Matrices positivas y aplicaciones

María Isabel García Planas
Profesora Titular de Universidad

Primera edición: Septiembre 2008
Editora: la autora
© Mª Isabel García Planas
ISBN: 978-84-612-6101-7
Depósito legal: B-41533-2008

Está rigurosamente prohibido, sin autorización escrita del titular del copyright, bajo sanciones establecidas por la ley, la reproducción total o parcial de esta obra por cualquier procedimiento, incluido la reprografía y el tratamiento informático

A mis hijos

El olvido de las Matemáticas perjudica a todo el conocimiento, ya que el que las ignora no puede conocer las otras ciencias ni las cosas de este mundo.

Roger Bacon (Inglaterra 1214-1294)

Presentación

El presente libro proporciona un tratamiento comprensivo de la teoría de matrices positivas y, más generalmente, de la teoría de matrices no negativas. Las matrices cuadradas de este tipo aparecen en una gran variedad de problemas como pueden ser estudios de procesos estocáscticos, cadenas de Markov, modelos económicos, teoría de la señal, entre otros.

La teoría desarrollada aquí es una parte del álgebra lineal que recientemente ha adquirido un nuevo auge debido a su alto grado de aplicabilidad, siendo así que algunos de los principales resultados han sido obtenidos en las últimas décadas.

Este libro está estructurado en dos partes, de manera que la primera incluye todos los resultados de álgebra matricial y lineal necesarios para hacer la obra autocontenida. La segunda parte está dedicada propiamente el estudio de las matrices no negativas y sus aplicaciones.

Este texto está pensado, pues, en general para aquellos estudiantes que necesitan utilizar este tipo de matrices, y da, por lo tanto, todos los conceptos básicos necesarios para que pueda, más adelante, comprender un texto científico en que éstas sean utilizadas, y en particular como apoyo para aquellos que cursen la asignatura "Introducción a la Teoría de Matricess Positivas"que se imparte en la ESTSEIB-UPC.

Aquellos estudiantes que, por su formación matemática, dominen los conceptos básicos de Álgebra lineal y estén familiarizados con el Álgebra matricial pueden directamente abordar el estudio de la segunda parte.
No obstante, puesto que no necesariamente ocurre así, hemos incluido una primera parte en que se abordan algunos tópicos de Álgebra lineal y ma-

tricial. Puede esta primera parte servir, además, como recordatorio para los estudiantes que previamente hayan estudiado esta materia, así como para referencia en caso de que algún concepto lo desconozca.

La Autora

Índice general

1. Matrices **11**

1.1. Notaciones y definiciones 11
1.2. Determinantes 19
1.3. Inversa de una matriz 21
1.4. Inversa generalizada de Moore-Penrose. 22
1.5. Ejercicios Resueltos 23
1.6. Ejercicios Propuestos 25

2. Aplicaciones lineales de $\mathbb{R}^n$ en $\mathbb{R}^m$ **27**

2.1. El espacio vectorial $\mathbb{R}^n$. Subespacios vectoriales 27
2.2. Dependencia lineal y rango de matrices 28
2.3. Bases. Matriz de cambio de base 30
2.4. Aplicación lineal y matriz asociada 37
2.5. Determinante de un endomorfismo de $\mathbb{R}^n$ 41
2.6. Subespacios invariantes por un endomorfismo 43
2.7. Ejercicios Resueltos 44
2.8. Ejercicios Propuestos 48

3. Propiedades espectrales **51**

3.1. Polinomio característico de una matriz 51
3.2. Valores y vectores propios de una matriz 52
3.3. Radio espectral 58
3.4. Valores singulares 59
3.5. Normas y cotas para valores propios 60
3.6. Ejercicios Resueltos 62
3.7. Ejercicios Propuestos 65

4. Matrices no negativas **69**
4.1. Definiciones y propiedades 69
4.2. Matrices irreducibles 72
4.3. Matrices primitivas 80
4.4. Matrices estocásticas 82
4.5. Matrices totalmente no negativas 85
4.6. Ejercicios Resueltos 86
4.7. Ejercicios Propuestos 88

5. Cadenas de Markov finitas **91**
5.1. Cadenas de Markov a variable discreta 91
5.2. Ejercicios Resueltos 99
5.3. Ejercicios Propuestos 102

6. Modelos de Leontief en economía **105**
6.1. Modelo abierto 108
6.2. Modelo cerrado 110
6.3. Ejercicio Resueltos 113
6.4. Ejercicios Propuestos 117

7. Métodos iterativos para sistemas lineales **119**
7.1. Métodos iterativos básicos 120
7.2. No negatividad y convergencia 124
7.3. Ejercicios 129

Bibliografía **131**

Capítulo 1

Matrices

1.1. Notaciones y definiciones

Trataremos tan sólo las matrices definidas sobre un cuerpo conmutativo K que será, o bien el cuerpo real $\mathbb{R}$, o bien el cuerpo complejo $\mathbb{C}$ (es decir, cuyos coeficientes pertenecen a $\mathbb{R}$ o a $\mathbb{C}$) si bien las matrices se pueden definir sobre cualquier cuerpo, o incluso sobre anillos.

Definición 1.1.1. Llamaremos matriz de orden $n \times m$ a coeficientes en el cuerpo K a un conjunto de $n \cdot m$ elementos del cuerpo distribuidos en n filas y m columnas y que notaremos por

$$A = \begin{pmatrix} a_{11} & \dots & a_{1m} \\ \vdots & & \vdots \\ a_{n1} & \dots & a_{nm} \end{pmatrix}, \qquad a_{ij} \in K.$$

El elemento a_{ij} está en la fila i y en la columna j. Algunas veces notaremos una matriz simplemente por $A = (a_{ij})$.

Ejemplo 1.1.1.

$$A_1 = \begin{pmatrix} 1{,}01 & 2 & 3{,}2 \\ -1{,}1 & 3 & 2{,}5 \end{pmatrix}, \qquad A_2 = \begin{pmatrix} \sqrt{2} & \sqrt{3} \\ 1 & 0 \\ \sqrt{3} & \sqrt{2} \end{pmatrix}$$

A_1 es una matriz de orden 2×3 y A_2 es una matriz de orden 3×2.

Las matrices de orden $1 \times m$ (es decir, con una sola fila) reciben el nombre de vectores fila y las matrices de orden $n \times 1$ (es decir, con una sola columna) reciben el nombre de vectores columna.

Ejemplo 1.1.2. Las matrices

$$\begin{pmatrix} 1 & 2 & 1 \end{pmatrix}, \qquad \begin{pmatrix} 4 \\ 3 \\ 2 \end{pmatrix}$$

son un vector fila y un vector columna, respectivamente.

Denotaremos por $M_{n\times m}(K)$ el conjunto de las matrices de orden $n \times m$ a coeficientes en el cuerpo K. En el caso particular en que $n = m$ denotaremos este conjunto simplemente por $M_n(K)$.

Operaciones con matrices

Suma

Definición 1.1.2. Dadas dos matrices $A = (a_{ij})$ y $B = (b_{ij})$, con $A, B \in M_{n\times m}(K)$ (esto es, del mismo orden) definimos la suma de estas dos matrices como la matriz

$$C = (c_{ij}) = (a_{ij} + b_{ij}).$$

Proposición 1.1.1. *En el conjunto $M_{n\times m}(K)$, la operación suma verifica las siguientes propiedades.*
a) Asociativa: $(A + B) + C = A + (B + C)$.
b) Conmutativa: $A + B = B + A$.
c) Existencia de elemento neutro, que es la matriz 0 (cuyos elementos son todos nulos), ya que $A + 0 = A$ para cualquier matriz $A \in M_{n\times m}(K)$. Esta matriz se llama matriz nula.
d) Existencia de elemento simétrico: dada una matriz $A = (a_{ij}) \in M_{n\times m}(K)$ existe $-A$ de forma que $A+(-A) = 0$ (en efecto, basta tomar $-A = (-a_{ij})$). Se dice entonces que el conjunto $M_{n\times m}(K)$ tiene estructura de grupo abeliano.

Notación. Dadas dos matrices $A, B \in M_{n\times m}(K)$, la operación $A + (-B)$ la notaremos simplemente por $A - B$.

Producto

Definición 1.1.3. Dadas dos matrices $A \in M_{n\times m}(K)$, $B \in M_{m\times p}(K)$, esto es, $A = (a_{ij})$ y $B = (b_{ij})$ de órdenes respectivos $n \times m$ y $m \times p$, definimos el producto de estas dos matrices por

$$C = (c_{ij}) = \left(\sum_{k=1}^{m} a_{ik} b_{kj} \right)$$

Proposición 1.1.2. *La operación producto verifica las siguientes propiedades.*
a) Asociativa: $(AB)C = A(BC)$.
b) Distributiva respecto de la suma: $A(B + C) = AB + AC$, $(A + B)C = AC + BC$.
c) En el caso $n = m$, *existe elemento unidad, que es la matriz* $I_n = (\delta_{ij})$ *con* $\delta_{ij} = 0$ *si* $i \neq j$ *y* $\delta_{ii} = 1$, *ya que* $AI_n = I_n A = A$ *para cualquier matriz* $A \in M_n(K)$. *La matriz* $I_n \in M_n(K)$ *recibe el nombre de matriz identidad de orden* n. *A veces, si no hay posibilidad de confusión, se denota simplemente por* I.
En el caso $n = m$ *y, puesto que dadas dos matrices cualesquiera siempre se puede obtener su producto, que cumple las propiedades anteriores, además de las propiedades mencionadas anteriormente con la operación suma, se dice que* $M_n(K)$ *es un anillo con unidad.*

Observación 1. La operación producto no es conmutativa. En realidad, si $A \in M_{n\times m}(K)$ y $B \in M_{m\times p}(K)$ podemos efectuar el producto AB pero el producto BA puede efectuarse sólo si $n = p$. En el caso de dos matrices cuadradas A, B del mismo orden se pueden efectuar los productos AB y BA, pero las matrices así obtenidas no tienen por qué coincidir. En efecto, podemos considerar las matrices

$$A = \begin{pmatrix} 1 & 1 & 1 \end{pmatrix} \in M_{1\times 3}(\mathbb{R}), \quad B = \begin{pmatrix} 1 & 0 \\ 0 & 1 \\ 1 & 4 \end{pmatrix} \in M_{3\times 2}(\mathbb{R})$$

Podemos hacer el producto AB,

$$AB = \begin{pmatrix} 1 & 1 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 \\ 0 & 1 \\ 1 & 4 \end{pmatrix} = \begin{pmatrix} 2 & 5 \end{pmatrix}$$

pero sin embargo no es posible efectuar el producto BA.
En el caso de dos matrices cuadradas, consideremos, por ejemplo,

$$A = \begin{pmatrix} 1 & 1 \\ 2 & 1 \end{pmatrix}, \qquad B = \begin{pmatrix} 1 & 0 \\ 0 & 2 \end{pmatrix}$$

En este caso existen AB y BA pero estos productos son dos matrices distintas:

$$AB = \begin{pmatrix} 1 & 1 \\ 2 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 \\ 0 & 2 \end{pmatrix} = \begin{pmatrix} 1 & 2 \\ 2 & 2 \end{pmatrix}$$
$$BA = \begin{pmatrix} 1 & 0 \\ 0 & 2 \end{pmatrix} \begin{pmatrix} 1 & 1 \\ 2 & 1 \end{pmatrix} = \begin{pmatrix} 1 & 1 \\ 4 & 2 \end{pmatrix}$$

Observación 2. Es posible que, sin que ni A ni B sean matrices nulas, el producto AB o BA (o ambos, en el caso en que tengan sentido) sean la matriz nula. Decimos entonces que las matrices A y B son divisores de cero.
Como ejemplos, podemos considerar los siguientes.

$$\begin{pmatrix} 1 & 1 & 0 \\ 0 & 0 & 0 \end{pmatrix} \begin{pmatrix} 1 & -2 \\ -1 & 2 \\ 3 & -5 \end{pmatrix} = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}$$

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

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

Observamos también que, sin ser $A = B$, puede ocurrir $AC = BC$:

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

y $CA = CB$:

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

Así pues, no es válida la "ley de simplificación".

Producto por un escalar

Dada una matriz $A = (a_{ij}) \in M_{n\times m}(K)$ y un elemento cualquiera del cuerpo $\lambda \in K$ definimos el producto del escalar λ por la matriz A del siguiente modo:

$$\lambda(a_{ij}) = (\lambda a_{ij})$$

Observación 3. Sea $A \in M_{n\times m}(K)$. Entonces

$$\lambda A = (\lambda I_n)A = A(\lambda I_m)$$

Las matrices de la forma λI_n reciben el nombre de *matrices escalares*.

Proposición 1.1.3. *La operación producto por escalares de K verifica las siguientes propiedades.*
a) Asociativa: $\lambda(\mu A) = (\lambda\mu)A$.
b) Distributiva respecto de la suma de matrices: $\lambda(A + B) = \lambda A + \lambda B$.
c) Distributiva respecto de la suma de escalares: $(\lambda + \mu)A = \lambda A + \mu A$.
d) $1 \cdot A = A$.

El conjunto $M_{m\times n}(K)$ con las operaciones suma y producto por escalares de K verificando estas propiedades se dice que tiene estructura de *espacio vectorial sobre K*.

Algunos tipos especiales de matrices

Definición 1.1.4. Llamaremos *matriz triangular superior* a una matriz $A = (a_{ij})$ tal que $a_{ij} = 0$ para todo $i > j$.
Llamaremos *matriz triangular inferior* a una matriz $A = (a_{ij})$ tal que $a_{ij} = 0$ para todo $i < j$.
Llamaremos *matriz diagonal* a una matriz $A = (a_{ij})$ tal que $a_{ij} = 0$ para todo $i \neq j$.
Llamaremos *matriz estrictamente triangular superior* a una matriz $A = (a_{ij})$ tal que $a_{ij} = 0$ para todo $i \geq j$.
Llamaremos *matriz estrictamente triangular inferior* a una matriz $A = (a_{ij})$ tal que $a_{ij} = 0$ para todo $i \leq j$.

Ejemplo 1.1.3. Consideremos las matrices

$$A_1 = \begin{pmatrix} 0{,}3 & 2{,}2 & 4{,}2 \\ 0 & 1{,}01 & 4{,}9 \\ 0 & 0 & 1 \end{pmatrix}, \; A_2 = \begin{pmatrix} 1 & 0 & 0 \\ 2{,}4 & 9{,}1 & 0 \\ 0{,}9 & 2 & 8{,}1 \end{pmatrix}, \; A_3 = \begin{pmatrix} 0{,}3 & 0 & 0 \\ 0 & 1{,}01 & 0 \\ 0 & 0 & 1 \end{pmatrix},$$

$$A_4 = \begin{pmatrix} 0 & 0{,}01 & 2{,}1 \\ 0 & 0 & 2 \\ 0 & 0 & 0 \end{pmatrix}, \quad A_5 = \begin{pmatrix} 0 & 0 & 0 \\ 2{,}4 & 0 & 0 \\ 0{,}9 & 2 & 0 \end{pmatrix}$$

La matriz A_1 es triangular superior, la matriz A_2 es triangular inferior, la matriz A_3 es diagonal, la matriz A_4 es estrictamente triangular superior y la matriz A_5 es estrictamente triangular inferior.
Las matrices escalares son casos especiales de matrices diagonales.

Trasposición

Definición 1.1.5. Dada una matriz $A = (a_{ij}) \in M_{n\times m}(K)$ llamaremos *matriz traspuesta* de A (y la notaremos por A^t) a la matriz definida de la forma

$$A^t = (a_{ji}) \in M_{m\times n}(K)$$

Es decir, A^t es la matriz que se obtiene a partir de A cambiando filas por columnas.

