

Dióscora Andrea Esteve González

Teoría Espectral de Grafos

Spectral Graph Theory

Trabajo Fin de Grado
Grado en Matemáticas
La Laguna, Marzo de 2023

DIRIGIDO POR
Ignacio García Marco
Irene Márquez Corbella

Ignacio García Marco
Matemáticas, Estadística e
Investigación Operativa
Universidad de La Laguna
38200 La Laguna, Tenerife

Irene Márquez Corbella
Matemáticas, Estadística e
Investigación Operativa
Universidad de La Laguna
38200 La Laguna, Tenerife

Resumen · Abstract

Resumen

En Teoría Espectral de Grafos se pretende estudiar las propiedades de un grafo mediante los autovalores de una matriz asociada al mismo. En este trabajo estudiaremos tres problemas desde esta perspectiva. El primero es la clasificación de los nodos de un grafo dirigido en función de su importancia. Para ello, asociamos al grafo una matriz estocástica y veremos cómo esta ordenación por importancias se corresponde con el autovector asociado al autovalor $\lambda = 1$ de la matriz. Esta es la idea sobre la que reposa el algoritmo PageRank de Google para ordenar las páginas web por importancia. En el segundo y tercer problema estudiaremos ciertos parámetros de grafos, esta vez no dirigidos, que, en general, son difíciles de calcular: el número de independencia y el número cromático. Veremos cómo los autovalores de la matriz de adyacencia del grafo nos aportan cotas superiores e inferiores para los parámetros citados anteriormente.

Palabras clave: *Teoría Espectral de Grafos – matriz estocástica – Teoría de Perron-Frobenius – método de las potencias – algoritmo PageRank – matriz de adyacencia – autovalores – número de independencia – número cromático.*

Abstract

Spectral Graph Theory aims to understand the properties of graphs by analyzing the eigenvalues of a matrix associated with them. In this document, we investigate three problems using this approach. The first one is the classification of nodes in a graph based on their importance. To achieve this objective, we associate a stochastic matrix to the graph and demonstrate that the eigenvector corresponding to the eigenvalue $\lambda = 1$ of the matrix represents the importance vector. This idea is used in the Google algorithm PageRank to rank websites based on their importance. In the second and third problems, we tackle the challenging task of computing the independence number and the chromatic number of undirected graphs. For this purpose, we analyze the adjacency matrix of the graph and prove that its eigenvalues provide upper and lower bounds for these parameters.

Keywords: *Spectral Graph Theory – stochastic matrix – Perron-Frobenius theory – iteration – PageRank algorithm – adjacency matrix – eigenvalues – independence number – chromatic number.*

Contenido

Resumen/Abstract	III
Introducción	VII
1. Clasificación de los nodos de una red por importancia: el algoritmo <i>PageRank</i>	1
1.1. Introducción al <i>PageRank</i>	1
1.2. Forma canónica de Jordan	4
1.3. Teorema de Perron-Frobenius para matrices estocásticas positivas	8
1.4. Adaptación del método de las potencias al algoritmo <i>PageRank</i> ..	13
2. Cotas para el número de independencia de un grafo	17
2.1. Teorema espectral y Teorema de entrelazamiento de matrices simétricas reales	17
2.2. Matriz de adyacencia de un grafo no dirigido	22
2.3. Cotas para el número de independencia de un grafo	25
3. Cotas para el número cromático de un grafo	33
3.1. Número cromático de un grafo	33
3.2. Cotas para el número cromático en función del número de independencia de un grafo	34
3.3. Caracterización espectral de grafos bipartitos	35
3.4. Una cota superior para el número cromático	38
Bibliografía	45
Poster	47

Introducción

La Teoría de Grafos se originó en el siglo XVIII con el famoso problema de los puentes de Königsberg, que consistía en encontrar un camino que recorriese los siete puentes del río Pregel, pasando una única vez por cada uno de ellos. El trabajo de Leonhard Euler dio solución a este problema: tal y como se sospechaba, era imposible de resolver. Este problema es considerado el origen de esta disciplina matemática. En la actualidad, podemos ver cómo los grafos ayudan a visualizar y modelizar fácilmente procesos que pueden darse en gran variedad de situaciones, como en el análisis de redes de transporte, telecomunicaciones, logística, planificación de proyectos, por lo que es un tema de gran interés y que, además, está en auge.

En este documento nos centraremos en la Teoría Espectral de Grafos, área que estudia las propiedades de un grafo mediante los autovalores y autovectores de ciertas matrices asociadas al grafo. Atacaremos tres problemas principales: la clasificación de los nodos de un grafo dirigido por importancia, cuyo estudio abordará todo el primer capítulo, y la estimación de parámetros importantes en un grafo no dirigido, como son el número de independencia y el número cromático del mismo, que abarcará los capítulos 2 y 3.

Sergei Brin y Larry Page dan solución al primer problema en 1998 con el algoritmo *PageRank* [6]. Este algoritmo está fundamentado por el Teorema de Perrón-Frobenius, que garantiza que, bajo ciertas hipótesis, hay una respuesta única al problema de ordenación. Además, el denominado método iterativo de las potencias proporciona una forma de aproximar rápidamente este vector de importancias. Este algoritmo es, quizás, el resultado más vistoso de la Teoría Espectral de Grafos, ya que se implementó en la primera versión del motor de búsqueda de *Google* para presentar las páginas web por orden de trascendencia al usuario que hace uso del buscador.

Para los problemas restantes haremos un estudio de otra matriz asociada a un grafo, esta vez no dirigido: la matriz de adyacencia. Más concretamente, buscaremos dar cotas para el número de independencia y el número cromático de un grafo mediante los autovalores de la misma. Este problema es de gran

interés debido a que, en general, determinar el número cromático de un grafo G es un problema difícil desde el punto de vista computacional, tanto es así que en Teoría de la complejidad han sido clasificados como problemas \mathcal{NP} -hard [5], y, por tanto, son problemas que no se espera que puedan ser resueltos en tiempo polinomial. Por supuesto que se disponen de algoritmos que calculan estos parámetros, pero no son eficientes. Es por ello que es de gran ayuda tener cotas y/o estimaciones.

La estructura de este trabajo es la siguiente. En el capítulo 1 se estudia el algoritmo *PageRank*, que nos permite identificar nodos importantes en una red, describiendo esta red como un grafo dirigido al que le asociamos una matriz estocástica. La idea principal de este algoritmo es que, bajo ciertas hipótesis, el generador del autoespacio asociado al autovalor máximo asociado a la matriz, es el vector que nos permite ordenar los nodos de la red por importancia. En la primera sección, basada en el artículo [6], utilizamos un ejemplo sencillo para introducir el algoritmo *PageRank*, y en las siguientes secciones daremos fundamento teórico del algoritmo. En la sección 3, demostraremos una versión del *Teorema de Perron-Frobenius* adaptado a matrices estocásticas positivas, resultado que data de principios del siglo XX. En la sección 4 presentamos el método de las potencias, un método iterativo eficiente que nos permite dar una aproximación rápida del vector de importancias. Acabamos el capítulo estudiando cómo se aplican todos los resultados anteriores para dar lugar al algoritmo *PageRank*.

En el Capítulo 2 daremos estimaciones para el número de independencia, que denotaremos por $\alpha(G)$, de un grafo no dirigido G , siendo este el número máximo de vértices no adyacentes dos a dos. Este capítulo está basado en el artículo [2]. En la primera sección demostramos el Teorema Espectral y el Teorema de Entrelazamiento de Cauchy para matrices simétricas reales, que son resultados de Álgebra Lineal que se remontan al siglo XIX. En la segunda sección definimos la matriz de adyacencia de un grafo no dirigido, y establecemos las primeras relaciones entre los autovalores de esta matriz y algunos parámetros del grafo como los grados máximo, mínimo y medio. En la tercera y última sección se incluyen los resultados principales de este capítulo: cotas para el número de independencia en función del espectro de la matriz de adyacencia. Aportamos tres estimaciones para este número: una de las estimaciones depende del número de autovalores positivos y negativos de la matriz, y las otras de los autovalores máximo y mínimo. Para este capítulo nos hemos basado en el artículo [2].

En el Capítulo 3 damos cotas para el número cromático $\chi(G)$ de un grafo no dirigido. Un problema fundamental en la Teoría de Grafos es determinar este valor, es decir, el número mínimo de colores que tenemos que utilizar para colorear un grafo de tal forma que no haya vértices adyacentes con el mismo color. En la primera sección introducimos resultados básicos sobre la coloración de grafos, demostrando que colorear un grafo es equivalente a dar una partición

en subconjuntos de vértices independientes. De aquí deducimos que $\chi(G)$ está acotado superiormente por el cociente entre el número de vértices y $\alpha(G)$ y, en consecuencia, las cotas inferiores para $\alpha(G)$ se traducen en cotas superiores para $\chi(G)$. Posteriormente, daremos una caracterización espectral de grafos bipartitos, es decir, los grafos 2-coloreables o coloreables con 2 colores. El resultado principal de la última sección es el Teorema de Wilf, que afirma que todo grafo se puede colorear con $\lambda_1 + 1$ colores, siendo λ_1 el mayor autovalor de la matriz de adyacencia del grafo G , es decir, $\chi(G) \leq \lambda_1 + 1$. Usando este resultado, se aporta una cota superior para $\chi(G)$ que depende exclusivamente del número de aristas y de vértices de G , dos parámetros que podemos conocer solo observando el grafo. Para esta sección, nos hemos basado en el artículo [9].

Finalizamos esta memoria con una serie de conclusiones y posibles líneas de trabajo futuro.

Clasificación de los nodos de una red por importancia: el algoritmo *PageRank*

Dado un grafo dirigido, es natural preguntarse cuáles son los vértices más importantes o relevantes del mismo. La respuesta a esta pregunta dependerá de la definición de importancia que tomemos. Si consideramos como definición de importancia la probabilidad de visitar un vértice en un recorrido aleatorio por el grafo, Brin y Page dieron una solución muy elegante a este problema (ver [6]). Más concretamente, proponen asociar a todo grafo una matriz cuadrada de forma que, bajo ciertas hipótesis, la importancia de cada vértice se pueda obtener a partir de cualquier autovector asociado al autovalor con mayor módulo de dicha matriz.

Esta es la idea en la que se sustenta el algoritmo *PageRank*, que fue incluido en la primera versión del motor de búsqueda *Google*. El objetivo de este capítulo es el de presentar las ideas principales de este algoritmo. Comenzamos con una primera sección donde incluimos algunos ejemplos ilustrativos. En la segunda y tercera sección incluimos las herramientas de Álgebra Lineal necesarias para justificar la validez del mismo; más concretamente nos introducimos en la Teoría de Perron-Frobenius. En la última sección discutiremos cómo aplicar esos resultados a *PageRank*.

La presentación de este capítulo está inspirada en [1].

1.1. Introducción al *PageRank*

PageRank es un algoritmo que fue desarrollado por los fundadores de *Google*, Sergei Brin y Larry Page, en el año 1998 en la Universidad de Standford. Este tiene como objetivo decidir en qué orden mostrar en el navegador los sitios web cuando realizamos una búsqueda. Para conseguir esta ordenación, se asigna un valor numérico positivo a cada página que representa su importancia para luego mostrar las páginas en orden decreciente de ese valor, es decir, de más importante a menos. Esta idea es extrapolable a cualquier grafo dirigido, y ordena los nodos (o vértices) en función de su relevancia. La relevancia de un nodo en una red es algo subjetivo, y que depende de la definición que se adopte. En esta

memoria presentamos dos definiciones alternativas que nos permiten modelizar de dos formas distintas el problema. Curiosamente, ambos modelos tienen la misma asignación de valores de importancia.

Mediante el Álgebra Lineal: se considera que la importancia de un sitio web está relacionada con desde dónde es enlazada esta página, y se tiene en cuenta no solo el número de enlaces que apuntan hacia ella, sino también la importancia de los sitios que la enlazan. De esta forma, si una página tiene como valor de importancia x_i y enlaza a otras k páginas, le aporta a cada una de esas páginas una importancia de $\frac{x_i}{k}$. Así, planteando una ecuación que describa la importancia para cada sitio web, obtenemos un sistema de ecuaciones de la forma $Ax^T = x^T$, y resolviendo el sistema lineal homogéneo de n ecuaciones y n incógnitas con matriz asociada $A - I$, obtenemos los valores de las importancias de cada página.

Ilustramos esta propuesta con un pequeño ejemplo.

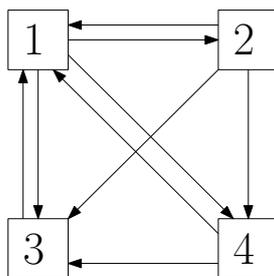


Figura 1.1. Grafo representativo de un hipotético modelo de Internet.

Ejemplo 1.1. El grafo dirigido de la Figura 1.1 representa un mini-modelo de Internet, donde cada nodo representa una página web, y cada arco conectando dos nodos representa un enlace de una página web hacia otra.

Sea x_i con $i \in \{1, 2, 3, 4\}$ el valor de relevancia de la página i -ésima. Utilizando el grafo y siguiendo el proceso descrito anteriormente, queda el siguiente sistema de ecuaciones:

$$\begin{cases} x_1 = \frac{1}{3}x_2 + x_3 + \frac{1}{2}x_4 \\ x_2 = \frac{1}{3}x_1 \\ x_3 = \frac{1}{3}x_1 + \frac{1}{3}x_2 + \frac{1}{2}x_4 \\ x_4 = \frac{1}{3}x_1 + \frac{1}{3}x_2 \end{cases}$$

Así, tendremos el siguiente sistema $Ax^T = x^T$, siendo $x = (x_1, x_2, x_3, x_4)$.

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

Por tanto, tenemos un sistema homogéneo con matriz asociada $A - I$, cuyo rango es 3, que tiene como conjunto de soluciones del sistema a:

$$\begin{cases} x_1 = 9\alpha \\ x_2 = 3\alpha \\ x_3 = 6\alpha \\ x_4 = 4\alpha \end{cases}, \text{ con } \alpha \in \mathbb{R}. \quad (1.1)$$

Podemos observar que las soluciones son proporcionales al vector $(9, 3, 6, 4)$ y, por tanto, la página 1 es la que más importancia tiene, seguida de la 3, de la 4 y finalmente de la 2, en ese orden. Como ya mencionamos, este método sugiere calcular la importancia de los nodos mediante la resolución de un sistema lineal homogéneo de n ecuaciones con n incógnitas. En el ejemplo, el sistema tiene como conjunto de soluciones un subespacio vectorial de dimensión uno que, además, está generado por un vector cuyas entradas son todas estrictamente positivas. Este hecho nos permite ordenar las páginas. En este capítulo presentaremos una versión del Teorema de Perron-Frobenius que asegura que, bajo ciertas hipótesis, esto es un hecho general.

Mediante un proceso estocástico (de Markov): suponemos que los usuarios de Internet navegan de forma aleatoria, es decir, un usuario está en un sitio web y este tiene varios enlaces, con la misma probabilidad de acceder a cualquiera de ellos. Una página será más importante que otra si tiene más probabilidad de ser visitada en este recorrido aleatorio.

Esta forma de modelizar el problema da lugar, de forma natural, a una matriz cuadrada con coeficientes reales positivos $A = (a_{i,j})$, donde $a_{i,j}$ es la probabilidad de pasar a la página i desde la página j . Así, si la página j -ésima enlaza a ciertas páginas i_1, \dots, i_k , entonces la columna j -ésima de la matriz A tendrá el valor $\frac{1}{k}$ en las posiciones i_1, \dots, i_k , y 0 en las posiciones restantes.

Observamos que la matriz A que se obtiene mediante esta aproximación al problema coincide con la obtenida anteriormente.

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

Observación 1.2. Si denotamos $A^2 = (c_{i,j})$, tenemos que $c_{i,j} = \sum_{k=1}^n a_{i,k}a_{k,j}$. Observamos que este valor coincide con la probabilidad de ir de la página j a la i en dos pasos; ya que para pasar de la j a la i en dos pasos, tendremos que pasar por una página k intermedia y, aplicando el teorema de la probabilidad total, se obtiene la citada expresión. Así, $A^m = (d_{i,j})$ tendrá entrada $d_{i,j}$ la probabilidad de ir de j a i en m pasos.

Ahora multiplicando una potencia de A por un vector tomado al azar obtenemos:

$$A^5 \begin{pmatrix} 4 \\ 5 \\ 11 \\ 2 \end{pmatrix} = \begin{pmatrix} 9,1687\dots \\ 2,8035\dots \\ 6,0555\dots \\ 3,9722\dots \end{pmatrix} \approx \begin{pmatrix} 9 \\ 3 \\ 6 \\ 4 \end{pmatrix}$$

Vemos que el resultado se aproxima al vector de importancias obtenido con el primer método. De la misma forma, cogiendo otro vector al azar y multiplicándolo por una potencia de A obtenemos:

$$A^7 \begin{pmatrix} 7 \\ 2 \\ -3 \\ 5 \end{pmatrix} = \begin{pmatrix} 4,5103\dots \\ 1,5104\dots \\ 2,9919\dots \\ 1,9875\dots \end{pmatrix} \approx \begin{pmatrix} 4,5 \\ 1,5 \\ 3 \\ 2 \end{pmatrix}$$

que también se aproxima a una solución del sistema definido en (1.1).

