = v = 1 \cdot v$$

Entonces, $\lambda_0 = 1$ es un valor propio de A^T y, por la Proposición 3.1.1, $\lambda_0 = 1$ también es un valor propio de A . Sea λ un valor propio cualquiera de A^T y sea $v \neq 0$ un vector propio asociado a λ . Definimos $|v_i| = \max_{1 \leq k \leq n} \{|v_k|\}$.

Como $A^T v = \lambda v$, $\sum_{j=1}^n a_{kj} v_j = \lambda v_k$, para $k = 1, \dots, n$. En particular, $\sum_{j=1}^n a_{ij} v_j = \lambda v_i$. Entonces, $|\lambda| |v_i| = \left| \sum_{j=1}^n a_{ij} v_j \right| \leq \sum_{j=1}^n |a_{ij} v_j| = \sum_{j=1}^n a_{ij} |v_j|$. Por lo tanto,

$$|\lambda| \leq \frac{\sum_{j=1}^n a_{ij} |v_j|}{|v_i|} \leq \sum_{j=1}^n a_{ij} \leq 1,$$

ya que $|v_i| = \max_{1 \leq k \leq n} \{|v_k|\}$ y las filas de A^T suman 1 por ser A estocástica. Como $|\lambda| \leq 1$, y $\lambda_0 = 1$ es un valor propio de A^T , el radio espectral de A^T es 1, y por la Proposición 3.1.1 el radio espectral de A es 1. ■

3.2. Fundamentos de Teoría de Grafos

3.2.1. Conceptos Básicos

En esta sección se tratará de contemplar aspectos fundamentales acerca de los grafos, dado que guardan una relación directa con la representación de cualquier red. Se evidenciará que los grafos constituyen una matriz conocida como matriz de adyacencia, la cual reviste interés significativo en el curso de la Teoría de Perron-Frobenius y se examinarán los diversos tipos de grafos posibles y propiedades relevantes, como la conectividad.

Definición 3.2.1.1. *Un grafo (no dirigido) se define como un par ordenado $G = (V, E)$, donde:*

- $V = \{v_1, v_2, \dots, v_n\}$ es un conjunto finito de vértices (también llamados nodos).
- $E = \{e_1, e_2, \dots, e_n\} \subseteq V \times V$. es un conjunto finito de aristas, donde cada arista es un par no ordenado de vértices.

Definición 3.2.1.2. *Un grafo dirigido es un par de conjuntos $G = (V, E)$, donde $V = \{1, \dots, n\}$ para algún $n \in \mathbb{N}$ y $E \subseteq V \times V$. Si $(i, j) \in E$, se dice que (i, j) es la arista dirigida que une los nodos $i, j \in V$, y se denota $i \rightarrow j$.*

Observación. *Todo grafo no dirigido $G = (X, E)$ puede verse como un grafo dirigido. Únicamente habría que tomar cada arista no dirigida $\{i, j\} \in E$ como dos aristas dirigidas (i, j) y (j, i) , es decir, obteniendo el grafo dirigido $G = (X, \tilde{E})$, donde*

$$\tilde{E} = \{(i, j), (j, i) \mid \{i, j\} \in E\}.$$

Definición 3.2.1.3. La matriz de adyacencia de un grafo dirigido $G = (V, E)$ con n nodos se denota como $A(G)$ y es una matriz de dimensión $n \times n$ definida por:

$$a_{ij} = \begin{cases} 1 & \text{si existe una arista } i \rightarrow j \text{ en } G, \\ 0 & \text{en otro caso.} \end{cases}$$

Observación. La matriz de adyacencia de un grafo es siempre una matriz no negativa ($\forall i, j \in \{1, \dots, n\}$ se tiene $a_{ij} \geq 0$).

Observación. Es inmediato comprobar que un grafo $G = (V, E)$ es no dirigido si y solamente si su matriz de adyacencia A es simétrica.

Ejemplo 3.2.1.1. Veamos ahora la representación de un Grafo no dirigido sencillo y su correspondiente Matriz de Adyacencia. Seguidamente, observamos las mismas características de un Grafo dirigido sencillo.

Para la representación de estos grafos se ha utilizado un programa desarrollado en Python (*graphs.py*), que usando la biblioteca *networkx* representa grafos dadas las aristas introducidas.

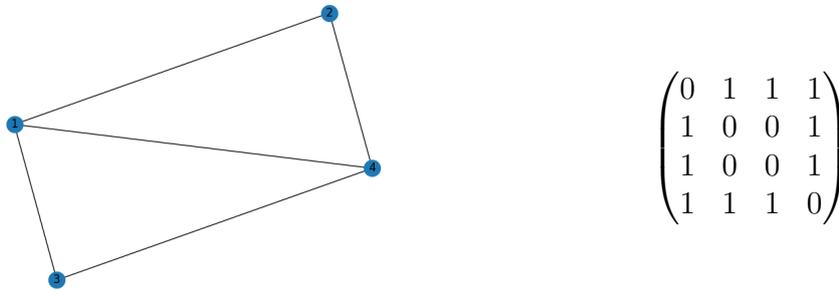


Figura 3.1: Ejemplo de un grafo no dirigido y su matriz de adyacencia

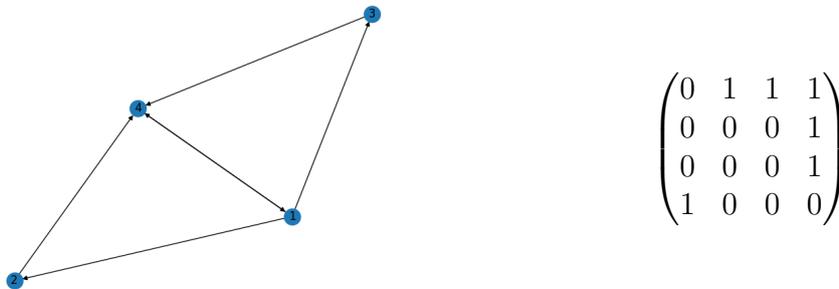


Figura 3.2: Ejemplo de un grafo dirigido y su matriz de adyacencia

Definición 3.2.1.4. Sea $G = (V, E)$ un grafo no dirigido y $k \in V$, llamamos grado de k al número de aristas de G que contienen a k , es decir, el valor:

$$gr(k) = |\{\{k, j\} \in E; j \in V\}|.$$

Definición 3.2.1.5. Sea $G = (V, E)$ un grafo dirigido y $k \in V$, llamamos grado de salida de k (denotado como $gr_{out}(k)$) al número de aristas de G que son de la forma $k \rightarrow j$ para algún $j \in V$, y grado de entrada (denotado como $gr_{in}(k)$) al número de aristas de G que son de la forma $j \rightarrow k$ para algún $j \in V$, es decir,

$$gr_{out}(k) = |\{(k, j) \in E; j \in V\}|,$$

$$gr_{in}(k) = |\{(j, k) \in E; j \in V\}|.$$

Observación. Es fácil comprobar que si tenemos un grafo no dirigido $G = (V, E)$ y lo interpretamos como un grafo dirigido, entonces, para cada $k \in V$:

$$gr(k) = gr_{in}(k) = gr_{out}(k).$$

Observación. Si A es la matriz de adyacencia de G , entonces:

$$gr_{in}(k) = \sum_{j=1}^n a_{jk},$$

$$gr_{out}(k) = \sum_{j=1}^n a_{kj}.$$

Ejemplo 3.2.1.2. El grado del nodo 4 en la Figura 3.1 es:

$$gr(4) = |\{\{1, 4\}, \{2, 4\}, \{3, 4\}\}| = 3$$

Sin embargo, el grado de entrada y salida del nodo 4 en la Figura 3.2 son:

$$gr_{in}(4) = |\{(1, 4), (2, 4), (3, 4)\}| = 3$$

$$gr_{out}(4) = |\{\}| = 0$$

Definición 3.2.1.6. Un camino en un grafo $G = (V, E)$ es una secuencia de vértices v_1, v_2, \dots, v_n en las que cada par consecutivo de vértices v_i, v_{i+1} está conectado por una arista en E . Es decir, un camino es una secuencia de aristas $(v_1, v_2), (v_2, v_3), \dots, (v_{n-1}, v_n)$ tal que $v_i \in V$ y $(v_i, v_{i+1}) \in E$ para $1 \leq i < n$. La longitud del camino es $n - 1$, es decir, el número de aristas que recorre.

Definición 3.2.1.7. Un camino dirigido es una secuencia de vértices v_1, v_2, \dots, v_n donde cada vértice v_i está conectado por una arista dirigida (v_i, v_{i+1}) en E para $1 \leq i < n$. Es decir, un camino dirigido en un grafo dirigido $G = (V, E)$ es similar a un camino, pero las aristas tienen una dirección específica. La longitud del camino es $n - 1$, es decir, el número de aristas que recorre.

Definición 3.2.1.8. Un grafo $G = (V, E)$ se dice que es conexo si, para cada par de vértices $i, j \in V$, existe al menos un camino que conecta i y j .

Definición 3.2.1.9. Un grafo dirigido $G = (V, E)$ se dice que es fuertemente conexo si, para cada par de vértices $i, j \in V$, existe al menos un camino dirigido que conecta i y j .

Ejemplo 3.2.1.3. El grafo de la Figura 3.1 es conexo pues

$$\forall i, j \in V = \{1, 2, 3, 4\}, \exists \text{ un camino de } i \text{ a } j$$

Por otro lado, el grafo de la Figura 3.2 es fuertemente conexo pues

$$\forall i, j \in V, \exists \text{ un camino dirigido de } i \text{ a } j$$

Por ejemplo, desde el nodo 4 se puede llegar al nodo 1 directamente, pero también al nodo 2 y al nodo 3 pasando antes por el nodo 1.

3.2.2. Relación entre Conectividad y Matrices Irreducibles

En esta sección se verá un resultado importante, pues aporta un modo sencillo de determinar si una matriz de adyacencia de un grafo dirigido es o no es irreducible estudiando la estructura, en concreto la conectividad, del grafo dirigido en cuestión. Primero, estudiaremos las siguientes proposiciones, siguiendo algunos resultados vistos en [3], [4] y [5]:

Proposición 3.2.2.1. Sea $A \in \mathbb{R}^{n \times n}$ tal que

$$A = \begin{bmatrix} B & C \\ 0 & D \end{bmatrix},$$

donde $B \in \mathbb{R}^{k \times k}$, $C \in \mathbb{R}^{k \times (n-k)}$, y $D \in \mathbb{R}^{(n-k) \times (n-k)}$, con $1 \leq k \leq n-1$. Entonces, el grafo dirigido asociado a A no es fuertemente conexo.

Demostración:

Cojamos un camino dirigido que comience en el nodo i tal que $i > k$. Como $i > k$, en la i -ésima fila de A , existen unos primeros $n - k$ elementos son 0. Por lo tanto, el nodo i solamente se puede conectar con un nodo j mediante una línea dirigida si $j > k$. Del mismo modo, como $j > k$, los primeros $n - k$ elementos de la j -ésima fila de A son 0. Luego, el nodo j solamente se puede conectar con el nodo z mediante una línea dirigida si $z > k$.

Siguiendo con este procedimiento, se llega a la conclusión de que si $i > k$ y $j \leq k$, nunca existirá un camino dirigido de i a j , lo que implica que el grafo dirigido asociado a A no es fuertemente conexo. ■

Proposición 3.2.2.2. *Sea $A \in \mathbb{R}^{n \times n}$ y sea $P \in \mathbb{R}^{n \times n}$ una matriz de permutación. Entonces, el grafo dirigido asociado a A es fuertemente conexo si y solo si el grafo dirigido asociado a PAP^T es fuertemente conexo.*

Demostración:

Supongamos que el grafo dirigido asociado a A es fuertemente conexo. Observamos que el grafo dirigido de PAP^T se obtiene reenumerando los nodos del grafo de A , pues la matriz de permutación P se encargará únicamente de reorganizar las filas o columnas de A . Es decir, si la matriz de permutación P pone la i -ésima fila de A en la j -ésima, entonces el nodo i del grafo de A será el nodo j del grafo de PAP^T . Por tanto, como el grafo dirigido asociado a A es fuertemente conexo, el grafo dirigido asociado a PAP^T también lo será.

Supongamos que el grafo dirigido asociado a PAP^T es fuertemente conexo. Como P es una matriz de permutación, P^T también lo es por la Proposición 3.1.1, y por tanto, recordando que $P^{-1} = P^T$ por la Proposición 3.1.1, el grafo dirigido asociado a $P^T P A P^T (P^T)^T = A$ (permutación revertida) es fuertemente conexo por el mismo procedimiento que el anterior. ■

Proposición 3.2.2.3. *Sea $A \in \mathbb{R}^{n \times n}$ una matriz no negativa e irreducible. Entonces, $(I_n + A)^{n-1} > 0$.*

Demostración:

Sea $y \in \mathbb{R}^n$ un vector tal que $y \geq 0$ y $y \neq 0$. Se define

$$z = (I_n + A)y = y + Ay.$$

Como $y \geq 0$ y por hipótesis $A \geq 0$, se tiene que $Ay \geq 0$. Por tanto, z tiene al menos tantos elementos no nulos como y . Vamos a ver que si algún elemento

de y es cero, entonces z tiene al menos un elemento no nulo más que y . Sea P una matriz de permutación tal que $Py = \begin{bmatrix} u \\ 0 \end{bmatrix}$ y $u > 0$. Utilizando la igualdad anterior, se tiene que

$$Pz = P(y + Ay) = Py + PAy = Py + PAP^T Py = \begin{bmatrix} u \\ 0 \end{bmatrix} + PAP^T \begin{bmatrix} u \\ 0 \end{bmatrix} \quad (1)$$

Se dividen Pz y PAP^T de acuerdo con $\begin{bmatrix} u \\ 0 \end{bmatrix}$. Es decir, $Pz = \begin{bmatrix} v \\ w \end{bmatrix}$ siendo u y v dos vectores del mismo tamaño y $PAP^T = \begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix}$ siendo A_{11} una matriz cuadrada del mismo tamaño que u . Debido a la ecuación (1),

$$v = u + A_{11}u \quad \text{y} \quad w = A_{21}u. \quad (2)$$

Como $A \geq 0$, $P \geq 0$ y $P^T \geq 0$, se tiene que $PAP^T \geq 0$. En particular, $A_{11} \geq 0$ y $A_{21} \geq 0$. Como A es irreducible, se tiene que $A_{21} \neq 0$. Por tanto, debido a las ecuaciones de (2) y que $u > 0$, se tiene que $v > 0$ y $w \geq 0$ pero $w \neq 0$. Luego $Pz = \begin{bmatrix} v \\ w \end{bmatrix}$ tendrá al menos un elemento no nulo más que y . Como $z = P^T \begin{bmatrix} v \\ w \end{bmatrix}$, z tiene los mismos elementos que Pz pero en distinto orden, y por tanto, z tiene al menos un elemento no nulo más que y .