Ejemplo 1.1.4. Dada la matriz

$$A = \begin{pmatrix} 1 & 1 & 0 \\ 0 & 1 & 2 \end{pmatrix} \in M_{2\times 3}(K),$$

su traspuesta es la matriz

$$A^t = \begin{pmatrix} 1 & 0 \\ 1 & 1 \\ 0 & 2 \end{pmatrix} \in M_{3\times 2}(K)$$

Ejemplo 1.1.5. Si v es un vector fila $\begin{pmatrix} v_1 & \dots & v_n \end{pmatrix} \in M_{1\times m}(K)$ su traspuesta es un vector columna, $v^t = \begin{pmatrix} v_1 \\ \vdots \\ v_n \end{pmatrix} \in M_{m\times 1}(K)$. Recíprocamente, la traspuesta de un vector columna es un vector fila.

Las siguientes propiedades son frecuentemente utilizadas.

Proposición 1.1.4. *Se verifican las siguientes igualdades.*
1.- $(A^t)^t = A$
2.- $(\lambda A)^t = \lambda A^t$
3.- $(A + B)^t = A^t + B^t$
4.- $(AB)^t = B^t A^t$

Vamos ahora a restringirnos al caso en que las matrices son cuadradas, es decir, $n = m$. En este caso A y A^t son ambas matrices cuadradas del mismo orden y tiene sentido preguntarse si coinciden, o bien si se relacionan de algún modo peculiar.

Definición 1.1.6. Dada una matriz $A \in M_n(\mathbb{R})$ diremos que es simétrica si y sólo si

$$A = A^t.$$

Ejemplo 1.1.6. La matriz

$$A = \begin{pmatrix} 1 & \sqrt{2} & 1 \\ \sqrt{2} & 0 & -1 \\ 1 & -1 & 2 \end{pmatrix}$$

es simétrica.

En general dada una matriz $A = (a_{ij}) \in M_n(\mathbb{C})$, se define la matriz A^* como la matriz traspuesta conjugada de la matriz A. Esto es,

$$A^* = (b_{ij}), \quad \text{con } b_{ij} = \overline{a}_{ji}$$

Ejemplo 1.1.7. Dada la matriz

$$A = \begin{pmatrix} 1+i & 2i \\ 2-i & 1 \end{pmatrix},$$

entonces

$$A^* = \begin{pmatrix} 1-i & 2+i \\ -2i & 1 \end{pmatrix}$$

Observamos que si A es a coeficientes reales $A^* = A^t$.

Definición 1.1.7. Dada una matriz $A \in M_n(\mathbb{R})$ diremos que es antisimétrica si y sólo si

$$A = -A^t$$

Ejemplo 1.1.8. La matriz

$$A = \begin{pmatrix} 0 & 1 & -0{,}5 \\ -1 & 0 & 2 \\ 0{,}5 & -2 & 0 \end{pmatrix}$$

es antisimétrica.

Proposición 1.1.5. *Dada una matriz $A \in M_n(\mathbb{R})$ cualquiera, existen una matriz $S \in M_n(\mathbb{R})$ simétrica y una matriz $T \in M_n(\mathbb{R})$ antisimétrica tales que*

$$A = S + T.$$

Demostración. Consideremos

$$S = \frac{1}{2}(A + A^t)$$
$$T = \frac{1}{2}(A - A^t)$$

Es fácil ver que S es simétrica, T es antisimétrica y $A = S + T$. □

También es fácil comprobar la afirmación de la siguiente proposición.

Proposición 1.1.6. *Sea $A \in M_{n\times m}(\mathbb{R})$ una matriz cualquiera. Entonces $S = A^tA \in M_m(\mathbb{R})$ y $R = AA^t \in M_n(\mathbb{R})$ son matrices simétricas.*

Para el caso de matrices reales

Definición 1.1.8. Sea $A \in M_n(\mathbb{R})$. Se dice que A es ortogonal cuando

$$AA^t = A^tA = I_n.$$

Ejemplo 1.1.9. Las matrices

$$A_1 = \begin{pmatrix} \frac{1}{3} & \frac{2}{3} & -\frac{2}{3} \\ \frac{2}{3} & -\frac{2}{3} & -\frac{1}{3} \\ \frac{2}{3} & \frac{1}{3} & \frac{2}{3} \end{pmatrix}, \quad A_2 = \begin{pmatrix} 0 & -1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & -1 \end{pmatrix},$$

son ortogonales.

En general, sea $A \in M_n(\mathbb{C})$ se dice que A es unitaria cuando

$$AA^* = A^*A = I_n$$

Ejemplo 1.1.10. Las matrices

$$A_1 = \begin{pmatrix} i & \\ & i \end{pmatrix}, \quad A_2 = \begin{pmatrix} 0 & i \\ -i & 0 \end{pmatrix}$$

son unitarias.

1.2. Determinantes

Definición 1.2.1. Dada una matriz cuadrada $A \in M_n(K)$ llamaremos determinante de $A = (a_{ij})$ (y lo notaremos por $\det A$) al escalar de K

$$\det A = \sum \varepsilon(h) a_{1h_1} a_{2h_2} \dots a_{nh_n}$$

donde el sumatorio se extiende a todas las permutaciones $h : (h_1, h_2, \dots, h_n)$ del conjunto $\{1, 2, \dots, n\}$ y $\varepsilon(h)$ vale 1 o -1 según si la permutación es par o impar.

Ejemplo 1.2.1. (a) Sea $A = (a) \in M_1(K)$. Entonces $\det A = a$.
(b) Sea $A = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix} \in M_2(K)$. Entonces $\det A = a_{11}a_{22} - a_{21}a_{12}$.

Propiedades.
1.- $\det A^t = \det A$.
2.- Si una matriz B se obtiene a partir A mediante intercambio de dos filas o dos columnas consecutivas, entonces $\det B = -\det A$.
3.- Si A es una matriz que tiene dos filas o dos columnas idénticas, $\det A = 0$.
4.- Si una matriz B se obtiene a partir de A añadiendo a una fila (o columna) de A una combinación lineal de las otras filas (o columnas) de A, entonces $\det B = \det A$.
5.- Si $A, B \in M_n(K)$, $\det AB = \det A \cdot \det B$.
6.- El determinante de una matriz triangular superior (o triangular inferior, o diagonal) es igual al producto de los elementos que están sobre la diagonal de la matriz.

De estas propiedades básicas, consecuencia de la definición de determinante, se deducen otras.
7.- Si una fila (o columna) de la matriz A es combinación lineal de las otras filas (o columnas) de A, entonces $\det A = 0$.
8.- Si A es una matriz estrictamente triangular superior (o estrictamente triangular inferior), $\det A = 0$.
9.- Para todo n, $\det I_n = 1$.

Menores y rango de una matriz

Definición 1.2.2. Dada una matriz $A \in M_{n\times m}(K)$, se denomina *menor de orden* r al determinante de una submatriz de A de orden r formada con los elementos que pertenecen simultáneamente a r filas y r columnas fijadas de la matriz A.

Definición 1.2.3. Dada una matriz $A \in M_{n\times m}(K)$ llamaremos *rango* de la matriz al orden del mayor menor no nulo de la matriz.

Definición 1.2.4. Una matriz $A \in M_n(K)$ se dice que es *regular* si su determinante es distinto de cero y, en caso contrario, se dice que es *singular*.

Ejemplo 1.2.2. Sea $A \in M_n(K)$ regular. Entonces $\text{rango}\, A = n$.

Adjunta de una matriz

Definición 1.2.5. Dada $A = (a_{ij}) \in M_n(K)$, se define el adjunto del término a_{ij} (que notaremos por α_{ij}) al determinante de la matriz que resulta al suprimir en A la fila y la columna en las que se encuentra el elemento a_{ij} (esto es, la fila i y la columna j) multiplicado por $(-1)^{i+j}$.

Definición 1.2.6. Dada una matriz $A = (a_{ij}) \in M_n(K)$ definimos la matriz adjunta de A, $\text{Adj}(A)$, como la matriz cuyos elementos son los adjuntos de los elementos de la matriz A,

$$\text{Adj}(A) = (\alpha_{ij}).$$

Ejemplo 1.2.3. Sea A la matriz

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

Los adjuntos de cada elemento son

$$\begin{array}{ccc} \alpha_{11} = -4 & \alpha_{12} = 0 & \alpha_{13} = 4 \\ \alpha_{21} = 0 & \alpha_{22} = 0 & \alpha_{23} = -2 \\ \alpha_{31} = 4 & \alpha_{32} = -2 & \alpha_{33} = -4 \end{array}$$

por lo que la matriz adjunta de A es

$$\mathrm{Adj}(A) = \begin{pmatrix} -4 & 0 & 4 \\ 0 & 0 & -2 \\ 4 & -2 & -4 \end{pmatrix}$$

1.3. Inversa de una matriz

Definición 1.3.1. Sea $A \in M_{n\times m}(K)$. Si existe una matriz $B \in M_{m\times n}(K)$ tal que

$$AB = I_n \in M_n(K)$$

se dice que la matriz B es una *inversa por la derecha* de A. Si existe una matriz $C \in M_{m\times n}(K)$ tal que

$$CA = I_m \in M_m(K)$$

se dice que la matriz C es una *inversa por la izquierda* de A.
En caso de que A sea una matriz cuadrada, si existe inversa por la derecha existe también por la izquierda y ambas coinciden. En dicho caso se dice que la matriz A es *inversible* y a una tal matriz se la denomina la *inversa* de A, que notaremos por A^{-1} (una tal matriz es única).

Proposición 1.3.1. *Las matrices regulares son las únicas matrices inversibles.*

Proposición 1.3.2. *Si $A \in M_n(K)$ es una matriz regular, entonces se verifican las siguientes propiedades.*
1.- $(A^{-1})^{-1} = A$.
2.- Si B es también una matriz regular, $(AB)^{-1} = B^{-1}A^{-1}$.
3.- Sea $\lambda \in K$, $\lambda \neq 0$. Entonces $(\lambda A)^{-1} = \dfrac{1}{\lambda}A^{-1}$.
4.- $(A^t)^{-1} = (A^{-1})^t$.
5.- $\det(A^{-1}) = \dfrac{1}{\det A}$.

Cálculo de la inversa.

La siguiente proposición proporciona un método para obtener la inversa de una matriz.

Proposición 1.3.3. *La inversa de una matriz regular $A \in M_n(K)$ es:*

$$A^{-1} = \frac{1}{\det A}(\mathrm{Adj}(A))^t.$$

1.4. Inversa generalizada de Moore-Penrose.

Definición 1.4.1. Dada una matriz $A \in M_{n\times m}(\mathbb{R})$, llamaremos matriz *inversa de Moore-Penrose* de A a una matriz de orden $m \times n$ (que notaremos por A^+) que verifica
1.- $AA^+A = A$
2.- $A^+AA^+ = A^+$
3.- $(AA^+)^t = AA^+$
4.- $(A^+A)^t = A^+A$

Proposición 1.4.1. *Sea $A \in M_{n\times m}(\mathbb{R})$ una matriz cualquiera. Entonces existe una única matriz $A^+ \in M_{m\times n}(\mathbb{R})$ que cumple las cuatro condiciones anteriores.*

Demostración. Sea $r = \mathrm{rango}\, A$. Entonces existen $P \in M_n(\mathbb{R})$ y $Q \in M_m(\mathbb{R})$ matrices inversibles tales que

$$A = P \begin{pmatrix} I_r & 0 \\ 0 & 0 \end{pmatrix} Q$$

Particionando las matrices P y Q según la partición de $\begin{pmatrix} I_r & 0 \\ 0 & 0 \end{pmatrix}$ tenemos que

$$A = \begin{pmatrix} P_{11} & P_{12} \end{pmatrix} \begin{pmatrix} I_r & 0 \\ 0 & 0 \end{pmatrix} \begin{pmatrix} Q_{11} \\ Q_{21} \end{pmatrix} = P_{11}Q_{11}$$

Observamos que P_{11} y Q_{11} son matrices de rango r. Luego $P_{11}^tP_{11}$ y $Q_{11}Q_{11}^t$ son matrices inversibles de orden r. Es fácil comprobar que la matriz

$$Q_{11}^t(Q_{11}Q_{11}^t)^{-1}(P_{11}^tP_{11})^{-1}P_{11}^t$$

verifica las cuatro condiciones dadas en la definición. □

Observación. Si $A \in M_n(\mathbf{R})$ es una matriz inversible, entonces $A^+ = A^{-1}$.

En el caso en que la matriz A tiene rango máximo es fácil obtener la matriz A^+, según se ve en el siguiente resultado.

Proposición 1.4.2. *Sea $A \in M_{n\times m}(\mathbb{R})$ una matriz que tiene rango máximo. Entonces,*
(a) si $n \geq m$, $A^+ = (A^tA)^{-1}A^t$.
(b) si $n \leq m$, $A^+ = A^t(AA^t)^{-1}$.

Ejemplo 1.4.1. Sea

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

La inversa de Moore-Penrose de la matriz A viene dada por

$$A^+ = \frac{1}{15}\begin{pmatrix} 1 & 1 & 1 \\ 2 & 2 & 2 \end{pmatrix}$$

lo que se comprueba fácilmente, puesto que la matriz A^+ verifica las cuatro condiciones dadas en la definición.

1.5. Ejercicios Resueltos

1.- Descomponer la matriz

$$A = \begin{pmatrix} 5 & 2 & 1 \\ 3 & 1 & 4 \\ 2 & 1 & 3 \end{pmatrix}$$

en suma de una matriz S simétrica y una matriz T antisimétrica.
Solución:

$$S = \frac{1}{2}\left(\begin{pmatrix} 5 & 2 & 1 \\ 3 & 1 & 4 \\ 2 & 1 & 3 \end{pmatrix} + \begin{pmatrix} 5 & 3 & 2 \\ 2 & 1 & 1 \\ 1 & 4 & 3 \end{pmatrix}\right) = \frac{1}{2}\begin{pmatrix} 10 & 5 & 3 \\ 5 & 2 & 5 \\ 3 & 5 & 6 \end{pmatrix}$$

$$T = \frac{1}{2}\left(\begin{pmatrix} 5 & 2 & 1 \\ 3 & 1 & 4 \\ 2 & 1 & 3 \end{pmatrix} - \begin{pmatrix} 5 & 3 & 2 \\ 2 & 1 & 1 \\ 1 & 4 & 3 \end{pmatrix}\right) = \frac{1}{2}\begin{pmatrix} 0 & -1 & -1 \\ 1 & 0 & 3 \\ 1 & -3 & 0 \end{pmatrix}$$

2.- Calcular el determinante de la matriz siguiente.

$$A = \begin{pmatrix} 1 & 0 & 3 \\ 4 & -1 & 7 \\ 7 & 2 & -1 \end{pmatrix}$$

Solución:

$$\det A = \begin{vmatrix} 1 & 0 & 3 \\ 4 & -1 & 7 \\ 7 & 2 & -1 \end{vmatrix} =$$
$$1 \cdot (-1) \cdot (-1) + 4 \cdot 2 \cdot 3 + 0 \cdot 7 \cdot 7 - 7 \cdot (-1) \cdot 3 - 2 \cdot 7 \cdot 1 - 4 \cdot 0 \cdot (-1) =$$
$$1 + 24 + 0 - (-21) - 14 - 0 = 32$$

3.- Determinar la inversa de la matriz del ejercicio anterior.
Solución:

$$A^{-1} = \frac{1}{32}\begin{pmatrix} \begin{vmatrix} -1 & 2 \\ 7 & -1 \end{vmatrix} & -\begin{vmatrix} 0 & 2 \\ 3 & -1 \end{vmatrix} & \begin{vmatrix} 0 & -1 \\ 3 & 7 \end{vmatrix} \\ -\begin{vmatrix} 4 & 7 \\ 7 & -1 \end{vmatrix} & \begin{vmatrix} 1 & 7 \\ 3 & -1 \end{vmatrix} & -\begin{vmatrix} 1 & 4 \\ 3 & 7 \end{vmatrix} \\ \begin{vmatrix} 4 & 7 \\ -1 & 2 \end{vmatrix} & -\begin{vmatrix} 1 & 7 \\ 0 & 2 \end{vmatrix} & \begin{vmatrix} 1 & 4 \\ 0 & -1 \end{vmatrix} \end{pmatrix} = \frac{1}{32}\begin{pmatrix} -13 & 6 & 3 \\ 53 & -22 & 5 \\ 15 & -2 & -1 \end{pmatrix}$$

4.- Determinar la inversa de Moore-Penrose de la matriz

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

Solución:
Observamos que rango $A = 2 = m < 3 = n$, entonces

$$A^t A = \begin{pmatrix} 1 & 1 & 1 \\ 1 & 2 & 3 \end{pmatrix}\begin{pmatrix} 1 & 1 \\ 1 & 2 \\ 1 & 3 \end{pmatrix} = \begin{pmatrix} 3 & 6 \\ 6 & 14 \end{pmatrix}$$

$$(A^tA)^{-1} = \frac{1}{6}\begin{pmatrix} 14 & -6 \\ -6 & 3 \end{pmatrix}$$

$$A^+ = (A^tA)^{-1}A^t = \frac{1}{6}\begin{pmatrix} 14 & -6 \\ -6 & 3 \end{pmatrix}\begin{pmatrix} 1 & 1 & 1 \\ 1 & 2 & 3 \end{pmatrix} = \frac{1}{6}\begin{pmatrix} 8 & 2 & -4 \\ -3 & 0 & 3 \end{pmatrix}$$

1.6. Ejercicios Propuestos

1.- Sea $A \in M_n(K)$ una matriz estrictamente triangular superior (inferior). Probar que $A^n = 0$.