Veamos ahora qué pasa con el vector $x = (-1, 1, 0, 0)$:

$$A^9 \begin{pmatrix} -1 \\ 1 \\ 0 \\ 0 \end{pmatrix} = \begin{pmatrix} 0,0001\dots \\ -0,0001\dots \\ 0 \\ 0 \end{pmatrix}$$

Si tomamos una potencia mayor de A y multiplicamos por x^T tenemos

$$A^{15} \begin{pmatrix} -1 \\ 1 \\ 0 \\ 0 \end{pmatrix} = \begin{pmatrix} 0,0000000069\dots \\ -0,0000000069\dots \\ 0 \\ 0 \end{pmatrix} \approx \begin{pmatrix} 0 \\ 0 \\ 0 \\ 0 \end{pmatrix}$$

En este caso podemos observar que $A^m x^T$ parece aproximarse al vector nulo cuando la potencia m es grande.

Cabe preguntarse si esto es un hecho general, y si elevando A a un m lo suficientemente grande y multiplicando A^m por un vector aleatorio, se obtiene siempre una aproximación de una solución del sistema. Además, sería interesante que este proceso converja a un vector distinto del $(0, 0, 0, 0)$. El estudio de la convergencia de este procedimiento, denominado método de las potencias, se llevará a cabo en la Sección 1.4. El interés principal de este método reside en el hecho de que es mucho más rápido multiplicar una potencia de A por un vector que resolver el sistema lineal de ecuaciones $Ax^T = x^T$.

1.2. Forma canónica de Jordan

Comenzamos esta sección recordando varios resultados relativos a la forma canónica de Jordan de una matriz cuadrada y de un endomorfismo. Para una

exposición más detallada y para completar las demostraciones de los resultados aquí expuestos, recomendamos consultar [3, Cap. IV].

Definición 1.3 (Polinomio característico). Sea $A \in M_{n \times n}(\mathbb{C})$ una matriz de tamaño $n \times n$ con coeficientes complejos. Se define el polinomio característico de la matriz A como $p(\lambda) = \det(A - \lambda I) \in \mathbb{C}[\lambda]$.

El polinomio característico de una matriz cuadrada A de tamaño $n \times n$ es un polinomio de grado n con coeficiente líder $(-1)^n$. Por el Teorema Fundamental del Álgebra tenemos que:

$$p(\lambda) = (-1)^n \prod_{i=1}^s (\lambda - \lambda_i)^{m(\lambda_i)}$$

con $\lambda_1, \dots, \lambda_s \in \mathbb{C}$, $\lambda_i \neq \lambda_j$ con $i \neq j$ autovalores de la matriz A , y donde $m(\lambda_i) \in \mathbb{Z}_{>0}$ representa la multiplicidad algebraica de λ_i . Recordemos que se tiene que $\sum_{i=1}^s m(\lambda_i) = n$.

Definición 1.4. Para cada $\lambda \in \mathbb{C}$ y cada $j \in \mathbb{Z}_{>0}$ se define el j -ésimo autoespacio generalizado asociado a λ como:

$$E^j(\lambda) = \{u \in \mathbb{C}^n \mid (A - \lambda I)^j u^T = \mathbf{0}\} = \ker((A - \lambda I)^j)$$

Se tiene que λ no es un autovalor si, y solo si, $E^j(\lambda) = \{0\}$ para todo $j \in \mathbb{Z}_{>0}$. Sean $\lambda_1, \dots, \lambda_s \in \mathbb{C}$ los autovalores, para todo $i \in \{1, \dots, s\}$, existe un $r_i \in \mathbb{Z}_{>0}$ tal que

$$E^1(\lambda_i) \subsetneq E^2(\lambda_i) \subsetneq \dots \subsetneq E^{r_i}(\lambda_i) = E^{r_i+1}(\lambda_i)$$

y, por tanto, $E^t(\lambda_i) = E^{t+1}(\lambda_i)$, para $t \geq r_i$. Además, sabemos que $1 \leq \dim(E^1(\lambda_i)) \leq \dim(E^{r_i}(\lambda_i)) = m(\lambda_i)$. Al espacio $E^1(\lambda_i)$ se le denomina autoespacio asociado a λ_i y se denota V_{λ_i} , al valor $\dim(E^1(\lambda_i))$ se le denomina multiplicidad geométrica de λ_i , y a $E^{r_i}(\lambda_i)$ el espacio máximo asociado a λ_i , que se denota $M(\lambda_i)$. Se tiene que la dimensión del espacio máximo coincide con la multiplicidad algebraica, es decir, $\dim(M(\lambda_i)) = m(\lambda_i)$.

Teorema 1.5. Sea $A \in M_{n \times n}(\mathbb{C})$ una matriz cuadrada y sean $\lambda_1, \dots, \lambda_s \in \mathbb{C}$ sus autovalores. Entonces $\mathbb{C}^n = M(\lambda_1) \oplus \dots \oplus M(\lambda_s)$. En particular, si $\mathcal{B}_i \subseteq \mathbb{C}^n$ es base del autoespacio máximo $M(\lambda_i)$ con $i \in \{1, \dots, s\}$, entonces $\mathcal{B}' = \mathcal{B}_1 \cup \dots \cup \mathcal{B}_s$ es base de \mathbb{C}^n .

Definición 1.6. Un bloque de Jordan $J(\lambda) = (j_{r,s})$ de tamaño $k \times k$ y autovalor λ es una matriz cuadrada tal que

$$j_{r,s} = \begin{cases} \lambda & \text{si } r = s, \\ 1 & \text{si } s = r + 1, \\ 0 & \text{en otro caso.} \end{cases}$$

de modo que

$$J(\lambda) = \begin{pmatrix} \lambda & 1 & 0 & \cdots & 0 & 0 \\ 0 & \lambda & 1 & \cdots & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & \lambda & 1 \\ 0 & 0 & 0 & \cdots & 0 & \lambda \end{pmatrix}$$

Una matriz de Jordan es una matriz diagonal por bloques, donde cada bloque que compone la diagonal es un bloque de Jordan.

Teorema 1.7 (Forma canónica de Jordan). *Toda matriz cuadrada compleja es semejante a una matriz de Jordan. Equivalentemente, para todo endomorfismo f de \mathbb{C}^n existe una base \mathcal{B} de \mathbb{C}^n tal que la matriz asociada a f respecto de \mathcal{B} , que denotamos por $M(f)_{\mathcal{B},\mathcal{B}}$, es una matriz de Jordan.*

Además, si $A \in M_{n \times n}(\mathbb{C})$ tiene como autovalores a $\lambda_1, \dots, \lambda_s \in \mathbb{C}$, entonces en la diagonal de la matriz de Jordan J que es semejante a A aparece $m(\lambda_i)$ veces el autovalor λ_i , distribuido en $\dim(E^1(\lambda_i))$ bloques de Jordan asociados a λ_i .

También, si $f : \mathbb{C}^n \rightarrow \mathbb{C}^n$ es un endomorfismo lineal, tomamos \mathcal{B}' una base de \mathbb{C}^n , $A = M(f)_{\mathcal{B}',\mathcal{B}'}$ la matriz asociada a f respecto de \mathcal{B}' y tenemos que $J = M(f)_{\mathcal{B},\mathcal{B}}$ es de Jordan, entonces

$$J = P^{-1}AP,$$

siendo P la matriz (invertible) de cambio de base de \mathcal{B} a \mathcal{B}' .

Para finalizar esta sección vamos a ver cómo se comportan las potencias de una matriz de Jordan. Como ya sabemos, si una matriz A es diagonal por bloques con bloques A_1, \dots, A_r , entonces A^m también es diagonal por bloques con bloques A_1^m, \dots, A_r^m . Por tanto, para describir la potencia de una matriz de Jordan bastará describir las potencias de sus bloques de Jordan. Esto último lo trataremos en el Lema 1.11.

Definición 1.8. Sean $n \in \mathbb{Z}_{>0}$, $\ell \in \mathbb{N}$ definimos: $N_\ell \in M_{n \times n}(\mathbb{R})$ como la matriz

$$N_\ell = (n_{i,j})_{1 \leq i,j \leq n} \text{ con } n_{i,j} = \begin{cases} 1 & \text{si } j = i + \ell, \\ 0 & \text{en otro caso.} \end{cases}$$

Observación 1.9. Si $\ell = 0$, entonces N_ℓ es la matriz identidad. Si $\ell \geq n$, entonces N_ℓ es la matriz nula.

Lema 1.10. Sea $\ell \in \mathbb{Z}_{>0}$, entonces $(N_1)^\ell = N_\ell$.

Demostración. Realizamos la demostración por inducción sobre ℓ . Para $\ell = 1$ el resultado es trivial. Supongamos que el resultado es cierto para ℓ , y veámoslo para $\ell + 1$ como sigue: sean

$$N_\ell = (a_{i,k}) \quad N_1 = (b_{k,j})$$

Usando la hipótesis de inducción tenemos que: $N_1^{\ell+1} = N_1^\ell N_1 = N_\ell N_1$. Entonces $N_\ell N_1 = (c_{i,j})$, donde $c_{i,j} = \sum_{k=1}^n a_{i,k} b_{k,j}$. Como $a_{i,k} \neq 0$ si, y solo si, $k = i + 1$ y $b_{k,j} \neq 0$ si, y solo si, $j = k + \ell$. Entonces, si $j \neq i + \ell + 1$ se tiene que $c_{i,j} = 0$. Si $j = i + \ell + 1$, entonces

$$c_{i,j} = c_{i,i+\ell+1} = \sum_{k=1}^n a_{i,k} b_{k,i+\ell+1} = \sum_{k \neq i+1} a_{i,k} b_{k,i+\ell+1} + a_{i,i+1} b_{i+1,i+\ell+1} = 1$$

pues $a_{i,i+1} = 1$, y $b_{i+1,i+\ell+1} = 1$. □

Lema 1.11. *Sea $J(\lambda)$ un bloque de Jordan de tamaño $k \times k$. Entonces*

$$J(\lambda)^m = \begin{pmatrix} \lambda^m \binom{m}{1} \lambda^{m-1} \binom{m}{2} \lambda^{m-2} \dots \binom{m}{m-1} \lambda \\ 0 \lambda^m \binom{m}{1} \lambda^{m-1} \dots \binom{m}{m-1} \lambda^2 \\ \vdots \quad \vdots \quad \quad \quad \ddots \quad \quad \quad \vdots \\ 0 \quad 0 \quad \quad 0 \quad \dots \binom{m}{1} \lambda^{m-1} \\ 0 \quad 0 \quad \quad 0 \quad \dots \lambda^m \end{pmatrix} \quad (1.2)$$

Demostración. Sabemos que $J(\lambda)$ podemos escribirlo de la siguiente manera:

$$J(\lambda) = \lambda I + N_1 \quad (1.3)$$

donde

$$N_1 = \begin{pmatrix} 0 & 1 & 0 & \dots & 0 & 0 \\ 0 & 0 & 1 & \dots & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & \dots & 0 & 1 \\ 0 & 0 & 0 & \dots & 0 & 0 \end{pmatrix}$$

Por tanto, teniendo en cuenta que I y N_1 conmutan, entonces:

$$J(\lambda)^m = (\lambda I + N_1)^m = \lambda^m I^m + \binom{m}{1} \lambda^{m-1} I^{m-1} N_1 + \dots + \binom{m}{m-1} \lambda I N_1^{m-1} + N_1^m$$

Aplicando el Lema 1.10

$$J(\lambda)^m = \sum_{j=0}^m \lambda^{m-j} N_j \binom{m}{j} = \sum_{j=0}^{k-1} \lambda^{m-j} N_j \binom{m}{j}$$

ya que, por la Observación 1.9, tenemos que N_j es la matriz nula para todo $j \geq k$. Por todo esto obtenemos el resultado buscado (1.2). □

1.3. Teorema de Perron-Frobenius para matrices estocásticas positivas

Comenzamos esta sección dando la definición de matriz estocástica con el objetivo de introducirnos en la teoría de Perron-Frobenius. Debido a que la demostración del Teorema de Perron-Frobenius para matrices estocásticas positivas es muy larga, la dividiremos en varios Lemas (Lema 1.20 y Lema 1.16), para finalmente completar su prueba (Teorema 1.13) para que su lectura sea más amena.

Definición 1.12. Sea $A = (a_{i,j}) \in M_{n \times n}(\mathbb{R})$ una matriz cuadrada de tamaño $n \times n$. Diremos que es estocástica si:

1. Todas sus entradas son no negativas, i.e., $a_{i,j} \geq 0$, con $i, j \in \{1, \dots, n\}$.
2. La suma de los elementos de cada columna vale 1, i.e., $\sum_{i=1}^n a_{i,j} = 1$ con $j \in \{1, \dots, n\}$.

El Teorema de Perrón-Frobenius para matrices estocásticas positivas es el siguiente:

Teorema 1.13 (Teorema de Perrón-Frobenius para matrices estocásticas positivas). Sea A una matriz estocástica con todas sus entradas positivas. Entonces:

1. $\lambda = 1$ es autovalor de multiplicidad algebraica 1.
2. El autoespacio asociado a $\lambda = 1$ es $V_{\lambda=1} = \langle u \rangle$ con $u = (u_1, \dots, u_n) \in \mathbb{R}^n$, $u_i > 0$ con $i \in \{1, \dots, n\}$.
3. Si γ es otro autovalor de A , entonces $|\gamma| < 1$.

No obstante, antes de demostrar este Teorema, presentamos una serie de definiciones y resultados que nos harán falta para la prueba. Hay múltiples demostraciones de este resultado en la literatura. La demostración aquí recogida está basada en [3, pág. 109] y [4, pág. 278].

Lema 1.14 (Desigualdad triangular). Sean $w_1, \dots, w_n \in \mathbb{C} \setminus \{0\}$. Entonces

$$\left| \sum_{j=1}^n w_j \right| \leq \sum_{j=1}^n |w_j|$$

Además, se tiene igualdad si, y solo si, $\frac{w_i}{w_j} \in \mathbb{R}_{\geq 0}$, para $i, j \in \{1, \dots, n\}$.

Teorema 1.15 (Teorema del punto fijo de Brouwer). Sea $K \subseteq \mathbb{R}^n$ un compacto convexo y $T : K \rightarrow K$ una aplicación continua. Entonces, T tiene un punto fijo, i.e., existe $x \in K$ tal que $T(x) = x$.

El siguiente resultado es consecuencia del teorema anterior.

Lema 1.16. *Sea $A \in M_{n \times n}(\mathbb{R})$ una matriz estocástica. Entonces existe un vector $u = (u_1, \dots, u_n) \in \mathbb{R}_{\geq 0}^n$ tal que $\sum_{j=1}^n u_j = 1$ y*

$$Au^T = u^T.$$

Además, si A tiene todas sus entradas positivas, i.e., $A \in M_{n \times n}(\mathbb{R}_{>0})$, entonces $u_j > 0, \forall j \in \{1, \dots, n\}$.

Demostración. Definimos la siguiente aplicación lineal:

$$T : \mathbb{R}^n \longrightarrow \mathbb{R}^n$$

$$x \longmapsto T(x) = \left(\sum_{j=1}^n a_{1,j}x_j, \dots, \sum_{j=1}^n a_{n,j}x_j \right)$$

T es aplicación continua pues es una aplicación lineal.

Consideramos K como el siguiente subconjunto de \mathbb{R}^n , convexo y compacto:

$$K = \{(x_1, \dots, x_n) \in \mathbb{R}^n : x_i \geq 0, \sum_{i=1}^n x_i = 1\}.$$

Vamos a probar que para $x \in K$ se tiene que $T(x) \in K$. Sea $x = (x_1, \dots, x_n) \in K$, entonces por definición:

$$T(x) = \left(\sum_{j=1}^n a_{1,j}x_j, \dots, \sum_{j=1}^n a_{n,j}x_j \right)$$

Para cada i tenemos que $\sum_{j=1}^n a_{i,j}x_j \geq 0$. Además,

$$\sum_{i=1}^n \sum_{j=1}^n a_{i,j}x_j = \sum_{j=1}^n x_j \sum_{i=1}^n a_{i,j} = \sum_{j=1}^n x_j = 1$$

y, por tanto, $T(x) \in K$. Aplicando el Teorema 1.15, obtenemos la existencia de algún $u = (u_1, \dots, u_n) \in K$ con $T(u) = u$, quedando así probada la primera parte del Lema.

Ahora supongamos que A tiene todas sus entradas positivas y veamos que $u_j > 0$, para $j \in \{1, \dots, n\}$. Como u no-nulo, entonces, para algún $i \in \{1, \dots, n\}$, tenemos que $u_i > 0$. Sea ahora $j \in \{1, \dots, n\}$, como $a_{j,i} > 0$ y $Au^T = u^T$, tenemos que:

$$0 < a_{j,i}u_i \leq a_{j,1}u_1 + \dots + a_{j,n}u_n = u_j.$$

□

Definición 1.17. *Sea $A = (a_{i,j}) \in M_{n \times n}(\mathbb{R})$. Se define la norma matricial infinita de A como el máximo de las sumas de los valores absolutos de los elementos de una fila, es decir, $\|A\|_\infty = \max_{1 \leq i \leq n} \sum_{j=1}^n |a_{i,j}|$. En particular, si $x = (x_1, \dots, x_n) \in \mathbb{R}^n$, entonces $\|x^T\|_\infty = \max_{1 \leq i \leq n} |x_i|$.*

Lema 1.18. Sea $A \in M_{n \times n}(\mathbb{R}_{\geq 0})$. Entonces, $\|A\|_{\infty} = \max_{\substack{x \in \mathbb{R}^n \\ \|x\|_{\infty} = 1}} \|Ax^T\|_{\infty}$.