Si $(I_n + A)y = z$ todavía tiene algún elemento nulo, se considera el vector $(I_n + A)y \neq 0$ y se repite el mismo proceso con $z = (I_n + A)((I_n + A)y) = (I_n + A)^2 y$ y se consigue que $(I_n + A)^2 y$ tiene al menos un elemento no nulo más que $(I_n + A)y$, es decir, que $(I_n + A)^2 y$ tiene al menos dos elementos no nulos más que y . Se repite el proceso hasta conseguir $k_0 \in \mathbb{N}$ tal que $(I_n + A)^{k_0} y > 0$. Como $y \in \mathbb{R}^n$, el proceso se repetirá como mucho $n - 1$ veces.

Vamos a diferenciar dos casos: $k_0 = n - 1$ y $k_0 < n - 1$. Si $k_0 = n - 1$, se tiene que $(I_n + A)^{n-1} y > 0$. Si $k_0 < n - 1$, se tiene que $(I_n + A)^{n-1} y = (I_n + A)^{n-1-k_0} (I_n + A)^{k_0} y > 0$, ya que los elementos de la diagonal principal de $(I_n + A)^{n-1-k_0}$ son positivos por ser $A \geq 0$ y $(I_n + A)^{k_0} y > 0$. Por tanto, $(I_n + A)^{n-1} y > 0$ para todo $y \in \mathbb{R}^n$ tal que $y \geq 0$, $y \neq 0$.

Por tanto, $(I_n + A)^{n-1} e_j > 0$ para todo $j \in \{1, \dots, n\}$. Es decir, los elementos de la j -ésima columna de $(I_n + A)^{n-1}$ son positivos para todo $j \in \{1, \dots, n\}$. Por tanto, $(I_n + A)^{n-1} > 0$. ■

Proposición 3.2.2.4. *Sea $A \in \mathbb{R}^{n \times n}$ una matriz no negativa e irreducible, donde $a_{ij}^{(q)}$ es el (i, j) -ésimo elemento de A^q . Entonces, para cada $i, j \in \{1, \dots, n\}$ existe $q \in \mathbb{N}$ tal que $a_{ij}^{(q)} > 0$.*

Demostración:

Como $A \in \mathbb{R}^{n \times n}$ es una matriz no negativa e irreducible, por la Proposición 3.2.2.3, se tiene que $(I_n + A)^{n-1} > 0$. Vamos a ver que $A(I_n + A)^{n-1} > 0$. Sea $A(I_n + A)^{n-1} = [c_{ij}]_{1 \leq i, j \leq n}$. Suponiendo por reducción al absurdo que $\exists i_0, j_0 \in \{1, \dots, n\}$ tales que $c_{i_0 j_0} = 0$. Sean $A = [a_{ij}]_{1 \leq i, j \leq n}$ y $(I_n + A)^{n-1} = [b_{ij}]_{1 \leq i, j \leq n}$. Como $A \geq 0$, se tiene que $a_{ij} \geq 0 \forall i, j \in \{1, \dots, n\}$ y como $(I_n + A)^{n-1} > 0$, se tiene que $b_{ij} > 0 \forall i, j \in \{1, \dots, n\}$. Dado que $c_{i_0 j_0} = 0$ es el (i_0, j_0) -ésimo elemento de $A(I_n + A)^{n-1}$, se tiene que

$$c_{i_0 j_0} = \sum_{k=1}^n a_{i_0 k} b_{k j_0} = 0. \quad (1)$$

Como $a_{i_0 k} \geq 0$ y $b_{k j_0} > 0 \forall k \in \{1, \dots, n\}$, para que se cumpla la igualdad (1), tiene que cumplirse que $a_{i_0 k} = 0 \forall k \in \{1, \dots, n\}$. Es decir, la i_0 -ésima fila de A está compuesta por ceros. Sea $P \in \mathbb{R}^{n \times n}$ la matriz de permutación que se obtiene al intercambiar la i_0 -ésima y la n -ésima filas de la matriz I_n .

$$PAP^T = \begin{bmatrix} a_{11} & \dots & a_{1n} & \dots & a_{1i_0} \\ \vdots & \ddots & \vdots & \ddots & \vdots \\ a_{n1} & \dots & a_{nn} & \dots & a_{ni_0} \\ \vdots & \ddots & \vdots & \ddots & \vdots \\ 0 & \dots & 0 & \dots & 0 \end{bmatrix} = \begin{bmatrix} B & C \\ 0 & D \end{bmatrix},$$

con $B \in \mathbb{R}^{(n-1) \times (n-1)}$, $C \in \mathbb{R}^{(n-1) \times 1}$ y $D \in \mathbb{R}^{1 \times 1}$. Entonces A es una matriz reducible y se obtiene una contradicción ya que A es una matriz irreducible por hipótesis. Por lo tanto, $c_{ij} > 0 \forall i, j \in \{1, \dots, n\}$, es decir, $A(I_n + A) > 0$. Entonces se utiliza la fórmula del binomio para calcular $(I_n + A)^{n-1}$.

$$\begin{aligned} A(I_n + A)^{n-1} &= A \left[\binom{n-1}{0} A^{n-1} + \binom{n-1}{1} I_n A^{n-2} + \dots + \binom{n-1}{n-1} I_n^{n-1} \right] \\ &= A^n + (n-1)A^{n-1} + \dots + (n-1)A^2 + A > 0. \end{aligned} \quad (2)$$

Por la igualdad (2), para cada $i, j \in \{1, \dots, n\}$, $\exists q \in \{1, \dots, n\}$ tal que $a_{ij}^{(q)} > 0$. ■

Proposición 3.2.2.5. *Sea $A \in \mathbb{R}^{n \times n}$ una matriz no negativa e irreducible, y sea $a_{ij}^{(q)}$ el (i, j) -ésimo elemento de A^q . Entonces, $a_{ij}^{(q)} > 0$ si y solo si existe una secuencia de q líneas dirigidas en el grafo dirigido asociado a A que conectan los*

nodos P_i y P_j .

Demostración:

Demostremos por inducción que

$$a_{ij}^{(q)} = \sum_{k_{q-1}=1}^n \sum_{k_{q-2}=1}^n \cdots \sum_{k_2=1}^n \sum_{k_1=1}^n a_{ik_1} a_{k_1 k_2} \cdots a_{k_{q-2} k_{q-1}} a_{k_{q-1} j}, \quad \forall q \geq 2.$$

o Caso $q = 2$, $a_{ij}^{(2)} = \sum_{k_1=1}^n a_{ik_1} a_{k_1 j}$

o Caso $q = 3$, $a_{ij}^{(3)} = \sum_{k_2=1}^n a_{ij}^{(2)} a_{k_2 j} = \sum_{k_2=1}^n \left(\sum_{k_1=1}^n a_{ik_1} a_{k_1 k_2} \right) a_{k_2 j} =$
 $= \sum_{k_2=1}^n \sum_{k_1=1}^n a_{ik_1} a_{k_1 k_2} a_{k_2 j}$

Supongamos como *hipótesis de inducción* que

$$a_{ij}^{(q-1)} = \sum_{k_{q-2}=1}^n \sum_{k_{q-3}=1}^n \cdots \sum_{k_2=1}^n \sum_{k_1=1}^n a_{ik_1} a_{k_1 k_2} \cdots a_{k_{q-3} k_{q-2}} a_{k_{q-2} j}.$$

Aplicando la hipótesis de inducción, se tiene que

$$\begin{aligned} a_{ij}^{(q)} &= \sum_{k_{q-1}=1}^n a_{ij}^{(q-1)} a_{k_{q-1} j} \\ &= \sum_{k_{q-1}=1}^n \left(\sum_{k_{q-2}=1}^n \sum_{k_{q-3}=1}^n \cdots \sum_{k_2=1}^n \sum_{k_1=1}^n a_{ik_1} a_{k_1 k_2} \cdots a_{k_{q-3} k_{q-2}} a_{k_{q-2} k_{q-1}} \right) a_{k_{q-1} j} \\ &= \sum_{k_{q-1}=1}^n \sum_{k_{q-2}=1}^n \cdots \sum_{k_2=1}^n \sum_{k_1=1}^n a_{ik_1} a_{k_1 k_2} \cdots a_{k_{q-2} k_{q-1}} a_{k_{q-1} j}. \end{aligned} \quad (1)$$

Por hipótesis, se asume que la matriz A es no negativa. Por lo tanto, al considerar la igualdad (1), se deduce que $a_{ij}^{(q)} > 0$ si y solo si existen índices $k_1, \dots, k_{q-1} \in \{1, \dots, n\}$ tales que $a_{ik_1} a_{k_1 k_2} \cdots a_{k_{q-1} j} > 0$. Dado que A es una matriz no negativa, esto es equivalente a $a_{ik_1} a_{k_1 k_2} \cdots a_{k_{q-1} j} > 0$ si y solo si $a_{ik_1}, a_{k_1 k_2}, \dots, a_{k_{q-1} j} > 0$. En otras palabras, existe una secuencia de q líneas dirigidas en el grafo dirigido de A que conectan los nodos P_i y P_j . ■

Ahora, hemos completado todos los preparativos necesarios para enunciar y demostrar el teorema principal de esta sección. Este teorema será fundamental para nuestro análisis, ya que nos permitirá establecer un criterio claro y preciso. A través de este teorema, lograremos una manera sencilla y efectiva de determinar

si una matriz es irreducible o reducible.

Teorema 3.2.2.1. *Sea $A \in \mathbb{R}_{n \times n}$ una matriz no negativa. Entonces A es irreducible si y solo si el grafo dirigido asociado a A es fuertemente conexo.*

Demostración:

\Rightarrow Supongamos que $A \in \mathbb{R}^{n \times n}$ es una matriz irreducible. Dado que A también es una matriz no negativa, según la Proposición 3.2.2.4, para cualquier $i, j \in \{1, \dots, n\}$ existe $q \in \mathbb{N}$ tal que $a_{ij}^{(q)} > 0$. Según la Proposición 3.2.2.5, entonces para todo $i, j \in \{1, \dots, n\}$, los nodos P_i y P_j están enlazados. Por lo tanto, el grafo dirigido asociado a M es fuertemente conexo.

\Leftarrow Supongamos que $A \in \mathbb{R}^{n \times n}$ es una matriz reducible. Entonces, existe $P \in \mathbb{R}^{n \times n}$ una matriz de permutación tal que $P^T A P = \begin{bmatrix} B & C \\ 0 & D \end{bmatrix}$, siendo B y D dos matrices cuadradas de orden menor que n . Según la Proposición 3.2.2.1, el grafo dirigido de $P^T A P$ no es fuertemente conexo. Dado que $P^T \in \mathbb{R}^{n \times n}$ es una matriz de permutación, según la Proposición 3.2.2.2, el grafo dirigido asociado a A no es fuertemente conexo. ■

Llegados a este punto ya contamos con una manera sencilla de determinar si una matriz es o no es irreducible. Comprobémoslo con algunos ejemplos:

Ejemplo 3.2.2.1 Sea A la siguiente matriz:

$$\begin{pmatrix} 0 & 4 & 7 & 0 & 0 \\ 0 & 0 & 0 & 2 & 0 \\ 1 & 0 & 0 & 0 & 0 \\ 0 & 6 & 5 & 0 & 2 \\ 0 & 0 & 0 & 1 & 0 \end{pmatrix}$$

Inicialmente, trataríamos de comprobar si la matriz es irreducible usando la definición. Sin embargo, probar que todas las matrices de permutación de dimensión 5×5 no satisfacen

$$P^T A P = \begin{bmatrix} B & C \\ 0 & D \end{bmatrix}$$

es una tarea complicada y tediosa. En su lugar, crearemos por tanto el grafo dirigido a partir de esta matriz de adyacencia, donde podremos observar que (3.3) es fuertemente conexo, y por el Teorema 3.2.2.1, la matriz A es irreducible.

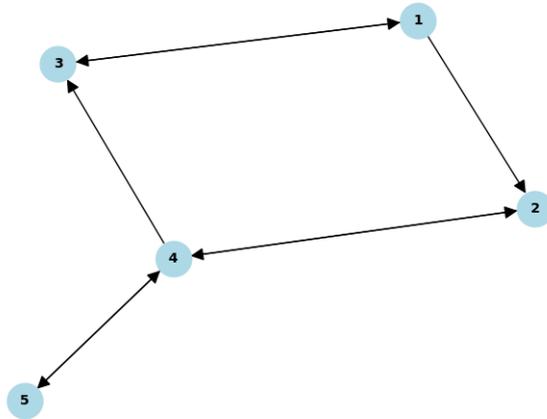


Figura 3.3: Grafo fuertemente conexo - Matriz de adyacencia irreducible

Ejemplo 3.2.2.2 Por el otro lado, sea B la siguiente matriz:

$$\begin{pmatrix} 0 & 5 & 2 & 0 & 0 \\ 0 & 0 & 0 & 3 & 0 \\ 3 & 0 & 0 & 0 & 0 \\ 0 & 1 & 7 & 0 & 0 \\ 0 & 0 & 0 & 9 & 0 \end{pmatrix}$$

En este caso, bastaría comprobar que el grafo dirigido correspondiente a esta matriz de adyacencia **no** es fuertemente conexo. Por ejemplo, no existe un camino desde cualquier nodo hasta el nodo 5. Por tanto, por el teorema 3.2.2.1, la matriz es reducible.

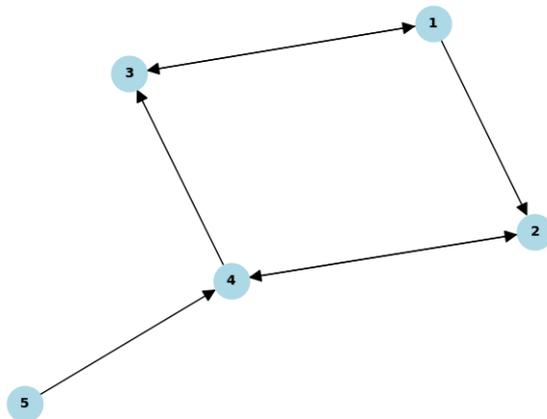


Figura 3.4: Grafo no fuertemente conexo - Matriz de adyacencia reducible

3.3. Teoría de Perron-Frobenius

3.3.1. Teorema de Perron-Frobenius

Una vez vistos los preliminares de Álgebra Lineal y Teoría de Grafos podemos empezar a desarrollar la Teoría de Perron-Frobenius. Para ello nos basaremos en los conceptos explicados en el artículo [2], con ayuda de otros como [6]. Recordemos primero que el principal interés de desarrollar esta teoría reside en las posibles aplicaciones que derivan de su resultado. Esto es, la obtención de **rankings**, ordenaciones por importancia o clasificaciones. Estas ordenaciones pueden hacerse sobre diferentes participantes, como por ejemplo artículos, autores, *páginas web*, equipos, y muchas otras cosas. Posteriormente se tratará también la idea de qué es exactamente ser *importante*.