2.- Calcular

$$\begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}^2.$$

Deducir de aquí A^n, siendo A la matriz

$$A = \begin{pmatrix} a & b \\ 0 & a \end{pmatrix}, \quad a, b \in K \text{ cualesquiera.}$$

3.- Sea A una matriz triangular superior (inferior) y regular. Probar que A^{-1} es triangular superior (inferior).

4.- Calcular el determinante de las siguientes matrices

$$A = \begin{pmatrix} 1 & 10^{-2} & 0 & 0 & 0 \\ 10^{-1} & 1 & 10^{-2} & 0 & 0 \\ 0 & 10^{-1} & 1 & 10^{-2} & 0 \\ 0 & 0 & 10^{-1} & 1 & 10^{-2} \\ 0 & 0 & 0 & 10^{-1} & 1 \end{pmatrix}, \quad B = \begin{pmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 1 & 2 & 3 & 4 \end{pmatrix}.$$

5.- Calcular el rango de las siguientes matrices

$$A = \begin{pmatrix} 1 & 1 & 2 & 1 \\ 0 & 1 & 0 & 1 \\ 1 & 2 & 2 & 3 \end{pmatrix}, \quad B = \begin{pmatrix} 0{,}1 & 0{,}2 & 0{,}3 & 0{,}4 \\ 0{,}2 & 0{,}3 & 0{,}4 & 0{,}5 \\ 0{,}3 & 0{,}4 & 0{,}5 & 0{,}6 \end{pmatrix}$$

6.- Calcular las matrices adjuntas de las siguientes matrices.

$$A = \begin{pmatrix} 0 & 0 & 1 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}, \qquad B = \begin{pmatrix} 0{,}3 & 0{,}2 & 0{,}1 & 0 \\ 0 & 0{,}3 & 0{,}2 & 0{,}1 \\ 0 & 0 & 0{,}3 & 0{,}2 \\ 0 & 0 & 0 & 0{,}3 \end{pmatrix}.$$

7.- Calcular las inversas de las matrices del ejercicio anterior.
8.- Calcular la inversa de Moore-Penrose de las siguientes matrices.

$$A_1 = \begin{pmatrix} 1 & 1{,}2 \\ -2{,}5 & 0 \\ 3{,}1 & 0{,}4 \end{pmatrix}, \quad A_2 = \begin{pmatrix} 10^{-6} & 0{,}5 \\ 6{,}10^{-6} & 0{,}4 \\ 5{,}10^{-6} & 0{,}3 \end{pmatrix}$$

9.- Probar que $(A^+)^+ = A$.
10.- Dadas las matrices

$$A_1 = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}, \qquad A_2 = \begin{pmatrix} 1 & 0 \\ 0 & 0 \\ 0 & 0 \end{pmatrix}$$

Probar que $A_1^+ = A_2$ y $A_2^+ = A_1$.

Capítulo 2

Aplicaciones lineales de $\mathbb{R}^n$ en $\mathbb{R}^m$

2.1. El espacio vectorial $\mathbb{R}^n$. Subespacios vectoriales

Denotamos por $\mathbb{R}^n$ al conjunto de vectores $(x_1, x_2, \ldots, x_n)$ con $x_1, x_2, \ldots, x_n \in \mathbb{R}$. La estructura de espacio vectorial de $\mathbb{R}^n$ es la generalización a "dimensión superior"de las propiedades y operaciones habituales con los vectores en $\mathbb{R}^2$ y $\mathbb{R}^3$.
En $\mathbb{R}^n$ tenemos definida una operación interna, la *suma de vectores*, que se corresponde con la suma de matrices de orden $1 \times n$, y que cumple las siguientes propiedades:

1. $(x + y) + z = x + (y + z)\ \forall\, x, y, z \in \mathbb{R}^n$ (propiedad asociativa)

2. $\forall\, x \in \mathbb{R}^n$, $x + 0 = 0 + x = x$ (existencia de elemento neutro)

3. $\forall\, x \in \mathbb{R}^n$ existe $-x \in \mathbb{R}^n$ tal que $x + (-x) = (-x) + x = 0$ (existencia de elemento simétrico)

4. $x + y = y + x\ \forall\, x, y \in \mathbb{R}^n$ (propiedad conmutativa)

y una operación externa, el *producto por escalares*, que cumple las siguientes propiedades:

5. $\lambda(x+y) = \lambda x+\lambda y \; \forall \, \lambda \in \mathbb{R}, \forall \, x, y \in \mathbb{R}^n$ (propiedad distributiva respecto de la suma de vectores)

6. $(\lambda + \mu)x = \lambda x + \mu x \; \forall \, \lambda, \mu \in \mathbb{R}, \; \forall \, x \in \mathbb{R}^n$ (propiedad distributiva respecto de la suma de escalares)

7. $\lambda(\mu x) = (\lambda\mu)x \; \forall \, \lambda, \mu \in \mathbb{R}, \; \forall \, x \in \mathbb{R}^n$ (propiedad asociativa)

8. $1x = x \; \forall \, x \in \mathbb{R}^n$

Algunos subconjuntos de $\mathbb{R}^n$ (aunque no todos) tienen estructura de espacio vectorial. Éstos son los llamados subespacios vectoriales.

Definición 2.1.1. Un subconjunto (no vacío) F de $\mathbb{R}^n$ se dice que es un *subespacio vectorial* de $\mathbb{R}^n$ si se cumplen las dos siguientes propiedades:

1. $\forall x, y \in F, \; x + y \in F$.

2. $\forall x \in F, \; \forall \lambda \in \mathbb{R}, \; \lambda x \in F$.

Ejemplo 2.1.1. (a) Los conjuntos $\{0\}$ y $\mathbb{R}^n$ son, evidentemente, subespacios vectoriales, que reciben el nombre de subespacios vectoriales triviales.
(b) $F = \{(x_1, x_2, x_3) \in \mathbb{R}^3 \,|\, x_1 + 2x_2 = 0\}$ es un subespacio vectorial. En efecto. Sean $(x_1, x_2, x_3), (y_1, y_2, y_3)$ dos vectores cualesquiera de F. Entonces el vector $(x_1, x_2, x_3)+(y_1, y_2, y_3) = (x_1+y_1, x_2+y_2, x_3+y_3)$ pertenece también a F:

$$x_1 + y_1 + 2(x_2 + y_2) = (x_1 + 2x_2) + (y_1 + 2y_2) = 0 + 0 = 0$$

Por otra parte, si (x_1, x_2, x_3) pertenece a F y λ es un escalar, el vector $\lambda(x_1, x_2, x_3) = (\lambda x_1, \lambda x_2, \lambda x_3)$ pertenece también a F:

$$\lambda x_1 + 2(\lambda x_2) = \lambda(x_1 + 2x_2) = 0$$

2.2. Dependencia lineal y rango de matrices

Definición 2.2.1. Se llama *combinación lineal de los vectores* $u_1, \ldots, u_m$ *de* $\mathbb{R}^n$ a todo vector de la forma $x = \lambda_1 u_1 + \cdots + \lambda_m u_m$ siendo $\lambda_1, \ldots, \lambda_m \in \mathbb{R}$.

Ejemplo 2.2.1. Los siguientes vectores $(2,-3,-1,-1),(-3,-6,15,-17)$, $(5,11,-14,-26)$ de $\mathbb{R}^4$ son combinación lineal de los vectores u_1=$(1,-1,0,0)$, $u_2=(0,1,1,1)$, $u_3=(1,2,-3,5)$. En efecto:

$$(2,-3,-1,-1)=2u_1-u_2$$

$$(-3,-6,15,-17)=u_1+3u_2-4u_3$$

$$(5,11,-14,26)=u_2+5u_3$$

Definición 2.2.2. Un conjunto de vectores $\{u_1,\ldots,u_m\}$ de $\mathbb{R}^n$ se dice que es *linealmente dependiente* si existe una combinación lineal de estos vectores, con no todos los coeficientes nulos, que es el vector 0:

$$\lambda_1u_1+\cdots+\lambda_mu_m=0$$

Equivalentemente, cuando al menos uno de estos vectores es combinación lineal de los restantes. En caso contrario diremos que los vectores son *linealmente independientes.*

Ejemplo 2.2.2. (i) Los vectores $(2,-3,-1,-1),(1,-1,0,0),(0,1,1,1)$ y $(1,2,-3,5)$ de $\mathbb{R}^4$ son linealmente dependientes, ya que

$$(2,-3,-1,-1)-2(1,-1,0,0)+(0,1,1,1)+0(1,2,-3,5)=0$$

(ii) Los vectores $(1,1,1),(1,2,0),(1,0,0)$ de R^3 son linealmente independientes. En efecto. Supongamos que existe una combinación lineal de estos vectores que es el vector cero,

$$\lambda_1(1,1,1)+\lambda_2(1,2,0)+\lambda_3(1,0,0)=(0,0,0)$$

Entonces los coeficientes de la combinación lineal anterior han de cumplir:

$$\lambda_1+\lambda_2+\lambda_3=0,\quad \lambda_1+2\lambda_2=0,\quad \lambda_1=0$$

lo que sólo es posible si

$$\lambda_1=\lambda_2=\lambda_3=0$$

Si tenemos un conjunto finito de vectores de $\mathbb{R}^n$, $\{u_1,\ldots,u_m\}$, para saber si estos vectores son, o no, linealmente independientes, basta con calcular el rango de la matriz A cuyas columnas coinciden con los vectores columna formados por las coordenadas de los vectores $u_1,\ldots,u_m$.

Proposición 2.2.1. *Con las notaciones anteriores, si el rango de la matriz A es igual a m, entonces los vectores dados son linealmente independientes. Si el rango de la matriz A es menor que m, estos vectores son linealmente dependientes.*

Demostración. Supongamos que existe una combinación lineal de los vectores dados que es el vector cero,

$$\lambda_1 u_1 + \cdots + \lambda_m u_m = 0$$

Basta observar que la igualdad anterior es equivalente al siguiente sistema lineal de ecuaciones:

$$A \begin{pmatrix} \lambda_1 \\ \vdots \\ \lambda_m \end{pmatrix} = 0$$

Como es bien sabido, la existencia de soluciones no triviales del sistema anterior equivale a que el rango de la matriz A sea estrictamente menor que m. □

Ejemplo 2.2.3. En el caso de las familias de vectores del ejemplo anterior, habría bastado con calcular los rangos de las matrices correspondientes a cada caso:

(i) $\operatorname{rg} \begin{pmatrix} 2 & 1 & 0 & 1 \\ -3 & -1 & 1 & 2 \\ -1 & 0 & 1 & -3 \\ -1 & 0 & 1 & 5 \end{pmatrix} = 3$

(ii) $\operatorname{rg} \begin{pmatrix} 1 & 1 & 1 \\ 1 & 2 & 0 \\ 1 & 0 & 0 \end{pmatrix} = 3$

2.3. Bases. Matriz de cambio de base

Sea F un subespacio vectorial de $\mathbb{R}^n$.

Definición 2.3.1. Se dice que un conjunto de vectores $\{u_1, \ldots, u_m\}$ es un *sistema de generadores* de F cuando todo vector de F puede escribirse como combinación lineal de los vectores $u_1, \ldots, u_m$ (de forma no necesariamente única).

Observación 1. Es obvio que todo subespacio vectorial de $\mathbb{R}^n$ admite un sistema de generadores.

Ejemplo 2.3.1. Consideremos los conjuntos de vectores de $\mathbb{R}^5$ siguientes:
(a) $\{(1,1,1,0,0),(2,1,3,0,0),(-1,3,1,0,0)\}$
(b) $\{(1,1,1,0,0),(2,1,3,0,0),(1,0,2,0,0)\}$
Como el lector puede comprobar fácilmente, el conjunto de vectores de $\mathbb{R}^5$

$$F = \{(x_1, x_2, x_3, x_4, x_5) \,|x_4 = x_5 = 0\}$$

es un subespacio vectorial de $\mathbb{R}^5$, del cual el conjunto (a) es un sistema de generadores, pero no así el conjunto (b).
Esto puede probarse del siguiente modo.
Un vector cualquiera de F es de la forma $(x_1, x_2, x_3, 0, 0)$ para unos ciertos escalares x_1, x_2, x_3 de $\mathbb{R}$. Podemos encontrar coeficientes $\lambda_1, \lambda_2, \lambda_3$ de $\mathbb{R}$ de modo que

$$(x_1, x_2, x_3, 0, 0) = \lambda_1(1,1,1,0,0) + \lambda_2(2,1,3,0,0) + \lambda_3(-1,3,1,0,0) \qquad (*)$$

En efecto, basta con tomar

$$\lambda_1 = \frac{8}{6}x_1 + \frac{5}{6}x_2 - \frac{7}{6}x_3$$

$$\lambda_2 = -\frac{1}{3}x_1 - \frac{1}{3}x_2 + \frac{2}{3}x_3$$

$$\lambda_3 = -\frac{1}{3}x_1 + \frac{1}{6}x_2 + \frac{1}{6}x_3$$

Estos valores se obtienen a partir de la resolución del sistema asociado a la igualdad $(*)$.
En cambio, si planteamos un sistema análogo, ahora con la igualdad

$$(x_1, x_2, x_3, 0, 0) = \lambda_1(1,1,1,0,0) + \lambda_2(2,1,3,0,0) + \lambda_3(1,0,2,0,0)$$

observamos que este sistema no tiene solución, por lo que deducimos que el conjunto de vectores (b) no es un sistema de generadores de F.
Habría, para este último caso, bastado con encontrar un vector concreto que no fuera posible escribir como combinación lineal de los tres vectores dados, por ejemplo, podríamos considerar el vector $(0,0,1,0,0)$.

Definición 2.3.2. Se llama *base* de F a todo conjunto de vectores que sean linealmente independientes y formen además un sistema de generadores de F.