Demostración. Sea $x \in \mathbb{R}^n$, con $\|x^T\|_{\infty} = 1$, entonces aplicando la desigualdad triangular, tenemos que:

$$\begin{aligned} \|Ax^T\|_{\infty} &= \max_{1 \leq i \leq n} \left| \sum_{j=1}^n a_{i,j} x_j \right| \leq \max_{1 \leq i \leq n} \sum_{j=1}^n |a_{i,j}| |x_j| \leq \max_{1 \leq i \leq n} \sum_{j=1}^n |a_{i,j}| \\ &= \|A\|_{\infty}. \end{aligned}$$

Para la otra desigualdad basta considerar $x = (1, \dots, 1)$. Así

$$\begin{aligned} \|A\|_{\infty} &= \max_{1 \leq i \leq n} \sum_{j=1}^n |a_{i,j}| = \max_{1 \leq i \leq n} \sum_{j=1}^n |a_{i,j}| |x_j| = \\ &= \max_{1 \leq i \leq n} \sum_{j=1}^n |a_{i,j} x_j| = \|Ax^T\|_{\infty} \end{aligned}$$

□

Lema 1.19. Sean $A, B \in M_{n \times n}(\mathbb{R})$, entonces $\|AB\|_{\infty} \leq \|A\|_{\infty} \|B\|_{\infty}$.

Demostración. Sea $x = (x_1, \dots, x_n) \in \mathbb{R}^n$ con $\|x^T\|_{\infty} = \max_{1 \leq i \leq n} |x_i| \leq 1$. Veamos que $0 \leq \|Ax^T\|_{\infty} \leq \|A\|_{\infty} \|x^T\|_{\infty}$. Si $x = (0, \dots, 0)$, entonces $\|Ax^T\|_{\infty} \leq \|A\|_{\infty} \|x^T\|_{\infty}$ de forma trivial.

Supongamos entonces que $x \neq (0, \dots, 0)$ y supongamos, por reducción al absurdo, que

$$\|Ax^T\|_{\infty} > \|A\|_{\infty} \|x^T\|_{\infty}$$

Por tanto, dividiendo entre $\|x^T\|_{\infty} \neq 0$ tenemos que:

$$\frac{1}{\|x^T\|_{\infty}} \|Ax^T\|_{\infty} > \|A\|_{\infty}$$

obteniendo así:

$$\left\| A \frac{x^T}{\|x^T\|_{\infty}} \right\|_{\infty} > \|A\|_{\infty}$$

pero como $\frac{x^T}{\|x^T\|_{\infty}}$ es un vector unitario para la norma infinito, esto contradice el Lema 1.18.

Ahora,

$$\begin{aligned} \|AB\|_{\infty} &= \max_{\|y^T\|_{\infty}=1} \|AB y^T\|_{\infty} \leq \max_{\|y^T\|_{\infty}=1} \|A\|_{\infty} \|B y^T\|_{\infty} = \\ &= \|A\|_{\infty} \max_{\|y^T\|_{\infty}=1} \|B y^T\|_{\infty} = \|A\|_{\infty} \|B\|_{\infty} \end{aligned}$$

□

En el siguiente Lema presentamos una parte de la demostración del Teorema de Perron-Frobenius para matrices estocásticas positivas (Teorema 1.13).

Lema 1.20. *Sea $A \in M_{n \times n}(\mathbb{R}_{>0})$ una matriz estocástica. Entonces, $\lambda = 1$ es autovalor de A con multiplicidad algebraica 1.*

Demostración. Sabemos que $\lambda = 1$ es autovalor de A por el Lema 1.16. Además, por el mismo Lema, sabemos que existe $u \in \mathbb{R}_{\geq 0}^n$ tal que $\sum_{j=1}^n u_j = 1$ y $Au^T = u^T$.

Ahora veamos que la multiplicidad geométrica de $\lambda = 1$ es 1. Procedemos por reducción al absurdo, suponiendo que es mayor o igual a 2. Sea $w \in \mathbb{R}^n$ otro vector que no es proporcional a u , y tal que $Aw^T = w^T$. Como todas las entradas de u son positivas, entonces existe $\alpha \in \mathbb{R}$ tal que $u + \alpha w$ tiene todas las entradas mayores o iguales a cero y, al menos, una de ellas es cero. Por tanto $A(u + \alpha w)^T = (u + \alpha w)^T$, pero A tiene todas sus entradas positivas, y $u + \alpha w$ tiene algunas entradas positivas y otras iguales a cero, y por el Lema 1.16 hemos llegado a una contradicción. Luego, la multiplicidad geométrica de $\lambda = 1$ es 1.

Veamos ahora que la multiplicidad algebraica de $\lambda = 1$ es también 1. Sabemos que si esto no fuese verdad, teniendo en cuenta el Teorema 1.7, la matriz de Jordan $J \in M_{n \times n}(\mathbb{C})$ de A tendría un bloque de Jordan de tamaño $k \times k$ con $k \geq 2$ de la forma:

$$J(1) = \begin{pmatrix} 1 & 1 & 0 & \cdots & 0 & 0 \\ 0 & 1 & 1 & \cdots & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & 1 & 1 \\ 0 & 0 & 0 & \cdots & 0 & 1 \end{pmatrix}$$

Como $J = P^{-1}AP$ con P invertible tenemos que $J^m = P^{-1}A^mP$ para todo $m \in \mathbb{Z}_{>0}$ y, por el Lema 1.19, tenemos que: $\|J^m\|_\infty \leq \|P^{-1}\|_\infty \|A^m\|_\infty \|P\|_\infty$ teniendo así $\|A^m\|_\infty \geq \frac{\|J^m\|_\infty}{\|P^{-1}\|_\infty \|P\|_\infty}$ donde $\|P^{-1}\|_\infty \|P\|_\infty = c \in \mathbb{R}$.

Por el Lema 1.11, tenemos que

$$J(1)^m = \begin{pmatrix} 1 & \binom{m}{1} & \binom{m}{2} & \cdots & \binom{m}{k-2} & \binom{m}{k-1} \\ 0 & 1 & \binom{m}{1} & \cdots & \binom{m}{k-3} & \binom{m}{k-2} \\ 0 & 0 & 1 & \cdots & \binom{m}{k-4} & \binom{m}{k-3} \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & 1 & \binom{m}{1} \\ 0 & 0 & 0 & \cdots & 0 & 1 \end{pmatrix}$$

Luego $\|J^m\|_\infty \geq \|J(1)^m\|_\infty > m$, por tanto $\lim_{m \rightarrow \infty} (\|J^m\|_\infty) = \infty$ y, en consecuencia, $\lim_{m \rightarrow \infty} \|A^m\|_\infty = \infty$. Tomando ahora $u = (u_1, u_2, \dots, u_n) \in \mathbb{R}^n$ tal que $Au^T = u^T$ tenemos que $A^m u^T = u^T$. Sea $m \in \mathbb{Z}_{>0}$, escribimos $A^m = (a_{i,j})$ y tomamos ℓ la fila tal que $\sum_{j=1}^n |a_{\ell,j}| = \|A^m\|_\infty$. Entonces, tenemos que:

$$\|u\|_\infty \geq u_\ell = \sum_{j=1}^n a_{\ell,j} u_j \geq \sum_{j=1}^n a_{\ell,j} \min(u_j) = \|A^m\|_\infty \min(u_j).$$

Sin embargo, esto no es posible porque $\frac{\|u\|_\infty}{\min(u_j)}$ es constante y $\lim_{m \rightarrow \infty} \|A^m\|_\infty = \infty$, llegando a un absurdo. \square

Una vez introducidos los resultados anteriores, procedemos a demostrar el Teorema 1.13 (la versión del Teorema de Perrón-Frobenius para matrices estocásticas positivas):

Demostración del Teorema 1.13. En el Lema 1.20 hemos demostrado que $\lambda = 1$ es autovalor de multiplicidad algebraica 1.

Además, por el Lema 1.16 se tiene que $V_{\lambda=1} = \langle u \rangle$, $u = (u_1, \dots, u_n) \in \mathbb{R}^n$.

Por último, veamos que si γ es otro autovalor de A , entonces $|\gamma| < 1$. Primero, observamos que si γ es autovalor de A , entonces γ es autovalor de A^T (pues A y A^T tienen el mismo polinomio característico). Así, existe $w = (w_1, \dots, w_n) \in \mathbb{C}^n \setminus \{0\}$, tal que $A^T w^T = \gamma w^T$, y por tanto $\sum_{j=1}^n a_{j,i} w_j = \gamma w_i$, con $i \in \{1, \dots, n\}$.

Ahora, consideremos k tal que $|w_k| = \max\{|w_i|, 1 \leq i \leq n\} \neq 0$. Entonces,

$$\sum_{j=1}^n a_{j,k} w_j = \gamma w_k$$

y de aquí,

$$\left| \sum_{j=1}^n a_{j,k} w_j \right| = |\gamma w_k| = |\gamma| |w_k|$$

Además, por la desigualdad triangular, como $a_{j,k} \in \mathbb{R}_{>0}$ y A es estocástica:

$$\begin{aligned} \left| \sum_{j=1}^n a_{j,k} w_j \right| &\leq \sum_{j=1}^n |a_{j,k} w_j| = \sum_{j=1}^n |a_{j,k}| |w_j| = \sum_{j=1}^n a_{j,k} |w_j| \leq \sum_{j=1}^n a_{j,k} |w_k| \\ &= |w_k| \left(\sum_{j=1}^n a_{j,k} \right) = |w_k| \end{aligned}$$

Así, tenemos que $|w_k| \geq |\gamma| |w_k|$, y por tanto $|\gamma| \leq 1$.

Además, si $|\gamma| = 1$ entonces

$$\left| \sum_{j=1}^n a_{j,k} w_j \right| = \sum_{j=1}^n a_{j,k} |w_j| = |w_k| \left(\sum_{j=1}^n a_{j,k} \right) = |w_k|$$

y por el Lema 1.14, esto solo pasa si $w = (w_k, \dots, w_k)$, teniendo así que $(1, \dots, 1)$ es proporcional a w y, por tanto, también cumple que

$$A^T(1, \dots, 1)^T = \gamma(1, \dots, 1)^T$$

pero al ser A estocástica también sabemos que

$$A^T(1, \dots, 1)^T = (1, \dots, 1)^T$$

lo que implica que $\gamma = 1$ quedando así completa la prueba. \square

1.4. Adaptación del método de las potencias al algoritmo PageRank

A continuación vamos a estudiar cómo, para toda matriz estocástica A con entradas positivas, con el método de las potencias podemos obtener el vector $u = (u_1, \dots, u_n) \in \mathbb{R}_{>0}^n$ tal que $Au^T = u^T$ descrito en el Teorema 1.13.

Previamente, nos centramos en demostrar que las potencias de los bloques de Jordan de autovalores con módulo menor que 1 tienden a la matriz nula.

Proposición 1.21. *Sea $J(\lambda)$ un bloque de Jordan de tamaño $k \times k$. Si $|\lambda| < 1$, entonces $\lim_{m \rightarrow \infty} J(\lambda)^m = 0$.*

Demostración. El resultado se obtiene teniendo en cuenta el Lema 1.11 y observando que cuando $|\lambda| < 1$:

$$\begin{aligned} \lim_{m \rightarrow \infty} \left| \binom{m}{j} \lambda^{m-j} \right| &= \lim_{m \rightarrow \infty} \left| \frac{m(m-1) \cdots (m-j+1)}{j!} \lambda^{m-j} \right| \leq \\ &\lim_{m \rightarrow \infty} \left| \frac{m^j}{j!} \lambda^{m-j} \right| = \lim_{m \rightarrow \infty} \frac{m^j}{j!} |\lambda|^{m-j} = 0 \end{aligned}$$

pues la exponencial domina sobre el polinomio y $|\lambda| < 1$. \square

Lema 1.22. *Sea $A \in M_{n \times n}(\mathbb{R})$ una matriz estocástica. Si tenemos $Ax^T = y^T$, con $x = (x_1, \dots, x_n)$, $y = (y_1, \dots, y_n)$, entonces $\sum_{i=1}^n x_i = \sum_{i=1}^n y_i$.*

Demostración. Como $Ax^T = y^T$, tenemos la siguiente igualdad:

$$y_i = \sum_{j=1}^n a_{i,j} x_j, \text{ para todo } i \in \{1, \dots, n\}$$

y, por consiguiente,

$$\sum_{i=1}^n y_i = \sum_{i=1}^n \sum_{j=1}^n a_{i,j} x_j = \sum_{j=1}^n \left(\sum_{i=1}^n a_{i,j} \right) x_j = \sum_{j=1}^n x_j.$$

En la última igualdad hemos usado que A es una matriz estocástica, por lo que sabemos que $\sum_{i=1}^n a_{i,j} = 1$ para todo $j \in \{1, \dots, n\}$. \square

A continuación demostraremos la convergencia del método de las potencias, para matrices estocásticas positivas.

Teorema 1.23 (Convergencia del método de las potencias). *Sea $A \in M_{n \times n}(\mathbb{R}_{>0})$ una matriz estocástica con todas sus entradas positivas, y sea $x \in \mathbb{R}^n$ un vector cualquiera. Entonces $\lim_{m \rightarrow \infty} A^m x^T = (x_1 + \cdots + x_n) u^T$ siendo $V_{\lambda=1} = \langle u \rangle$, con $u = (u_1, \dots, u_n) \in \mathbb{R}_{>0}^n$ y $\sum_{i=1}^n u_i = 1$.*

Demostración. Sea $f : \mathbb{C}^n \rightarrow \mathbb{C}^n$ el endomorfismo con matriz asociada A respecto a la base canónica de \mathbb{C}^n . Sean $\lambda_1 = 1, \lambda_2, \dots, \lambda_s \in \mathbb{C}$ los autovalores de f y consideremos $\mathcal{B}' = \mathcal{B}_{M(\lambda_1)} \cup \dots \cup \mathcal{B}_{M(\lambda_s)} = \{u, v_2, \dots, v_n\}$ una base de \mathbb{C}^n tal que:

- $\mathcal{B}_{M(\lambda_1)} = \{u\}$ es base de $M(\lambda_1) = V_{\lambda_1}$, siendo $u = (u_1, \dots, u_n) \in \mathbb{R}_{>0}^n$ el vector descrito en el Teorema 1.13 del que, además, suponemos que $\sum_{i=1}^n u_i = 1$.
- $\mathcal{B}_{M(\lambda_i)}$ es base de $M(\lambda_i)$ para todo $i \in \{2, \dots, s\}$.
- $M(f)_{\mathcal{B}', \mathcal{B}'}$ es de Jordan.

Si consideramos la composición $f \circ f \circ \dots \circ f = f^m$ tenemos que $M_{\mathcal{B}_C, \mathcal{B}_C}(f^m) = A^m$, con $M_{\mathcal{B}', \mathcal{B}'}(f^m) = J^m$.

Sea $x = (x_1, \dots, x_n) \in \mathbb{R}^n$. Veamos cuánto vale $f^m(x)$. Para ello, escribimos x en función de la base \mathcal{B}' :

$$x = \alpha_1 u + \alpha_2 v_2 + \cdots + \alpha_n v_n$$

Entonces, como los autovectores γ asociados a un autoespacio distinto del $\lambda = 1$ cumplen que $|\gamma| < 1$ (por el Teorema 1.13), usando la Proposición 1.21 y que la multiplicidad algebraica de $\lambda = 1$ es 1, tenemos que:

$$\lim_{m \rightarrow \infty} J^m(\alpha_1, \dots, \alpha_n)^T = (\alpha_1, 0, \dots, 0)^T$$

Por tanto

$$\begin{aligned} \lim_{m \rightarrow \infty} f^m(x) &= \lim_{m \rightarrow \infty} f^m((x_1, \dots, x_n)_{\mathcal{B}'}) = \lim_{m \rightarrow \infty} f^m((\alpha_1, \dots, \alpha_n)_{\mathcal{B}'}) = \\ &(\alpha_1, 0, \dots, 0)_{\mathcal{B}'} = \alpha_1 u \end{aligned}$$

Así, tenemos que $\lim_{m \rightarrow \infty} A^m x = \alpha_1 u$ y por el Lema 1.22 tenemos que:

$$\sum_{i=1}^n x_i = \sum_{i=1}^n \alpha u_i = \alpha \sum_{i=1}^n u_i = \alpha$$

□

Veamos ahora cómo combinar el Teorema 1.13 y el Teorema 1.23 para obtener el algoritmo *PageRank*. Dado un grafo dirigido con n nodos $\{v_1, \dots, v_n\}$ construimos la siguiente matriz cuadrada $A = (a_{i,j}) \in M_{n \times n}(\mathbb{R})$ con $a_{i,j} = \frac{1}{d_j}$ si

hay un arco del nodo v_j al nodo v_i , y donde d_j denota el número de arcos que salen de v_j (el llamado *grado saliente* del nodo v_j), y $a_{i,j} = 0$ en caso contrario. Por construcción, la matriz A es estocástica. El vector de importancias buscado es el vector $u \in \mathbb{R}_{>0}^n$ tal que $Au^T = u^T$, y su existencia y unicidad (salvo producto por escalares) está garantizada por el Teorema de Perron-Frobenius.

Además, para obtener una estimación de este vector, se toma un vector $x = (x_1, \dots, x_n)$ aleatorio, $x_0 = x$, y se define $x_{i+1}^T = Ax_i^T$, y x_m con m lo suficientemente grande.

Como $x_m^T = A^m x^T$ para todo $m \in \mathbb{N}$, el Teorema 1.23 garantiza que $A^m x^T$ converge a un vector proporcional a u^T . Para asegurar que este proceso no converge al vector nulo, basta con tomar $x = (x_1, \dots, x_n)$ tal que $x_1 + \dots + x_n \neq 0$. En particular, esto explica por qué en la Sección 1.1, dependiendo del vector x utilizado, el proceso a veces convergía al vector nulo y otras veces no.