Entonces, nuestro problema se podría basar en ordenar de mayor a menor cierta clase de participantes según su importancia. Para ello, se supondrá que la importancia a repartir conforma un total de 1, y que dicha importancia se repartirá entre los participantes asignando a cada uno de ellos un valor entre 0 y 1. Este coeficiente asignado a cada uno medirá el valor de importancia en relación con el resto, es decir, si quiero conocer qué participante es más importante bastaría con ver qué coeficiente es mayor. Además, toda la importancia (el valor total de 1) se ha repartido entre los participantes. Por tanto, lo que se busca es encontrar un vector $c = (c_1, c_2, \dots, c_n)$ tal que cada c_i represente la importancia de cada participante i , es decir;

$$\forall i \in \{1, \dots, n\}, 0 \leq c_i \leq 1 \quad \text{y} \quad \sum_{i=1}^n c_i = 1. \quad (1)$$

Una vez obtenidos estos valores, sólo quedaría ordenar estos coeficientes de mayor a menor para saber la importancia de cada participante.

Ahora bien, ¿qué es ser importante? Para explicarlo centrémonos en el ejemplo de los artículos. Un artículo no sólo ganará importancia por su propia calidad o influencia, si no también por la importancia (o calidad, influencia, etc) de aquellos artículos que lo incluyen en sus referencias. Más tarde se explicará por qué debe seguirse esta línea con respecto a la importancia.

Observación. *No es difícil ver la relación de esto último con los preliminares de álgebra lineal y teoría de grafos previamente tratados. Supongamos que existen n artículos, entonces el grafo dirigido G es aquel en el que los nodos son los n artículos y existirá una arista $i \rightarrow j$ entre los nodos i y j si el artículo j aparece entre las referencias del artículo i .*

En este sentido, sea G el grafo generado por los participantes y sus conexio-

nes. Teniendo en cuenta que *la importancia de un participante i es proporcional a la importancia de los participantes que le referencian*, tendremos que cada coeficiente de importancia además vendrá dada por:

$$c_i = \lambda \sum_{j \rightarrow i} c_j \Rightarrow \frac{1}{\lambda} c_i = \sum_{j \rightarrow i} c_j \quad (2)$$

donde $\sum_{j \rightarrow i} c_j$ indica la suma de los coeficientes de importancia de aquellos nodos de los que sale un arista que llega al nodo i .

Buscamos entonces un vector $c = (c_1, c_2, \dots, c_n) \in \mathbb{R}^n$ tal que se cumpla:

- $c_i \geq 0$ para todo $1 \leq i \leq n$ (expresado mediante $c \geq 0$)
- $c_1 + \dots + c_n = \|c\|_1 = 1$
- existe $K > 0$ (el mismo para todo $1 \leq i \leq n$) tal que

$$c_i = K \sum_{j \rightarrow i} c_j = K \sum_{j=1}^n a_{ji} c_j,$$

que expresado de manera matricial queda

$$\frac{1}{K} c = A^T c$$

y reescribiendo la ecuación anterior con $\lambda = \frac{1}{K}$

$$\lambda c = A^T c$$

es decir, c es un autovector de A^T , siendo esta matriz la transpuesta de la matriz de adyacencia de G , y está asociado al autovalor λ .

En este punto, vamos a enunciar el teorema de Perron, demostrado por *Oskar Perron* en 1907 y aplicable a todas las matrices positivas (coeficientes todos positivos, $a_{ij} > 0$).

Teorema 3.3.1.1. (Perron) *Si $A = (a_{ij})$ es una matriz cuadrada positiva, entonces se cumplen las siguientes afirmaciones:*

- (i) $\rho(A) \in \sigma(A)$ es real y positivo y su multiplicidad como autovalor es 1.
- (ii) Existe un vector $c \in \mathbb{R}^n$ tal que $c > 0$, $\|c\|_1 = 1$ y es un autovector de A asociado al autovalor $\rho(A)$.
- (iii) Este vector c es el único vector no negativo y de norma $\|c\|_1 = 1$ que satisface las condiciones anteriores, es decir, que si existe otro autovector v

de A que cumple que $v \geq 0$ y $\|v\|_1 = \sum_{i=1}^n c_i = 1$, entonces necesariamente $v = c$.

Sin embargo, volviendo al ejemplo de los artículos, no siempre se puede garantizar que la matriz A sea positiva, pues un artículo i puede no estar referenciado por otro artículo j , lo que implicaría que la posición a_{ji} de la matriz de adyacencia tendría un valor de 0 (ya que no todos los artículos están referenciados por todos los demás). No obstante, sí que se puede garantizar que la matriz A sea no negativa.

Observación. Si la matriz A es primitiva, entonces el Teorema de Perron también es válido para la matriz A , ya que una de sus potencias será una matriz positiva.

Afortunadamente, tres años más tarde en 1912, *Ferdinand Frobenius*, un matemático alemán, generalizó este teorema para matrices no necesariamente positivas, pero sí no negativas. Para conseguirlo, tuvo que incluir una condición adicional que se ha tratado en la Sección 3.2.2 (3.2) a fondo, esto es, la matriz A también debe ser irreducible. Sabiendo esto, estamos preparados para formular el Teorema de Perron-Frobenius y su posterior demostración, que haremos de una forma geométrica siguiendo [7]. Al ser una demostración geométrica, necesitaremos establecer de antemano aspectos utilizados en ella como la definición de *rayo* o el teorema del punto fijo de Brouwer, que se utilizará en la demostración:

Definición 3.3.1.1. Un rayo en \mathbb{R}^n (en la dirección de $v \in \mathbb{R}^n \setminus \{0\}$) es el conjunto $r[v] = \{\mu v / \mu > 0\}$. La colección de rayos en \mathbb{R}^n se denota como S^{n-1} .

Teorema 3.3.1.2. (Teorema del Punto Fijo de Brouwer) Sea K un conjunto no vacío, compacto y convexo en \mathbb{R}^n . Toda función continua de K en sí misma tiene un punto fijo.

Teorema 3.3.1.3. (Perron-Frobenius) Si $A = (a_{ij})$ es una matriz cuadrada no negativa e irreducible, entonces se cumplen las siguientes afirmaciones:

- (i) $\rho(A) \in \sigma(A)$ es real y positivo y su multiplicidad como autovalor es 1.
- (ii) Existe un vector $c \in \mathbb{R}^n$ tal que $c > 0$, $\|c\|_1 = 1$ y es un autovector de A asociado al autovalor $\rho(A)$.
- (iii) Este vector c es el único vector no negativo y de norma $\|c\|_1 = 1$ que satisface las condiciones anteriores, es decir, que si existe otro autovector v de A que cumple que $v \geq 0$ y $\|v\|_1 = \sum_{i=1}^n c_i = 1$, entonces necesariamente $v = c$. Además, le denominaremos vector de Perron o vector propio dominante.

Demostración:

- (i) En primer lugar, A es una matriz no negativa, por lo que actúa en el conjunto de rayos no negativos de \mathbb{R}^+ , es decir, semirrectas que emanan del origen y contenidas en el cuadrante

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

En este contexto, no hay rayo en \mathbb{R}^+ que pueda dirigirse al cero, porque en ese caso al ser A no negativa necesariamente existiría una columna de ceros. Sin embargo, A es irreducible, por lo que no es posible.

- (ii) En segundo lugar, ningún rayo en $\partial\mathbb{R}^+$ queda invariante por A . Para demostrarlo, supongamos que hay un rayo $r = r[v]$ que yace en ∂C^+ para el cual $A(r) = r$. Entonces, existe un k tal que $0 < k < n$, para el que precisamente las primeras k coordenadas de v son cero (después de una ordenación posiblemente necesaria). La condición $A(r) = r$ ahora implica que A tiene la forma $A = \begin{bmatrix} B & C \\ 0 & D \end{bmatrix}$, siendo B de orden $k \times k$ y D de orden $(n - k) \times (n - k)$, pero eso no es posible porque A es irreducible.

- (iii) Ahora bien, como $A(\mathbb{R}^+) \subseteq \mathbb{R}^+$ por (i), \mathbb{R}^+ es un espacio convexo, compacto y no vacío, podemos aplicar el Teorema del punto fijo de Brouwer, que nos asegura que hay un punto fijo, o equivalentemente, un rayo r invariante en \mathbb{R}^+ . Por (ii) conocemos que ningún rayo perteneciente al borde del cuadrante es fijo, lo que implica que r es un rayo positivo. Es decir, si consideramos un vector v apuntando en la dirección del rayo r (siendo v un autovector), entonces v no presenta ninguna de sus coordenadas como nula. Además, el autovalor vinculado a v , representado por Π , es positivo debido a que tanto v como su transformación pertenecen al cuadrante positivo, C^+ . Por tanto, ya hemos demostrado que existe el **autovector** de estas características (vector c en los apartados (ii) e (iii) del enunciado). Queda demostrar que el autovalor asociado tiene multiplicidad 1 y que es de hecho $\rho(A)$.

- (iv) Supongamos que Π es un plano invariante que contiene a r , sea L el arco $R^+ \cap S^1$ en Π . Debido a que L pertenece al plano invariante y por (i), se cumple que $A(L) \subseteq L$:

- (1) Por (ii), S^1 no es punto fijo bajo la acción de A .
- (2) El conjunto de puntos fijos de la acción de A^2 en S^1 no consiste sólo de r y $-r$. De lo contrario, la dinámica de la acción de A sobre S^1 ocurriría de forma que $A^2(L)$ no estaría incluido en L , lo cual no es posible:

De (1) se deduce la simplicidad geométrica del autovalor λ correspondiente a r , y de (2) la simplicidad algebraica. Por lo tanto, el autovalor λ correspondiente a r es simple (es decir, tiene **multiplicidad 1**) positivo y tiene un autovector asociado positivo \mathbf{v}_1 , quedando sólo pendiente demostrar que de hecho este autovalor λ es $\rho(A)$.

- (v) Se busca demostrar para finalizar que el autovalor $\lambda \geq |\mu|$ para otro autovalor cualquiera μ . Razonamos mediante un argumento de reducción al absurdo. Supongamos que el autovalor μ satisface que $\lambda < |\mu|$, consideramos dos situaciones distintas:

Caso 1 μ es real:

Sea v_μ el autovector asociado al autovalor μ tal que su rayo $r_\mu \notin \mathbb{R}^+$ (en el caso de que $r_\mu \in \mathbb{R}^+$ ya estaría demostrado el teorema). Entonces A^n actúa en la circunferencia S^1 que representa todos los rayos del plano generado por v_π y v_μ , y fija al conjunto de rayos $\{\pm r_\lambda, \pm r_\mu\}$. Como $\lambda < |\mu|$, el movimiento que realiza A_n en S^1 tiene dos puntos de atracción, $\{\pm r_\mu\}$, y dos puntos de repulsión, $\{\pm r_\lambda\}$. Entonces, r_μ atrae a uno de los dos puntos de $\partial\mathbb{R}^+ \cap S^1$ fuera de \mathbb{R}^+ , lo cual no es posible.

Caso 2 μ es complejo:

Supongamos que $|\lambda| < |\mu|$. Llamamos E^3 al espacio generado por \mathcal{P} y por v_λ , siendo \mathcal{P} el plano de la intersección generada por un autovector complejo v , denotado como $\mathcal{P} = \langle \operatorname{re}(v), \operatorname{im}(v) \rangle$. Dado que $|\lambda| < |\mu|$, la acción de A en E^3 muestra repulsión hacia la línea generada por v_λ , y una atracción al plano \mathcal{P} . Esto implica que los rayos de $\mathbb{R}^+ \cap r_\lambda$ en E^3 se aproximan a \mathcal{P} tanto como queramos si aplicamos repetidamente la aplicación de A . Por lo tanto, concluimos que todos los rayos salen de \mathbb{R}^+ después de aplicar un número de veces, lo cual no es posible porque sabemos que $A(\mathbb{R}^+) \subseteq \mathbb{R}^+$. Una demostración formal de esto es la siguiente: sea $r[x]$ un rayo de E_3 en $\mathbb{R}^+ \setminus r_{\mathcal{P}}$ con $x = v_{\mathcal{P}} + w$, siendo $w \in \mathcal{P}$ y $w \neq 0$.

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

Cuando $n \rightarrow \infty$, entonces $\frac{\lambda^n}{|\mu|^n} \rightarrow 0$, y por tanto el rayo $A^n(r[x])$ se acerca a \mathcal{P} tanto como queramos, ya que el módulo de $\frac{\mu^n}{|\mu|^n} w \in \mathcal{P}$ es constante.

Por tanto, la hipótesis no es verdadera y podemos afirmar que $\lambda \geq |\mu|$, es decir, $\lambda = \rho(A)$ es **el autovalor máximo**.

■

Observación. *Es importante destacar que según la teoría desarrollada al principio de la sección, estamos buscando un autovector de la matriz **traspuesta** A^T , mientras que el teorema desarrollado nos devuelve un autovector de la matriz A .*

Resulta que las propiedades de irreducibilidad y no negatividad son invariantes bajo la trasposición, es decir, si una matriz A es irreducible y no negativa, su **traspuesta** A^T **también es irreducible y no negativa** porque el grafo dirigido traspuesto es el mismo pero invirtiendo las direcciones. Además, el **radio espectral** de cualquier matriz A **es el mismo** $\rho(A)$ que de la matriz A^T , porque los autovalores de A y A^T son los mismos según la Proposición 3.1.1.

En otras palabras, si se dan las condiciones buscadas anteriormente (la importancia de un participante es proporcional a la importancia de los participantes que le referencian), el teorema se podría formular a favor de esa situación y seguiría siendo aplicable. De hecho, la condición $\|c\|_1 = 1$ también se ha añadido de forma adicional al teorema original, ya que en todos los ejemplos de este trabajo se busca representar la importancia de cada componente en una escala del 0 al 1.

Bajo nuestras condiciones y siguiendo nuestros intereses, el teorema por tanto podría reformularse tal que si $A = (a_{ij})$ es una matriz cuadrada no negativa e irreducible, entonces existe un vector c tal que es el único que cumple $c > 0$, $\|c\|_1 = 1$ y es un **autovector de** A^T asociado al autovalor $\rho(A)$.