Observación 2. Siempre es posible, para todo subespacio vectorial, encontrar una base. Basta con suprimir, de un sistema de generadores del subespacio, todos aquellos vectores que sean combinación lineal de los anteriores. Se enuncia a continuación el Teorema de Steinitz que tiene una especial importancia por las múltiples consecuencias que de él se derivan.

Teorema 2.3.1 (Steinitz)**.** *Sea $\{u_1, \ldots, u_m\}$ una base de un subespacio F de $\mathbb{R}^n$ y sean $v_1, \ldots, v_r$ r vectores de F linealmente independientes. Entonces se pueden substituir r vectores de la base $\{u_1, \ldots, u_m\}$ de F por $v_1, \ldots, v_r$ obteniéndose de este modo una nueva base de F.*

Ejemplo 2.3.2. Sea $\{(1,0,0,0),(0,1,0,0),(0,0,1,0),(0,0,0,1)\}$ la base canónica de $\mathbb{R}^4$ y sean $(1,2,-2,1),(2,1,-1,-2)$ una colección de vectores independientes de $\mathbb{R}^4$. Entonces

$$\{(1,2,-2,1),(2,1,-1,-2),(0,0,1,0),(0,0,0,1)\}$$

es una nueva base de $\mathbb{R}^4$.
Hemos cambiado los dos primeros vectores de la base por los nuevos vectores y hemos obtenido una nueva base. Hubieramos podido cambiar los dos últimos y también obendríamos una uueva base.

$$\{(1,0,0,0),(0,1,0,0),(1,2,-2,1),(2,1,-1,-2)\}$$

es una base de $\mathbb{R}^4$. Sin embargo no podemos cambiar el segundo y tercero por los nuevos pues en este caso los vectores son dependientes

$$\{(1,0,0,0),(1,2,-2,1),(2,1,-1,-2),(0,0,0,1)\}$$

es una familia de vectores dependientes.

Como corolario del Teorema de Steinitz se obtiene, entre otros, el siguiente resultado, cuya demostración no se incluye, pero que el lector interesado puede encontrar en la mayor parte de textos de Álgebra lineal.

Proposición 2.3.1. *Toda base de F consta del mismo número de vectores.*

Este resultado da sentido a la siguiente definición.

Definición 2.3.3. Se denomina *dimensión de* F al número de vectores de que consta un conjunto cualquiera de vectores que es base de F.

Observación 3. Es fácil comprobar que el conjunto

$$\{(1,0,\ldots,0),(0,1,0,\ldots,0),\ldots,(0,\ldots,0,1)\}$$

es un sistema de generadores de $\mathbb{R}^n$ y que los vectores que lo forman son linealmente independientes. Así pues, estos vectores forman una base de $\mathbb{R}^n$. Esta base recibe el nombre de *base natural* de $\mathbb{R}^n$ (también base canónica o base ordinaria).
En particular, deducimos que la dimensión de $\mathbb{R}^n$ es igual a n.
Algunos de los enunciados de la siguiente observación son especialmente útiles a la hora de trabajar con bases de subespacios vectoriales.

Observación 4.
(1) Un conjunto de vectores $u = \{u_1,\ldots,u_n\}$ es una *base* del espacio vectorial $\mathbb{R}^n$ si estos vectores forman un conjunto linealmente independiente.
(2) Un conjunto de vectores $u = \{u_1,\ldots,u_n\}$ es una *base* del espacio vectorial $\mathbb{R}^n$ si estos vectores forman un sistema de generadores de $\mathbb{R}^n$.
(3) Más en general: si conocemos la dimensión del subespacio vectorial F, que denotamos por m, un conjunto de m vectores linealmente independientes de F es una base de F. También un sistema de generadores de F formado por m vectores constituye una base de F.
(4) La dimensión de F coincide con el número máximo de vectores linealmente independientes de F que podemos seleccionar.
(5) La dimensión de F coincide con el mínimo número de vectores que ha de contener cualquier sistema de generadores de F.
(6) La dimensión de todo subespacio vectorial F de $\mathbb{R}^n$ es menor o igual que n, e igual a n sólo en el caso en que F coincida con todo $\mathbb{R}^n$. Y si F, G son dos subespacios vectoriales de $\mathbb{R}^n$ tales que F está contenido en G, entonces la dimensión de F es menor o igual que la dimensión de G, y sus respectivas dimensiones coinciden sólamente en el caso en que coincidan ambos subespacios.

Convendremos en considerar que la dimensión del subespacio vectorial trivial $\{0\}$ es igual a cero.

Si el conjunto de vectores $u = \{u_1, \ldots, u_m\}$ es una base del subespacio vectorial F, todo vector de F puede escribirse, de forma única, como combinación lineal de estos vectores (en el mismo orden en que han sido dados).

Definición 2.3.4. A los coeficientes de esta combinación lineal se les llama *coordenadas del vector* en la base dada.

Ejemplo 2.3.3. Consideremos la base de $\mathbb{R}^4$ dada por los vectores

$$u_1 = (1,1,1,0),\ u_2 = (1,0,-1,-1),\ u_3 = (0,1,0,1),\ u_4 = (0,0,0,1)$$

Las coordenadas del vector $v = (0,2,2,3)$ en esta base son $(1,-1,1,1)$, puesto que

$$v = 1\cdot(1,1,1,0) + (-1)\cdot(1,0,-1,-1) + 1\cdot(0,1,0,1) + 1\cdot(0,0,0,1)$$

Observemos que si cambiamos el orden de los vectores de la base considerada obtenemos una nueva base (son generadores y linealmente independientes) y las coordenadas de un vector cualquiera cambian. Así, los vectores

$$u_3 = (0,1,0,1), u_1 = (1,1,1,0), u_2 = (1,0,-1,-1), u_4 = (0,0,0,1)$$

forman también una base de $\mathbb{R}^4$ y las coordenadas del vector $v = (0,2,2,3)$ en esta nueva base son $(1,1,-1,1)$, puesto que

$$v = 1\cdot(0,1,0,1) + 1\cdot(1,1,1,0) + (-1)\cdot(1,0,-1,-1) + 1\cdot(0,0,0,1)$$

Si $v = (v_1, \ldots, v_n)$ es una base de $\mathbb{R}^n$ y $y_1, \ldots, y_n$ son las coordenadas de x en la base v, es decir,

$$x = y_1 v_1 + \cdots + y_n v_n$$

(sabemos ya que estas coordenadas están unívocamente determinadas) entonces la relación entre las coordenadas de x en la base u y sus coordenadas en esta base viene dada por:

$$\begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} = S \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix} \Longleftrightarrow \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix} = S^{-1} \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}$$

siendo S la matriz cuyas columnas son las coordenadas de los vectores de la base v en la base u.

Definición 2.3.5. Esta matriz se llama *matriz del cambio de base* (de la base v a la base u).

Las matrices de cambio de base son siempre matrices inversibles.

Ejemplo 2.3.4. Consideremos las bases de $\mathbb{R}^3$

$$u = \{u_1 = (1,0,0), u_2 = (0,1,0), u_3 = (0,0,1)\}$$

$$v = \{v_1 = (1,1,1), v_2 = (1,1,0), v_3 = (1,0,0)\}$$

Puesto que

$$\begin{aligned} v_1 &= 1 \cdot u_1 + 1 \cdot u_2 + 1 \cdot u_3 \\ v_2 &= 1 \cdot u_1 + 1 \cdot u_2 + 0 \cdot u_3 \\ v_3 &= 1 \cdot u_1 + 0 \cdot u_2 + 0 \cdot u_3 \end{aligned}$$

la matriz del cambio de base, de la base v a la base u es:

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

Dado el vector x cuyas coordenadas, en la base v son $(2,1,3)$, sus coordenadas, en la base u, son:

$$S \begin{pmatrix} 2 \\ 1 \\ 3 \end{pmatrix} = \begin{pmatrix} 1 & 1 & 1 \\ 1 & 1 & 0 \\ 1 & 0 & 0 \end{pmatrix} \begin{pmatrix} 2 \\ 1 \\ 3 \end{pmatrix} = \begin{pmatrix} 6 \\ 3 \\ 2 \end{pmatrix}$$

lo que se comprueba fácilmente, pues $x = 6 \cdot v_1 + 3 \cdot v_2 + 2 \cdot v_3$.

Definición 2.3.6. Una matriz de cambio de base, obtenida por reordenación de los vectores de la base inicial, recibe el nombre de *matriz permutación*.

Ejemplo 2.3.5. Así, si consideramos las bases de $\mathbb{R}^4$ en el Ejemplo (2.3.2), la matriz del cambio de base, de la base $\{u_1, u_2, u_3, u_4\}$ respecto a la base $\{u_3, u_1, u_2, u_4\}$, es la matriz

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

Dados dos subespacios vectoriales F y G de $\mathbb{R}^n$, su intersección (es decir, el conjunto formado por los vectores que pertenecen a ambos conjuntos) es a su vez un subespacio vectorial de $\mathbb{R}^n$. En cambio, la unión de estos dos conjuntos no es un subespacio vectorial.
Consideremos, por ejemplo, los siguientes subespacios vectoriales de $\mathbb{R}^2$ (es muy fácil comprobar que, en efecto, se trata de subespacios):

$$F = \{(x_1, x_2) \in \mathbb{R}^2 \,|\, x_1 = 0\}$$

$$G = \{(x_1, x_2) \in \mathbb{R}^2 \,|\, x_2 = 0\}$$

El vector $(1, 1)$ es suma de los vectores $(1, 0) \in F$ y $(0, 1) \in G$ y, sin embargo, no pertenece a $F \cup G$.
Esto lleva a la introducción de un subespacio vectorial que es importante en muchos ámbitos, el denominado *subespacio vectorial suma* de los subespacios F y G, que viene definido del modo siguiente:

$$F + G = \{x \in \mathbb{R}^n \,|\, x = x_F + x_G \text{ para todos los vectores } x_F \in F \text{ y } x_G \in G\}$$

Este concepto puede generalizarse al caso de considerar un número arbitrario de subespacios vectoriales (incluso infinito, aunque aquí no lo consideraremos).

Definición 2.3.7. El conjunto

$$F_1 + F_2 + \cdots + F_r = \left\{ x \in \mathbb{R}^n \;\middle|\; \begin{array}{l} x = x_1 + x_2 + \cdots + x_r \text{ para todos los} \\ \text{vectores } x_1 \in F_1,\ x_2 \in F_2, \ldots, x_r \in F_r \end{array} \right\}$$

es un subespacio vectorial de $\mathbb{R}^n$, llamado el *subespacio vectorial suma* de los subespacios $F_1, F_2, \ldots, F_r$.

Cuando existe una única posibilidad de escribir un vector x de $F_1 + F_2 + \cdots + F_r$ de la forma $x = x_1 + x_2 + \cdots + x_r$ con $x_1 \in F_1,\ x_2 \in F_2, \ldots, x_r \in F_r$ se dice que la suma $F_1 + F_2 + \cdots + F_r$ es *directa*.
En el caso particular en que sólo tenemos dos subespacios vectoriales, se tiene la siguiente caracterización de suma directa:

$$F_1 + F_2 \text{ es directa si, y sólo si, } F_1 \cap F_2 = \{0\}$$

Ejemplo 2.3.6. Consideremos los subespacios vectoriales

$$\begin{aligned} F_1 &= \{(x,y,z) \in \mathbb{R}^3 \,|\, x+y+z=0\} \\ F_2 &= \{(x,y,z) \in \mathbb{R}^3 \,|\, x-y+z=0\} \end{aligned}$$

La suma de los subespacios F_1, F_2 no es directa, ya que

$$\begin{aligned} F_1 \cap F_2 &= \{(x,y,z) \in \mathbb{R}^3 \,|\, x+y+z=x-y+z=0\} \\ &= \{(x,y,z) \in \mathbb{R}^3 \,|\, x+z=y=0\} = [(1,0,-1)] \end{aligned}$$

Consideremos ahora los subespacios vectoriales

$$\begin{aligned} G_1 &= \{(x,y,z) \in \mathbb{R}^3 \,|\, x+y+z=0\} \\ G_2 &= \{(x,y,z) \in \mathbb{R}^3 \,|\, x-y=x-z=0\} \end{aligned}$$

La suma de los subespacios G_1, G_2 es directa, ya que

$$\begin{aligned} G_1 \cap G_2 &= \{(x,y,z) \in \mathbb{R}^3 \,|\, x+y+z=x-y=x-z=0\} \\ &= \{(0,0,0)\} \end{aligned}$$

2.4. Aplicación lineal y matriz asociada

Definición 2.4.1. Una *aplicación* entre dos conjuntos A y B es una asociación de elementos del conjunto B a los elementos del conjunto A, de modo que a todo elemento del conjunto A le corresponda un solo elemento del conjunto B, que denominamos *imagen* del elemento dado (también se dice que el elemento de A que se corresponde con un elemento determinado de B es su *antiimagen*).
Cuando no hay dos elementos distintos de A con la misma imagen, se dice que la aplicación es *inyectiva*.
Cuando todo elemento de B tiene alguna antiimagen, se dice que la aplicación es *exhaustiva*.
Cuando se cumplen ambas condiciones, inyectiva y exhaustiva (es decir, todo elemento de B tiene una única antiimagen) se dice que la aplicación es *biyectiva*. En este caso, tiene sentido considerar la aplicación inversa de la dada, como aquella aplicación del conjunto B al conjunto A que a cada elemento de B le hace corresponder su antiimagen.

Pasamos ahora a introducir el concepto de aplicación lineal, de gran importancia en Álgebra lineal.

Definición 2.4.2. Una *aplicación lineal* de $\mathbb{R}^n$ en $\mathbb{R}^m$ es una aplicación f entre estos dos conjuntos de modo que se cumplen las dos siguientes propiedades:
1. $f(x+y) = f(x) + f(y)\ \forall x, y \in \mathbb{R}^n$
2. $f(\lambda x) = \lambda f(x)\ \forall \lambda \in \mathbb{R}, x \in \mathbb{R}^n$

En particular, para toda aplicación lineal se cumple que $f(0) = 0$.

Ejemplo 2.4.1. Consideremos la aplicación

$$\begin{array}{rl} f : \mathbb{R}^3 & \longrightarrow \mathbb{R}^5 \\ (x_1, x_2, x_3) & \longrightarrow (x_1 - x_2, x_2, x_1 + x_2, x_2, x_3) \end{array}$$

Dejamos al lector la comprobación de que esta aplicación es lineal, así como de que es inyectiva pero no exhaustiva.
En cambio, la aplicación

$$\begin{array}{rl} g : \mathbb{R}^3 & \longrightarrow \mathbb{R}^5 \\ (x_1, x_2, x_3) & \longrightarrow (x_1 - x_2 + 1, x_2, x_1 + x_2 - 1, x_2, x_3) \end{array}$$

no es lineal. Basta con observar que $g(0,0,0) = (1,0,-1,0,0)$.

Proposición 2.4.1. *Sea A una matriz con m filas y n columnas. La aplicación*

$$f : \mathbb{R}^n \longrightarrow \mathbb{R}^m$$

que viene definida del siguiente modo: $f(x)$ es el vector de $\mathbb{R}^m$ cuyas coordenadas, expresadas en una matriz columna, coinciden con AX, siendo X la matriz de orden $1 \times n$ cuyos elementos son las coordenadas de x, es una aplicación lineal.

La demostración es consecuencia inmediata de las propiedades de las operaciones con matrices.
El recíproco viene dado por el siguiente resultado.

Proposición 2.4.2. *Si f es una aplicación lineal de $\mathbb{R}^n$ en $\mathbb{R}^m$, la aplicación lineal que se obtiene a partir de la matriz A cuyas columnas son las coordenadas de las imágenes de los vectores de la base natural de $\mathbb{R}^n$ es precisamente f. Indicaremos por $M(f)$ a esta matriz.*

La demostración se omite, por tratarse de una simple comprobación.