Ahora bien, hemos pasado por alto una hipótesis del Teorema 1.13, y es que este resultado requiere que la matriz A tenga entradas estrictamente positivas. ¿Qué podemos hacer si A tiene entradas nulas? En este caso, podemos perturbar ligeramente las entradas de A . Es decir, transformamos nuestra matriz inicial $A = (a_{i,j})$, siendo A una matriz estocástica, a la matriz $A' = \left(\frac{a_{i,j} + \varepsilon}{1 + n\varepsilon}\right)$ consiguiendo dicha condición.

Lema 1.24. *Sea $A = (a_{i,j})$ una matriz estocástica. Entonces, para $\varepsilon > 0$ tenemos que $\frac{a_{i,j} + \varepsilon}{1 + n\varepsilon} > 0$. Además, la matriz $A' = \left(\frac{a_{i,j} + \varepsilon}{1 + n\varepsilon}\right)$ es estocástica.*

Demostración. Veamos primero que todas sus entradas son positivas. Sea $\varepsilon > 0$ y sean $a_{i,j} \geq 0$ las entradas de nuestra matriz A . Entonces, sabemos que $\frac{a_{i,j} + \varepsilon}{1 + n\varepsilon} > 0$, y esto se da porque $\varepsilon > 0$ y $n \in \mathbb{N}$.

Veamos pues que A' es estocástica (ver Definición 1.12). Es decir, veamos que $\sum_{i=1}^n \frac{a_{i,j} + \varepsilon}{1 + n\varepsilon} = 1$. En virtud de que $\sum_{i=1}^n a_{i,j} = 1$ se tiene que:

$$\sum_{i=1}^n \frac{a_{i,j} + \varepsilon}{1 + n\varepsilon} = \sum_{i=1}^n \frac{a_{i,j}}{1 + n\varepsilon} + \sum_{i=1}^n \frac{\varepsilon}{1 + n\varepsilon} = \frac{1}{1 + n\varepsilon} \sum_{i=1}^n a_{i,j} + \frac{n\varepsilon}{1 + n\varepsilon} = \frac{1 + n\varepsilon}{1 + n\varepsilon} = 1$$

□

Cotas para el número de independencia de un grafo

El número de independencia de un grafo es el número máximo de vértices no adyacentes dos a dos. El objetivo de este capítulo es dar cotas para este parámetro. Para ello, asociaremos a todo grafo su matriz de adyacencia (una matriz simétrica real) y acotaremos el valor del número de independencia en términos de los autovalores de esta matriz. Para justificar estas cotas, usaremos herramientas de Álgebra Lineal tales como el Teorema espectral y el Teorema del entrelazamiento de Cauchy, que serán introducidos y demostrados en la primera sección de este capítulo. En la segunda sección definimos la matriz de adyacencia de un grafo, y establecemos las primeras relaciones entre autovalores y ciertos parámetros del grafo, como son el grado mínimo, el grado medio y el grado máximo. Finalmente, en la tercera sección se incluyen los resultados principales de este capítulo, donde se presentan varias cotas para el número de independencia en términos del espectro de la correspondiente matriz de adyacencia.

2.1. Teorema espectral y Teorema de entrelazamiento de matrices simétricas reales

Como ya se indicó, el objetivo de esta sección es introducir y demostrar el Teorema espectral (Teorema 2.6) y el Teorema del entrelazamiento (Teorema 2.9) de matrices simétricas reales. Comenzamos esta sección recordando algunas nociones.

Definición 2.1. *Un endomorfismo lineal $f : \mathbb{R}^n \rightarrow \mathbb{R}^n$ es simétrico si cumple que*

$$\langle f(u), v \rangle = \langle u, f(v) \rangle$$

para cualquier $u, v \in \mathbb{R}^n$, donde $\langle \cdot, \cdot \rangle$ denota el producto escalar usual de \mathbb{R}^n ; es decir $\langle x, y \rangle = xy^T = \sum_{i=1}^n x_i y_i$ siendo $x = (x_1, \dots, x_n)$, $y = (y_1, \dots, y_n) \in \mathbb{R}^n$.

Definición 2.2. *Una base $\mathcal{B} = \{e_1, \dots, e_n\}$ de \mathbb{R}^n es ortonormal si*

$$\langle e_i, e_j \rangle = \begin{cases} 0 & \text{si } i \neq j, \\ 1 & \text{si } i = j. \end{cases}$$

Lema 2.3. Sean $\mathcal{B} = \{e_1, \dots, e_n\}$ una base ortonormal de \mathbb{R}^n y $x \in \mathbb{R}^n$. Entonces, $x = \sum_{i=1}^n \langle x, e_i \rangle e_i$.

Demostración. Como \mathcal{B} es base, existen unos únicos $\alpha_1, \dots, \alpha_n \in \mathbb{R}$ tales que $x = \sum_{i=1}^n \alpha_i e_i$. Como \mathcal{B} es ortonormal, se tiene que:

$$\langle x, e_i \rangle = \langle \alpha_1 e_1 + \dots + \alpha_n e_n, e_i \rangle = \alpha_1 \langle e_1, e_i \rangle + \dots + \alpha_i \langle e_i, e_i \rangle + \dots + \alpha_n \langle e_n, e_i \rangle = \alpha_i.$$

Así, $\langle x, e_i \rangle = \alpha_i$, como queríamos demostrar. \square

Proposición 2.4. Sea $f : \mathbb{R}^n \rightarrow \mathbb{R}^n$ un endomorfismo, y sea A la matriz asociada a f respecto de una base ortonormal de \mathbb{R}^n . Entonces f es un endomorfismo simétrico si, y solo si, A es una matriz simétrica.

Demostración. Sea $\mathcal{B} = \{e_1, \dots, e_n\}$ una base ortonormal. Como la columna j -ésima $A = (a_{i,j})$ está compuesta por las coordenadas del vector $f(e_j)$ respecto de \mathcal{B} y, por el Lema 2.3, $f(e_j) = \sum_{i=1}^n \langle f(e_j), e_i \rangle e_i$, se tiene que:

$$a_{i,j} = \langle f(e_j), e_i \rangle = \langle e_j, f(e_i) \rangle = \langle f(e_i), e_j \rangle = a_{j,i}$$

por lo que A es simétrica.

Ahora, supongamos que A es simétrica, y sean $x, y \in \mathbb{R}^n$, con $x = \sum_{i=1}^n x_i e_i$, $y = \sum_{i=1}^n y_i e_i$. Utilizando que \mathcal{B} ortonormal tenemos que $\langle x, y \rangle = xy^T$ con $x = (x_1, \dots, x_n)$, $y = (y_1, \dots, y_n)$, y por tanto:

$$\langle f(x), y \rangle = (Ax^T)^T y^T = (xA^T)y^T = x(A^T y^T) = x(Ay^T) = \langle x, f(y) \rangle,$$

llegando a que f es un endomorfismo simétrico. \square

Proposición 2.5. Sea $f : \mathbb{R}^n \rightarrow \mathbb{R}^n$ un endomorfismo simétrico. Entonces, todos los autovalores de f son reales.

Demostración. Sea A la matriz asociada al endomorfismo f respecto de una base ortonormal de \mathbb{R}^n . Por el Teorema Fundamental del Álgebra, f tiene n autovalores complejos (contando cada uno tantas veces como indica su multiplicidad algebraica). Veamos que se trata de autovalores reales. Sean $\lambda = x + iy \in \mathbb{C}$ un autovalor de f y sea $\omega = u + iv \in \mathbb{C}^n$ un autovector no nulo asociado a λ , con $x, y \in \mathbb{R}$, $u, v \in \mathbb{R}^n$. Entonces:

$$A\omega = Au + iAv$$

y como sabemos que

$$A\omega = \lambda\omega = (x + iy)(u + iv) = (xu - yv) + i(yu + xv)$$

separando partes real e imaginaria tenemos lo siguiente:

$$\begin{cases} Au = xu - yv \\ Av = yu + xv \end{cases}$$

Además,

$$\begin{aligned} \langle Au, v \rangle &= \langle xu - yv, v \rangle = x\langle u, v \rangle - y\langle v, v \rangle \\ \langle u, Av \rangle &= \langle u, yu + xv \rangle = x\langle u, u \rangle + y\langle u, v \rangle \end{aligned}$$

y si restamos $\langle Av, v \rangle - \langle u, Av \rangle$, nos queda:

$$0 = y(\langle u, u \rangle + \langle v, v \rangle)$$

lo que implica que $y = 0$ puesto que $\langle u, u \rangle + \langle v, v \rangle \neq 0$ ya que habíamos supuesto que $\omega = u + iv$ es no nulo, luego $\lambda \in \mathbb{R}$. \square

Ya estamos en condiciones de enunciar y demostrar el Teorema espectral.

Teorema 2.6 (Teorema espectral). *Toda matriz simétrica real es diagonalizable en \mathbb{R} .*

Demostración. Por la Proposición 2.4 sabemos que la matriz A es la asociada a un endomorfismo simétrico $f : \mathbb{R}^n \rightarrow \mathbb{R}^n$ respecto de una base ortonormal, y según la Proposición 2.5, todos sus autovalores $\lambda_1, \dots, \lambda_s$ son reales. Veamos que \mathbb{R}^n es suma directa de los autoespacios de f , es decir, $\mathbb{R}^n = V_{\lambda_1} \oplus \dots \oplus V_{\lambda_s}$. Vamos a demostrarlo por reducción al absurdo. Llamemos $U = V_{\lambda_1} \oplus \dots \oplus V_{\lambda_s}$ y supongamos que $\mathbb{R}^n \neq U$, entonces $U^\perp = \{x \in \mathbb{R}^n \mid \langle x, u \rangle = 0, \text{ con } u \in U\} \neq \emptyset$.

Veamos ahora que $f(U^\perp) \subseteq U^\perp$. Sea $x \in U^\perp$. Veamos que $f(x) \in U^\perp$. Sea $\{u_1, \dots, u_r\}$ una base de U formada por autovectores y veamos que $\langle f(x), u_i \rangle = 0, \forall i \in \{1, \dots, r\}$. Por ser f endomorfismo simétrico tenemos $\langle f(x), u_i \rangle = \langle x, f(u_i) \rangle$, y por ser u_i autovector tenemos que existe $\lambda \in \mathbb{R}$ tal que $f(u_i) = \lambda u_i$ y, por tanto, $\langle x, f(u_i) \rangle = \langle x, \lambda u_i \rangle = \lambda \langle x, u_i \rangle$, y como $x \in U^\perp$ tenemos que $\langle x, u_i \rangle = 0$.

Considerando la restricción de f a U^\perp obtenemos un endomorfismo $g : U^\perp \rightarrow U^\perp$ que, al menos, tiene un autovector no nulo $0 \neq v \in U^\perp$. Así, tenemos que $f(v) = g(v) = \lambda v$, luego v es autovector de f y, por tanto, $v \in U$. Entonces hemos deducido que $U \cap U^\perp \neq \emptyset$, pero esto no es posible. \square

Como consecuencia directa del resultado anterior se tiene lo siguiente:

Corolario 2.7. *Sea $A \in M_{n \times n}(\mathbb{R})$ una matriz simétrica. Entonces, existe una base de \mathbb{R}^n ortonormal formada por autovalores de A .*

Demostración. Por el Teorema 2.6, la matriz A es diagonalizable y sus autovalores $\lambda_1, \dots, \lambda_s$ son reales. Sea \mathcal{B}_i una base ortonormal del autoespacio V_{λ_i} . Si consideramos $\mathcal{B} = \mathcal{B}_1 \cup \dots \cup \mathcal{B}_s$, como A es diagonalizable entonces \mathcal{B} es base de \mathbb{R}^n . Para probar que \mathcal{B} además es ortonormal solo falta comprobar que para $u \in \mathcal{B}_i, v \in \mathcal{B}_j$, con $i \neq j$, se tiene que $\langle u, v \rangle = 0$. Esto sucede porque $\lambda_i \langle u, v \rangle = \langle f(u), v \rangle = \langle u, f(v) \rangle = \lambda_j \langle u, v \rangle$, por tanto $(\lambda_i - \lambda_j) \langle u, v \rangle = 0$, y como $\lambda_i - \lambda_j \neq 0$ se tiene que $\langle u, v \rangle = 0$. \square

Nuestro siguiente objetivo es demostrar el Teorema de entrelazamiento, que obtendremos como consecuencia del llamado Principio de Rayleigh. La demostración recogida en esta memoria está sacada de [2].

Proposición 2.8 (Principio de Rayleigh). Sean $A \in M_{n \times n}(\mathbb{R})$ una matriz simétrica real con autovalores $\lambda_1 \geq \dots \geq \lambda_n$ y $\mathcal{B} = \{u_1, \dots, u_n\}$ una base ortonormal de \mathbb{R}^n , donde u_i es un vector propio asociado a λ_i , con $i \in \{1, \dots, n\}$. Entonces:

$$\begin{cases} \frac{uAu^T}{uu^T} \geq \lambda_i & \text{si } u \in \langle u_1, u_2, \dots, u_i \rangle \\ \frac{uAu^T}{uu^T} \leq \lambda_i & \text{si } u \in \langle u_1, u_2, \dots, u_{i-1} \rangle^\perp \end{cases}$$

Demostración. Sea $x \in \mathbb{R}^n$. Entonces, tenemos lo siguiente:

$$\begin{aligned} xAx^T &= (\alpha_1 u_1 + \dots + \alpha_n u_n)A(\alpha_1 u_1 + \dots + \alpha_n u_n)^T = \\ &= (\alpha_1 u_1 + \dots + \alpha_n u_n)(\alpha_1 Au_1 + \dots + \alpha_n Au_n)^T = \\ &= (\alpha_1 u_1 + \dots + \alpha_n u_n)(\lambda_1 \alpha_1 u_1 + \dots + \lambda_n \alpha_n u_n)^T = \\ &= \sum_{i,j=1}^n \lambda_j \alpha_i \alpha_j \langle u_i, u_j \rangle = \lambda_1 \alpha_1^2 + \dots + \lambda_n \alpha_n^2 \end{aligned}$$

Además, por la misma razón:

$$xx^T = (\alpha_1 u_1 + \dots + \alpha_n u_n)(\alpha_1 u_1 + \dots + \alpha_n u_n)^T = \alpha_1^2 + \dots + \alpha_n^2$$

Así

$$\frac{xAx^T}{xx^T} = \frac{\lambda_1 \alpha_1^2 + \dots + \lambda_n \alpha_n^2}{\alpha_1^2 + \dots + \alpha_n^2}$$

Sea $x \in \langle u_1, \dots, u_i \rangle$. Por hipótesis, tenemos que $\lambda_1 \geq \dots \geq \lambda_n$, entonces:

$$\frac{xAx^T}{xx^T} = \frac{\lambda_1 \alpha_1^2 + \dots + \lambda_i \alpha_i^2}{\alpha_1^2 + \dots + \alpha_i^2} \geq \frac{\lambda_i \alpha_1^2 + \dots + \lambda_i \alpha_i^2}{\alpha_1^2 + \dots + \alpha_i^2} = \frac{\lambda_i (\alpha_1^2 + \dots + \alpha_i^2)}{\alpha_1^2 + \dots + \alpha_i^2} = \lambda_i$$

Análogamente, si $x \in \langle u_1, \dots, u_{i-1} \rangle^\perp$, o lo que es lo mismo $x \in \langle u_i, u_{i+1}, \dots, u_n \rangle$. Entonces, tenemos que

$$\frac{xAx^T}{xx^T} = \frac{\lambda_i \alpha_i^2 + \dots + \lambda_n \alpha_n^2}{\alpha_i^2 + \dots + \alpha_n^2} \leq \frac{\lambda_i \alpha_i^2 + \dots + \lambda_i \alpha_n^2}{\alpha_i^2 + \dots + \alpha_n^2} = \frac{\lambda_i (\alpha_i^2 + \dots + \alpha_n^2)}{\alpha_i^2 + \dots + \alpha_n^2} = \lambda_i$$

□

Si $A \in M_{n \times n}(\mathbb{R})$ simétrica y $P \in M_{m \times n}(\mathbb{R})$, con $m < n$, entonces $B = P^T A P$ también es simétrica y, por la Proposición 2.5, los autovalores de A y B son todos reales. El siguiente resultado relaciona estos autovalores cuando $P^T P$ es la matriz identidad de tamaño $m \times m$, que denotaremos por I .

Teorema 2.9 (Teorema de entrelazamiento). *Sea $A \in M_{n \times n}(\mathbb{R})$ una matriz simétrica con autovalores $\lambda_1 \geq \dots \geq \lambda_n$ reales. Sean $P \in M_{n \times m}(\mathbb{R})$ tal que $P^T P = I$ y sea $B = P^T A P \in M_{m \times m}(\mathbb{R})$, con autovalores $\mu_1 \geq \dots \geq \mu_m$. Entonces, los autovalores de B se entrelazan con los de A , esto es, para $i \in \{1, \dots, m\}$:*

$$\lambda_i \geq \mu_i \geq \lambda_{n-m+i}$$

Demostración. Sea $A \in M_{n \times n}(\mathbb{R})$ una matriz simétrica. Entonces A es diagonalizable y todos sus autovalores $\lambda_1 \geq \dots \geq \lambda_n$ son reales por el Teorema espectral (Teorema 2.6). Por el Corolario 2.7, podemos considerar $\mathcal{B} = \{u_1, \dots, u_n\}$ una base ortonormal formada por autovalores de A , tal que u_i es el vector asociado al autovalor λ_i . Sea $\mathcal{B}' = \{v_1, \dots, v_m\}$ una base ortonormal formada por los autovalores de la matriz B , tal que v_i es el vector asociado al autovalor μ_i , con $i \in \{1, \dots, m\}$. Consideremos los subespacios de \mathbb{R}^m siguientes:

$$U = \langle v_1, \dots, v_i \rangle, W = \langle P^T u_1, \dots, P^T u_{i-1} \rangle^\perp, \text{ con } i \in \{1, \dots, m\}.$$

Vamos a estudiar el espacio $U \cap W$, en particular queremos ver que $U \cap W \neq \{0\}$. Observamos que $\dim(U) = i$ porque $U = \langle v_1, \dots, v_i \rangle$ y $\dim(W) = m - \dim(\langle P^T u_1, \dots, P^T u_{i-1} \rangle) \geq m - i + 1$. Además, $\dim(U + W) \leq m$ ya que $U + W \subseteq \mathbb{R}^m$. Por tanto,

$$\dim(U \cap W) = \dim(U) + \dim(W) - \dim(U + W) \geq i + m - i + 1 - m = 1$$

Es decir, existe $s_i \in U \cap W$ con $s_i \neq 0$.