Una vez demostrado y entendido el Teorema de Perron-Frobenius, se presentarán una serie de ejemplos que permitirán profundizar en la comprensión del teorema, llegando a una aplicación práctica en un problema real.

Para resolver estos ejercicios sencillos se ha desarrollado un programa en Python (*perron-frobenius.py*) que calcula los autovalores y autovectores dada una matriz y devuelve el autovalor máximo junto con su vector de Perron asociado. Este programa se comentará con más detalle en la *Sección 3.3.2*.

Ejemplo 3.3.1.1 En este ejemplo, consideremos la matriz de la figura 3.5 y supongamos que queremos obtener el vector de importancias entre tres participantes:

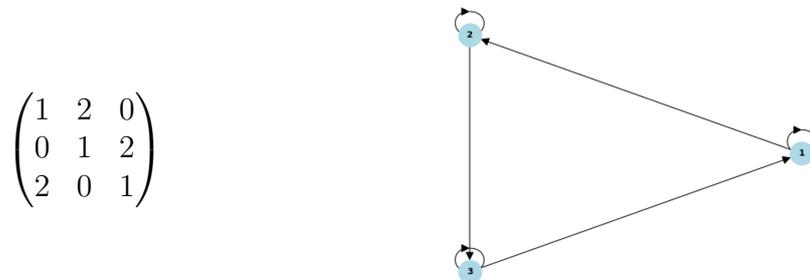


Figura 3.5: Ejemplo de un grafo dirigido y su matriz de adyacencia

En este caso, esta matriz es no negativa por definición, y además el grafo dirigido resultante (ignorando los pesos) que observamos en 3.5 es fuertemente conexo. Por tanto, por lo visto anteriormente, la matriz es irreducible. En este punto, se están cumpliendo los requisitos para aplicar el Teorema de Perron-Frobenius.

Además, queremos seguir la teoría explicada en el inicio de la sección, en la que el vector de Perron es un autovector de A^T , matriz en la que ya sabemos que es equivalente aplicar el teorema (Observación 3.3.1) Es por ello que calculamos ahora los autovalores de A^T , calculando primero su polinomio característico:

$$|A^T - \lambda I| = \begin{vmatrix} 1 - \lambda & 0 & 2 \\ 2 & 1 - \lambda & 0 \\ 0 & 2 & 1 - \lambda \end{vmatrix} = 0$$

$$|\lambda I - A^T| = -(\lambda - 3) \cdot (\lambda^2 + 3) = 0$$

Lo cual nos lleva a los siguientes valores para λ :

$$\lambda_1 = 3$$

$$\lambda_2 = -\sqrt{3}i$$

$$\lambda_3 = \sqrt{3}i$$

Ahora, el radio espectral o autovalor máximo es

$$\rho(A^T) = \max_{\lambda_i \in \sigma(A^T)} \{|\lambda_i|\} = \lambda_1 = 3$$

Llegado a este punto, el autovector correspondiente al autovalor máximo se obtendrá como solución al siguiente sistema.

$$\begin{pmatrix} 1 - 3 & 0 & 2 \\ 2 & 1 - 3 & 0 \\ 0 & 2 & 1 - 3 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}$$

Por tanto,

$$v = \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}$$

Ahora bien, como hemos explicado antes y siguiendo las condiciones del enunciado del teorema, el vector de Perron buscado cumple que $c > 0$ y $\|c\|_1 = \sum_{i=1}^n c_i = 1$, por lo que dividiendo al vector $v = (1, 1, 1)$ por la suma de sus

componentes se obtiene:

$$c = \left(\frac{1}{3}, \frac{1}{3}, \frac{1}{3} \right)$$

es el vector de Perron y en este caso los tres nodos presentan la misma importancia con un monto total de 1.

Ejemplo 3.3.1.2 Sea otro ejemplo formado por la siguiente matriz:

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

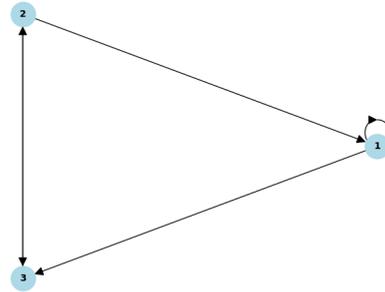


Figura 3.6: Ejemplo de un grafo dirigido y su matriz de adyacencia

Con este ejemplo remarcaremos la importancia de que, en las condiciones establecidas, el vector de Perron es único. En caso contrario, no tendríamos un criterio para seleccionar un vector con las características que buscamos para resolver nuestro problema.

En este caso, de nuevo la matriz es no negativa y el grafo dirigido resultante que observamos en 3.6 es fuertemente conexo. Por tanto, la matriz es irreducible. En este punto, se están cumpliendo los requisitos para aplicar el Teorema de Perron-Frobenius. Por ello, calculemos ahora los distintos autovalores resolviendo el polinomio característico de la misma forma, sin olvidar que nos interesa un autovector de la matriz traspuesta de A:

$$|A^T - \lambda I| = \begin{vmatrix} 1 - \lambda & 1 & 0 \\ 0 & -\lambda & 1 \\ 1 & 1 & -\lambda \end{vmatrix} = 0$$

$$|\lambda I - A^T| = -\lambda^3 + \lambda^2 + \lambda = 0$$

Lo cual nos lleva a los siguientes valores para λ :

$$\begin{aligned}\lambda_1 &= -\frac{-1 + \sqrt{5}}{2} \\ \lambda_2 &= \frac{1 + \sqrt{5}}{2} \\ \lambda_3 &= 0\end{aligned}$$

Ahora, el radio espectral o autovalor dominante es

$$\rho(A^T) = \max_{\lambda_i \in \sigma(A^T)} \{|\lambda_i|\} = \lambda_1, \lambda_2 = \frac{1 + \sqrt{5}}{2}$$

Se puede ver que en este caso la dimensión del autoespacio asociado a $\rho(A)$ es dos. Sin embargo, el autovector asociado a λ_1 , que obtenemos del programa *perron-frobenius.py* es aproximadamente $(0,465, -0,753, 0,465)$, mientras que el autovector asociado λ_2 es $(-0,648, -0,4, -0,648)$, que en este contexto es proporcional a $(0,648, 0,4, 0,648)$ (no importa el signo absoluto de los elementos del vector, sino más bien sus relaciones proporcionales y el hecho de que sean todos positivos o todos negativos).

Por el teorema de Perron-Frobenius, sólo existe un único vector que cumple las características deseadas, y aunque en este caso la dimensión del autoespacio asociado a $\rho(A)$ es dos, sólo uno de los vectores es positivo ($c > 0$), en concreto el asociado a λ_2 . Por tanto este autovector es el vector que, transformándolo de forma que la suma de sus componentes sea 1 (dividiendo cada componente entre la suma de ellas), será el vector de Perron. Según el programa, el vector de Perron es aproximadamente $(0,382, 0,236, 0,382)$, que efectivamente muestra la importancia de cada componente con un monto total de 1.

3.3.2. Aplicación a un Problema Real

Veamos ahora la aplicación de este teorema a un caso real como es la estructura de algunos **campeonatos deportivos**, en especial el baloncesto en otros países como Estados Unidos. Cabe destacar que este ejemplo es distinto a lo mostrado al inicio de la sección anterior, y por tanto, al contrario que los dos ejemplos anteriores, se hará uso del teorema de Perron-Frobenius sin trasponer la matriz A , pues como veremos ahora, no nos interesa.

En una liga europea, se suelen jugar dos partidos entre cada par de equipos de la competición. Sin embargo, en países como Estados Unidos, debido a las grandes distancias los equipos suelen dividirse en conferencias y divisiones, lo que provoca que todos los equipos acaben jugando el mismo número de partidos pero contra rivales distintos, es decir, el número de veces que un equipo se enfrenta a otro no es el mismo (que en general es ida y vuelta en Europa), si no que depende

de cada equipo porque se suele jugar más veces contra aquellos equipos de tu misma conferencia.

Esto sucede durante la temporada regular, y posteriormente se clasifican los *mejores* equipos para las eliminatorias finales. No obstante, el problema que se presenta es determinar con exactitud quién es el mejor equipo acabada la temporada regular. Actualmente, el equipo que tiene el mejor porcentaje de victorias ocupa el primer lugar en su respectiva conferencia. Ahora bien, siguiendo la línea que comentábamos sobre la importancia, la duda reside en si es más importante el **número** de victorias o la **calidad** de las victorias.

Para exponer de manera sencilla este problema veremos un ejemplo reducido en el que se tomarán 4 equipos distintos. Estos serán los siguientes: *Los Angeles Lakers*, *Golden State Warriors*, *Boston Celtics*, y *New York Knicks*. En la siguiente tabla se puede observar a información de las victorias conseguidas por cada equipo en la temporada regular:

	LA Lakers	GS Warriors	Boston Celtics	NY Knicks
LA Lakers	-	6	0	3
GS Warriors	0	-	2	2
Boston Celtics	4	2	-	2
NY Knicks	1	2	4	-

Tabla 3.1: Tabla de clasificación del campeonato por victorias en los partidos

Como se puede observar en la tabla 3.1, cada equipo juega un total de 14 partidos. Además, por distancia, los *Lakers* y los *Warriors* pertenecen a la Conferencia Oeste y, por otro lado, los *Celtics* y los *Knicks* pertenecen a la Conferencia Este, lo cual implica que los equipos juegan 6 partidos contra el equipo de su propia conferencia y 4 partidos contra cada equipo de los de la otra.

Transformamos ahora la tabla anterior en la Matriz A, donde cada a_{ij} representa las victorias de un Equipo i sobre un Equipo j. Además, dividiremos por el número total de partidos como medida de normalización. Así obtenemos la siguiente matriz:

$$A = \begin{pmatrix} 0 & \frac{6}{14} & 0 & \frac{3}{14} \\ 0 & 0 & \frac{2}{14} & \frac{2}{14} \\ \frac{4}{14} & \frac{2}{14} & 0 & \frac{2}{14} \\ \frac{1}{14} & \frac{2}{14} & \frac{4}{14} & 0 \end{pmatrix}$$

Sumando el número de victorias de cada equipo (cada fila), obtenemos los siguientes resultados:

Siguiendo el modelo tradicional pero ignorando la fase eliminatoria, el mayor porcentaje de victorias lo tienen *Los Angeles Lakers*, con 9 victorias en 14

Equipos	Nº Victorias
LA Lakers	9
GS Warriors	4
Boston Celtics	8
NY Knicks	7

Tabla 3.2: Tabla de victorias de cada equipo

partidos. Por tanto, la clasificación sería:

Los Angeles Lakers → Boston Celtics → New York Knicks → Golden State Warriors

Sin embargo, podemos observar que el líder *Los Angeles Lakers* acumuló muchas victorias contra el peor equipo de la clasificación, los *Golden State Warriors*. Además, el equipo que quedó en segundo lugar, los *Boston Celtics*, ganaron todos los partidos que jugaron contra los campeones. Esto nos hace preguntarnos si el resultado final es verdaderamente justo, pues posiblemente la 'calidad' de las victorias de los *Celtics* sea mayor. Afortunadamente, podemos resolver nuestras dudas y comprobarlo aplicando el **Teorema de Perron-Frobenius**.

Ahora bien, es importante darse cuenta que en este caso a cada equipo le interesa la importancia de cada victoria contra cada equipo contrario. Dicho de otra manera, esta vez buscamos un vector de importancias de la forma:

$$\begin{aligned}
 c_{Lakers} &= K \sum_{j \in V}^n a_{(Lakers\ j)} c_j, \\
 &\vdots \\
 c_{Knicks} &= K \sum_{j \in V}^n a_{(Knicks\ j)} c_j,
 \end{aligned}$$

donde $V = \{Lakers, Warriors, Celtics, Knicks\}$, y que en general:

$$c_i = K \sum_{j \in V}^n a_{ij} c_j,$$

$$\lambda c = Ac$$

En otras palabras, en este caso no nos interesa trasponer la matriz A , ya que por ejemplo, mirando la tabla 3.1, la importancia final c_{Lakers} será una proporción de las importancias de sus victorias (aquellas que aparecen en su fila de la matriz y no en la columna), al contrario que el modelo anterior en el que nos interesábamos por quién nos referenciaba (columnas de la matriz de adyacencia). En el caso de

los *Lakers* serían 6 contra los *Warriors*, 0 contra los *Celtics* y 3 contra los *Knicks*. Esto es un claro ejemplo de cómo aplicar el teorema según las necesidades de cada problema.