Ejemplo 2.4.2. En las bases naturales de $\mathbb{R}^3$ y $\mathbb{R}^5$, la matriz de la aplicación f del Ejemplo (2.4.1) es:

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

Una generalización del resultado anterior nos la da la siguiente proposición.

Definición 2.4.3. Si $\{u_1, \ldots, u_n\}$ es una base cualquiera de $\mathbb{R}^n$ y $\{v_1, \ldots, v_m\}$ una base cualquiera de $\mathbb{R}^m$, la matriz B cuyas columnas están constituidas por las coordenadas, en la base $\{v_1, \ldots, v_m\}$, de las imágenes por la aplicación lineal f de los vectores $u_1, \ldots, u_m$ recibe el nombre de *matriz de la aplicación lineal f en las bases $\{u_1, \ldots, u_n\}$ y $\{v_1, \ldots, v_m\}$*. Habitualmente se denota esta matriz de la forma siguiente: $B = M_{u,v}(f)$, donde u y v representan las bases $\{u_1, \ldots, u_n\}$ y $\{v_1, \ldots, v_m\}$, respectivamente. En el caso de una aplicación lineal de $\mathbb{R}^m$ en $\mathbb{R}^m$ podemos considerar la misma base $u = \{u_1, \ldots, u_n\}$ en el espacio de salida que en el espacio de llegada, denotándose en este caso a la matriz de la aplicación en esta base simplemente por $M_u(f)$.

El siguiente resultado es de gran importancia, puesto que permite obtener la relación existente entre las matrices de una misma aplicación lineal en distintas bases.

Proposición 2.4.3. *Con las notaciones de los resultados precedentes,*

$$B = T^{-1}AS$$

donde S representa la matriz del cambio de base con respecto a la base $\{u_1, \ldots, u_n\}$ y T representa la matriz del cambio de base con respecto a la base $\{v_1, \ldots, v_m\}$.
En el caso de una aplicación lineal de $\mathbb{R}^m$ en $\mathbb{R}^m$, si consideramos las mismas bases en el espacio de salida y en el espacio de llegada, la relación anterior es de la forma:

$$B = S^{-1}AS$$

Ejemplo 2.4.3. Consideremos la aplicación lineal $f : \mathbb{R}^3 \longrightarrow \mathbb{R}^5$ del Ejemplo (2.4.1), que viene definida por $f(x_1, x_2, x_3) = (x_1 - x_2, x_2, x_1 + x_2, x_2, x_3)$. En

el Ejemplo (2.4.2) se ha visto que la matriz de esta aplicación, en las bases naturales de $\mathbb{R}^3$ y $\mathbb{R}^5$ es:

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

Consideremos ahora las bases $u = \{(1,0,1),(1,2,1),(0,1,1)\}$ de $\mathbb{R}^3$ y $v = \{(1,0,-1,1,0),(0,1,1,1,0),(1,-1,1,0,0),(0,0,1,0,1),(0,0,1,1,-1)\}$ de $\mathbb{R}$. En estas bases la matriz de la aplicación f es

$$B = T^{-1}AS = \frac{1}{3}\begin{pmatrix} 0 & 0 & 1 \\ 3 & 3 & -1 \\ 3 & -1 & -4 \\ 0 & 6 & 6 \\ -3 & 3 & 3 \end{pmatrix}$$

Observación. Una consecuencia especialmente importante de este resultado es el hecho de que los rangos de las matrices de una misma aplicación lineal deben coincidir necesariamente.
Damos a continuación las definiciones de dos subespacios vectoriales íntimamente asociados a una aplicación lineal $f : \mathbb{R}^n \longrightarrow \mathbb{R}^m$.

Definición 2.4.4. Se denomina *núcleo de* f al subespacio vectorial de $\mathbb{R}^n$

$$\operatorname{Ker} f = \{x \in \mathbb{R}^n \mid f(x) = 0\}$$

Aquí la nomenclatura de este subespacio corresponde al término alemán núcleo (*kern*).

Ejemplo 2.4.4. En el caso de la aplicación lineal del Ejemplo (2.4.1),

$$\operatorname{Ker} f = \{(x_1,x_2,x_3) \in \mathbb{R}^3 \mid f(x) = 0\} = \{(0,0,0)\}$$

Este ejemplo ilustra la siguiente caracterización de las aplicaciones lineales inyectivas.
Una aplicación lineal $f : \mathbb{R}^n \longrightarrow \mathbb{R}^m$ es inyectiva si, y sólo si, $\operatorname{Ker} f = \{0\}$.

Definición 2.4.5. Se denomina *imagen de* f al subespacio vectorial de $\mathbb{R}^m$

$$\operatorname{Im} f = \{y \in \mathbb{R}^m \mid \text{existe un vector } x \in \mathbb{R}^n \text{ con } f(x) = y\}$$

Ejemplo 2.4.5. En el caso de la aplicación lineal del Ejemplo (2.4.1),

$$\operatorname{Im} f = \{(x_1 - x_2, x_2, x_1 + x_2, x_2, x_3) \,|\, x_1, x_2, x_3 \in \mathbb{R}\}$$

Es fácil comprobar que los vectores $\{(1,0,1,0,0), (-1,1,1,1,0), (0,0,0,0,1)\}$ forman una base del subespacio vectorial $\operatorname{Im} f$. Así pues, su dimensión es 3.

De hecho, se tiene la siguiente caracterización de las aplicaciones exhaustivas: Una aplicación lineal $f : \mathbb{R}^n \longrightarrow \mathbb{R}^m$ es exhaustiva si, y sólo si, $\dim \operatorname{Im} f = m$. El siguiente resultado nos da la relación existente, y de gran utilidad, entre las dimensiones de los dos subespacios vectoriales anteriores.

Proposición 2.4.4. *Dada una aplicación lineal $f : \mathbb{R}^n \longrightarrow \mathbb{R}^m$ se cumple la siguiente relación:*

$$\dim \operatorname{Ker} f + \dim \operatorname{Im} f = n$$

Finalmente, damos los nombres que suelen recibir las aplicaciones lineales según sus propiedades.
Morfismo = aplicación lineal
Endomorfismo = aplicación lineal en el caso particular en que $n = m$
Monomorfismo = aplicación lineal inyectiva
Epimorfismo = aplicación lineal exhaustiva
Isomorfismo (automorfismo) = aplicación lineal biyectiva

Ejemplo 2.4.6. La aplicación lineal considerada en el Ejemplo (2.4.1) es un monomorfismo, pero no es un epimorfismo (por lo tanto, tampoco podría ser un isomorfismo. Las observaciones de los Ejemplos (2.4.4) y (2.4.5) indican que sólo las aplicaciones lineales $f : \mathbb{R}^n \longrightarrow \mathbb{R}^n$ pueden ser isomorfismos.

Observación. Dada una aplicación lineal f de $\mathbb{R}^n$ en $\mathbb{R}^m$, si u y v son dos bases cualesquiera de $\mathbb{R}^n$ y de $\mathbb{R}^m$, respectivamente, entonces $\operatorname{rango}(f) = \operatorname{rango} M_{u,v}(f)$. Este escalar recibe el nombre de rango de f y coincide con la dimensión del subespacio vectorial $\operatorname{Im} f$ de $\mathbb{R}^m$.

2.5. Determinante de un endomorfismo de $\mathbb{R}^n$

Observamos que las matrices asociadas a un endomorfismo son siempre matrices cuadradas, en cualesquiera bases que se considere, por lo que tiene sentido calcular su determinante. Además, si B es la matriz de un endomorfismo f

de $\mathbb{R}^n$ en una base u formada por los vectores $u_1, \ldots, u_n$ (consideramos la misma base en el espacio de salida que en el de llegada), se cumple que:

$$B = S^{-1}AS$$

si hemos indicado por A a la matriz, en la base natural de $\mathbb{R}^n$, del mismo endomorfismo.
Entonces puede comprobarse fácilmente el siguiente resultado.

Proposición 2.5.1. *El determinante de la matriz de f no depende de la base elegida.*

Demostración. Puesto que $B = S^{-1}AS$, se cumple que

$$\det B = \det S^{-1} \cdot \det A \cdot \det S = \frac{1}{\det S^{-1}} \cdot \det A \cdot \det S = \det A$$

□

Esto da sentido a la siguiente definición.

Definición 2.5.1. El determinante del endomorfismo f, que denotaremos por $\det f$, es igual a $\det A$, con las notaciones anteriores.

Ejemplo 2.5.1. Consideremos la aplicación lineal

$$\begin{array}{rl} f : \mathbb{R}^2 & \longrightarrow \mathbb{R}^2 \\ (x_1, x_2) & \longrightarrow (x_1 + x_2, x_1 - x_2) \end{array}$$

La matriz de esta aplicación lineal, en la base canónica de $\mathbb{R}^2$, es:

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

Luego

$$\det f = \det \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix} = -2$$

Además, este determinante permite caracterizar los automorfismos.

Proposición 2.5.2. *f es un automorfismo si y sólo si $\det f \neq 0$.*

Para finalizar este apartado, podemos relacionar el determinante de un automorfismo y el de su inverso, que es también un automorfismo (la comprobación de esta afirmación se deja como ejercicio para el lector).

Proposición 2.5.3. *Sea f un automorfismo de $\mathbb{R}^n$. Indicamos por f^{-1} a su aplicación inversa. Entonces se tiene que*

$$\det f^{-1} = \frac{1}{\det f}$$

2.6. Subespacios invariantes por un endomorfismo

Sea f un endomorfismo de $\mathbb{R}^n$ y sea F un subespacio vectorial.

Definición 2.6.1. Se dice que el subespacio F es *invariante* por el endomorfismo f cuando la imagen por f de todos los vectores de F pertenece también a F.

Ejemplo 2.6.1. Consideremos la aplicación lineal

$$\begin{aligned} f : \mathbb{R}^3 &\longrightarrow \mathbb{R}^3 \\ (x_1, x_2, x_3) &= (2x_1 + x_2, x_1 + 2x_2, x_3) \end{aligned}$$

El subespacio vectorial

$$F = \{(x_1, x_2, x_3) \,|\, x_1 - x_2 = x_3 = 0\}$$

es invariante por el endormorfismo f. En efecto. Dado un vector cualquiera del subespacio vectorial F, que es de la forma $(a, a, 0)$ para una cierta $a \in \mathbb{R}$, su imagen por el endomorfismo f es $f(a, a, 0) = (3a, 3a, 0) = 3(a, a, 0)$ que, obviamente, pertenece también a F.

El siguiente resultado nos proporciona una forma fácil de comprobar si un subespacio vectorial es, o no, invariante por un endomorfismo.

Proposición 2.6.1. *Sea $\{u_1, \ldots, u_n\}$ un sistema de generadores de F. Entonces F es invariante por f si y sólo si las imágenes de todos los vectores del conjunto $\{u_1, \ldots, u_n\}$ pertenece a F.*

Observación. En particular, tenemos que si $\{u_1, \ldots, u_n\}$ es una base de F, F es invariante por f si y sólo si las imágenes de todos los vectores del conjunto $\{u_1, \ldots, u_n\}$ pertenece a F.
Observemos que en el Ejemplo (2.6.1) habría bastado con considerar una base del subespacio vectorial F, por ejemplo, $\{(1,1,0)\}$ y haber comprobado que $f(1,1,0) = (3,3,0) = 3(1,1,0)$ pertenece a F.
Como forma práctica de averiguar si un subespacio vectorial es, o no, invariante por un endomorfismo, tenemos el siguiente resultado.

Proposición 2.6.2. *Sea $\{u_1, \ldots, u_n\}$ un sistema de generadores de F. Llamamos A a la matriz con vectores columna las coordenadas de los vectores de este sistema de generadores y B a la matriz con vectores columna las coordenadas de los vectores anteriores y también de sus imágenes por f. Entonces F es invariante por f si y sólo si sólo si*

$$rg\, A = rg\, B$$

Ejemplo 2.6.2. Consideremos el caso del endomorfismo del Ejemplo (2.6.1). Ya hemos indicado que $\{(1,1,0)\}$ es una base de F y que $f(1,1,0) = (3,3,0)$. Bastaría en este caso con comprobar que las matrices

$$A = \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix}, \qquad B = \begin{pmatrix} 1 & 3 \\ 1 & 3 \\ 0 & 0 \end{pmatrix}$$

tienen el mismo rango (en este caso, ambas tienen rango igual a 1).

2.7. Ejercicios Resueltos

1.- Probar que el subconjunto de $\mathbb{R}^4$

$$F = \{(x_1, x_2, x_3, x_4 \mid x_1 + x_2 = 0, x_3 + x_4 = 0\}$$

es un subespacio vectorial. Dar una base y su dimensión.
Solución:
Sean $x = (x_1, x_2, x_3, x_4), y = (y_1, y_2, y_3, y_4) \in F$, entonces $x+y = (x_1+y_1, x_2+y_2, x_3+y_3, x_4+y_4)$ es tal que $(x_1+y_1)+(x_2+y_2) = (x_1+x_2)+(y_1+y_2) =$

$0+0=0$ y $(x_3+y_3)+(x_4+y_4)=(x_3+x_4)+(y_3+y_4)=0+0=0$, por lo que $x+y\in F$.
Sean ahora $\lambda\in\mathbb{R}$ y $x=(x_1,x_2,x_3,x_4)$, entonces $\lambda x=(\lambda x_1.\lambda x_2,\lambda x_3.\lambda x_4)$ es tal que $\lambda x_1+\lambda x_2=\lambda(x_1+\lambda x_2)=\lambda\cdot 0=0$ y $\lambda x_3+\lambda x_4=\lambda(x_3+\lambda x_4)=\lambda\cdot 0=0$, por lo que $\lambda x\in F$
Puesto que $\forall x,y\in F$, $\forall\lambda\in\mathbb{R}$, se tiene $x+y\in F, \lambda x\in F$. se concluye que F es un subespacio.
Busquemos una base.
$\forall x=(x_1,x_2,x_3,x_4)\in F$ es $x_2=-x_1$ y $x_4=-x_3$, por lo que $x=(x_1,-x_1,x_3,-x_3)=x_1(1,-1,0,0)+x_3(0,0,1,-1)$.
Observamos que $(1,-1,0,0),(0,0,1,-1)\in F$ generan F y son independientes, luego determinan una base.
Puesto que la base obtenida tiene dos elementos, la dimensión del subespacio es 2.