Veamos ahora que $P s_i \in \langle u_1, \dots, u_{i-1} \rangle^\perp$. Sea $j \in \{1, \dots, i-1\}$, se tiene que

$$\langle u_j, P s_i \rangle = \langle P^T u_j, s_i \rangle = 0$$

y, por la Proposición 2.8, tenemos que

$$\frac{(P s_i) A (P s_i)^T}{(P s_i) (P s_i)^T} \leq \lambda_i \text{ y } \frac{s_i B s_i^T}{s_i s_i^T} \geq \mu_i.$$

Como $P P^T = I$ y $P A P^T = B$ por hipótesis, tenemos la siguiente igualdad

$$\frac{(P s_i) A (P s_i)^T}{(P s_i) (P s_i)^T} = \frac{s_i P A P^T s_i^T}{s_i P P^T s_i^T} = \frac{s_i B s_i^T}{s_i s_i^T}$$

Por tanto, ya tenemos que

$$\lambda_i \geq \frac{(P s_i) A (P s_i)^T}{(P s_i) (P s_i)^T} = \frac{s_i B s_i^T}{s_i s_i^T} \geq \mu_i$$

Si denotamos $\lambda'_1 \geq \dots \geq \lambda'_n$, $\mu'_1 \geq \dots \geq \mu'_m$ los autovalores de las matrices opuestas $-A$ y $-B$ respectivamente, tenemos que $\lambda'_i = -\lambda_{n+1-i}$, $\mu'_i = -\mu_{m+1-i}$, y la condición $\lambda'_i \geq \mu'_i$, para cualquier $i \in \{1, \dots, m\}$ es equivalente a $\mu_i \geq \lambda_{n-m+i}$, con $i \in \{1, \dots, m\}$. \square

Sea $A \in M_{n \times n}(\mathbb{R})$, una submatriz de A es aquella que se obtiene al eliminar ciertas filas y columnas de A . Se dice que una submatriz cuadrada es principal si los índices de filas y columnas que se eliminan son los mismos. Claramente, las submatrices principales de una matriz simétrica también son simétricas.

Teorema 2.10. *Sea $A \in M_{n \times n}(\mathbb{R})$ tal que $A = A^T$ con autovalores $\lambda_1, \dots, \lambda_n$ y sea B una submatriz principal de tamaño $m \times m$ con autovalores μ_1, \dots, μ_m . Entonces los autovalores de B se entrelazan con los de A , es decir:*

$$\lambda_i \geq \mu_i \geq \lambda_{n-m+i}$$

Demostración. Supongamos, sin pérdida de generalidad, que B es la submatriz principal formada por las primeras m filas y m columnas de A . Así, $B = P^T A P$, siendo $P \in M_{n \times m}(\mathbb{R})$ la matriz

$$P = \begin{pmatrix} 1 & 0 & \cdots & 0 \\ 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 1 \\ 0 & 0 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 0 \end{pmatrix}$$

se observa que $P^T P = I$ y, por el Teorema 2.9, se sigue el resultado. \square

2.2. Matriz de adyacencia de un grafo no dirigido

A continuación veremos algunas aplicaciones en Teoría de Grafos de los resultados vistos en las secciones anteriores. En esta sección y en las posteriores, solo consideraremos grafos no dirigidos y simples, es decir, sin aristas múltiples ni lazos. Usualmente, n denotará el número de vértices del grafo y m el número de aristas.

Definición 2.11. *Sea $G = (V, E)$ un grafo, se define la matriz de adyacencia de G , como la matriz $A(G) = (a_{i,j}) \in M_{n \times n}(\mathbb{R})$, donde*

$$a_{i,j} = \begin{cases} 1 & \text{si } \{v_i, v_j\} \in E, \\ 0 & \text{en caso contrario.} \end{cases}$$

Observación 2.12. Se observa que $A(G)$ es una matriz simétrica y real para todo grafo G . Por tanto, en virtud del Teorema espectral, $A(G)$ es diagonalizable y tiene todos sus autovalores reales. Además, como el grafo G no tiene lazos, la diagonal principal está compuesta de ceros y, en particular, $\text{tr}(A) = 0$, su traza¹ vale 0.

¹ La traza de una matriz $A \in M_{n \times n}(\mathbb{R})$ se define como $\text{tr}(A) = \sum_{i=1}^n a_{i,i}$, i.e., la suma de los elementos de la diagonal.

Ejemplo 2.13. En la Figura 2.1 presentamos un grafo con 9 vértices y 15 aristas, y su correspondiente matriz de adyacencia. Esta matriz, como hemos mencionado

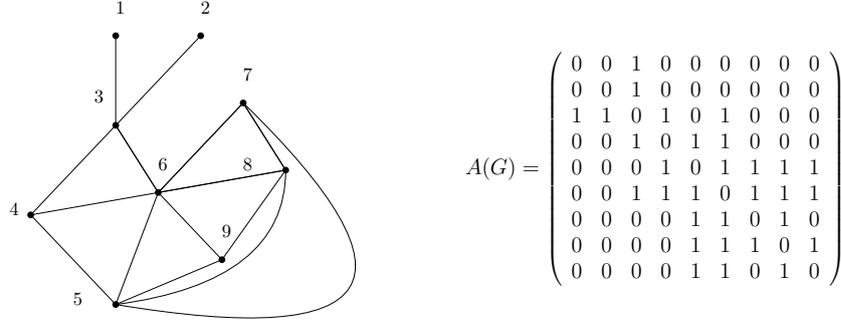


Figura 2.1. Un grafo y su matriz de adyacencia.

anteriormente, representa las conexiones que hay entre cada par de vértices. Así, si dos vértices i, j están conectados habrá un 1 tanto en la posición $a_{i,j}$ como en la posición $a_{j,i}$. Por contra, si no están conectados, habrá un 0 en ambas posiciones.

Definición 2.14. Definimos como el grado de un vértice v en G al número de aristas incidentes al vértice y se denota por $\deg(v)$. Denotamos por $\delta(G)$ al menor grado de los vértices de G y por $\Delta(G)$ al mayor grado de los vértices de G .

El grado medio de G , se define como $\bar{\delta}(G) = \sum_{i=1}^n \frac{\deg(v_i)}{n} = \frac{2m}{n}$.

Ejemplo 2.15. Seguimos con el Ejemplo 2.13. Observamos que en la Figura 2.1 el grado mínimo de G es $\delta(G) = 1$, ya que $\deg(v_1) = \deg(v_2) = 1$ y no hay vértices de grado 0. El grado máximo es $\Delta(G) = 6$ y se alcanza en el vértice v_6 , que está conectado con v_3, v_4, v_5, v_7, v_8 y v_9 . Y el grado medio es

$$\bar{\delta}(G) = \sum_{i=1}^9 \frac{\deg(v_i)}{9} = \frac{1 + 1 + 4 + 3 + 5 + 6 + 3 + 4 + 3}{9} = \frac{2m}{n} = \frac{30}{9} \approx 3,3.$$

Observación 2.16. En la fila i -ésima de $A(G)$ hay tantos 1 como vértices adyacentes a v_i , es decir, tantos como $\deg(v_i)$ (grado del vértice v_i). En consecuencia,

$$A(G) \begin{pmatrix} 1 \\ \vdots \\ 1 \end{pmatrix} = \begin{pmatrix} \deg(v_1) \\ \vdots \\ \deg(v_n) \end{pmatrix} \text{ y } (1, \dots, 1)A(G) \begin{pmatrix} 1 \\ \vdots \\ 1 \end{pmatrix} = \sum_{i=1}^n \deg(v_i) = 2m = n\bar{\delta}(G).$$

Como ya mencionamos, $A(G)$ es de tamaño $n \times n$, diagonalizable y tiene todos sus autovalores reales. Cuando decimos que $\lambda_1 \geq \dots \geq \lambda_n$ son los autovalores de $A(G)$, estamos escribiendo cada autovalor tantas veces como indica su multiplicidad. En ocasiones nos referiremos a λ_1 como el autovalor dominante.

A continuación tenemos un resultado en el que se muestran algunas relaciones entre los autovalores de mayor y menor valor de $A(G)$ y algunos parámetros del grafo G .

Teorema 2.17. *Sea G un grafo y sean $\lambda_1 \geq \dots \geq \lambda_n$ los autovalores de la matriz de adyacencia $A(G)$. Entonces,*

$$\bar{\delta}(G) \leq \lambda_1 \leq \Delta(G)$$

Además, $\lambda_n \leq 0$, y $\lambda_n = 0$ si, y solo si, el grafo no tiene aristas.

Demostración. Primero vamos a demostrar que $\bar{\delta}(G) \leq \lambda_1$. Sea $A = A(G)$, y sea $P = \left(\frac{1}{\sqrt{n}}, \dots, \frac{1}{\sqrt{n}}\right) \in M_{n \times 1}(\mathbb{R})$. Sabemos que

$$P^T P = \left(\frac{1}{\sqrt{n}}, \dots, \frac{1}{\sqrt{n}}\right)^T \left(\frac{1}{\sqrt{n}}, \dots, \frac{1}{\sqrt{n}}\right) = 1.$$

Además,

$$P^T A P = \frac{1}{\sqrt{n}}(1, \dots, 1)^T A(G) \frac{1}{\sqrt{n}}(1, \dots, 1) = \frac{1}{n}(1, \dots, 1)^T A(G)(1, \dots, 1)$$

Además, por la Observación 2.16 sabemos que

$$\begin{aligned} \frac{1}{n}(1, \dots, 1)^T A(G)(1, \dots, 1) &= \frac{1}{n}(1, \dots, 1)^T (\deg(v_1), \dots, \deg(v_n)) \\ &= \frac{1}{n} \left(\sum_{i=1}^n \deg(v_i) \right) = \bar{\delta}(G) \end{aligned}$$

Además, por el Teorema 2.9, tenemos que:

$$\lambda_1 \geq \bar{\delta}(G)$$

Ahora solo nos falta demostrar que

$$\Delta(G) \geq \lambda_1$$

Sabemos que si λ_1 es autovalor de $A(G)$ tenemos que existe $x = (x_1, \dots, x_n) \in \mathbb{R}^n$ tal que

$$A(G)x^T = \lambda_1 x^T$$

Supongamos, sin pérdida de generalidad, que $|x_1| \geq |x_i|$, con $i \in \{1, \dots, n\}$, y que $x_1 > 0$. Entonces tenemos que:

$$a_{1,1}x_1 + \dots + a_{1,n}x_n = \lambda_1 x_1$$

y, en consecuencia

$$\begin{aligned} \lambda_1 x_1 &= a_{1,1}x_1 + \cdots + a_{1,n}x_n \leq a_{1,1}x_1 + \cdots + a_{1,n}x_1 = x_1(a_{1,1} + \cdots + a_{1,n}) = \\ &= \deg(v_1)x_1 \leq \Delta(G)x_1 \end{aligned}$$

Por tanto, tenemos que $\lambda_1 \leq \Delta(G)$ por la Observación 2.2.

Vamos a demostrar ahora que $\lambda_n \leq 0$. Sabemos que $\lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_n$ y $\text{tr}(A(G)) = \lambda_1 + \lambda_2 + \cdots + \lambda_n$. De aquí se tiene que $\lambda_n \leq 0$. Además, en las mismas condiciones, tenemos que $\lambda_n = 0$ si, y solo si, $\lambda_{n-1} = \cdots = \lambda_1 = 0$, y en este caso tenemos que $0 = \lambda_1 \geq \bar{\delta}(G) = 0$, concluyendo así que G no tiene aristas. \square

Ejemplo 2.18. Volviendo al Ejemplo 2.13, el autovalor dominante de su matriz de adyacencia es $\lambda_1 \approx 4,066$. Además, por el Ejemplo 2.15 sabemos que el grado medio es $\bar{\delta}(G) \approx 3,3$ y que el grado máximo es $\Delta(G) = 6$ y, como podemos observar, se verifica que

$$\bar{\delta} \leq \lambda_1 \leq \Delta(G).$$

2.3. Cotas para el número de independencia de un grafo

En esta sección acotaremos el número de independencia de un grafo G en función del mayor y menor autovalor de $A(G)$.

Definición 2.19. Sea $G = (V, E)$ un grafo y $V' \subset V$ un subconjunto de vértices. Decimos que V' es independiente si, y solo si, para todo $u, v \in V'$ se tiene que $\{u, v\} \notin E$.

El máximo número de vértices en un conjunto de vértices independientes se denomina número de independencia de G y se denota por $\alpha(G)$.

Ejemplo 2.20. Volviendo al Ejemplo 2.13, podemos ver que un conjunto independiente es $V' = \{v_1, v_2, v_4, v_7, v_9\}$ porque para cualquier $i, j \in \{1, 2, 4, 7, 9\}$, $i \neq j$, $\{v_i, v_j\} \notin E$. Esto demuestra que $\alpha(G) \geq 5$. Además, en este ejemplo se puede comprobar que no hay conjuntos independientes de 6 elementos y, por tanto, $\alpha(G) = 5$.

En general, calcular el grado de $\alpha(G)$ es una tarea complicada. Vamos a ver que, como hemos mencionado anteriormente, este valor se puede acotar superiormente por una expresión que depende únicamente del mayor y menor autovalor de $A(G)$. Primero veamos la relación que hay entre los independientes y los subgrafos inducidos.

Definición 2.21. Sea $G = (V, E)$ un grafo. Al grafo $H = (V', E')$ se le dice subgrafo inducido por H si $V' \subseteq V$ y $E' = \{e \in E \mid e \subseteq V'\}$.

Lema 2.22. Sea $G = (V, E)$ un grafo, y sea $V' \subseteq V$. Entonces V' es independiente si, y solo si, el subgrafo inducido por V' no tiene aristas.

Demostración. Supongamos que V' es independiente. Entonces, $\{e \in E \mid e \subseteq V'\} = \emptyset$ y, por tanto, el subgrafo inducido por V' no tiene aristas.

Por otra parte, si el subgrafo H inducido por V' no tiene aristas, entonces H solo consta de vértices aislados, de donde se deduce que V' es independiente. \square

Ejemplo 2.23. Continuando con el Ejemplo 2.13, el subgrafo H inducido por $V' = \{v_3, v_6, v_7, v_8\}$ es el coloreado en violeta en la Figura 2.2.

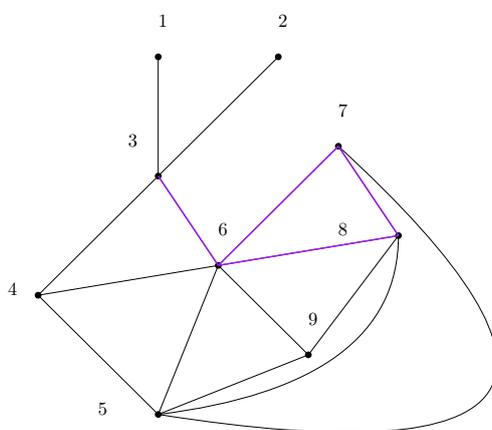


Figura 2.2. Subgrafo inducido por $V' = \{v_3, v_6, v_7, v_8\}$.

Además, su matriz de adyacencia será la correspondiente a tomar la submatriz principal formada por las filas y columnas 3, 6, 7 y 8 de la matriz de adyacencia del grafo G , es decir,

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

Como $A(H)$ es no nula, se tiene que H tiene aristas y, por tanto, $V' = \{v_3, v_6, v_7, v_8\}$ no es un independiente.

En general, si H es un subgrafo inducido por un conjunto de vértices V' de G , entonces $A(H)$ es una submatriz principal de $A(G)$ y, por el Teorema 2.9, los autovalores de $A(G)$ se entrelazan con los de $A(H)$. Una consecuencia directa del Lema 2.22 es que V' es independiente si, y solo si, el subgrafo inducido por V' (que hemos llamado H) no tiene aristas, por lo que

$$A(H) = \begin{pmatrix} 0 & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & 0 \end{pmatrix}$$

y por tanto su único autovalor es $\lambda = 0$. Esta es la idea principal de la demostración del siguiente resultado.