Ahora bien, para resolver computacionalmente el problema usaremos un programa desarrollado en Python (*perron-frobenius.py*) que, utilizando la biblioteca NumPy, se encargará de ello. Primero se carga la matriz y, en segundo lugar, debemos comprobar si la matriz es cuadrada, no negativa e irreducible, pues son los requisitos del teorema. Para ello, aunque el programa comprueba que la matriz es irreducible analizando si esta matriz es primitiva (Observación 3.1), podemos establecer un grafo dirigido ignorando los pesos que, en efecto, es fuertemente conexo y por tanto la matriz de adyacencia es irreducible. Dicho grafo se puede observar en la figura 3.7:

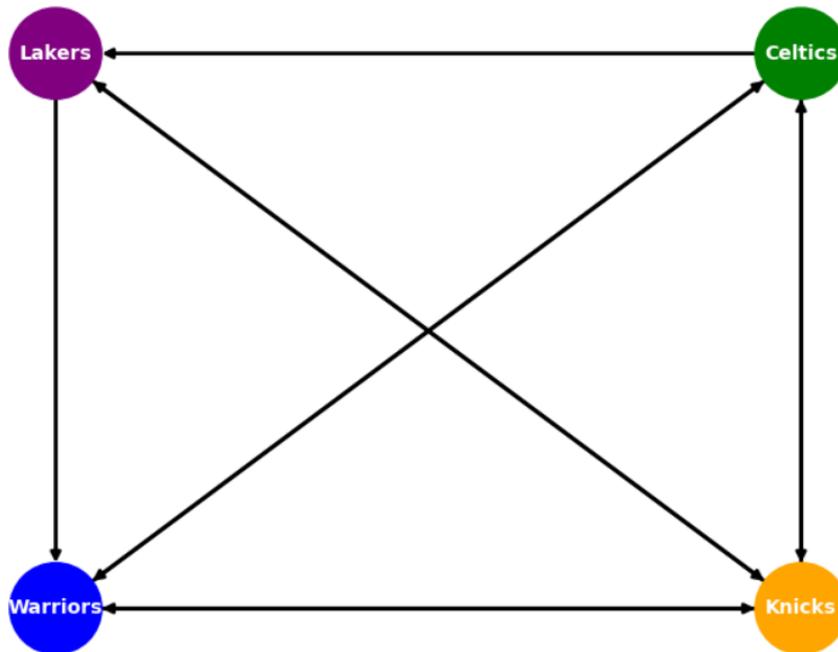


Figura 3.7: Ejemplo Partidos NBA Fuertemente Conexo

En tercer lugar, el programa se encargará de calcular todos los autovalores y autovectores. A continuación, se observan en 3.1, 3.2, 3.3 y 3.4, los fragmentos de código relevantes para el **cálculo del vector de Perron** y la visualización de los resultados obtenidos:

Código 3.1: Métodos de Perron-Frobenius.py

```

1 # Verificar si la matriz es cuadrada
2 if matrix.shape[0] != matrix.shape[1]:
3     raise ValueError("La matriz no es cuadrada.")
4
5 # Verificar si la matriz es no negativa
6 if np.any(matrix < 0):
7     raise ValueError("La matriz contiene elementos negativos.")
8
9 # Verificar si la matriz es primitiva, y por tanto irreducible
10 if not np.all(np.linalg.matrix_power(matrix, matrix.shape[0]) >
11              0):
12     raise ValueError("La matriz no es primitiva.")
13
14 # Obtener autovalores y autovectores de la matriz
15 eigenvalues, eigenvectors = np.linalg.eig(matrix)

```

dando el siguiente resultado aproximado:

Código 3.2: Datos devueltos por Perron-Frobenius.py

```

1 Autovalor 1: (0.47629186134276125+0j)
2 Autovector 1: [0.533+0.j 0.33+0.j 0.576+0.j 0.5245+0.j]
3
4 Autovalor 2: (-0.148+0.234j)
5 Autovector 2: [-0.679+0.j 0.047-0.3j 0.067+0.53j
6              0.37513405-0.14082163j]
7
8 Autovalor 3: (-0.148-0.234j)
9 Autovector 3: [-0.68-0.j 0.0475+0.3j 0.067-0.53j
10              0.37513405+0.14082163j]
11
12 Autovalor 4: (-0.18+0j)
13 Autovector 4: [-0.0126+0.j -0.419+0.j -0.32+0.j 0.849+0.j]

```

Como podemos observar, sólo uno de los autovalores devueltos es real y positivo, es decir, el autovalor dominante es $\rho(A) = \lambda_1 = 0,47629186134276125 \approx 0,4763$. Por el Teorema de Perron-Frobenius el autovector que buscamos es el único autovector de A asociado al autovalor $\rho(A)$ tal que $c \in \mathbb{R}^n$, $c > 0$ y $\|c\|_1 = 1$. Es decir, el **autovector asociado a λ_1** normalizado para que $\|c\|_1 = 1$ es el vector de Perron que buscamos, lo cual se obtiene con el siguiente fragmento de código:

Código 3.3: Más Métodos de Perron-Frobenius.py

```

1 # Obtener el índice del valor propio de Perron (maximo valor
  propio)
2 perron_index = np.argmax(np.real(eigenvalues))
3 # Obtener el autovector correspondiente al valor propio de
  Perron
4 perron_vector = np.real(eigenvectors[:, perron_index])
5 # Normalizar el autovector para que la suma de sus componentes
  sea 1
6 perron_vector /= np.sum(perron_vector)
7 # Redondear el resultado a 4 decimales
8 perron_vector = np.round(perron_vector, decimals=4)

```

Finalmente, obtenemos que el **vector de Perron** resultante, aproximado a 4 decimales, es:

Código 3.4: Resultado Perron-Frobenius.py

```

1 Vector propio de Perron:
2 [0.2714 0.1681 0.2934 0.2671]

```

Cumpliendo nuestras sospechas, tras realizar estos cálculos hemos obtenido un vector de Perron que propone un orden de importancia de la siguiente forma:

Boston Celtics → **Los Angeles Lakers** → New York Knicks → Golden State Warriors

Esta ordenación difiere de la anterior, pues ahora los *Boston Celtics* son el mejor equipo. De hecho, la importancia entre los tres primeros clasificados es bastante cercana debido a que aunque los *Boston Celtics* ganaron los 4 partidos contra los *Lakers*, que son los que más victorias presentan, los *New York Knicks* ganaron 4 de 6 partidos contra los *Boston Celtics*, lo que aporta también mucha calidad a las victorias del tercer clasificado.

De esta manera queda demostrado que fijándonos en la ponderación de la importancia de los equipos a los que vences y no sólo en la proporción de victorias, se llega a un modelo de clasificación distinto gracias al teorema de Perron-Frobenius.

4

Utilidad en el Algoritmo PageRank

Ya hemos visto los fundamentos teóricos del Teorema de Perron-Frobenius, así como su enunciado, demostración y cómo aplicarlo. Es ahora cuando se nos brinda la oportunidad de sumergirnos completamente en cómo Perron-Frobenius encuentra una utilidad crucial en el algoritmo PageRank, con el que Google ordena los resultados de las billones de búsquedas diarias.

Este algoritmo, desarrollado por *Sergei Brin* y *Lawrence Page* en 1998, representa un hito significativo en el ámbito de los motores de búsqueda. PageRank se ha convertido en un pilar fundamental para medir la importancia relativa de las páginas web, pues es un algoritmo revolucionario que utiliza la estructura de enlaces entre páginas para determinar su relevancia, y el Teorema de Perron-Frobenius proporciona los fundamentos matemáticos necesarios para comprender y garantizar su funcionamiento.

Dentro de este contexto, nos interesa examinar cómo se modeló la red, buscando obtener una matriz que funcione como la matriz de adyacencia de un grafo que represente toda la web y que satisfaga los requisitos del Teorema de Perron-Frobenius. Este proceso nos permitirá obtener el vector deseado que proporciona la clasificación de las páginas, conocido como *vector PageRank*. Para abordar este objetivo, exploraremos paso a paso este proceso, basándonos en el desarrollo teórico de estudios como [8], [9], y [2].

4.1. El Modelo de la Red

En primera instancia, relacionaremos toda la red de la misma manera que hacemos con los artículos. Al fin y al cabo, la red se compone de muchos contenidos o sitios web y entre ellos existen enlaces de una página a otra. Es fácil ver que si tenemos un grupo de $n \in \mathbb{N}$ páginas web, o más bien, toda la **World Wide Web** (WWW) con sus correspondientes hiperenlaces, se podría ver como un **grafo dirigido** $G = (X, E)$ de forma que cada una de las páginas web se identifica como un nodo en el grafo, y si entre dos páginas i y j hay un hiperenlace (por ejemplo, de la página i hay un hiperenlace a la página j), entonces el grafo G tiene una arista dirigida de la forma $i \rightarrow j$.

En este contexto, volvemos al punto principal, cómo darle una determinada importancia a cada página del grafo para que el modelo funcione correctamente. Como venimos comentando anteriormente en este trabajo, la opción sencilla de suponer que una página es más importante que otra si recibe más enlaces de las demás (es decir, está referenciada por un mayor número de páginas y por tanto existen más aristas dirigidas en el grafo de la red que la apuntan), tiene una **debilidad**, y es que un usuario podría crear un gran número de páginas que tengan enlaces a una página en concreto, provocando que todo este sistema fuese fácil de influir.

La siguiente figura muestra un grafo simple con el que facilitar el entendimiento de esto último. En él, representando la importancia según el grosor de los nodos, seis páginas web apuntan a otra, convirtiendo a esta en la más importante. Sin embargo, la siguiente más importante es P_1 , que está apuntada por P_3 y P_7 , páginas a las cuales nadie apunta. Por otro lado, hay que prestar atención a P_5 , la cuál es apuntada por la página más importante, pero es igual de importante que el resto de nodos siguiendo este modelo.

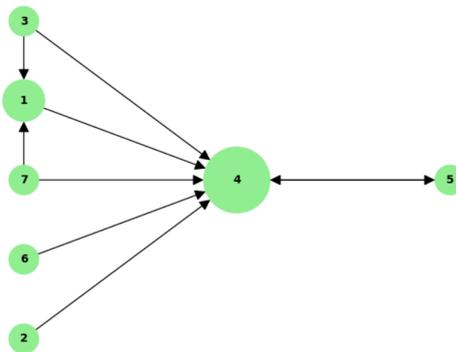


Figura 4.1: Ejemplo 1 - Importancia Páginas Web

Obviamente, debemos tomar el modelo en el que la 'importancia' de cada enlace tiene prioridad sobre el 'número' de enlaces. De esta forma, si una página

P_i obtiene 1000 enlaces de páginas sin importancia alguna (es decir, aquellas en las que no existen otras páginas con un enlace hacia ellas), dicha página será menos importante que una página P_j que reciba un número mucho menor de enlaces, pero vengan de páginas importantes como *Amazon.com* o *Facebook.com*. La figura siguiente representa de manera realista la importancia de dichas páginas, pues ahora P_5 es la segunda página más importante al ser apuntada por aquella a la que todos apuntan, obteniendo importancia de manera proporcional a la **calidad** del enlace recibido.

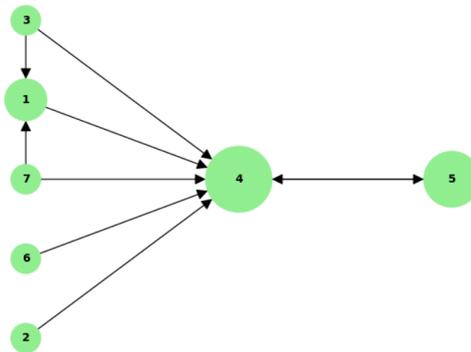


Figura 4.2: Ejemplo 2 - Importancia Páginas Web

Matricialmente, nos interesará representar la información que obtenemos de la World Wide Web formando una matriz M de dimensiones $n \times n$ donde las filas y las columnas serán etiquetadas de la forma P_1, \dots, P_n , tomando cada entrada m_{ij} el valor de 1 si hay un enlace de la página P_j a la página P_i y el valor de 0 en caso contrario. Esta matriz es justamente **la matriz traspuesta de la matriz de adyacencia** del grafo dirigido. Nos interesa tomar la matriz M de esta manera puesto que, en la matriz de adyacencia, las filas indican cuántos enlaces salen de una página dada y en las columnas aparecen tantos *unos* como enlaces haya hacia la página indexada por esa columna. Para utilizar la matriz en el problema de PageRank debemos por tanto trasponerla, para poder calcular la proporción de importancia de cada fila (columna traspuesta) fijándonos en cuántos enlaces llegan a ella. Numéricamente se calcularía la importancia de cada fila como:

$$\begin{aligned}
 c_1 &= K(a_{11}c_1 + a_{12}c_2 + \dots + a_{1n}c_n) \\
 c_2 &= K(a_{21}c_1 + a_{22}c_2 + \dots + a_{2n}c_n) \\
 &\vdots \\
 c_n &= K(a_{n1}c_1 + a_{n2}c_2 + \dots + a_{nn}c_n)
 \end{aligned}$$

De lo cual se obtiene que:

$$\begin{pmatrix} c_1 \\ \vdots \\ c_n \end{pmatrix} = K \begin{pmatrix} m_{11} & \dots & m_{1n} \\ \vdots & \ddots & \vdots \\ m_{n1} & \dots & m_{nn} \end{pmatrix} \begin{pmatrix} c_1 \\ \vdots \\ c_n \end{pmatrix}$$

$$\frac{1}{K} \begin{pmatrix} c_1 \\ \vdots \\ c_n \end{pmatrix} = \begin{pmatrix} m_{11} & \dots & m_{1n} \\ \vdots & \ddots & \vdots \\ m_{n1} & \dots & m_{nn} \end{pmatrix} \begin{pmatrix} c_1 \\ \vdots \\ c_n \end{pmatrix}$$

Obteniendo, de igual que manera que en la sección 3.3, tomando $\lambda = \frac{1}{K}$:

$$Mc = \lambda c$$

ó

$$A^t c = \lambda c$$

donde A es la matriz de adyacencia del grafo dirigido ($M = A^t$), λ es un valor propio de M y $c = (c_1, \dots, c_n)$ es un vector propio asociado al valor propio λ . Además, este vector c expresa la importancia de cada página a partir de las importancias de las demás páginas que están conectadas a ella.

De esta manera, hemos conseguido transformar el problema de calcular las importancias en un problema de **autovalores y autovectores**, acercándonos a la teoría de Perron-Frobenius. Sin embargo, con esta aproximación surgen **problemas** como, por un lado, la dificultad de calcular los autovectores de la matriz M, que puede llegar a estar formada por millones de páginas, y por otro lado, no se cumplen todos los requisitos del Teorema de Perron-Frobenius, por lo que no se asegura la unicidad del autovector con coordenadas positivas. Explicado de otro modo, ¿y si creo una página sin enlazarla a ninguna otra?, en ese caso estoy forzando que el grafo de la red no sea fuertemente conexo, por lo que la matriz de adyacencia no es irreducible y no se puede aplicar el teorema de Perron-Frobenius. Es por ello que ahora veremos una nueva aproximación que nos proporciona una solución alternativa y aproximada, pero mucho más rápida y con mejores características.

4.2. El Surfista Aleatorio

En esta nueva aproximación, nos basaremos en la **Hipótesis del Surfista Aleatorio**. En ella, supondremos que un individuo, al que se le denomina el *surfista*, está navegando por la red. Se plantea entonces que su forma de navegar por la World Wide Web es completamente aleatoria. En concreto, si en un cierto instante el individuo se encuentra en una página P_i , entonces selecciona al

azar alguna de las páginas que están conectadas a ella para saltar siguiendo una probabilidad uniforme.

En otras palabras, tenemos un modelo **probabilístico** en el que no sabemos donde estará un instante después, pero sí con qué probabilidad estará en cada uno de los posibles destinos, es decir, en cada uno de los posibles enlaces desde la página en la que se encuentra actualmente.