2.- Estudiar la dependencia o independencia lineal de las familias de vectores de $\mathbb{R}^4$ siguientes:
(a) $\{(1,-1,2,2),(-2,-1,3,4),(7,-1,0,-2)\}$
(b) $\{(1,1,-2,3),(2,-1,0,4),(5,-1,0,0),(1,0,1,0)\}$
(c) $\{(0,-1,3,2),(2,7,-3,4),(5,1,2,-2),(8,9,3,4),(6,3,2,4)\}$
Solución:
Sabemos que una colección de vectores $\{u_1,\ldots,u_4\}$, estos son independientes si y sólo si

$$\lambda_1u_1+\ldots+\lambda_ru_r=0 \;\Leftrightarrow\; \lambda_1=\ldots=\lambda_r=0$$

a) $\lambda_1(1,-1,2,2)+\lambda_2(-2,-1,3,4)+\lambda_3(7,-1,0,-2)=(0,0,0,0)$.

$$\left.\begin{array}{rl}\lambda_1-2\lambda_2+7\lambda_3 &=0\\ -\lambda_1-\lambda_2-\lambda_3 &=0\\ 2\lambda_1+3\lambda_2 &=0\\ 2\lambda_1+4\lambda_2-2\lambda_3 &=0\end{array}\right\} \Rightarrow \left.\begin{array}{rl}\lambda_1 &=-3\lambda_3\\ \lambda_2 &=2\lambda_3\end{array}\right\}$$

Luego los tres vectores son dependientes. Concretamente $(7-1,0,-)=3(1,-1,2,2)-2(-2,-1,3,4)$.

b) $\lambda_1(1,1,-2,3)+\lambda_2(2,-1,0,4)+\lambda_3(5,-1,0,0)+\lambda_4(1,0,1,0)=(0,0,0,0)$

$$\left.\begin{array}{rl} \lambda_1 + 2\lambda_2 + 5\lambda_3 + \lambda_4 & = 0 \\ \lambda_1 - \lambda_2 - \lambda_3 & = 0 \\ -2\lambda_1 + \lambda_4 & = 0 \\ 3\lambda_1 + 4\lambda_2 & = 0 \end{array}\right\} \Rightarrow \left.\begin{array}{rl} \lambda_1 & = 0 \\ \lambda_2 & = 0 \\ \lambda_3 & = 0 \\ \lambda_4 & = 0 \end{array}\right\}$$

Luego los cuatro vectores son linealmente independientes.

c) Como que tenemos cinco vectores en un espacio de dimensión cuatro, podemos asegurar que al menos uno de ellos es dependiente de los cuatro restantes.

3.- En $\mathbb{R}^5$ consideremos los subespacios vectoriales seguientes:

$$\begin{array}{rl} F & = [(1,0,2,1,3)] \\ G & = [(1,0,0,1,0),(0,1,0,0,0),(0,0,1,1,0)] \end{array}$$

Encontrar $F+G$. ¿Es directa la suma?
Solución:
La reunión de bases de F y G proporcionan un sistema de generadores de $F+G$ si además son independientes pasarán a ser una base por lo que además la suma será directa.
Estudiemos pues la independencia

$$\lambda_1(1,0,2,1,3)+\lambda_2(1,0,0,1,0)+\lambda_3(0,1,0,0,0)+\lambda_4(0,0,1,1,0) = (0,0,0,0,0)$$

Resolviendo el sistema tenemos $\lambda_1 = \lambda_2 = \lambda_3 = \lambda_4 = 0$. Por lo que los vectores son independientes.
Tenemos que una base de $F+G$ es

$$\{(1,0,2,1,3),(1,0,0,1,0),(0,1,0,0,0),(0,0,1,1,0)\}$$

y la suma es directa.

4.- Discutir si las siguientes aplicaciones son, o no, lineales:

a) $f : \mathbb{R}^3 \longrightarrow \mathbb{R}^4$

$(x_1, x_2, x_3) \longmapsto (3x_1 + 5x_2, x_2 + x_3, x_2 - 5x_3, x_1 + x_2 + x_3)$.

b) $f : \mathbb{R}^3 \longrightarrow \mathbb{R}^2$

$(x_1, x_2, x_3) \longmapsto (x_1 + a, bx_2)$ con $a, b \in \mathbb{R}$.

Solución:

a) Sean (x_1, x_2, x_3), (y_1, y_2, y_3),
$(x_1, x_2, x_3) + (y_1, y_2, y_3) = (x_1 + y_1, x_2 + y_2, x_3 + y_3)$
$\lambda(x_1, x_2, x_3) = (\lambda x_1, \lambda x_2, \lambda x_3)$
$f((x_1 + y_1, x_2 + y_2, x_3 + y_3) =$
$(3(x_1 + y_1) + 5(x_2 + y_2), (x_2 + y_2) + (x_3 + y_3), (x_2 + y_2) - 5(x_3 + y_3),$
$(x_1 + y_1) + (x_2 + y_2) + (x_3 + y_3)) =$
$(3x_1 + 5x_2, x_2 + x_3, x_2 - 5x_3, x_1 + x_2 + x_3) + (3y_1 + 5y_2, y_2 + y_3, y_2 - 5y_3,$
$y_1 + y_2 + y_3) =$
$f(x_1, x_2, x_3) + f(y_1, y_2, y_3)$,
$f(\lambda x_1, \lambda x_2, \lambda x_3) =$
$(3\lambda x_1 + 5\lambda x_2, \lambda x_2 + \lambda x_3, \lambda x_2 - 5\lambda x_3, \lambda x_1 + \lambda x_2 + \lambda x_3) =$
$\lambda(3x_1 + 5x_2, x_2 + x_3, x_2 - 5x_3, x_1 + x_2 + x_3) =$
$\lambda f(x_1, x_2, x_3)$.

b) Al igual que en el apartado a)
$f(x_1 + y_1, x_2 + y_2, x_3 + y_3) =$
$((x_1 + y_1) + a, b(x_2 + y_2)) = (x_1 + a, bx_2) + (y_1 + a, by_2) - (a, 0) =$
$f(x_1, x_2, x_3) + f(y_1, y_2, y_3) - (a, 0)$.
Luego la primera condición se verifica si y sólo si $a = 0$.
Impongamos $a = 0$,
$f(\lambda x_1, \lambda x_2, \lambda x_3) = (\lambda x_1, b\lambda x_2) =$
$\lambda(x_1, bx_2)$.
Luego la aplicación es lineal para $a = 0$ y para todo b.

5.- Sea f la aplicación lineal de $\mathbb{R}^3$ en $\mathbb{R}^3$ tal que

$$\operatorname{Ker} f = [(2, -1, 0), (2, 0, 1)], \quad f(1, 0, 0) = (2, 0, -1).$$

Determinar la matriz de f en la base natural de $\mathbb{R}^3$.

Solución:

Puesto que $(2, -1, 0), (2, 0, 1) \in \operatorname{Ker} f$, se tiene que
$f(0, 1, 0) = 2f(1, 0, 0)$, $f(0, 0, 1) = -2f(1, 0, 0)$,
por lo que la matriz de la aplicación es

$$A = \begin{pmatrix} 2 & 4 & -4 \\ 0 & 0 & 0 \\ -1 & -2 & 2 \end{pmatrix}.$$

6.- Calcular el determinante del endomorfismo siguiente.

$$\begin{aligned} f : \mathbb{R}^4 &\longrightarrow \mathbb{R}^4 \\ (x_1, x_2, x_3, x_4) &\longrightarrow (x_1 + x_2 - x_3, x_1 - x_2 + x_3, x_1 - x_4, x_3 + x_4) \end{aligned}$$

Solución: Escribamos la matriz de la aplicación en la base canónica

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

Calculemos el determinante de A.

$$\begin{aligned} \det A = \begin{vmatrix} 1 & 1 & -1 & 0 \\ 1 & -1 & 1 & 0 \\ 1 & 0 & 0 & -1 \\ 0 & 0 & 1 & 1 \end{vmatrix} = \begin{vmatrix} 1 & 1 & -1 & 0 \\ 0 & -2 & 2 & 0 \\ 0 & -1 & 1 & -1 \\ 0 & 0 & 1 & 1 \end{vmatrix} = \\ = \begin{vmatrix} -2 & 2 & 0 \\ -1 & 1 & -1 \\ 0 & 1 & 1 \end{vmatrix} = \begin{vmatrix} -2 & 0 & 0 \\ -1 & 0 & -1 \\ 0 & 1 & 1 \end{vmatrix} = -2 \begin{vmatrix} 0 & -1 \\ 1 & 1 \end{vmatrix} = -2. \end{aligned}$$

Por lo que $\det f = -2$.

7.- Sea f una aplicación lineal de $\mathbb{R}^4$ en $\mathbb{R}^4$ tal que $f(e_1) = e_1 + e_2$, $e_1 - e_2 \in \operatorname{Ker} f$. Probar que el subespacio vectorial $F = [e_1, e_2]$ es invariante por f y dar la matriz de la aplicación lineal restricción $f_{|F}$ de F en F, en la base (e_1, e_2) de F.

Solución:

Puesto que $e_1 - e_2 \in \operatorname{Ker} f$ tenemos que $f(e_2) = f(e_1)$ y puesto que $f(e_1) = e_1 + e_2$ tenemos que $f(e_1), f(e_2) \in [e_1, e_2]$ y el subespacio es invariante.

La matriz de la aplicación restricción es

$$A_F = \begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix}.$$

2.8. Ejercicios Propuestos

1.- Considerar el subconjunto de $\mathbb{R}^4$:

$$F = \{(x_1, x_2, x_3, x_4) \,|\, x_2 + x_3 + x_4 = 0\}$$

¿Es F un subespacio vectorial? Justificar la respuesta. En caso afirmativo, encontrar una base y deducir cuál es su dimensión.

2.- Estudiar la dependencia o independencia lineal de la familia de vectores de $\mathbb{R}^4$ siguiente:

$$\{(1,1,2,2),(2,1,3,4),(-1,1,0,-2)\}$$

Dar la dimensión del subespacio que engendran, así como una base de este subespacio.

3.- Encontrar el rango de las siguientes matrices:

$$A=\begin{pmatrix} 0 & 3 & 9 \\ \frac{1}{3} & 0 & 3 \\ \frac{1}{9} & \frac{1}{3} & 0 \end{pmatrix}, \quad B=\begin{pmatrix} 1 & 0 & 2 \\ 1 & 1 & 3 \\ 2 & 0 & 4 \end{pmatrix}$$

4.- Determinar si es posible que las matrices del ejercicio anterior sean las matrices de una misma aplicación lineal, en distintas bases.

5.- En $\mathbb{R}^5$ se consideran los subespacios vectoriales:

$$\begin{aligned} F &= [(1,0,2,1,3)] \\ G &= [(1,0,0,1,0),(0,1,0,0,0),(0,0,1,1,0)] \\ H &= [(1,0,2,1,0),(1,0,1,0,1)] \end{aligned}$$

Encontrar $F+G$, $F+H$, $G+H$ y $F+G+H$.

6.- Considerar la aplicación lineal

$$\begin{aligned} f:\mathbb{R}^4 &\longrightarrow \mathbb{R}^5 \\ (a,b,c,d) &\longrightarrow (a-b,a-d,b-c,b-d,c-d) \end{aligned}$$

Determinar la matriz de f en las bases canónicas de $\mathbb{R}^4$ y $\mathbb{R}^5$.

7.- Determinar la matriz de la aplicación lineal del ejercicio anterior en las bases $\{(1,1,2,3),(2,1,0,0),(0,0,1,1,),(1,0,0,1)\}$ de $\mathbb{R}^4$ y $\{(1,1,0,0,0),(1,-1,0,0,0),(0,0,0,1,1,1),(0,0,0,1,2),(0,0,0,1,1)\}$ de $\mathbb{R}^5$.

8.- Sea f el endomorfismo de $\mathbb{R}^3$ que viene determinado por $f(1,0,0)=(2,3,1)$, $f(1,1,0)=(2,2,-3)$ y $f(1,1,1)=(4,1,2)$. Encontrar la matriz de este endomofismo en la base natural de $\mathbb{R}^3$.

9.- Sea f el endomorfismo de $\mathbb{R}^3$ definido por

$$f(x_1,x_2,x_3)=(x_1-2x_2+2x_3,x_1+3x_2-x_3,2x_1+x_2+x_3)$$

Calcular rango f y $\det f$.

10.- Considerando el endomorfismo definido en el ejercicio anterior, comprobar que el subespacio vectorial $F=[(1,0,1),(1,-1,0)]$ es invariante por f.

Capítulo 3

Propiedades espectrales

3.1. Polinomio característico de una matriz

El problema espectral consiste, dada una matriz cuadrada A real o compleja, de orden n, en encontrar aquellos valores $\lambda \in \mathbb{C}$ y vectores no nulos v de $\mathbb{C}^n$ tales que $AV = \lambda V$ si V es la matriz de orden $n \times 1$ cuyos elementos son las coordenadas del vector v. El escalar λ recibe el nombre de valor propio de A y v el de vector propio asociado.
Para hallar dichos valores, es útil calcular el llamado polinomio característico de la matriz.
Sea A una matriz cuadrada de orden n.

Definición 3.1.1. Se llama *polinomio característico de la matriz* A a:

$$Q_A(t) = \det(A - tI_n)\,.$$

Ejemplo 3.1.1. Consideremos las matrices

$$A = \begin{pmatrix} 2 & 1 \\ 1 & 2 \end{pmatrix}, \qquad B = \begin{pmatrix} 1 & 1 & 0 \\ 0 & 1 & 1 \\ 1 & 0 & 1 \end{pmatrix}$$

Entonces:

$$Q_A(t) = t^2 - 4t + 3, \qquad Q_B(t) = -t^3 + 3t^2 - 3t + 2$$

Proposición 3.1.1. *Sea A una matriz cuadrada de orden n. Entonces:*

$$Q_A(t) = Q_{A^t}(t)$$

Demostración. Para probar este resultado, basta con aplicar la primera propiedad de los determinantes: $\det(A-tI) = \det(A-tI)^t$ y comprobar que $\det(A-tI)^t$ es igual a $\det(A^t - tI)$. □

La siguiente fórmula puede resultar útil a la hora de calcular el polinomio característico de una determinada matriz. Su demostración no se incluye, por ser una ardua comprobación.

Proposición 3.1.2. *Sea A una matriz cuadrada de orden n. Entonces:*

$$Q_A(t) = (-1)^n t^n + (-1)^{n-1}\Sigma_1 t^{n-1} + \cdots + (-1)^{n-k}\Sigma_k t^{n-k} + \cdots + \Sigma_n\,,$$

con $\Sigma_1, \ldots, \Sigma_k, \ldots, \Sigma_n$ las sumas de los menores de orden k de la matriz $A = (a^i_j)_{1\leq i,j\leq n}$ "sobre la diagonal principal"; es decir, de los menores de la matriz A formados por los elementos pertenecientes a las filas y columnas $\ell_1 \ldots \ell_k$, que podemos denotar por $A^{\ell_1\ldots\ell_k}_{\ell_1\ldots\ell_k}$, con $1 \leq \ell_1 < \ell_2 < \ldots < \ell_k \leq n$.

Observación. En particular:

$$\begin{aligned} \Sigma_1 &= \operatorname{tr} A = a_{11} + \cdots + a_{nn} \\ \Sigma_n &= \det A \end{aligned}$$

3.2. Valores y vectores propios de una matriz

Las matrices diagonales son especialmente sencillas, desde un punto de vista operativo, para trabajar con ellas.
Sea A una matriz cuadrada de orden n. Se dice que una matriz A es diagonalizable cuando es posible encontrar una matriz inversible S de manera que $S^{-1}AS$ es una matriz diagonal D; es decir, de la forma:

$$D = \begin{pmatrix} \lambda_1 & 0 & \ldots & 0 \\ 0 & \lambda_2 & & \vdots \\ \vdots & & \ddots & 0 \\ 0 & \ldots & 0 & \lambda_n \end{pmatrix}$$

El primer paso para saber si existe una matriz S que permite diagonalizar una matriz dada es encontrar los llamados valores propios y los correspondientes vectores propios asociados a la matriz, que a continuación pasamos a definir.

La herramienta que nos permite hallar los valores propios es el polinomio característico de la matriz, puesto que los valores propios coinciden con sus raíces.
Pasamos ahora a dar, con más precisión, las definiciones.
Sea A una matriz cuadrada de orden n.

Definición 3.2.1. Un escalar $\lambda \in \mathbb{C}$ recibe el nombre de *valor propio* de A si $Q_A(\lambda) = 0$, es decir, si λ es una raíz del polinomio característico de la matriz A. Y se llama *espectro* de la matriz A a

$$\operatorname{Spec} A = \{\lambda \in \mathbb{C} \,|\, \lambda \text{ es valor propio de } A\}.$$

Observamos que una matriz cuadrada real de orden n tiene n valores propios en $\mathbb{C}$ que pueden ser iguales o distintos y que en $\mathbb{R}$ puede tener n, menos de n o incluso no tener ninguno.