Teorema 2.24. *Sea G un grafo y sean $\lambda_1 \geq \cdots \geq \lambda_n$ los autovalores de $A(G)$. Entonces*

$$\alpha(G) \leq \min \{ |\{i \mid \lambda_i \leq 0\}|, |\{i \mid \lambda_i \geq 0\}| \}.$$

Demostración. Sean $\lambda_1 \geq \cdots \geq \lambda_r > 0 = \lambda_{r+1} = \cdots = \lambda_{r+s} > \lambda_{r+s+1} \geq \cdots \geq \lambda_n$ los autovalores de $A(G)$ y sea $V' \subseteq V$ conjunto independiente con máximo número de vértices posible. Entonces H , el subgrafo inducido por V' tiene $\alpha = \alpha(G)$ vértices (es un independiente maximal). Como $A(H)$ es la matriz nula, entonces tenemos que $\mu_1 = \cdots = \mu_\alpha = 0$ y, por el Teorema 2.9, tenemos lo siguiente:

$$\lambda_1 \geq \mu_1 = 0 \geq \lambda_{1+(n-\alpha)}$$

entonces $0 \geq \lambda_{1+(n-\alpha)}$ de aquí $n - \alpha + 1 \geq r + 1$ y despejando tenemos $\alpha(G) = \alpha \leq n - r$. De forma similar, por el Teorema 2.9 tenemos que $\lambda_\alpha \geq 0 = \mu_\alpha \geq \lambda_n$ en particular $\lambda_\alpha \geq 0$ y, por tanto, $\alpha(G) = \alpha \leq r + s$. \square

Definición 2.25. *Un grafo G se le dice regular cuando todos sus vértices tienen el mismo grado. Un grafo regular cuyos vértices son de grado k se denomina k -regular.*

Lema 2.26. *Si G es k -regular, entonces $\lambda_1 = k$, donde λ_1 representa el autovalor principal asociado a la matriz $A(G)$.*

Demostración. Como G es k -regular, todos sus vértices son de grado k , con lo que el grado medio y el grado máximo de G coinciden, i.e., $\bar{\delta}(G) = \Delta(G) = k$. Además, por el Teorema 2.17 sabemos que $\bar{\delta}(G) \leq \lambda_1 \leq \Delta(G)$, por tanto, $\lambda_1 = k$. \square

Teorema 2.27. *Sea G un grafo k -regular. Entonces*

$$\alpha(G) \leq n \frac{-\lambda_n}{k - \lambda_n}$$

donde λ_n representa el menor autovalor asociado a la matriz $A(G)$.

Demostración. Sea $V' = \{v_1, \dots, v_\alpha\}$ un independiente con $|V'| = \alpha = \alpha(G)$. Analizamos ahora la matriz de adyacencia de $A(G)$:

$$A(G) = \left(\begin{array}{c|c} 0 & A_1 \\ \hline A_1^T & A_2 \end{array} \right)$$

con $A_1 \in M_{\alpha \times (n-\alpha+1)}(\mathbb{R})$ y $A_2 \in M_{(n-\alpha+1) \times (n-\alpha+1)}(\mathbb{R})$ y $0 \in M_{\alpha \times \alpha}(\mathbb{R})$ matriz nula pues $\{v_1, \dots, v_\alpha\}$ es independiente. En cada fila de A_1 hay k unos, entonces en total hay αk unos. Ahora sea $d_{\alpha+1}, \dots, d_n$ el número de unos en la columna $\alpha + 1, \dots, n$ de A_1 ; así $d_{\alpha+1} + \dots + d_n = \alpha k$. Tomamos ahora

$$P = \left(\begin{array}{cc} 1 & 0 \\ \vdots & \vdots \\ 1 & 0 \\ 0 & 1 \\ \vdots & \vdots \\ 0 & 1 \end{array} \right) \left. \begin{array}{l} \vphantom{\begin{pmatrix} 1 \\ \vdots \\ 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix}} \\ \vphantom{\begin{pmatrix} 1 \\ \vdots \\ 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix}} \end{array} \right\} \begin{array}{l} \alpha \\ n - \alpha \end{array}$$

y

$$S = P \begin{pmatrix} \frac{1}{\sqrt{\alpha}} & 0 \\ 0 & \frac{1}{\sqrt{n-\alpha}} \end{pmatrix}.$$

Se observa que

$$P^T P = \begin{pmatrix} \alpha & 0 \\ 0 & n - \alpha \end{pmatrix}$$

y, por tanto, $S^T S = I$. Vamos a aplicar el Teorema 2.9 con la matriz S . Para ello, calculemos antes $P^T A P$ en dos pasos:

$$P^T A = \begin{pmatrix} 0 \cdots 0 & d_{\alpha+1} & \cdots & d_n \\ k \cdots k & k - d_{\alpha+1} & \cdots & k - d_n \end{pmatrix}$$

$$(P^T A)P = \begin{pmatrix} 0 & d_{\alpha+1} + \dots + d_n \\ \alpha k & (n - \alpha)k - d_{\alpha+1} - \dots - d_n \end{pmatrix} = \begin{pmatrix} 0 & \alpha k \\ \alpha k & (n - 2\alpha)k \end{pmatrix}.$$

Por tanto

$$\begin{aligned} S^T A S &= \begin{pmatrix} \frac{1}{\sqrt{\alpha}} & 0 \\ 0 & \frac{1}{\sqrt{n-\alpha}} \end{pmatrix} \begin{pmatrix} 0 & \alpha k \\ \alpha k & (n - 2\alpha)k \end{pmatrix} \begin{pmatrix} \frac{1}{\sqrt{\alpha}} & 0 \\ 0 & \frac{1}{\sqrt{n-\alpha}} \end{pmatrix} = \\ &= \begin{pmatrix} 0 & \frac{\sqrt{\alpha} k}{\sqrt{n-\alpha}} \\ \frac{\alpha}{\sqrt{n-\alpha}} k & \frac{n-2\alpha}{\sqrt{n-\alpha}} k \end{pmatrix} \begin{pmatrix} \frac{1}{\sqrt{\alpha}} & 0 \\ 0 & \frac{1}{\sqrt{n-\alpha}} \end{pmatrix} = \begin{pmatrix} 0 & \frac{\sqrt{\alpha}}{\sqrt{n-\alpha}} k \\ \frac{\sqrt{\alpha}}{\sqrt{n-\alpha}} k & \frac{n-2\alpha}{n-\alpha} k \end{pmatrix} \end{aligned}$$

cuyo polinomio característico es

$$-\lambda \left(\frac{n-2\alpha}{n-\alpha} k - \lambda \right) - \frac{\alpha}{n-\alpha} k^2 = \lambda^2 - \left(\frac{n-2\alpha}{n-\alpha} k \right) \lambda - \frac{\alpha}{n-\alpha} k^2$$

y sus raíces son

$$\begin{aligned}\lambda &= \frac{\frac{n-2\alpha}{n-\alpha}k \pm \sqrt{\frac{(n-2\alpha)^2}{(n-\alpha)^2k^2 + 4\frac{\alpha}{n-\alpha}k^2}}}{2} = \frac{n-2\alpha}{2(n-\alpha)}k \pm \frac{k}{2} \sqrt{\frac{n^2 - 4\alpha n + 4\alpha^2 + 4\alpha(n-\alpha)}{(n-\alpha)^2}} \\ &= \frac{(n-2\alpha \pm n)k}{2(n-\alpha)} \pm \frac{k}{2(n-\alpha)} \sqrt{n^2} = \frac{(n-2\alpha \pm n)k}{2(n-\alpha)},\end{aligned}$$

obteniendo así dos raíces $\mu_1 = k$ y $\mu_2 = -\frac{\alpha k}{n-\alpha}$ y, por el Teorema 2.9, tenemos que

$$-\frac{\alpha k}{n-\alpha} = \mu_2 \geq \lambda_n.$$

Esto implica que

$$-\alpha k \geq n\lambda_n - \alpha\lambda_n,$$

o lo que es lo mismo

$$-n\lambda_n \geq \alpha(k - \lambda_n).$$

Despejando α como sigue, tenemos el resultado buscado:

$$\alpha \leq -\frac{n\lambda_n}{k - \lambda_n}.$$

□

El resultado anterior se puede extender para grafos no regulares de la siguiente forma:

Teorema 2.28. *Sea G un grafo. Si el grado mínimo de G es δ , entonces*

$$\alpha(G) \leq n \frac{-\lambda_1 \lambda_n}{\delta^2 - \lambda_1 \lambda_n}$$

donde λ_1, λ_n representan el mayor y menor autovalor asociado a la matriz $A(G)$, respectivamente.

Demostración. Sea $V = \{v_1, \dots, v_n\}$ y $V' = \{v_1, \dots, v_\alpha\}$ con $\alpha < n$ conjuntos independientes maximales. Definimos $k := \sum_{i=1}^{\alpha} \frac{\deg(v_i)}{\alpha}$. Analizamos ahora la matriz de adyacencia de $A(G)$:

$$A(G) = \left(\begin{array}{c|c} 0 & A_1 \\ \hline A_1^T & A_2 \end{array} \right).$$

Ahora sea $d_{\alpha+1}, \dots, d_n$ el número de ceros que hay en la columna $\alpha + 1, \dots, n$, respectivamente. Como el número total de unos que hay en las filas de A_1 es el mismo que el total del número de unos que hay en las columnas, se cumple que $\sum_{\alpha+1}^n d_i = \sum_{i=1}^{\alpha} \deg(v_i)$, y en cada fila i de la matriz A_1 hay $\deg(v_i)$ unos, entonces en la matriz $A(G)$ hay en total $\sum_{i=1}^{\alpha} \deg(v_i) = k\alpha$ unos.

Sea

$$P = \left(\begin{array}{cc} 1 & 0 \\ \vdots & \vdots \\ 1 & 0 \\ 0 & 1 \\ \vdots & \vdots \\ 0 & 1 \end{array} \right) \left. \begin{array}{l} \} \alpha \\ \\ \} n - \alpha \end{array} \right.$$

y sea

$$S = \left(\begin{array}{cc} \frac{1}{\sqrt{\alpha}} & 0 \\ \vdots & \vdots \\ \frac{1}{\sqrt{\alpha}} & 0 \\ 0 & \frac{1}{\sqrt{n-\alpha}} \\ \vdots & \vdots \\ 0 & \frac{1}{\sqrt{n-\alpha}} \end{array} \right), \text{ o lo que es lo mismo } S = P \left(\begin{array}{cc} \frac{1}{\sqrt{\alpha}} & 0 \\ 0 & \frac{1}{\sqrt{n-\alpha}} \end{array} \right).$$

Entonces tenemos que $S^T S = I$.

Calculemos ahora el producto de matrices SAS^T como sigue:

$$\begin{aligned} (P^T A)P &= P^T \left(\begin{array}{c|c} 0 & A_1 \\ \hline A_1^T & A_2 \end{array} \right) P = \\ & \left(\begin{array}{cccc} 0 & \cdots & 0 & d_{\alpha+1} & \cdots & d_n \\ \deg(v_1) & \cdots & \deg(v_n) & \deg(v_{\alpha+1}) - d_{\alpha+1} & \cdots & \deg(v_n) - d_n \end{array} \right) P = \\ & \left(\begin{array}{cc} 0 & d_{\alpha+1} + \cdots + d_n \\ \deg(v_1) + \cdots + \deg(v_\alpha) & \sum_{i=\alpha+1}^n (\deg(v_i) - d_i) \end{array} \right) = \left(\begin{array}{cc} 0 & k\alpha \\ k\alpha & \sum_{i=\alpha+1}^n \deg(v_i) - k\alpha \end{array} \right). \end{aligned}$$

Ahora hacemos el producto $S^T AS$ y queda:

$$\begin{aligned} (S^T A)S &= \left(\begin{array}{cc} \frac{1}{\sqrt{\alpha}} & 0 \\ 0 & \frac{1}{\sqrt{n-\alpha}} \end{array} \right) \left(\begin{array}{cc} 0 & k\alpha \\ k\alpha & \sum_{i=\alpha+1}^n \deg(v_i) - k\alpha \end{array} \right) \left(\begin{array}{cc} \frac{1}{\sqrt{\alpha}} & 0 \\ 0 & \frac{1}{\sqrt{n-\alpha}} \end{array} \right) = \\ & \left(\begin{array}{cc} 0 & k\sqrt{\alpha} \\ k\frac{\alpha}{\sqrt{n-\alpha}} & \frac{\sum_{i=\alpha+1}^n \deg(v_i) - k\alpha}{\sqrt{n-\alpha}} \end{array} \right) \left(\begin{array}{cc} \frac{1}{\sqrt{\alpha}} & 0 \\ 0 & \frac{1}{\sqrt{n-\alpha}} \end{array} \right) = \left(\begin{array}{cc} 0 & k\frac{\sqrt{\alpha}}{\sqrt{n-\alpha}} \\ k\frac{\sqrt{\alpha}}{\sqrt{n-\alpha}} & \frac{\sum_{i=\alpha+1}^n \deg(v_i) - k\alpha}{n-\alpha} \end{array} \right) = B. \end{aligned}$$

Una vez obtenida esta matriz, vamos a demostrar que $\alpha(G) \leq n \frac{-\lambda_1 \lambda_n}{\delta^2 - \lambda_1 \lambda_n}$. Sean μ_1, μ_2 los autovalores de B . Por el Teorema 2.9 sabemos que $\lambda_1 \geq \mu_1 \geq \mu_2 \geq \lambda_n$, y por la Proposición 2.8 sabemos que $\lambda_1 \geq 0 \geq \lambda_n$.

Veamos que $-\lambda_1 \lambda_n \geq -\mu_1 \mu_2$ estudiando los diferentes casos:

- Si $\mu_2 \geq 0$ o $\mu_1 \leq 0$, el resultado es evidente.
- Si $\mu_1 \geq 0 \geq \mu_2$, tenemos que $\lambda_1 \geq \mu_1 \geq 0 \geq \mu_2 \geq \lambda_n$. Entonces, obtenemos que $-\lambda_n \geq -\mu_2 \geq 0$, y a su vez $-\lambda_1 \lambda_n \geq -\mu_1 \mu_2$.

Como $\mu_1\mu_2 = -\det(B) = \frac{k^2\alpha}{n-\alpha}$ y, por tanto, $k = \sum_{i=1}^{\alpha} \frac{\deg(v_i)}{\alpha} > \delta$, entonces $-\lambda_1\lambda_n \geq \frac{\delta^2\alpha}{n-\alpha}$. Despejando, obtenemos $\alpha = \alpha(G) \leq n \frac{-\lambda_1\lambda_n}{\delta^2 - \lambda_1\lambda_n}$. \square

Ejemplo 2.29. Sea el grafo bipartito G de la Figura 2.3 obtenido al quitarle dos aristas adyacentes al vértice v_4 del grafo bipartito completo $K_{3,5}$ (ver Figura 2.3).

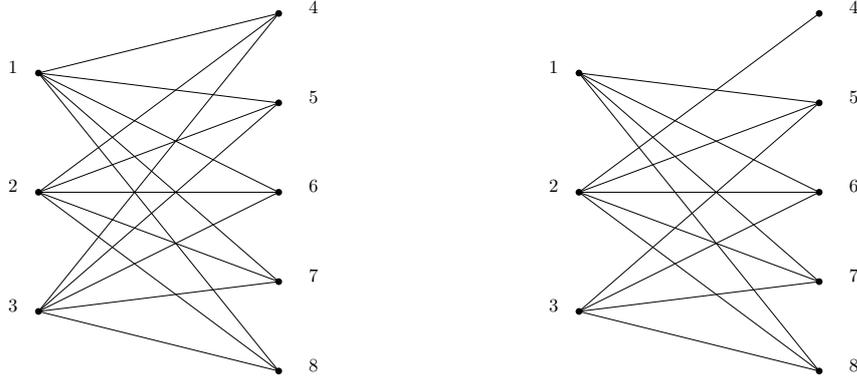


Figura 2.3. Grafo bipartito completo $K_{3,5}$ y un subgrafo $G_{K_{3,5}}$.

y sus respectivas matrices de adyacencia:

$$A(K_{3,5}) = \begin{pmatrix} 0 & 0 & 0 & 1 & 1 & 1 & 1 & 1 & 1 \\ 0 & 0 & 0 & 1 & 1 & 1 & 1 & 1 & 1 \\ 0 & 0 & 0 & 1 & 1 & 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 1 & 1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 1 & 1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 1 & 1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 1 & 1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 1 & 1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix} \quad A(G_{K_{3,5}}) = \begin{pmatrix} 0 & 0 & 0 & 0 & 1 & 1 & 1 & 1 & 1 \\ 0 & 0 & 0 & 1 & 1 & 1 & 1 & 1 & 1 \\ 0 & 0 & 0 & 0 & 1 & 1 & 1 & 1 & 1 \\ 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 1 & 1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 1 & 1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 1 & 1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 1 & 1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 1 & 1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix}$$

Entonces, los autovalores de $A(G_{K_{3,5}})$ ordenados de mayor a menor son los siguientes: $\lambda_1 = \sqrt{\frac{13+\sqrt{137}}{2}}$, $\lambda_2 = \sqrt{\frac{13-\sqrt{137}}{2}}$, $\lambda_3 = \lambda_4 = \lambda_5 = \lambda_6 = 0$ (es decir, 0 es autovalor con multiplicidad 4), $\lambda_7 = -\lambda_2$, y $\lambda_8 = -\lambda_1$. Así, obtenemos que la cota del Teorema 2.24 es $\alpha(G_{K_{3,5}}) \leq \min(|\{i \mid \lambda_i \geq 0\}|, |\{i \mid \lambda_i \leq 0\}|) = 6$ y la del Teorema 2.28 es $\alpha(G_{K_{3,5}}) \leq n \frac{-\lambda_1\lambda_n}{\delta^2 - \lambda_1\lambda_n} \approx 7,4$

Podemos observar que las cotas de ambos teoremas son distintas. Además, el valor real para el número de independencia de nuestro grafo es 5, pues, al ser G un grafo de $K_{3,5}$, entonces $\alpha(G) \geq \alpha(K_{3,5}) = 5$ y no hay subconjuntos independientes con, al menos, 6 vértices. En efecto, si V' tiene al menos 6 vértices, entonces $W_1 := V' \cap \{v_1, v_2, v_3\} \neq \emptyset$ y $W_2 := V' \cap \{v_5, v_6, v_7, v_8\} \neq \emptyset$, y si

tomamos $w_1 \in W_1$, $w_2 \in W_2$, entonces $\{w_1, w_2\} \in E$ y, por tanto, V' no sería un conjunto independiente.