De hecho, el modelo es *dinámico*, porque se puede aplicar este argumento a cada movimiento de manera recursiva, resultando en un paseo aleatorio por el grafo. Según la Hipótesis del Surfista Aleatorio, si se navega de forma aleatoria, entonces **pasaremos con mayor frecuencia por los nodos de mayor importancia**. Una posible situación del modelo probabilístico para este paseo aleatorio sería la siguiente:

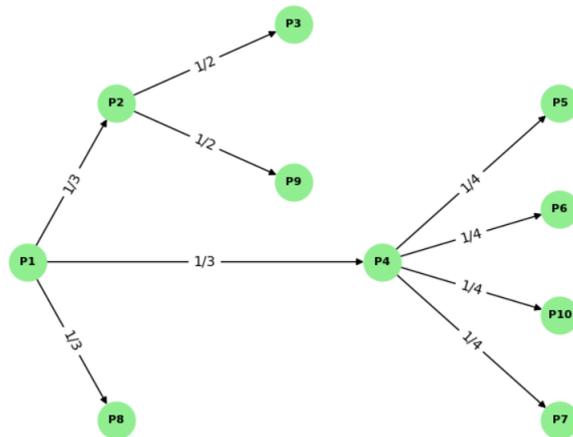


Figura 4.3: Ejemplo Paseo Aleatorio

Como se puede observar, si el navegante se encontrase en P_1 , una página que enlaza con otras tres, el navegante sortearía su siguiente movimiento con una probabilidad uniforme de $1/3$ para cada una de ellas. En su siguiente movimiento, si hubiese avanzado hacia P_2 , se repetiría el sorteo con probabilidad $1/2$, y si hubiese avanzado a P_4 lo haría con probabilidad $1/4$. En general, se formaliza este proceso de la siguiente manera:

Sea N_j el **número de enlaces salientes** desde la página P_j . Recordando que la matriz M era la traspuesta de la matriz de adyacencia del grafo, entonces N_j es la **suma de las entradas de cada columna** de M . Expresado en términos del grafo de la red asociado, $N_j = gr_{out}(j)$. Ahora, construiremos una nueva matriz M' sustituyendo cada entrada m_{ji} tal que:

$$m'_{ij} = \frac{m_{ij}}{N_j}$$

Esta nueva matriz está formada por números no negativos entre el 0 y el 1, y en concreto, la suma de los registros de **cada columna suma 1**. Esto puede comprobarse imaginando que fijamos una columna j (página j), entonces esa

columna $\begin{pmatrix} m_{1j} \\ \vdots \\ m_{nj} \end{pmatrix}$ contiene tantos *unos* como enlaces salgan de dicha página, es decir, la suma de dichos *unos* es justamente N_j , por lo que si dividimos cada componente de la columna entre N_j , la nueva suma de la columna será:

$$\frac{1 + \dots + 1(N_j \text{ veces})}{N_j} = \frac{N_j}{N_j} = 1$$

Dicho de otra manera, M' es una **matriz estocástica**.

Llegados a este punto, hemos realizado nuestro primer ajuste y convertido M' en una matriz especial diseñada para abordar los **camino aleatorios** en una red. En realidad, los procesos de paseo aleatorio representan un ejemplo específico de procesos más amplios denominados **cadenas de Márkov**. Por lo tanto, en la próxima sección, aunque de manera superficial, se explorará la relevancia de los procesos estocásticos de Márkov y las llamadas **matrices de transición** en el contexto de las redes.

4.3. Cadenas de Márkov y Matrices de transición

Durante esta sección conoceremos los conceptos básicos de las cadenas de Márkov gracias a una revisión de artículos como [10] y [11] y mostraremos cómo afectan al modelo web que desarrollábamos en la sección anterior.

4.3.1. Conceptos Básicos

Definición 4.3.1.1. *Un proceso estocástico es una sucesión de observaciones X_1, \dots, X_t tal que los valores de estas observaciones no se pueden predecir exactamente, pero se pueden especificar las probabilidades para los distintos valores posibles en cualquier instante de tiempo.*

Definición 4.3.1.2. *Una cadena de Márkov es un proceso estocástico en el que si el estado actual X_t y los estados previos X_1, \dots, X_{t-1} son conocidos, entonces la probabilidad del estado futuro X_{t+1} no depende de los estados anteriores X_1, \dots, X_{t-1} y solamente depende del estado actual X_t . Es decir, para todo t instante de tiempo y para cualquier sucesión de estados s_1, \dots, s_{n+1} ,*

$$P(X_{t+1} = s_{n+1} \mid X_1 = s_1, X_2 = s_2, \dots, X_t = s_n) = P(X_{t+1} = s_{n+1} \mid X_t = s_n)$$

Definición 4.3.1.3. Una probabilidad de transición se define como cada probabilidad condicionada de la forma:

$$P(X_{n+1} = s_j \mid X_n = s_i)$$

Definición 4.3.1.4. Una cadena de Márkov tiene probabilidades de transición estacionarias si, para cualquier par de estados s_i y s_j , existe una probabilidad de transición p_{ij} tal que

$$P(X_{n+1} = s_j \mid X_n = s_i) = p_{ij} \quad \text{para todo instante de tiempo } t,$$

es decir, las probabilidades de pasar de un estado a otro se mantienen constantes a lo largo del tiempo. En este caso, se dice entonces que la cadena de Márkov es homogénea o estacionaria. En nuestro caso, la matriz de transición es una matriz estocástica constante $P = [p_{ij}]$. En lo sucesivo, asumiremos que nuestras cadenas son estacionarias.

Definición 4.3.1.5. Dada una cadena de Márkov con n estados posibles s_1, \dots, s_n y probabilidades de transición estacionarias, donde $p_{ij} = P(X_{t+1} = s_j \mid X_t = s_i)$, la matriz de transición P es:

$$P = \begin{bmatrix} p_{11} & p_{12} & \cdots & p_{1n} \\ p_{21} & p_{22} & \cdots & p_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ p_{n1} & p_{n2} & \cdots & p_{nn} \end{bmatrix}$$

Además, la matriz de transición P de cualquier cadena de Márkov finita con probabilidades de transición estacionarias es una matriz estocástica.

Definición 4.3.1.6. Dada una cadena de Márkov con n posibles estados s_1, \dots, s_n y matriz de transición P , notamos $p_{ij}^{(2)} = P(X_{t+2} = s_j \mid X_t = s_i)$, siendo $p_{ij}^{(2)}$ el elemento de la i -ésima fila y j -ésima columna de la matriz P^2 . Generalizando, P^m es la matriz de probabilidades $p_{ij}^{(m)}$ de que la cadena pase del estado s_i al estado s_j en m pasos. En otras palabras, para cualquier valor de m , P^m es la matriz de transición de m pasos de la cadena de Márkov.

Definición 4.3.1.7. Una cadena de Márkov irreducible es una cadena cuya matriz de probabilidad de transición P es una matriz irreducible. Este tipo de cadenas implican que es posible alcanzar cualquier estado desde cualquier otro en un número finito de pasos.

Definición 4.3.1.8. Una cadena de Márkov periódica es aquella en la que existe un número entero $k > 1$ tal que es posible regresar a cualquier estado inicial i únicamente en múltiplos de k pasos. El mayor valor de k que cumple esta condición para todos los estados (máximo común divisor) se denomina el periodo de la cadena. Es conocido que las cadenas periódicas son cadenas **irreducibles** cuya matriz de probabilidad de transición P es **no primitiva**. Un ejemplo detallado de este comportamiento se proporcionará posteriormente.

Definición 4.3.1.9. Por otro lado, una cadena de Márkov aperiódica (no periódica) es una cadena en la que cada estado tiene un periodo de 1. Además, las cadenas aperiódicas son cadenas irreducibles cuya matriz de probabilidad de transición P es una **matriz primitiva**.

Definición 4.3.1.10. Una cadena de Markov se llama ergódica si es irreducible, aperiódica y estacionaria. Esto asegura que, independientemente del estado inicial, la cadena eventualmente converge a una distribución de probabilidad.

Para entender esto último, nos enfocaremos en los **vectores de estado**, una herramienta esencial para comprender la distribución de probabilidades en cada instante específico de una cadena de Márkov.

Definición 4.3.1.11. El vector $\mathbf{v} = (v_1, \dots, v_n)$ se llama vector de probabilidades si cumple con las siguientes condiciones:

1. $v_i \geq 0$ para $i = 1, \dots, n$.
2. $\|\mathbf{v}\|_1 = 1$, es decir, $\sum_{i=1}^n v_i = 1$.

Definición 4.3.1.12. El Vector de Estado Inicial (\mathbf{v}_0) es un vector de probabilidades que representa las probabilidades iniciales de estar en cada estado al comienzo del proceso. Es un vector de dimensión \mathbf{n} , donde \mathbf{n} es el número de estados en la cadena de Márkov.

Observación. El vector de probabilidades iniciales y la matriz de transición determinan la probabilidad para el estado de la cadena de Márkov en el **segundo instante de tiempo**. De hecho, dicha probabilidad viene dada por el vector $\mathbf{v} = P \cdot \mathbf{v}_0$.

Definición 4.3.1.13. Dado un estado n , el vector de distribución de estado t (\mathbf{v}_t) representa la distribución de probabilidades de estar en cada estado después de t pasos. Se calcula mediante la multiplicación del vector de estado inicial por la matriz de transición elevada a la n -ésima potencia, es decir, $\mathbf{v}_t = \mathbf{P}^t \mathbf{v}_0$.

Definición 4.3.1.14. El vector de estado estacionario (\mathbf{v}_{est}) es aquel vector que contiene las probabilidades estacionarias, es decir, las probabilidades límite a las que la cadena de Márkov converge a medida que t tiende a infinito. Este vector satisface la ecuación $\mathbf{v}_{est} = \lim_{t \rightarrow \infty} \mathbf{v}_t = \lim_{t \rightarrow \infty} \mathbf{P}^t \mathbf{v}_0$, donde P es la matriz de transición y \mathbf{v}_0 el vector de estado inicial.

Es por ello que el *vector estacionario* cumple que $Pv_{est} = v_{est}$, por lo que es un **autovector** de la matriz de transición **asociado al autovalor $\lambda = 1$** , y que además al ser P estocástica se sabe que entonces dicho autovalor λ además es el **radio espectral** de P por el teorema 3.1.1.

Observación. *Es muy importante destacar que esta propiedad del vector de estado estacionario se verifica en particular para cadenas de Márkov ergódicas (como se ha explicado en la Definición 4.3.1.10, es decir, para cadenas estacionarias, irreducibles y aperiódicas. Además, bajo estas condiciones dicha convergencia ocurre para cualquier vector inicial. Recordando que la no periodicidad sumada a la irreducibilidad implican la primitividad, entonces estamos hablando de matrices de transición estocásticas y primitivas.*

Relacionando esta teoría de las cadenas de Márkov, matrices de transición y vectores de estado con el modelo que comentábamos anteriormente, no es difícil darse cuenta de que, en efecto, nuestra matriz M' es una **matriz de transición**. En ella, cada m'_{ij} representa la probabilidad de pasar del estado P_j al estado P_i (recordemos que había P_1, \dots, P_n páginas web y que la matriz estaba traspuesta).

Imaginemos que un surfista aleatorio se encuentra en la página P_k en el primer instante de tiempo, es decir, el vector de estado inicial tendría todas sus entradas a 0, excepto la posición k, que sería 1. El surfista entonces sortea, como comentábamos en la sección anterior, su siguiente salto entre las N_k páginas destino, asignando una probabilidad de $1/N_k$ a cada una. Multiplicando la matriz M' por el vector inicial se obtendrá:

$$\begin{pmatrix} \dots & \dots & m'_{1k} & \dots \\ \vdots & \ddots & \vdots & \vdots \\ \dots & \dots & m'_{kk} & \dots \\ \vdots & \vdots & \ddots & \vdots \\ \dots & \dots & m'_{nk} & \dots \end{pmatrix} \cdot \begin{pmatrix} 0 \\ \vdots \\ 1 \\ \vdots \\ 0 \end{pmatrix} = \begin{pmatrix} m'_{1k} \\ \vdots \\ m'_{kk} \\ \vdots \\ m'_{nk} \end{pmatrix}$$

El vector resultante también es un vector de probabilidades porque las entradas son o bien 0, o bien $1 / N_k$, y hay justamente N_k entradas no nulas, por lo que el vector contiene componentes que son números entre 0 y 1 y además suman 1 en total. Como se explica en la Observación 4.3.1, el vector resultante determina exactamente la **probabilidad** con la que **el surfista podría encontrarse**, tras una unidad de tiempo, en cada una de las **páginas de la red**.

En este punto observamos que este modelo es muy manejable e interesante porque, según la 4.3.1.6, como M' es nuestra matriz de transición, bastaría con multiplicar $(M')^2$ por el vector inicial v_0 para conocer con qué probabilidad se encontraría el surfista en cada una de las páginas tras **dos instantes** de tiempo. Así, se repite de manera análoga, y por tanto ya tenemos una forma de calcular las probabilidades de transición entre páginas web para varias unidades de tiempo.

Ahora bien, según la Hipótesis del Surfista Aleatorio, si se navega por un tiempo suficientemente prolongado, acabaremos pasando con mayor frecuencia por los nodos de mayor importancia, es decir, nos interesa la estacionariedad de este vector. En términos de cadenas de Márkov, la estacionariedad en este contexto implica que, dado un tiempo suficientemente largo, la distribución de probabilidad de estar en una página web particular converge a una distribución estacionaria, que es independiente del estado inicial. Esto significa que **las páginas web alcanzan una importancia relativa constante a medida que el usuario sigue navegando**, y estará representada por un vector límite de distribución de probabilidad que resultará en el vector PageRank que buscamos.

4.3.2. Ejemplo

A continuación, se propone un ejemplo sencillo para ilustrar lo que hemos aprendido. Sea el siguiente conjunto de páginas web $P = \{P_1, P_2, P_3, P_4\}$. Podemos observar la figura 4.4 el grafo correspondiente:

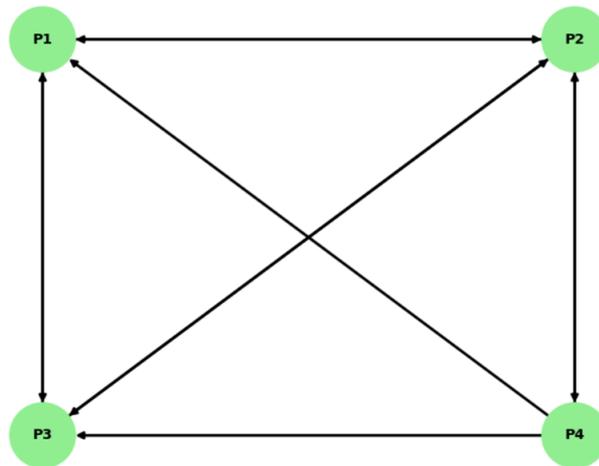


Figura 4.4: Ejemplo Grafo Web - Matrices de transición

Siguiendo los pasos aprendidos hasta ahora, en primer lugar debemos construir la matriz de adyacencia del grafo dado y trasponerla para conseguir la matriz M . Como observación, la matriz de adyacencia en este caso es no negativa e irreducible porque el grafo dado es fuertemente conexo.