Ejemplo 3.2.1. Consideremos las matrices del Ejemplo (3.1.1),

$$A = \begin{pmatrix} 2 & 1 \\ 1 & 2 \end{pmatrix}, \qquad B = \begin{pmatrix} 1 & 1 & 0 \\ 0 & 1 & 1 \\ 1 & 0 & 1 \end{pmatrix}$$

de las cuales ya sabemos que

$$\begin{aligned} Q_A(t) &= t^2 - 4t + 3 = (t-1)(t-3), \\ Q_B(t) &= -t^3 + 3t^2 - 3t + 2 = (t-2)\left(t - \tfrac{1+\sqrt{3}i}{2}\right)\left(t - \tfrac{1-\sqrt{3}i}{2}\right) \end{aligned}$$

Luego los valores propios de la matriz A son: 1 y 3 (reales), mientras que los valores propios de la matriz B son $2, \dfrac{1+\sqrt{3}i}{2}, \dfrac{1-\sqrt{3}i}{2}$ (sólo uno real y los otros dos complejos),

El siguiente resultado nos da una condición equivalente a la enunciada en la definición anterior y que es utilizada con mucha frecuencia.

Proposición 3.2.1. *Un escalar $\lambda \in \mathbb{C}$ recibe el nombre de valor propio de A si y sólo si existe $x \in M_{n\times 1}(\mathbb{C})$, $x \neq 0$, tal que $Ax = \lambda x$.*

La idea de la demostración es considerar las siguientes equivalencias, ninguna de las cuales es difícil de probar.

Son equivalentes:

$$\begin{aligned}\lambda \text{ es valor} \quad & \text{propio de } A \\ & \iff x \neq 0 \text{ y } Ax = \lambda x \\ & \iff x \neq 0 \text{ y } (A - \lambda I_n)x = 0 \\ & \iff \text{el sistema } (A - \lambda I_n)x = 0 \text{ es compatible indeterminado} \\ & \iff \operatorname{rg}(A - \lambda I_n) < n\end{aligned}$$

Definición 3.2.2. Un vector columna $x \in M_{n\times 1}(\mathbb{C})$ se dice que es un *vector propio* de A si $x \neq 0$ y existe $\lambda \in \mathbb{C}$ tal que $Ax = \lambda x$.
Se dice entonces que x *es un vector propio asociado al valor propio* λ y que λ *es un valor propio asociado al vector propio* x.

Proposición 3.2.2. *Sean A, B matrices cuadradas de orden n. Si $B = S^{-1}AS$, entonces: $Q_B(t) = Q_A(t)$.*

Demostración. Basta con observar que:

$$\begin{aligned}Q_B(t) = Q_{S^{-1}AS}(t) \quad & = \det(S^{-1}AS - tI) = \det(S^{-1}AS - S^{-1}(tI)S) \\ & = \det(S^{-1}(A - tI)S) = \det(S^{-1})\det(A - tI)\det(S) \\ & = \frac{1}{\det S}\det(A - tI)\det(S) = \det(A - tI) = Q_A(t)\end{aligned}$$

□

Proposición 3.2.3. *En general, dada una matriz A cuadrada de orden n, se cumple que:*

$$\begin{cases} tr\ A = \lambda_1 + \cdots + \lambda_n \\ \det A = \lambda_1 \cdot \cdots \cdot \lambda_n \end{cases}$$

si $\lambda_1, \ldots, \lambda_n$ son los valores propios de la matriz A.

La demostración no se incluye puesto que para ella se necesita estudiar la llamada "triangulación"de una matriz, no incluida en este texto.
Así pues, para calcular los valores propios y los vectores propios de una matriz puede procederse como sigue.

1. Calcular $Q_A(t)$.

2. Buscar las raíces $\lambda \in \mathbb{C}$ de $Q_A(t)$ (son los valores propios).

3. Para cada valor propio λ, los vectores propios asociados a λ son los vectores no nulos cuyas componentes $x_1, \ldots, x_n$ satisfacen:

$$(A - \lambda I)x = 0 \quad \text{con} \quad x = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}.$$

Observamos que este sistema es siempre compatible indeterminado, ya que $\det(A - \lambda I) = 0$.

Las siguientes observaciones permiten calcular valores y vectores propios de determinados tipos especiales de matrices.

Observación 1. Observamos que si $A \in M_n(\mathbb{R})$ es una matriz triangular (superior o inferior) con elementos $a_{11}, \ldots, a_{nn}$ en la diagonal,

$$\det(A - tI) = (a_{11} - t) \ldots (a_{nn} - t) = (-1)^n (t - a_{11}) \ldots (t - a_{nn})$$

Es decir, los valores propios de esta matriz son: $a_{11}, \ldots, a_{nn}$. Observamos que en este caso son todos reales.

Observación 2. Los vectores propios de A asociados a λ son los vectores no nulos de $\mathrm{Ker}(A - \lambda I)$, puesto que son equivalentes:
(i) Existe $x \neq 0$ y $(A - \lambda I)x = 0$
(ii) Existe un vector $x \neq 0$ en $\mathrm{Ker}(A - \lambda I)$

Observación 3. Vectores propios asociados a valores propios diferentes son linealmente independientes.

Observación 4. Si la suma de todos los vectores columna de A es un vector columna del tipo

$$\begin{pmatrix} \lambda \\ \lambda \\ \lambda \\ \vdots \\ \lambda \end{pmatrix}$$

(es decir el vector $\lambda u_1 + \ldots + \lambda u_n$) entonces λ es un valor propio de A. En efecto:

$$A(u_1) + \cdots + A(u_n) = A(u_1 + \cdots + u_n) = \lambda u_1 + \cdots + \lambda u_n = \lambda(u_1 + \cdots + u_n)$$

es decir,

$$A(u_1 + \cdots + u_n) = \lambda(u_1 + \cdots + u_n)$$

Hemos obtenido al mismo tiempo que $u_1 + \cdots + u_n$ es un vector propio asociado al valor propio λ. Lo mismo ocurre si todas las filas de la matriz suman lo mismo: este valor es un valor propio de la matriz.

Ejemplo 3.2.2. Dada la matriz

$$M_1 = \begin{pmatrix} 2 & 1 & 1 \\ 1 & 2 & 1 \\ 0 & 1 & 3 \end{pmatrix}$$

el vector $(1, 1, 1)$ es un vector propio de M, asociado al valor propio 4.
Dada la matriz

$$M_2 = \begin{pmatrix} 2 & 1 & 0 \\ 1 & 2 & 1 \\ 1 & 1 & 3 \end{pmatrix}$$

4 es valor propio de esta matriz, (sumando las tres filas obtenemos el vector $\begin{pmatrix} 4 & 4 & 4 \end{pmatrix}$) si bien ahora $(1, 1, 1)$ no es vector propio de dicha matriz.

Observación 5. De forma análoga, si al sumar el vector columna i-simo con el vector columna j-simo obtenemos un nuevo vector columna en que todas sus componentes son el mismo escalar λ, en las filas "i" y "j" y el resto son "ceros", este escalar λ es un valor propio y $u_i + u_j$ es un vector propio asociado a este valor propio, ya que

$$A(u_i) + A(u_j) = A(u_i + u_j) = \lambda u_i + \lambda u_j = \lambda(u_i + u_j)\,.$$

Este resultado puede generalizarse a otros criterios similares. Proponemos al lector esta generalización como ejercicio.
Si nos interesamos en el problema de cuando una matriz A es diagonalizable nos interesamos por los valores propios distintos de la matriz.
Dado un valor propio λ de A llamaremos *multiplicidad geométrica* g de λ al máximo número de vectores propios independientes que se pueden asociar a dicho valor propio, esto es $g = \dim \mathrm{Ker}(A - \lambda I)$. Su *multiplicidad algebraica* m es igual a la multiplicidad en tanto que raíz del polinomio característico. Se puede demostrar que $g \leq m$. Un valor propio con $m = 1$ recibe el nombre de *simple*, en caso contrario de *múltiple*. Un valor propio múltiple

se dice *semisimple* si y sólo si $m = g$, es decir admite m vectores propios independientes. En caso contrario se dice *defectivo*.
Finalmente, enunciamos un criterio que permite saber cuándo una matriz es diagonalizable.

Teorema 3.2.1. *Una matriz A cuadrada de orden n es diagonalizable si, y sólo si, existen n vectores linealmente independientes de $\mathbb{C}^n$ que son vectores propios de la matriz A.*

Es decir, cuando podemos encontrar una base $v = \{v_1, \ldots, v_n\}$ de $\mathbb{C}^n$ de modo que todos los vectores de esta base sean vectores propios de la matriz A.
La matriz A es pues diagonalizable si y sólo si sus valores propios son semisimples. Si la matriz no es diagonalizable recibe el nombre de *defectiva*.
Si interpretamos la matriz A como la matriz de un endomorfismo f_A de $\mathbb{R}^n$ cuya matriz, en la base natural de $\mathbb{R}^n$ es la matriz A, la matriz de f_A en la base v es $S^{-1}AS$ siendo S la matriz del cambio de base, de la base v a la base natural de $\mathbb{R}^n$ y esta matriz es, obviamente, una matriz diagonal. Así pues, una matriz A es diagonalizable si existe una matriz inversible S de orden n de modo que $S^{-1}AS$ es una matriz diagonal.

En el caso en que A es una matriz defectiva se tiene que si λ es un valor propio de multiplicidad algebraica m, entonces $\dim \mathrm{Ker}\,(A - \lambda I) < m$.
Es fácil probar que

$$\mathrm{Ker}\,(A-\lambda I) \subset \mathrm{Ker}\,(A-\lambda I)^2 \subset \ldots \subset \mathrm{Ker}\,(A-\lambda I)^i = \mathrm{Ker}\,(A-\lambda I)^{i+1} = \ldots$$

la igualdad se consigue con n pasos a lo sumo ya que la dimensión de un subespacio es menor o igual que la de todo el espacio. Se tiene el siguiente resultado

Proposición 3.2.4. *Sea A una matriz cuadrada de orden n y λ un valor propio de multiplicidad algebraica m. Entonces el menor i para el cual la cadena de subespacios estabiliza es tal que*

$$\dim Ker\,(A - \lambda I)^i = m$$

Este valor i recibe el nombre de *índice* del valor propio.
Observación. Si el valor propio λ es semisimple entonces $i = 1$ y es estrictamente mayor que 1 en caso de ser defectivo.
Sea A una matriz compleja con valores propios $\lambda_1, \ldots, \lambda_n$ e índices respectivos $i_1, \ldots, i_r$, tenemos la siguiente definición.

Definición 3.2.3. Llamaremos *polinomio mínimo* a

$$(t-\lambda_1)^{i_1}\dots(t-\lambda_r)^{i_r}$$

Observación. El polinomio mínimo es un divisor del polinomio característico y ambos tienen las mismas raíces.
Observación. Si la matriz A es diagonalizable los factores del polinimio mínimo son todos lineales.

3.3. Radio espectral

Sea A una matriz cuadrada de orden n.

Definición 3.3.1. Se llama radio espectral de la matriz A a

$$\rho(A)=\max\{|\lambda_1|,\dots,|\lambda_r|\}$$

si $\lambda_1,\dots,\lambda_r\in\mathbb{C}$ son los valores propios de la matriz A.

Ejemplo 3.3.1. Consideremos las matrices del Ejemplo (3.1.1) y la matriz M del Ejemplo (3.2.2)

$$A=\begin{pmatrix}2&1\\1&2\end{pmatrix},\qquad B=\begin{pmatrix}1&1&0\\0&1&1\\1&0&1\end{pmatrix},\qquad M=\begin{pmatrix}2&1&1\\1&2&1\\0&1&3\end{pmatrix}$$

cuyos valores propios son:

$$\begin{array}{rl} 1 \text{ y } 3 & \text{en el caso de la matriz } A\\ 2,\ \frac{1+\sqrt{3}i}{2} \text{ y } \frac{1-\sqrt{3}i}{2} & \text{en el caso de la matriz } B\\ 1,\ 2,\ \text{y } 4, & \text{en el caso de la matriz } M\end{array}$$

Entonces

$$\rho(A)=3,\qquad \rho(B)=2,\qquad \rho(M)=4$$

Propiedad. Para todo valor propio λ de una matriz A se cumple:

$$|\lambda|\le\max\{|a_{11}|+\cdots+|a_{1n}|,\dots,|a_{n1}|+\cdots+|a_{nn}|\}$$

$$|\lambda|\le\max\{|a_{11}|+\cdots+|a_{n1}|,\dots,|a_{1n}|+\cdots+|a_{nn}|\}$$

En particular,

$$\rho(A)\le\max\{|a_{11}|+\cdots+|a_{1n}|,\dots,|a_{n1}|+\cdots+|a_{nn}|\}$$

$$\rho(A)\le\max\{|a_{11}|+\cdots+|a_{n1}|,\dots,|a_{1n}|+\cdots+|a_{nn}|\}$$

3.4. Valores singulares

Como aplicación, veremos en esta sección una descomposición de una matriz A, la llamada *descomposición a valores singulares*, de gran importancia práctica.
Sea A una matriz de m filas y n columnas a coeficientes reales.

Definición 3.4.1. Se llaman *valores singulares* de la matriz A a la raíz cuadrada (positiva) de los valores propios de la matriz A^tA.

La siguiente proposición nos da la relación entre el rango de una matriz y sus valores singulares.

Proposición 3.4.1. *El número de valores singulares no nulos de la matriz A coincide con el rango de dicha matriz.*

No es difícil probar las siguientes propiedades relativas a los valores singulares de una matriz.

Proposición 3.4.2. *(a) Si A es una matriz cuadrada, el cuadrado del producto de los valores singulares de la matriz A es igual al cuadrado del determinante de dicha matriz.*
(b) Los valores singulares no nulos de una matriz coinciden con los de su traspuesta.

Pasamos ahora a enunciar la descomposición a valores singulares de una matriz.

Teorema 3.4.1. *Dada una matriz A con m filas y n columnas, cuyos valores singulares son $s_1, \ldots, s_r$, existen matrices S y T tales que*

$$A = S \cdot \begin{pmatrix} s_1 & & \\ & \ldots & \\ & & s_r \\ 0 & \ldots & 0 \\ \ldots & & \ldots \\ 0 & \ldots & 0 \end{pmatrix} \cdot T \qquad si\ m > n,\ y$$

$$A = S \cdot \begin{pmatrix} s_1 & & & 0 & \ldots & 0 \\ & \ldots & & \ldots & & \ldots \\ & & s_r & 0 & \ldots & 0 \end{pmatrix} \cdot T \qquad si\ m < n$$

Ésta es la llamada *descomposición a valores singulares* de la matriz A.

Ejemplo 3.4.1. Los valores singulares de la matriz

$$A = \begin{pmatrix} 10 & 0{,}1 & 0{,}01 \\ 0{,}1 & 0{,}01 & 0 \\ 0{,}01 & 0 & 0 \end{pmatrix}$$

son: 10.00101090, 0.009000212019, 0.00001110972967.

3.5. Normas y cotas para valores propios

Norma

Consideremos un espacio vectorial E sobre el cuerpo K de los números reales o complejos. Una función $\| \ \|$ definida sobre E a valores en $\mathbb{R}$, se dice que es una norma sobre E si y sólo si verifica las siguientes condiciones

1. $\|x\| \geq 0$, para todo $x \in E$ y $\|x\| = 0$ si y sólo si $x = 0$
2. $\|\lambda x\| = |\lambda|\|x\|$, para todo $x \in E$ y $\lambda \in K$
3. $\|x + y\| \leq \|x\| + \|y\|$, para todo $x, y \in E$

Ejemplo 3.5.1. Sea $E = \mathbb{R}^n$, para todo $x = (x_1, \ldots, x_n)$

$$\|x\| = \sqrt{x_1^2 + \ldots x_n^2}$$

es una norma llamada *norma euclídea*.