Cotas para el número cromático de un grafo

En este capítulo estudiaremos la coloración de grafos. Una coloración de un grafo no dirigido consiste en una asignación de colores a los vértices del grafo de tal forma que no haya vértices adyacentes con el mismo color. Un problema central de la Teoría de Grafos es la determinación del menor número de colores necesarios para la coloración de un grafo, lo que se conoce como número cromático de un grafo. El objetivo de este capítulo es dar cotas tanto superiores como inferiores para este valor, en términos de los autovalores de su matriz de adyacencia. También, en este capítulo, estudiaremos una caracterización espectral de los grafos bipartitos, es decir, aquellos que se pueden colorear con 2 colores.

3.1. Número cromático de un grafo

Comenzamos definiendo el número cromático de un grafo, que es el parámetro objeto de estudio en este capítulo. Incluimos un ejemplo para su mejor comprensión.

Definición 3.1. *Dado un grafo simple no dirigido $G = (V, E)$ y $k \in \mathbb{Z}_{>0}$. Una k -coloración es una aplicación $f : V \rightarrow \{1, \dots, k\}$ que verifica que $f(u) \neq f(v)$ para todo $\{u, v\} \in E$. Decimos que un grafo G es k -coloreable si existe una k -coloración. Además, el número cromático de G se define como el menor número $k \in \mathbb{Z}_{>0}$ tal que G es k -coloreable, y se denota por $\chi(G)$.*

Ejemplo 3.2. En la Figura 3.1 se presenta una 3-coloración del grafo de Petersen, que denotamos por P , que es un grafo no dirigido con 10 vértices y 15 aristas. En efecto, basta hacer corresponder los enteros $\{1, 2, 3\}$ con los colores rosa, naranja y cian, y observar que los vértices adyacentes tienen asignados colores diferentes. Como el grafo admite una 3-coloración se tiene que $\chi(P) \leq 3$. Además, el grafo de Petersen tiene caminos cerrados de longitud 5 y, como veremos en el Lema 3.6, no puede ser 2-coloreable. En consecuencia se deduce que $\chi(P) = 3$.

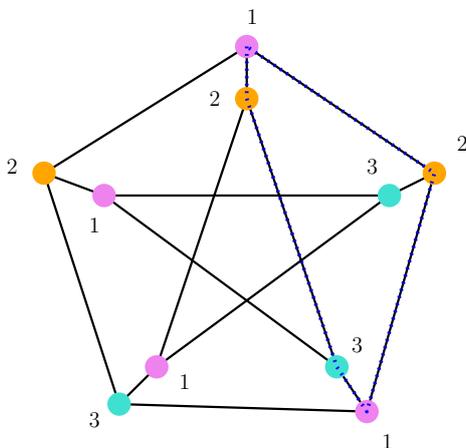


Figura 3.1. Una 3-coloración para el grafo de Petersen utilizando los colores rosa (1), naranja (2) y cian (3), y un camino cerrado de longitud 5.

3.2. Cotas para el número cromático en función del número de independencia de un grafo

Comenzamos esta sección mostrando que colorear un grafo es equivalente a particionar su conjunto de vértices en subconjuntos independientes. Así, se puede dar una cota del número cromático, que depende exclusivamente del número de vértices y del número de independencia del mismo, como veremos a continuación.

Proposición 3.3. *Un grafo G es k -coloreable si, y solo si, su conjunto de vértices se puede particionar en, a lo sumo, k subconjuntos independientes de vértices. Por tanto,*

$$\chi(G) \geq \frac{n}{\alpha(G)}$$

donde $\alpha(G)$ denota el número de independencia de G , es decir, el máximo número de vértices en un conjunto de vértices independientes.

Demostración. Sea $f : V \rightarrow \{1, 2, \dots, k\}$ una k -coloración de G . Definimos k subconjuntos de V dados por $V_i = \{v \in V : f(v) = i\}$ con $i \in \{1, 2, \dots, k\}$. El conjunto $P = \{V_i \mid V_i \neq \emptyset\}$ es una partición de V en, a lo sumo, k conjuntos independientes. En efecto, como f es aplicación, entonces P es una partición, además si tomamos $u, v \in V_i$, tenemos que $f(u) = f(v) = i$, y como f es coloración, tenemos que $\{u, v\} \notin E$. Luego V_i es independiente.

Por otro lado, sea $P = \{V_1, \dots, V_r\}$ una partición de V en r conjuntos independientes con $r \leq k$. Definimos $f : V \rightarrow \{1, 2, \dots, r\}$ como $f(u) = i$ si, y solo si, $u \in V_i$. Tenemos que f es aplicación porque P es una partición de V y, por tanto, todo $u \in V$ pertenece a un único V_i con $i \in \{1, \dots, r\}$. Veamos ahora que f es una r -coloración. Sean $u, v \in V$ tales que $f(u) = f(v) = i$. Como $u, v \in V_i$ es independiente, entonces $\{u, v\} \notin E$. Luego f es una r -coloración. \square

Sea ahora $P = \{V_1, \dots, V_r\}$ una partición de V en r conjuntos independientes con $r \leq \chi(G)$. Entonces

$$n = |V| = |\cup_{i=1}^r V_i| = \sum_{i=1}^r |v_i| \leq \sum_{i=1}^r \alpha(G) \leq r\alpha(G) \leq \chi(G)\alpha(G)$$

y, de aquí, se deduce que $\chi(G) \geq \frac{n}{\alpha(G)}$. Como consecuencia de las cotas para $\alpha(G)$ vistas en el capítulo anterior, podemos formular el siguiente resultado:

Corolario 3.4. *Sea G un grafo y sean $\lambda_1 \geq \dots \geq \lambda_n$ los autovalores de su matriz de adyacencia $A(G)$. Entonces se tienen las siguientes cotas para $\chi(G)$:*

$$\chi(G) \geq \begin{cases} \frac{n}{\min(|\{i \mid \lambda_i \leq 0\}|, |\{i \mid \lambda_i \geq 0\}|)} \\ \frac{\delta^2 - \lambda_1 \lambda_n}{-\lambda_1 \lambda_n} \end{cases}$$

En particular, si G es un grafo k -regular, entonces $\chi(G) \geq \frac{k - \lambda_n}{-\lambda_n}$.

Demostración. Aplicando la Proposición 3.3 el Teorema 2.24 tenemos la primera cota.

Del mismo modo y por el Teorema 2.28 tenemos la segunda cota.

Finalmente, aplicando nuevamente la Proposición anterior, la tercera cota viene del Teorema 2.27. □

3.3. Caracterización espectral de grafos bipartitos

Un grafo $G = (V, E)$ es bipartito si existe una partición de su conjunto de vértices en, a lo sumo, dos conjuntos independientes.

De la definición de grafo bipartito y de la Proposición 3.3 se deduce que G es bipartito si, y solo si, $\chi(G) \leq 2$.

En esta sección se va a dar una caracterización espectral de los grafos bipartitos. Más concretamente, vamos a demostrar que el grafo es bipartito si, y solo si, el espectro de la matriz de adyacencia es simétrico.

Para demostrar este enunciado usaremos los siguientes lemas:

Lema 3.5. *Sea $A(G)$ la matriz de adyacencia de G y sea $k \in \mathbb{Z}_{>0}$. Entonces, $A(G)^k = (a_{i,j}^{(k)})$ donde $a_{i,j}^{(k)}$ representa el número de caminos de v_i a v_j de longitud k .*

Demostración. Denotamos por $(b_{i,j}^{(k)})$ a las entradas de la matriz $A(G)^k$. Veamos que $a_{i,j}^{(k)} = b_{i,j}^{(k)}$ con $i, j \in \{1, \dots, n\}$. Vamos a demostrarlo por inducción sobre $k \in \mathbb{Z}_{>0}$.

Para $k = 1$ se da trivialmente, puesto que un camino de longitud uno es exactamente una arista.

Supongamos cierto el resultado para k y veamos ahora qué ocurre para $k + 1$. Observamos que hay tantos caminos de longitud $k + 1$ entre v_i y v_j como caminos de longitud k entre v_i y un vértice adyacente a v_j . Por lo tanto, tenemos que $a_{i,j}^{(k+1)}$ viene dado por:

$$a_{i,j}^{(k+1)} = \sum_{\{v_\ell, v_j\} \in E(G)} a_{i,\ell}^{(k)} = \sum_{\ell=1}^n a_{i,\ell}^{(k)} a_{\ell,j}^{(1)} = \sum_{i=1}^n b_{i,\ell}^{(k)} b_{\ell,j}^{(1)} = b_{i,j}^{(k+1)}$$

de donde se deduce el resultado. \square

Lema 3.6. *Un grafo G es bipartito si, y solo si, G no tiene caminos cerrados impares.*

Demostración. Sea $V = V_1 \cup V_2$ el conjunto de vértices del grafo bipartito G , con V_1 y V_2 independientes. Supongamos por reducción al absurdo que G tiene un ciclo $C = (v_0, \dots, v_k)$, $v_k = v_0$ de longitud k impar. Podemos suponer que $v_0 \in V_1$, entonces $v_1 \in V_2$, $v_2 \in V_2$ y, así, consecutivamente. En particular, $v_{k-1} \in V_1$, y la arista $\{v_0, v_{k-1}\} \in E$, lo cual es imposible porque V_1 es un conjunto independiente.

Para la otra implicación, supongamos que G es un grafo sin caminos cerrados impares. Supongamos además que G es conexo, en caso contrario, hacemos este razonamiento para cada una de sus componentes conexas. Fijamos un vértice v_0 y dividimos el conjunto de vértices de G , en dos conjuntos: en V_1 están los vértices $v_i \in V$ tales que el camino más corto desde v_0 a v_i tiene longitud par, y en V_2 los vértices $v_i \in V$ tales que el camino más corto de v_0 a v_i tiene longitud impar. Claramente V_1 y V_2 forma una partición de G (por ser G conexo), y $v_0 \in V_1$. Comprobemos ahora que no hay aristas entre los vértices de V_1 (ni entre los de V_2). Supongamos que hay una arista v_0 entre dos vértices v_1 y v_2 de V_1 , es decir, $\{v_1, v_2\} \in E$. Entonces, por el Lema 3.6 existen dos caminos pares $c_1 = (v_0, w_1, \dots, w_{2s} = v_1)$ de v_0 a v_1 y $c_2 = (v_0, x_1, \dots, x_{2r} = v_2)$ de v_0 a v_2 . Si $\{v_1, v_2\} \in E$, entonces podemos definir $c = (v_0, w_1, \dots, w_{2s} = v_1, v_2 = x_{2r}, \dots, x_1, v_0)$, camino cerrado de longitud $2s + 1 + 2r$, impar, lo que es un absurdo. De forma análoga se demuestra, entonces, que V_2 es un conjunto independiente. \square

Teorema 3.7. *Sea G un grafo, $A(G)$ su matriz de adyacencia y $\lambda_1 \geq \dots \geq \lambda_n$ los autovalores de $A(G)$. Entonces G es bipartito si, y solo si, el espectro de $A(G)$ es simétrico, i.e., $\lambda_i = -\lambda_{n+1-i}$, para cualquier $i \in \{1, \dots, n\}$.*

Demostración. Veamos primero que, si el espectro de $A(G)$ es simétrico, entonces G es bipartito. Sabemos que la matriz $A(G)$ con autovalores $\lambda_1, \dots, \lambda_n$

es diagonalizable, entonces, $A(G)^k$ también es diagonalizable con autovalores $\lambda_1^k, \dots, \lambda_n^k$. Además, como hipótesis tenemos que el espectro de nuestra matriz $A(G)$ es simétrico, por tanto $\sum_{i=1}^n \lambda_i^k = 0$, con $k \in \mathbb{Z}_{>0}$ impar. Así, $\text{tr}(A(G)^k) = \sum_{i=1}^n \lambda_i^k = 0$, con $k \in \mathbb{Z}_{>0}$ impar. De aquí, y sabiendo que, como nuestra matriz de adyacencia tiene entradas $\{0, 1\}$, las entradas de $A(G)^k$ serán no negativas, tenemos que $a_{i,i}^{(k)} = 0$, con $i \in \{1, \dots, n\}$ y k impar. Del Lema 3.5 se deduce que no hay caminos cerrados impares, y como consecuencia de esto y, por Lema 3.6, concluimos que G es bipartito.

Para la otra implicación queremos demostrar que si G es bipartito, entonces el espectro de $A(G)$ es simétrico. Supongamos entonces que G es bipartito. Entonces sabemos que existe una partición de su conjunto de vértices en dos subconjuntos independientes de vértices, $V_1 = \{v_1, \dots, v_r\}$ y $V_2 = \{v_{r+1}, \dots, v_n\}$. Por tanto, $A(G)$ es de la forma:

$$A(G) = \begin{pmatrix} 0 & B \\ B^T & 0 \end{pmatrix}$$

con $B \in M_{r \times (n-r)}(\mathbb{R})$.

Si suponemos ahora que $u = (x, y)^T$ es un autovector de $A(G)$ con autovalor λ , y con $x \in \mathbb{R}^r, y \in \mathbb{R}^{n-r}$. Entonces

$$\begin{pmatrix} 0 & B \\ B^T & 0 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = \lambda \begin{pmatrix} x \\ y \end{pmatrix}$$

desarrollando tenemos que $B^T x = \lambda y$ y $By = \lambda x$.

Así, teniendo en cuenta lo anterior, obtenemos el siguiente resultado:

$$\begin{pmatrix} 0 & B \\ B^T & 0 \end{pmatrix} \begin{pmatrix} x \\ -y \end{pmatrix} = \begin{pmatrix} -By \\ B^T x \end{pmatrix} = \begin{pmatrix} -\lambda x \\ \lambda y \end{pmatrix} = -\lambda \begin{pmatrix} x \\ -y \end{pmatrix}$$

o lo que es lo mismo, $(x, -y)^T$ es autovector asociado al autovalor $-\lambda$.

Ahora, si $\{(x_1, y_1), \dots, (x_k, y_k)\}$ es una base del autoespacio V_λ , entonces $\{(x_1, -y_1), \dots, (x_k, -y_k)\}$ es base del autoespacio $V_{-\lambda}$. Esto es así ya que se da la igualdad $\sum_{i=1}^k a_i(x_i, y_i) = 0$ si, y solo si, $(\sum_{i=1}^k a_i x_i, \sum_{i=1}^k a_i y_i) = 0$, y de aquí obtenemos que $(\sum_{i=1}^k a_i x_i, \sum_{i=1}^k a_i(-y_i)) = 0$. Como consecuencia tenemos que los autoespacios asociados al autovalor λ y $-\lambda$ tienen la misma dimensión. \square

El siguiente teorema se puede demostrar con una versión mejorada del Teorema 1.13 para grafos conexos. Para la demostración acudir a [8, pág. 455].

Teorema 3.8. *Sea G un grafo conexo. G es bipartito si, y solo si, $\lambda_1 = -\lambda_n$, donde λ_1 y λ_n representan el mayor y menor autovalor asociado a la matriz $A(G)$, respectivamente.*

3.4. Una cota superior para el número cromático

A continuación vamos a demostrar el Teorema de Wilf, que acota superiormente el número cromático de un grafo cualquiera G , obteniendo así un valor k tal que G es k -coloreable. En su demostración utilizaremos el Teorema del entrelazamiento (Teorema 2.9), el Teorema 2.17 y el siguiente Lema:

Lema 3.9. *Sea G un grafo y $v \in V(G)$. Si $\chi(G) = k$ y $\chi(G - \{v\}) < k$, entonces $\deg(v) \geq k - 1$.*

Demostración. Vamos a demostrarlo por reducción al absurdo. Sea G un grafo y $v \in V(G)$ tal que $\chi(G) = k$, $\chi(G - \{v\}) < k$ y supongamos que $\deg(v) \leq k - 2$. Como $\chi(G - \{v\}) < k$, entonces sabemos que $G - \{v\}$ es $(k - 1)$ -coloreable. Es decir, existe $p : V - \{v\} \rightarrow \{1, \dots, k - 1\}$ una $(k - 1)$ -coloración de $G - \{v\}$. Vamos a extender esta $(k - 1)$ -coloración de $G - \{v\}$ a una $(k - 1)$ -coloración de G .