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

En segundo lugar, debemos obtener la matriz de transición M' con la trans-

formación

$$m'_{ij} = \frac{m_{ij}}{N_j}$$

donde N_j indica la suma de las entradas de cada columna de M , o expresado de otra forma, $gr_{out}(j)$. Sea entonces M' la matriz de transición:

$$M' = \begin{pmatrix} 0 & 1/3 & 1/2 & 1/3 \\ 1/2 & 0 & 1/2 & 1/3 \\ 1/2 & 1/3 & 0 & 1/3 \\ 0 & 1/3 & 0 & 0 \end{pmatrix}$$

Como hemos comentado en la sección anterior, esta matriz indica la probabilidad de pasar del estado P_j al estado P_i para cada m'_{ij} . Además, se comprueba que esta matriz es estocástica ya que $\forall j$ se cumple $\sum_i m'_{ij} = 1$.

Ahora, supongamos en este ejemplo que deseamos conocer las probabilidades de estar en otra página partiendo de la página 2. El vector de estado inicial \mathbf{v}_0 que usaremos entonces es el vector columna

$$v_0 = \begin{pmatrix} 0 \\ 1 \\ 0 \\ 0 \end{pmatrix}$$

Para conocer estas probabilidades debemos entonces multiplicar la matriz de transición obtenida por el vector de estado inicial:

$$M' \cdot v_0 = \begin{pmatrix} 0 & 1/3 & 1/2 & 1/3 \\ 1/2 & 0 & 1/2 & 1/3 \\ 1/2 & 1/3 & 0 & 1/3 \\ 0 & 1/3 & 0 & 0 \end{pmatrix} \cdot \begin{pmatrix} 0 \\ 1 \\ 0 \\ 0 \end{pmatrix} = \begin{pmatrix} 1/3 \\ 0 \\ 1/3 \\ 1/3 \end{pmatrix}$$

Así, observamos cómo podemos determinar la probabilidad de estar en cada página de la red en el segundo instante de tiempo. En este caso, existe una probabilidad de $1/3$ de saltar a las páginas P_1 , P_3 y P_4 , mientras que no hay posibilidad de permanecer en la página 2 (no existe enlace a ella misma en el grafo de la red).

Ahora bien, si deseamos seguir navegando por la red podemos seguir obteniendo *vectores de estado* t (v_t) con las respectivas probabilidades de encontrarnos en cada una de las páginas después de t instantes o saltos. Multiplicando la matriz de transición una y otra vez por estos vectores de estado nos irá dando información sobre cada salto, pero el interés reside en observar que en algún momento el vector de estado tiende a un valor, el cual será el *vector de estado estacionario*

que buscamos. Como hemos explicado en la Observación 4.3.1, la matriz debe ser también primitiva para asegurar dicha convergencia.

Comprobémoslo con un pequeño programa desarrollado en Python (*transition-matrix.py*) que se encarga de, tras comprobar que la matriz es primitiva, multiplicar en bucle la matriz de transición por los respectivos vectores de estado. Se observa entonces que este vector acaba convergiendo, como podemos ver en el siguiente código 4.1:

Código 4.1: Resultado de transition-matrix.py

```

1 Matriz de transicion:
2 [0, 0.3333, 0.5, 0.3333]
3 [0.5, 0, 0.5, 0.3333]
4 [0.5, 0.3333, 0, 0.3333]
5 [0, 0.3333, 0, 0]
6
7 Vector Inicial:
8 [0, 1, 0, 0]
9
10 t   Vectores de estado t
11 1   [0.3333 0 0.3333 0.3333]
12 2   [0.2778 0.4444 0.2778 0]
13 3   [0.287 0.2778 0.287 0.1481]
14 4   [0.2855 0.3364 0.2855 0.0926]
15 5   [0.2858 0.3164 0.2858 0.1121]
16 6   [0.2857 0.3231 0.2857 0.1055]
17 7   [0.2857 0.3209 0.2857 0.1077]
18 8   [0.2857 0.3216 0.2857 0.1070]
19 9   [0.2857 0.3214 0.2857 0.1072]
20 10  [0.2857 0.3214 0.2857 0.1071]
21 11  [0.2857 0.3214 0.2857 0.1071]

```

Se puede ver que a partir del vector de estado 10 (\mathbf{v}_{10}), el vector no cambia. En otras palabras, hemos encontrado el vector estacionario \mathbf{v}_{est} tal que $M'\mathbf{v}_{est} = \mathbf{v}_{est}$, es decir, el vector estacionario es un vector propio de la matriz de transición M' asociado al autovalor $\lambda = 1$, y de hecho dicho autovalor λ es el radio espectral de M' por ser dicha matriz estocástica.

Destacamos de nuevo que esta convergencia se da gracias a que la matriz es estocástica y primitiva, siguiendo la Observación 4.3.1. Bajo estas condiciones, una de las propiedades más interesantes del vector estacionario es que se puede demostrar que llegaremos a dicho resultado sin importar cuál sea el vector inicial. Dicho de otra manera, si cambiamos el vector inicial (\mathbf{v}_0) y esta vez empezamos desde otra página, por ejemplo, la página P_4 , entonces

$$v_0 = \begin{pmatrix} 0 \\ 0 \\ 0 \\ 1 \end{pmatrix}$$

y el vector estacionario se sigue obteniendo de la misma manera:

```

1 1 [0.3333 0.3333 0.3333 0]
2 2 [0.2778 0.3333 0.2778 0.1111]
3 3 [0.2870 0.3148 0.2870 0.1111]
4 4 [0.2855 0.3241 0.2855 0.1049]
5 5 [0.2858 0.3205 0.2858 0.108 ]
6 6 [0.2857 0.3218 0.2857 0.1068]
7 7 [0.2857 0.3213 0.2857 0.1073]
8 8 [0.2857 0.3215 0.2857 0.1071]
9 9 [0.2857 0.3214 0.2857 0.1072]
10 10 [0.2857 0.3214 0.2857 0.1071]
11 11 [0.2857 0.3214 0.2857 0.1071]
```

Gracias a esto, podemos **ordenar por importancia** las páginas web. Sin embargo, en este caso en concreto, obtenemos dos importancias iguales entre las páginas P_1 y P_3 . Además, tomando más decimales y mirando en las iteraciones siguientes se comprueba que el resultado sigue siendo el mismo:

```

1 17 [0.28571429 0.32142857 0.28571429 0.10714286]
```

En este caso, habría otros elementos en el buscador de Google que terminarían decidiendo que página debería mostrarse por encima de la otra. En conclusión, se obtiene así una ordenación de las páginas de manera que conocemos cuál es la página más importante y cuál es la menos importante. Al usar el buscador, el resultado por orden sería el siguiente:

PÁGINA 2 → PÁGINA 1/PÁGINA 3 → PÁGINA 4

4.4. PageRank

4.4.1. Problema con los Nodos Dangling y su Ajuste Correspondiente

Según lo que acabamos de ver en la sección anterior, ya habríamos explicado cómo obtener el vector de ordenación de páginas web que buscábamos. Posteriormente formalizaremos el método usado, pues ha sido un primer vistazo al usado por Google y conocido como *método de las potencias*. Aunque hemos demostrado la eficacia del modelo mediante un ejemplo específico, es importante señalar que el éxito se logró cuando la matriz de adyacencia del grafo era primitiva (y por tanto, no negativa e irreducible). Aunque generalizamos que cualquier matriz será no negativa con este modelo, surge la pregunta de qué ocurriría **si la matriz no fuera irreducible**, es decir, si **el grafo asociado no fuera fuertemente conexo**.

En efecto, si tomamos la World Wide Web en su totalidad, es un hecho que el grafo formado por las más de mil millones de páginas web no es fuertemente conexo. La complicación radica en las páginas que se identifican como '*nodos dangling*', que son aquellos nodos en un grafo dirigido que **carecen de enlaces salientes**. En términos sencillos, se trata de un nodo que no posee conexiones hacia otros nodos mediante aristas salientes, lo que implica que, una vez alcanzado dicho nodo, la transición a otras partes de la red resulta imposible. Asimismo, es posible encontrarse con la situación en la que se accede a un nodo que induce un bucle interno, impidiendo así regresar a las demás páginas una vez se ha ingresado a una componente conexa específica de la red. Por ejemplo, observando la siguiente matriz:

$$\begin{pmatrix} 0 & 1/3 & 1/2 & 0 \\ 1/2 & 0 & 1/2 & 0 \\ 1/2 & 1/3 & 0 & 0 \\ 0 & 1/3 & 0 & 0 \end{pmatrix}$$

Podemos observar que el nodo 4, cuyas aristas salientes vienen dadas por la columna 4, no tiene ningún enlace saliente, es decir, una vez llegado a este nodo no es posible volver al nodo 1, 2 o 3. Esta ausencia de conexiones salientes imposibilita la aplicación de la hipótesis del surfista aleatorio, que normalmente facilitaría la convergencia hacia el vector estacionario.

En este punto, surge la pregunta fundamental: ¿Cuál fue la acción de Google?, o mejor dicho, ¿qué estrategia idearon Brin y Page para conseguir una matriz irreducible?. Básicamente, la idea y último ajuste para su matriz fue la siguiente:

Se fijó un valor $q \in (0, 1)$, que mediría la probabilidad de que un navegante aleatorio **no cambie radicalmente su curso de navegación** y se mueva a

otra página web que esté conectada con la anterior, en lugar de saltar a otra que no tenga nada que ver con la página web en la que se encuentra. En el caso de Google, este valor es típicamente $q = 0,85$ y se denomina **factor de salto**.

En otras palabras, si en el instante $t = 0$ elegimos un nodo al azar j , en el instante $t + 1$ nos movemos, con **probabilidad q** y de forma equiprobable, a uno de los **vecinos** de j , mientras que con **probabilidad $1 - q$** saltamos a otra **página aleatoria** que puede no tener relación alguna con j (en este caso, la página destino también la elegimos de forma equiprobable). Es decir, para cada $1 \leq i \leq n$, la probabilidad de que pasemos de la página j a la página i es:

$$p_{ji} = q \frac{m_{ij}}{N_j} + \frac{1 - q}{n}$$

Añadiendo probabilidades de transición a todos los vértices, consideramos entonces la matriz:

$$M'' = qM' + (1 - q) \begin{pmatrix} v_1 \\ \vdots \\ v_n \end{pmatrix} (1, \dots, 1)$$

donde (v_1, \dots, v_n) es un vector de probabilidades, es decir, $v_i \geq 0$ para $i = 1, \dots, n$ y $\|v\|_1 = 1$, y que en este caso cumple que $v_i = \frac{1}{n}$ para $i = 1, \dots, n$. Este vector representa una distribución de probabilidad equiprobable y es el utilizado para obtener el *PageRank Clásico*. No obstante, podría utilizarse un vector de probabilidad que represente una distribución de probabilidad distinta, por lo que se le llama **vector de personalización**. En conclusión, estamos sumando una matriz asociada a la probabilidad q a otra asociada a la probabilidad $1 - q$.

Gracias a esta última modificación en el modelado de la matriz que representará la web, el modelo PageRank consigue que **todas las páginas** estén **conectadas**, de manera que aunque un nodo no tenga ningún enlace saliente, el surfista aleatorio tiene una probabilidad $1 - q$ de saltar desde ese nodo a otro cualquiera. En otras palabras, se consigue un efecto estadístico que consiste en introducir saltos aleatorios que no dependen de las propiedades de enlace de la página, creando aristas desde una página hacia todas las demás representando al menos la pequeña probabilidad que hay de saltar hacia ellas.

A efectos matriciales, se fuerza la irreducibilidad de M'' al conectar todos los nodos y se obtiene una matriz resultante positiva, pues todos sus elementos son no nulos porque tienen una probabilidad de $(1 - q)$ de realizar el salto aleatorio. Esto implica a su vez que la matriz es primitiva, lo cual asegura la convergencia del vector estacionario, como vimos en la anterior sección. En general, hemos obtenido una **matriz estocástica, positiva, irreducible y primitiva**.

4.4.2. Teorema de Perron-Frobenius y PageRank: Una Conexión Fundamental

Ahora bien, estas propiedades de la matriz construida son muy importantes, pues ha llegado el momento **crucial** de este trabajo, ya que todavía no hemos hablado sobre la **relación entre el Teorema de Perron-Frobenius y PageRank**. Hemos demostrado en la sección anterior que el vector PageRank u ordenación de páginas de una red estará determinado por el vector estacionario de una cadena de Márkov. De hecho, este vector estacionario v_{est} es un vector propio asociado al autovalor $\lambda = 1$ ya que $M' \cdot v_{est} = v_{est}$. Con M'' sucederá lo mismo ya que sigue siendo una matriz de transición, y de hecho, al ser estocástica se conoce también que dicho autovalor $\lambda = 1$ será el radio espectral, es decir,

$$v_{est} = M'' \cdot v_{est}$$

Sin embargo, ¿y si este vector no existe?, ¿como sabemos que el vector es único? Hasta el momento, nada nos garantizaba la existencia y unicidad del autovector de M'' que buscamos, es decir, aunque asegurábamos que la cadena de Márkov tendría una convergencia, nada garantizaba que dicha convergencia fuese única o estable. Aquí entra en escena el **Teorema de Perron-Frobenius**, pues al ser M'' una matriz no negativa e irreducible (de hecho, es positiva y primitiva) el teorema nos **garantiza la existencia y unicidad** de un autovector positivo $c > 0$ y con $\|c\|_1 = 1$ asociado al radio espectral. Como en este caso la matriz es estocástica entonces el radio espectral es 1:

$$1 \cdot c = M'' \cdot c$$

En otras palabras, justamente $c = v_{est}$. Esto significa que el **vector de Perron** es justamente el vector estacionario de la cadena de Márkov, y por tanto es el **vector PageRank** que nos devolverá la ordenación de páginas web. En conclusión, el teorema de Perron-Frobenius aporta la **robustez y fiabilidad** fundamental para asegurar que el modelo proporcione resultados coherentes.

4.4.3. Definición Formal de PageRank

Tras esta explicación, se va a proporcionar finalmente una definición formal de PageRank. Previamente, recapitularemos de manera concisa: En resumen, el PageRank de una página web (utilizado por Google en sus motores de búsqueda) se establece a través de una cadena de Márkov. La posición que ocupará una página en los resultados del buscador estará determinada por su peso en la distribución estacionaria de esta cadena.

En este contexto, nuestro objetivo ha sido siempre obtener el **vector PageRank**, que representa las ponderaciones asignadas a cada página web en función de su importancia relativa. Este vector es esencial para clasificar las páginas en los resultados de búsqueda, y su cálculo implica considerar las conexiones entre las páginas y la probabilidad de navegación entre ellas. Durante las secciones anteriores, se ha estado estudiando cómo se modela esta interconexión de manera efectiva en el contexto del algoritmo PageRank.

También es importante destacar que el algoritmo PageRank fue el algoritmo más influyente de clasificación de páginas web, pero ha evolucionado y se han desarrollado otros algoritmos complejos con los que Google mejora la calidad de los resultados de búsqueda cada vez más. Sin embargo, la idea básica detrás del PageRank, es decir, la evaluación de la importancia de una página en función de la calidad de los enlaces que recibe, sigue siendo un concepto fundamental en la búsqueda en Internet.

Definición 4.4.3.1. (PageRank) Si $G = (V, E)$ es una red dirigida de n nodos, $q \in (0, 1)$ y $v = (v_1, \dots, v_n) \in \mathbb{R}^n$ tal que $v \geq 0$ y $\|v\|_1 = 1$, entonces se llama **vector PageRank** de G con **factor de salto** q y **vector de personalización** v al único $PR \in \mathbb{R}^n$ tal que

1. $PR \geq 0$ y $\|PR\|_1 = 1$, es decir, PR es un vector de probabilidades.
2. PR es un autovector propio (asociado al autovalor 1) de la matriz $M'' = (M''_{ij})$ dada por

$$M''_{ij} = q \frac{m_{ij}}{N_j} + (1 - q)v_i.$$

Observación. La manera de ordenar los nodos de una red $G = (V, E)$ de acuerdo con el algoritmo PageRank será la siguiente: si tenemos dos nodos $i, j \in V$, si $PR_i > PR_j$ entonces la página P_i aparece por delante de la página P_j en la ordenación de las páginas.

Observación. Si en la definición anterior tomamos $v = \frac{1}{n}(1, \dots, 1)$, obtenemos el llamado PageRank clásico (como hemos explicado anteriormente).

Observación. En general, debemos considerar paseos aleatorios con factor de salto positivo (es decir, tomar $q < 1$), ya que en caso contrario la matriz M'' solo sería no negativa. Aunque los valores próximos a 1 ofrecen comportamientos más realistas es importante preservar la irreducibilidad y primitividad de la matriz M'' . Para terminar, veremos en una última sección cómo la primitividad tiene otra especial utilidad para Google, pues aparece un método que facilitará en gran medida el cálculo del PageRank, conocido como el método de las potencias.

4.5. Método de las potencias

Hemos observado que el Teorema de Perron-Frobenius asegura la existencia y unicidad del vector PageRank. No obstante, nos enfrentamos nuevamente a un problema de escala, pues hemos tratado con un ejemplo de una red pequeña, lo cual no refleja la realidad de la World Wide Web. En este punto, la cuestión pendiente es determinar **cómo** podemos **calcular de manera eficiente el vector PageRank**.

Es en este instante cuando se usa el **método de las potencias**, que ya usado en el ejemplo de la Sección 4.3.2 (4.3), se generalizará ahora de manera breve y sencilla. El método de las potencias tiene una **alta eficiencia computacional**, especialmente cuando se trabaja con una matriz gigantesca como la de Google. La eficiencia del método de las potencias se debe a que permite obviar la necesidad de calcular todos los autovalores y autovectores de la matriz, un proceso que resultaría considerablemente más ineficiente. En su lugar, este método se basa en la multiplicación recursiva de la matriz por un vector, lo que resulta ser un procedimiento mucho más rápido y computacionalmente menos demandante.

Para aplicar el método de las potencias, se requiere que la matriz A cuadrada tenga n autovalores $\lambda_1, \lambda_2, \dots, \lambda_n$ con un conjunto asociado de autovectores linealmente independientes v_1, v_2, \dots, v_n . Entre ellos, es necesario que A tenga exactamente un **autovalor λ_1 cuya magnitud es la mayor**, es decir,

$$|\lambda_1| > |\lambda_2| \geq |\lambda_3| \dots \geq |\lambda_n| \geq 0.$$

Además, **el método asegura la convergencia si la matriz A es primitiva**. De hecho, asegura que converge a un único vector sin depender del vector inicial.

Es evidente que, en el caso analizado, la matriz M'' es primitiva, irreducible y no negativa. De acuerdo con el teorema de Perron-Frobenius, estas características aseguran la existencia del autovalor dominante, que corresponde al **radio espectral** de la matriz. Además, se confirma la **primitividad** de la matriz, por lo que el método convergerá hacia el vector PageRank.

Para explicarlo por pasos, supongamos entonces que tenemos una matriz A en las condiciones que aseguran la existencia del autovalor λ_1 , positivo y estrictamente mayor que el resto de autovalores. Llamamos entonces v_1 al autovector asociado a λ_1 . Los autovectores restantes $\{v_2, \dots, v_n\}$ corresponden a cada autovalor $\lambda_2, \dots, \lambda_n$ en orden descendiente de tamaños tal y como se propuso. Dichos autovectores forman una base de \mathbb{R}^n . Partiendo entonces de un vector de estado inicial v_0 , se tiene que:

$$v_0 = c_1 v_1 + c_2 v_2 + \dots + c_n v_n$$

donde c_1, \dots, c_n son las coordenadas de v_0 en la base considerada. Es necesario

mencionar que dichos autovalores no necesitan ser calculados, nos basta con saber que tales números existen. Ahora, multiplicamos el vector inicial v_0 por la matriz A , obteniendo:

$$Av_0 = c_1\lambda_1v_1 + c_2\lambda_2v_2 + \cdots + c_n\lambda_nv_n$$

puesto que v_1, \dots, v_n son autovectores de A . Como acabamos de comentar, sólo realizaremos el producto de la izquierda. Repitiendo esta operación k veces se llega a:

$$\begin{aligned} A^2v_0 &= c_1\lambda_1^2v_1 + c_2\lambda_2^2v_2 + \cdots + c_n\lambda_n^2v_n \\ &\vdots \\ A^kv_0 &= c_1\lambda_1^kv_1 + c_2\lambda_2^kv_2 + \cdots + c_n\lambda_n^kv_n \end{aligned}$$

Como $\lambda_1 \neq 0$, entonces

$$\frac{1}{\lambda_1^k}A^kv_0 = c_1v_1 + c_2\left(\frac{\lambda_2}{\lambda_1}\right)^k v_2 + \cdots + c_n\left(\frac{\lambda_n}{\lambda_1}\right)^k v_n$$

Sin embargo, como $|\lambda_i/\lambda_1| < 1$ para cada $i = 2, \dots, n$ porque λ_1 es el autovalor dominante, entonces cuando k tienda a infinito

$$\frac{1}{\lambda_1^k}A^kv_0 \xrightarrow{k \rightarrow \infty} c_1v_1$$

En resumen, al **multiplicar reiteradamente el vector inicial por la matriz A** , se va determinando cada vez con más precisión la dirección que nos interesa, la que determina el valor v_1 .

En general, la convergencia del método de las potencias será más rápida o lenta dependiendo del tamaño de los autovalores, específicamente del cociente o razón entre autovalores. La elección del vector inicial también influirá en la convergencia del método, aunque siempre devolverá un resultado que será el mismo independiente del vector inicial escogido. También, se recomienda **normalizar** el vector de cada iteración para **prevenir** el **desbordamiento** (tanto como para números demasiado grandes como para números demasiado pequeños).

Además, la convergencia a v_1 se da independientemente del valor específico de λ_1 , que en nuestro ejemplo con el algoritmo PageRank es $\lambda_1 = 1$. Llegado a este punto, hemos visto cómo la **existencia y unicidad** de v_1 , que es el autovector de la matriz M'' asociado al autovalor dominante $\rho(M'') = 1$, queda asegurada por el Teorema de Perron-Frobenius y cómo se realiza de manera eficiente su **cálculo** u obtención a través del método de las potencias.

5

Conclusiones y trabajos futuros

Al finalizar esta extensa exposición sobre el Teorema de Perron-Frobenius y su integración en el algoritmo PageRank, me gustaría **concluir** destacando la **trascendencia** del Teorema de Perron-Frobenius en el marco de este estudio, que radica en su capacidad para establecer la **existencia y unicidad** del vector propio dominante, proporcionando así una base matemática sólida para el algoritmo PageRank.

De esta manera, el teorema garantiza la estabilidad del algoritmo y fundamenta la fiabilidad del proceso de clasificación de páginas web. La aplicación de Perron-Frobenius no solo valida matemáticamente el enfoque de PageRank, sino que también resalta su utilidad en la resolución de problemas complejos relacionados con la relevancia y jerarquía de la información en la web.

No obstante, como era previsible, el teorema presenta diversas aplicaciones adicionales que justifican un estudio más detenido y podrían ser exploradas como direcciones prometedoras para **investigaciones futuras**. A continuación, me gustaría comentar algunas de las más interesantes:

- Ya sabemos que el teorema puede aplicarse en el ámbito de **Redes Complejas** como la World Wide Web. Sin embargo, podría realizarse un estudio de cualquier otra red compleja, como por ejemplo una *red social*. En este contexto, el teorema podría emplearse para analizar la importancia relativa de distintos usuarios, revelando nodos centrales y patrones de interconexión que influyen en la dinámica general de la red social. Este enfoque proporcionaría una perspectiva matemática valiosa para comprender la estructura y

la influencia de individuos dentro de la red, contribuyendo así al desarrollo de estrategias más efectivas en campos como el marketing, la difusión de información y la detección de tendencias.

- En **Economía**, el teorema se utiliza en modelos de entrada-salida y en modelos de equilibrio general en la teoría económica. El teorema podría utilizarse para evaluar la importancia relativa de diferentes sectores, empresas o variables económicas, proporcionando una base matemática para la *asignación de recursos* y la *toma de decisiones*.
- En **Biología**, el teorema se aplica en modelos matemáticos de poblaciones, especialmente en el análisis de matrices conocidas como *matrices de Leslie*, que describen la dinámica de las poblaciones en biología. El teorema de Perron-Frobenius es esencial aquí para estudiar la estabilidad y las tendencias a largo plazo en el crecimiento de dichas poblaciones.
- En la **Teoría de Juegos**, el teorema de Perron-Frobenius se aplica al análisis de matrices estocásticas y garantiza la existencia de un equilibrio de probabilidad estacionario. Al emplear este teorema en el contexto de juegos estocásticos, se puede determinar la distribución de *estrategias óptimas*, identificando la importancia relativa de cada estrategia, lo que proporciona una perspectiva valiosa para entender la dinámica y la estabilidad en interacciones estratégicas.
- Como hemos visto previamente en este trabajo en la Sección 3.3, también es posible emplear el teorema en el contexto de las **Clasificaciones Deportivas**. Hemos evidenciado que el teorema de Perron-Frobenius proporciona una perspectiva única en competiciones deportivas, como la NBA, donde se puede evaluar la importancia relativa de cada victoria para obtener una clasificación distinta de manera más precisa que con enfoques convencionales. De hecho, este tema capturó mi interés en gran medida y se acabó convirtiendo en el enfoque principal de mi trabajo de fin de grado en Ingeniería del Software [12].

Por otra parte, es relevante subrayar que el algoritmo PageRank continúa evolucionando año tras año, y este estudio se fundamenta en la versión clásica de PageRank. Dada la constante innovación tecnológica, resulta necesario explorar mejoras continuas en este algoritmo para mantener su eficacia en el contexto de avances tecnológicos. Por lo tanto, una dirección prometedora para **investigaciones futuras** podría centrarse en analizar los progresos y otros algoritmos que se aplican para perfeccionar el rendimiento del PageRank.

Algunas de las mejoras más notables que se han aplicado al algoritmo desde que se creó en 1998 han sido, entre muchas otras:

- La inclusión de **Algoritmos de Personalización**. Es decir, se han desarrollado variantes del PageRank que incorporan elementos de personalización, adaptando las clasificaciones según las preferencias y comportamientos específicos del usuario.
- La inclusión de **Algoritmos de Aprendizaje Automático**, pues la red necesita adaptarse continuamente a patrones de comportamiento cambiantes, por lo que se han explorado enfoques de aprendizaje automático para mejorar la capacidad de adaptación de PageRank.
- La **Integración de PageRank con otros algoritmos**, pues aunque PageRank es un indicador valioso de la importancia de una página, el algoritmo puede beneficiarse de la complementariedad con otros métodos. Se han usado desde *algoritmos de relevancia de contenido* como TF-IDF, *algoritmos de detección de comunidades*, *filtros de seguridad*, y hasta *algoritmos de análisis de sentimientos*.

En conclusión, tanto la aplicabilidad del teorema de Perron-Frobenius como las evoluciones del algoritmo PageRank ofrecen buenas opciones para futuras investigaciones. Mi experiencia de trabajo en el estudio del teorema de Perron-Frobenius y su utilidad en el algoritmo PageRank ha sido enriquecedora y reveladora. Este aprendizaje no sólo ha fortalecido mis habilidades analíticas y matemáticas, sino que también ha consolidado mi comprensión de cómo los principios teóricos pueden traducirse en aplicaciones prácticas con impacto significativo en diversos ámbitos.

Bibliografía

- [1] R. Criado and A. Gallinari, *Fundamentos Matemáticos I: Álgebra y ecuaciones diferenciales con coeficientes constantes*. Ed. Centro de Estudios Ramón Areces, 1998.
- [2] R. Criado, M. Romance, and L. E. Solá, *Teoría de Perron-Frobenius: Importancia, poder y centralidad*. La Gaceta de la RSME, 2014.
- [3] A. Berman and R. Plemmons, *Nonnegative matrices in mathematical sciences*. SIAM, Filadelfia, Pensilvania, Estados Unidos, 1994.
- [4] R. B. Bapat and T. E. S. Raghavan, *Nonnegative matrices and applications*. Cambridge University Press, 1997.
- [5] R. A. Horn and C. R. Johnson, *Matrix analysis*. Cambridge University Press, 1985.
- [6] G. P. Barker and H. Schneider, *Algebraic Perron-Frobenius Theory*. Linear Algebra and its Applications, 1975.
- [7] A. Borobia and U. R. Trías, *A geometric proof of the Perron-Frobenius theorem*. Rev. Mat. Univ. Complut. Madrid, 1992.
- [8] P. F. Gallardo, *El secreto de Google y el álgebra lineal*. Universidad Autónoma de Madrid, 2006.
- [9] F. Pedroche, *Métodos de cálculo del vector PageRank*. Bol. Soc. Esp. Mat. Apl., 2013.
- [10] J. S. Cruz, *Uso de las Cadenas de Markov en el Estudio del Algoritmo PageRank*, 2014.
- [11] F. Mahfuz, *Markov Chains And Their Applications*. University of Texas at Tyler, 2021.
- [12] G. G. Santi, *Desarrollo de una Aplicación Web: Una perspectiva innovadora sobre la Clasificación en la NBA mediante la utilización del Teorema de Perron-Frobenius*, 2024.