Definimos $f : V \rightarrow \{1, \dots, k - 1\}$ donde para $x \in V - \{v\}$, $f(x) = p(x)$ si $x \neq v$. Denotamos por y_1, \dots, y_s los vértices adyacentes al vértice v , necesariamente $s = \deg(v) \leq k - 2$. Tomando $f(v)$ cualquier valor en el conjunto $\{1, \dots, k - 1\} - \{p(y_1), \dots, p(y_s)\}$, tenemos que f es una $(k - 1)$ -coloración. En efecto, $\{1, \dots, k - 1\} - \{p(y_1), \dots, p(y_s)\} \neq \emptyset$ dado que $\#\{1, \dots, k - 1\} = k - 1$ y $\#\{p(y_1), \dots, p(y_s)\} \leq k - 2$. Por tanto podemos concluir que $\chi(G) \leq k - 1$, que contradice la hipótesis de que $\chi(G) = k$. \square

Teorema 3.10 (Teorema de Wilf). *Sea $A(G)$ la matriz de adyacencia de un grafo G con autovalor principal λ_1 . Entonces $\chi(G) \leq 1 + \lambda_1$.*

Demostración. Procedemos a quitar vértices y todas las aristas adyacentes sin que el número cromático de G disminuya. Hacemos esto repetidas veces hasta obtener un *grafo crítico*, esto es, un grafo tal que $\chi(G - \{v\}) < \chi(G)$, para todo $v \in V(G)$. Al grafo obtenido tras este proceso lo llamaremos G_c y tiene $k \leq n$ vértices. Al ser G_c crítico, deducimos que todo vértice de G_c tiene grado mayor o igual a $\chi(G_c) - 1$ y, en particular $\delta(G_c) \geq \chi(G_c) - 1$. Sea $A(G_c) \in M_{k \times k}(\mathbb{R})$ la matriz de adyacencia del grafo G_c . Claramente, G_c es un subgrafo inducido de G y, por tanto, $A(G_c)$ es una submatriz principal de $A(G)$. Denotamos por μ_1 el autovalor dominante de G_c . Entonces, en virtud del Teorema 2.9, tenemos que:

$$\mu_1 \leq \lambda_1$$

Teniendo en cuenta el Lema 3.9 y que, por el Teorema 2.17, tenemos que $\mu_1 \geq \bar{\delta}(G_c) \geq \delta(G_c)$, obtenemos:

$$\lambda_1 \geq \mu_1 \geq \delta(G_c) \geq \chi(G_c) - 1 = \chi(G) - 1$$

\square

A continuación, vamos a usar el Teorema de Wilf para acotar el valor de $\chi(G)$ únicamente en términos de n (el número de vértices) y de m (el número de aristas). Para ello, incluimos varios resultados que buscan acotar λ_1 en función de n y m .

Lema 3.11. Sean $\alpha_1, \dots, \alpha_s \in \mathbb{R}$. Entonces

$$\left(\sum_{i=1}^s \alpha_i \right)^2 \leq s \left(\sum_{i=1}^s \alpha_i^2 \right)$$

Como consecuencia, si $\lambda_1, \dots, \lambda_n \in \mathbb{R}$ cumplen que $\lambda_1 + \dots + \lambda_n = 0$, entonces $\lambda_1^2 \leq \left(1 - \frac{1}{n}\right) \left(\sum_{i=1}^n \lambda_i^2\right)$.

Demostración. Usando que $(\alpha_i - \alpha_j)^2 \geq 0$, y por tanto $2\alpha_i\alpha_j \leq \alpha_i^2 + \alpha_j^2$, se tiene que:

$$\begin{aligned} (\alpha_1 + \dots + \alpha_s)^2 &= \sum_{i=1}^s \alpha_i^2 + \sum_{1 \leq i < j \leq s} 2\alpha_i\alpha_j \leq \sum_{i=1}^s \alpha_i^2 + \sum_{1 \leq i < j \leq s} (\alpha_i^2 + \alpha_j^2) \\ &\leq s(\alpha_1^2 + \dots + \alpha_s^2) \end{aligned}$$

Como $-\lambda_1 = \lambda_2 + \dots + \lambda_n$. De aquí tenemos que

$$\left(\sum_{i=2}^n \lambda_i \right)^2 = \lambda_1^2$$

Así,

$$\lambda_1^2 \leq (n-1) \sum_{i=2}^n \lambda_i^2$$

a su vez

$$\frac{1}{n} \lambda_1^2 \leq \frac{n-1}{n} \sum_{i=2}^n \lambda_i^2$$

sumando $\frac{n-1}{n} \lambda_1^2$ a ambos lados, obtenemos el resultado buscado:

$$\lambda_1^2 \leq \frac{n-1}{n} \sum_{i=1}^n \lambda_i^2$$

□

Corolario 3.12. Sea G un grafo, $A(G)$ su matriz de adyacencia y λ_1 su autovalor principal. Entonces $\lambda_1^2 \leq 2 \left(1 - \frac{1}{n}\right) m$.

Demostración. Sean $\lambda_1, \dots, \lambda_n$ los autovalores de $A(G)$. Como $\lambda_1 + \dots + \lambda_n = \text{tr}(A(G)) = 0$, usando el Lema 3.11 sabemos que

$$\lambda_1^2 \leq \left(1 - \frac{1}{n}\right) \sum_{i=1}^n \lambda_i^2$$

Veamos ahora que $\sum_{i=1}^n \lambda_i^2 = 2m$. Por el Lema 3.5 sabemos que $A(G)^2 = (a_{i,j}^{(2)})$, donde $a_{i,j}$ indica el número de caminos de longitud 2 que van de v_i a v_j , por lo que $a_{i,i}^{(2)} = \text{deg}(v_i)$, con $i \in \{1, \dots, n\}$.

Como los autovalores de A^2 son $\lambda_1^2, \dots, \lambda_n^2$, entonces

$$\sum_{i=1}^n \lambda_i^2 = \text{tr}(A^2) = \sum_{i=1}^n a_{i,i}^{(2)} = \sum_{i=1}^n \text{deg}(v_i) = 2m$$

obteniendo así el resultado esperado

$$\lambda_1^2 \leq 2 \left(1 - \frac{1}{n}\right) m$$

□

Utilizando estos resultados previos conseguimos una cota superior del número cromático de un grafo que depende exclusivamente del número de vértices y aristas del mismo.

Corolario 3.13. *Sea $G = (V, E)$ con n vértices y m aristas. Entonces*

$$\chi(G) \leq 1 + \sqrt{2 \left(1 - \frac{1}{n}\right) m}$$

Demostración. Por el Teorema de Wilf (Teorema 3.10) sabemos que $\chi(G) \leq 1 + \lambda_1$ donde λ_1 es el autovalor principal de la matriz $A(G)$ y, por el Corolario 3.12, sabemos que existe una cota para λ_1^2 que depende exclusivamente del número de aristas y de vértices del grafo G , y es $\lambda_1^2 \leq \left\{2 \left(1 - \frac{1}{n}\right) m\right\}$, con lo que $\lambda_1 \leq \sqrt{2 \left(1 - \frac{1}{n}\right) m}$. Ahora, sustituyendo en la cota del Teorema de Wilf tenemos el resultado buscado:

$$\chi(G) \leq 1 + \lambda_1 = 1 + \sqrt{2 \left(1 - \frac{1}{n}\right) m}.$$

□

Ejemplo 3.14. Volviendo al ejemplo del grafo de Petersen 3.1, teniendo en cuenta que su matriz de adyacencia es:

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

tenemos los autovalores ordenados de mayor a menor: $\lambda_1 = 3$, $\lambda_2 = \lambda_3 = \lambda_4 = \lambda_5 = 1$ (con multiplicidad 4), $\lambda_6 = \frac{\sqrt{5}-1}{2}$, $\lambda_7 = \frac{-\sqrt{5}-1}{2}$ y $\lambda_8 = \lambda_9 = \lambda_{10} = -2$ (con multiplicidad 3).

Partiendo del Corolario 3.4, una cota superior para el número cromático es, según el Teorema 2.24,

$$\chi(P) \geq \frac{5}{\alpha(P)} \geq \frac{5}{\min \{ |\{i \mid \lambda_i \leq 0\}|, |\{i \mid \lambda_i \geq 0\}| \}} = \frac{5}{4} = 1,25.$$

Ahora, como P es un grafo regular, podemos utilizar la cota del Teorema 2.27 y tenemos que $\chi(P) \geq \frac{(k-\lambda_n)}{-\lambda_n} = \frac{3-(-2)}{-(-2)} = 2,5$. Observemos que esta cota coincide con la cota del Teorema 2.28 por ser G un grafo regular.

Según el Teorema de Wilf (Teorema 3.10) tenemos que una cota superior para el número cromático de P es $\chi(P) \leq 1 + \lambda_1 = 1 + 3 = 4$.

Además, según el Corolario 3.13, otra cota superior es

$$\chi(P) \leq 1 + \sqrt{2 \left(1 - \frac{1}{n}\right) m} = 1 + \sqrt{2 \left(1 - \frac{1}{10}\right) 15} = 1 + 3\sqrt{3} \approx 6,1961\dots$$

Cabe observar que la cota del Corolario 3.13 se obtiene a partir de la de Wilf y, por tanto, el Teorema de Wilf siempre da una mejor estimación que la cota del Corolario 3.13. No obstante, el interés del Corolario 3.13 es que solo depende de los parámetros n y de m (número de aristas y de vértices de un grafo, respectivamente).

Estas cotas nos indican que el número cromático de nuestro grafo (que sabemos de antemano que es 3 por el Ejemplo 3.2) se encuentra entre 2,4 y 4.

Conclusiones

La Teoría Espectral de Grafos busca utilizar herramientas del Álgebra Lineal para el estudio de grafos, que son objetos de naturaleza combinatoria. En esta memoria hemos estudiado tres problemas de Teoría de Grafos mediante esta perspectiva. De esta forma, en el Capítulo 1 estudiamos cómo ordenar los nodos de una red según su importancia. En el Capítulo 2 estimamos el número de independencia de un grafo no dirigido utilizando el espectro de la matriz de adyacencia del grafo, es decir, en función de los autovalores de dicha matriz. En el Capítulo 3 hacemos un estudio similar, pero con el número cromático de un grafo no dirigido, es decir, el número mínimo de colores que podemos utilizar para colorear un grafo.

Este trabajo solo supone una introducción a la Teoría Espectral de Grafos y deja el camino hacia futuras líneas de trabajo como son las siguientes:

En el Capítulo 1, la elección de nodos importantes en una red depende de la definición de importancia que se elija. Sería interesante cambiar esta definición y estudiar si en la literatura existen respuestas con esta nueva definición que relacione la ordenación de nodos según la nueva importancia y la teoría espectral de matrices.

También sería interesante estudiar otras matrices asociadas al grafo. En este trabajo hemos utilizado dos matrices asociadas a un grafo: una matriz estocástica y su matriz de adyacencia. Es interesante preguntarse qué información del grafo aportan los autovalores de otras matrices asociadas al mismo, como puede ser la matriz laplaciana.

En el Capítulo 2 y 3 estudiamos diferentes cotas para dos parámetros del grafo: $\alpha(G)$ que representa el número de independencia de G y $\chi(G)$. Sin embargo, las cotas que se definen dependen de dos autovalores: el autovalor dominante y el autovalor más pequeño. Una posible línea de trabajo futuro sería estudiar otras cotas más precisas que involucren al resto de autovalores de la matriz. Así como otra vía natural de continuar este trabajo sería el estudio de algoritmos para el cálculo de $\alpha(G)$ y $\chi(G)$.

Además, hay otros parámetros como, por ejemplo, el número *clique* de un grafo, $\omega(G)$, del que se podría hacer un estudio similar. El número *clique* de un grafo G es el tamaño del mayor subgrafo completo de G .

Bibliografía

- [1] P. Fernández, *El secreto de Google y el Álgebra Lineal*. Sociedad Española de Matemática Aplicada 30, 2004, 115-141.
- [2] W. H. Haemers, *Interlacing eigenvalues and graphs*. Linear Algebra Appl. 226 \ 228, 1995, 593-616.
- [3] L. Merino, E. Santos, *Álgebra Lineal con métodos elementales*. Ediciones Paraninfo, 2006.
- [4] C. D. Meyer, *Matrix analysis and applied linear algebra. With solutions to problems*. Society of Industrial and Applied Mathematics, 2001.
- [5] R. M. Karp, *Reducibility among combinatorial problems*. In R. E. Miller; J. W. Thatcher; J. D. Bohlinger, Complexity of Computer, 1972.
- [6] L. Page, S. Brin, R. Motwani and T. Winograd, *The PageRank Citation Ranking: Bringing Order to the Web. Technical Report*. Stanford InfoLab, 1999.
- [7] D. A. Spielman. *Spectral and Algebraic Graph Theory*. National Science Foundation, 2019.
- [8] D. B. West, *Introduction to Graph Theory*. Second edition, University of Illinois-Urbana, 2002.
- [9] H. S. Wilf, *The eigenvalues of a graph and its chromatic number*. Journal of the London Mathematical Society 42, 1967, 330-332.

Spectral Graph Theory

Dióscora Andrea Esteve González

Facultad de Ciencias · Sección de Matemáticas

Universidad de La Laguna

alu0101107042@ull.edu.es

Abstract

SPECTRAL GRAPH THEORY studies graph properties through the eigenvalues of a matrix associated to the graph. In this document we study the following problems from this perspective:

- Classification of the nodes of a directed graph based on their relevance.
- Upper and lower bounds for the independence number and the chromatic number for non directed graphs.

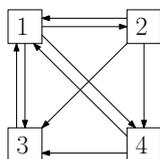
1. Network nodes classification based on their relevance: PageRank.

Given a directed graph, it is natural to ask which vertices are the most relevant. If we consider the importance of a vertex as the probability of being visited in a random walk in the graph, Brin and Page gave an elegant solution to this problem. More precisely, they associated a square matrix to the graph such that, under certain hypotheses, the importance of the vertices is given by any eigenvector associated to the eigenvalue $\lambda = 1$.

We consider the $n \times n$ matrix (where n is the number of vertices) $A = (a_{i,j})$, with

$$a_{i,j} = \begin{cases} \frac{1}{k}, & \text{if } i \in \{i_1, \dots, i_k\} \\ 0, & \text{otherwise.} \end{cases}$$

whenever the j -th vertex is connected the vertices i_1, \dots, i_k . The importance order is given by any vector $x \in \mathbb{R}_{>0}^n$ such that $Ax^T = x^T$.



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

Figure 1: Directed graph and its associated matrix.

In the example of Figure 1, any solution to $Ax^T = x^T$ is proportional to the vector $(x_1, x_2, x_3, x_4) = (9, 3, 6, 4)$. We observe that $x_1 > x_3 > x_4 > x_2$, thus, vertex 1 is the most important, followed by the 3, the 4 and, finally, the 2. The validity of this method is justified by the next result, which is a consequence of Perron-Frobenius Theorem:

Theorem 1: Let A be a matrix such that the addition of its rows is $(1, \dots, 1)$, and with positive entries. Then:

- $\lambda = 1$ is an eigenvalue with eigenspace $V_{\lambda=1} = \langle u \rangle$, where $u = (u_1, \dots, u_n) \in \mathbb{R}^n$, and $u_i > 0$ for all $i \in \{1, \dots, n\}$.
- $\lambda = 1$ is an eigenvalue with algebraic multiplicity 1. If γ is another eigenvalue from the matrix A , then $|\gamma| < 1$.

Also, from 2 one can derive an efficient method to approximate the importances vector based on the iterated power method. The resulting algorithm for vertex sorting is known as PageRank and it was included in the first version of Google.

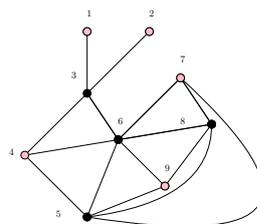
2. Bounds on the independence number.

Definition 1: For a non directed graph $G = (V, E)$, with $V = \{v_1, \dots, v_n\}$, its adjacency matrix is defined as $A(G) = (a_{i,j}) \in M_{n \times n}(\mathbb{R})$, with

$$a_{i,j} = \begin{cases} 1 & \text{if } \{v_i, v_j\} \in E, \\ 0 & \text{otherwise.} \end{cases}$$

Note that $A(G)$ is symmetric and, thus, their eigenvalues are all real numbers. We denote $\lambda_1 \geq \dots \geq \lambda_n$ the ordered sequence of (possibly repeated) eigenvalues of $A(G)$.

Definition 2: The independence number of a graph G is the maximum number of pairwise nonadjacent vertices of G .



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

Figure 2: A graph G and its adjacency matrix $A(G)$.

For the graph in Figure 2, we have an independent vertex set in pink colour, $V' = \{v_1, v_2, v_4, v_7, v_9\}$. Thus, for this example, $\alpha(G) \geq 5$. In fact, one can check that $\alpha(G) = 5$.

Theorem 2:

$$\alpha(G) \leq n \frac{-\lambda_1 \lambda_n}{\delta^2 - \lambda_1 \lambda_n},$$

with δ the smallest degree of the vertices of G .

3. Bounds on the chromatic number.

Definition 3: A graph coloring is an assignment of labels (colors) to its vertices such that vertices are assigned different colors. The minimum number of colors needed to color a graph is denoted by $\chi(G)$.

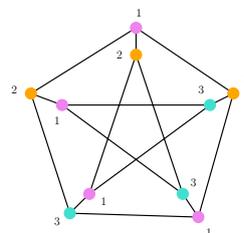


Figure 3: A coloring of Petersen graph P using pink (1), orange (2) and cyan (3) colors. Thus, $\chi(P) \leq 3$. In fact, it can be proved that $\chi(P) = 3$.

Theorem 3: Let G be a non directed graph with n vertices, m edges and minimum degree δ , then

$$\frac{\delta^2 - \lambda_1 \lambda_n}{-\lambda_1 \lambda_n} \leq \frac{n}{\alpha(G)} \leq \chi(G) \leq 1 + \lambda_1 \leq 1 + \sqrt{2 \left(1 - \frac{1}{n}\right) m}.$$

References

- W. H. Haemers, *Interlacing eigenvalues and graphs*. Linear Algebra Appl. 226 \ 228 (1995), 593-616.
- L. Page, S. Brin, R. Motwani and T. Winograd, *The PageRank Citation Ranking: Bringing Order to the Web*. Technical Report. Stanford InfoLab (1999).
- H. S. Wilf, *The eigenvalues of a graph and its chromatic number*. Journal of the London Mathematical Society 42 (1967), 330-332.