Normas de matrices

El conjuto de matrices cuadradas es un espacio vectorial sobre el que se puede definir una norma. La posibilidad de multiplicar dos matrices $A, B \in M_n(K)$ nos plantea la pregunta sobre la relación entre las normas de cada una de las matrices y la norma del producto.

Definición 3.5.1. Diremos que una norma $\| \ \|$ definida sobre $M_n(K)$ es una norma de matriz (también llamada norma submultiplicativa) si y sólo si se verifica

$$\|AB\| \leq \|A\| \, \|B\|$$

Ejemplo 3.5.2. (Norma Euclídea o de Frobenius). En $M_n(K)$ la relación

$$\|A\| = \sqrt{\sum_{i,j=1}^{n} |a_{ij}|^2}, \quad A = (a_{ij})$$

es una norma de matriz.

Teorema 3.5.1. *Sea A una matriz cuadrada y $\rho(A)$ su radio espectral. Entonces*

$$\|A\| \geq \rho(A)$$

para cualquier norma de matriz.

Como ejemplo de matriz para la cual no se da la igualdad podemos considerar

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

dicha matriz es no nula por lo que su norma no puede ser nula, sin embargo el radio espectral es 0.

A menudo se presenta el problema de calcular la norma de un vector $x \in E$ que viene dado de la forma Ax, nos gustaría que se verificara que

$$\|Ax\|_v \leq \|A\| \|x\|_v$$

($\| \ \|_v$ denota la norma defiida sobre los vectores y $\| \ \|$ la norma definida sobre ls matrices).
Cuando esta condición se verifica para todo $x \in E$ y para toda matriz $A \in M_n(K)$, se dice que la norma del vector $\| \ \|_v$ y la norma de matriz $\| \ \|$ son *compatibles*.

Ejemplo 3.5.3. la norma de vector definida en 3.5.1 y la norma de matriz definida en 3.5.2 son compatibles.

3.6. Ejercicios Resueltos

1.- Sea f la aplicación lineal de $\mathbb{R}^6$ en $\mathbb{R}^6$ cuya matriz en la base canónica de $\mathbb{R}^6$, es

$$M = \begin{pmatrix} 1 & 2 & 1 & 0 & 2 & -5 \\ 0 & -1 & 0 & 0 & 0 & 0 \\ -1 & 3 & -1 & 0 & 8 & 9 \\ 0 & 0 & 0 & 2 & 0 & 0 \\ 0 & 0 & 0 & 0 & 4 & 12 \\ 0 & 0 & 0 & 0 & 1 & 3 \end{pmatrix}$$

Encontrar el polinomio característico.
Solución:

$$\begin{aligned} \det(M - tI) &= \begin{vmatrix} 1-t & 2 & 1 & 0 & 2 & -5 \\ 0 & -1-t & 0 & 0 & 0 & 0 \\ -1 & 3 & -1-t & 0 & 8 & 9 \\ 0 & 0 & 0 & 2-t & 0 & 0 \\ 0 & 0 & 0 & 0 & 4-t & 12 \\ 0 & 0 & 0 & 0 & 1 & 3-t \end{vmatrix} = \\ &= \begin{vmatrix} 1-t & 2 & 1 \\ 0 & -1-t & 0 \\ -1 & 3 & -1-t \end{vmatrix} \cdot (2-t) \cdot \begin{vmatrix} 4-t & 12 \\ 1 & 3-t \end{vmatrix} = \\ &= (-1-t) \begin{vmatrix} 1-t & 1 \\ -1 & -1-t \end{vmatrix} \cdot (2-t) \cdot \begin{vmatrix} 4-t & 12 \\ 1 & 3-t \end{vmatrix} \end{aligned}$$

Por lo tanto el polinomio característico es $Q(t) = (t+1)t^3(t-2)(t-7)$.

2.- Encontrar los valores propios del endomorfismo de $\mathbb{R}^4$ que viene definido por

$$f(x_1, x_2, x_3, x_4) = (x_1 + 5x_3, 2x_2 - x_4, -x_3 + 8x_4, 2x_4)$$

Solución:
Escribamos la matriz de la aplicación en la base canónica

$$A = \begin{pmatrix} 1 & 0 & 5 & 0 \\ 0 & 2 & 0 & -1 \\ 0 & 0 & -1 & 8 \\ 0 & 0 & 0 & 2 \end{pmatrix}$$

Esta matriz es triangular por lo que los valores propios son los valores de la diagonal.
Ésto es, los valores propios son: $-1, 1,$ y 2 de multiplicidad 2.

3.- Consideremos el endomorfismo de $\mathbb{R}^4$ definido por

$$\begin{gathered} f(e_1+e_2)=2(e_1-e_2) \\ f(e_1-e_2+e_3)=e_1-e_2+3e_3 \\ f(2e_2+e_3)=2e_1-2e_2+3e_3 \\ f(e_4)=0 \end{gathered}$$

Encontrar los valores propios de este endomorfismo, y, para cada uno de ellos, encontrar el subespacio de vectores propios.
Solución:
La matriz de la aplicación en la base $u_1=e_1+e_2, u_2=e_1-e_2+e_3, u_3=2e_2+e_3, u_4=e_4$ en el espacio de salida y la base canónica en el espacio de llegada es

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

Por lo que la matriz en la base canónica es:

$$A=\bar{A}S^{-1}$$

con

$$S=\begin{pmatrix} 1 & 1 & 0 & 0 \\ 1 & -1 & 2 & 0 \\ 0 & 1 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \quad \text{por lo tanto} \quad S^{-1}=\frac{1}{4}\begin{pmatrix} 3 & 1 & -2 & 0 \\ 1 & -1 & 2 & 0 \\ -1 & 1 & 2 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}$$

Haciendo pues el producto tenemos

$$A=\frac{1}{4}\begin{pmatrix} 5 & 3 & 2 & 0 \\ -5 & -3 & -2 & 0 \\ 0 & 0 & 12 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}$$

Calculemos el polinomio característico

$$\det(A-tI)=\begin{vmatrix} \frac{5}{4}-t & \frac{3}{4} & \frac{1}{2} & 0 \\ -\frac{5}{4} & -\frac{3}{4}-t & -\frac{1}{2} & 0 \\ 0 & 0 & 3-t & 0 \\ 0 & 0 & 0 & -t \end{vmatrix} =$$
$$(3-t)t\begin{vmatrix} \frac{5}{4}-t & \frac{3}{4} \\ -\frac{5}{4} & -\frac{3}{4}-t \end{vmatrix} = (3-t)t^2(\tfrac{1}{2}-t)$$

Valores propios $\dfrac{1}{2}$, 3 y 0 de multiplicidad 2.
Busquemos los subespacios de vectores propios
Para el valor propio 0:

$$\begin{aligned} \mathrm{Ker}A &= \{(x,y,z,t) \mid 5x+3y+2z=0, 12z=0,\} = \\ &= [(3,-5,0,0),(0,0,0,1)] \end{aligned}$$

Para el valor propio $\dfrac{1}{2}$:

$$\begin{aligned} \mathrm{Ker}(A-\frac{1}{2}I) &= \{(x,y,z,t) \mid 3x+3y+2z=0, \frac{5}{2}z=0, -\frac{1}{2}t=0\} = \\ &= [(1,-1,0,0)] \end{aligned}$$

Para el valor propio 3:

$$\begin{aligned} \mathrm{Ker}(A-3tI) &= \{(x,y,z,t) \mid -7x+3y+2z=0, -5x-15y-2z=0, -3t=0\} = \\ &= [(1,-1,5,0)]. \end{aligned}$$

4.- Hallar el radio espectral de la matriz

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

Solución:

$$\det(A-tI) = -(t-1)(t-1-i)(t-1+i)$$

de donde los valores propios son

$$\lambda_1 = 1,\ \lambda_2 = 1 + i,\ \lambda_2 = 1 - i$$

y

$$|\lambda_1| = 1,\ |\lambda_2| = |\lambda_3| = \sqrt{2}$$

Por lo que

$$\rho = \sqrt{2}$$

5.- Encontrar los valores singulares de la matriz

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

Solución:

$$A^tA = \begin{pmatrix} 1 & 2 & 1 \\ 1 & 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & 1 \\ 2 & 0 \\ 1 & 1 \end{pmatrix} = \begin{pmatrix} 6 & 2 \\ 2 & 2 \end{pmatrix}$$

$$\det(A^tA - tI) = t^2 - 8t + 8 = (t - 4 - 2\sqrt{2})(t - 4 + 2\sqrt{2})$$

Luego

$$\sigma_1 = \sqrt{4 + 2\sqrt{2}},\quad \sigma_2 = \sqrt{4 - 2\sqrt{2}}$$

3.7. Ejercicios Propuestos

1.- Determinar el polinomio característico de las matrices siguientes:

$$M = \begin{pmatrix} 2 & 1 & 0 & 1 \\ 1 & 2 & 0 & 1 \\ -1 & 1 & 3 & 1 \\ 2 & -2 & 0 & 4 \end{pmatrix}, \qquad N = \begin{pmatrix} 1 & 0 & -1 & 0 & 0 \\ 2 & -1 & -3 & 0 & 0 \\ 1 & 0 & -1 & 0 & 0 \\ 2 & 0 & 8 & 4 & 1 \\ -5 & 0 & 9 & 12 & 3 \end{pmatrix}$$

2.- Determinar los valores propios y un vector propio, para cada valor propio, de las matrices

$$A = \begin{pmatrix} -9 & 12 & 12 \\ 6 & -5 & -9 \\ 6 & -9 & 4 \end{pmatrix}, \qquad B = \begin{pmatrix} 2 & -1 & 1 \\ -1 & 2 & -1 \\ 1 & -1 & 2 \end{pmatrix}$$

3.- Encontrar el radio espectral de la matriz

$$A = \begin{pmatrix} 1 & 0 & 1 & 0 \\ 0 & 1 & 1 & 0 \\ 1 & 0 & 1 & 0 \\ 0 & 0 & 0 & 13 \end{pmatrix}$$

4.- Considerar el endomorfismo de $\mathbb{R}^4$ cuya matriz, en la base natural de $\mathbb{R}^4$, es:

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

Comprobar que los vectores $(1,0,0,1), (0,1,1,0), (1,0,0,-1), (0,1,-1,0)$ son vectores propios de f y forman una base de $\mathbb{R}^4$. Encontrar los valores propios asociados a estos vectores propios.

5.- Decir cuales de las siguientes matrices son diagonalizables y cuales son defectivas.

$$A = \begin{pmatrix} 1 & 1 & 1 \\ 0 & 1 & 1 \\ 0 & 0 & 1 \end{pmatrix}, \qquad B = \begin{pmatrix} 1 & 1 & 1 & \\ 0 & 2 & 1 & \\ & 0 & 0 & 3 \end{pmatrix},$$

$$C = \begin{pmatrix} 0{,}1 & 0{,}2 & 0{,}3 \\ 0{,}2 & 0{,}3 & 0{,}4 \\ 0{,}3 & 0{,}4 & 0{,}5 \end{pmatrix}, \qquad D = \begin{pmatrix} 1 & -1 & 1 \\ 1 & -1 & 1 \\ 0 & 0 & 2 \end{pmatrix}$$

6.- Encontrar los valores singulares de la matriz

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

7. Demostrar que la norma de Hölder definida sobre el espacio de matrices $M_n(K)$, de la manera

$$\|A\|_p = \left(\sum_{ij}^{n} |a_{ij}|^p \right)^{1/p}, \quad A = (a_{ij})$$

es una norma de matriz, si y sólo si $1 \leq p \leq 2$.

8. Probar que para cualquier norma de matriz, se tiene

$$\|I\| \geq 1, \quad \|A^n\| \leq \|A\|^n, \quad \|A^{-1}\| \geq \frac{1}{\|A\|} \ \ (\text{con } \det A \neq 0)$$

Capítulo 4

Matrices no negativas

4.1. Definiciones y propiedades

Definición 4.1.1. Una matriz $A \in M_{n\times m}(\mathbb{R})$ se dice que es *no negativa* (y lo notaremos por $A \geq 0$) si todos los elementos de A son no negativos. Esto es, la matriz

$$A = \begin{pmatrix} a_{11} & \dots & a_{1m} \\ \vdots & & \vdots \\ a_{n1} & \dots & a_{nm} \end{pmatrix}$$

es no negativa cuando $a_{ij} \geq 0$ para todo $i \in \{1, \dots, n\}$, $j \in \{1, \dots, m\}$.

Ejemplo 4.1.1. Consideramos las matrices

$$A_1 = \begin{pmatrix} 1 & 0 & 3 \\ 0 & 0 & 1 \\ 1 & 1 & 0 \end{pmatrix}, \qquad A_2 = \begin{pmatrix} -1 & 0 & -3 \\ -1 & 0 & -2 \end{pmatrix}, \qquad A_3 = \begin{pmatrix} -1 & 0 & 1 \\ 1 & 3 & -2 \\ 1 & 0 & -3 \end{pmatrix}$$

La matriz A_1 es no negativa. La matriz A_2 no lo es; sin embargo, lo es $-A_2$. Finalmente, observamos que ni A_3 ni $-A_3$ son matrices no negativas.

Ejemplo 4.1.2. Las matrices

$$v_1 = \begin{pmatrix} 1 & 1 & 0 & 1 \end{pmatrix} \in M_{1\times 4}(\mathbb{R}), \qquad v_2 = \begin{pmatrix} 1 \\ 1 \\ 0 \\ 1 \end{pmatrix} \in M_{4\times 1}(\mathbb{R})$$

son no negativas. La matriz v_1 se identifica con un vector fila y v_2 con un vector columna. En estos casos se habla de vectores fila no negativos y vectores columna no negativos (o, simplemente, de vectores no negativos si no hay lugar a confusión).

Definición 4.1.2. Una matriz $A \in M_{n\times m}(\mathbb{R})$ se dice que es *positiva* (y lo notaremos por $A > 0$) si todos los elementos de A son positivos. Esto es, la matriz

$$A = \begin{pmatrix} a_{11} & \dots & a_{1m} \\ \vdots & & \vdots \\ a_{n1} & \dots & a_{nm} \end{pmatrix}$$

se dice que es positiva cuando $a_{ij} > 0$ para todo $i \in \{1, \dots, n\}$, $j \in \{1, \dots, m\}$.

Ejemplo 4.1.3. La matriz

$$A = \begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 9 \end{pmatrix}$$

es positiva.

Ejemplo 4.1.4. Las matrices

$$v_1 = \begin{pmatrix} 1 & 2 & 1 & 2 \end{pmatrix} \in M_{1\times 4}(\mathbb{R}), \qquad v_2 = \begin{pmatrix} 1 \\ 2 \\ 1 \\ 3 \end{pmatrix} \in M_{4\times 1}(\mathbb{R})$$

son matrices positivas. Análogamente al ejemplo (4.1.2), diremos que v_1 es un vector fila positivo y v_2 un vector columna positivo (o, simplemente, vectores positivos si no hay lugar a confusión).

Operaciones con matrices no negativas y positivas

Designaremos por $MN_{n\times m}(\mathbb{R})$ al conjunto de matrices no negativas de orden $n\times m$ con coeficiente reales y por $MP_{n\times m}(\mathbb{R})$ al conjunto de matrices positivas de orden $n \times m$ con coeficiente reales. En el caso de matrices cuadradas de orden n, notaremos estos conjuntos simplemente por $MN_n(\mathbb{R})$ y $MP_n(\mathbb{R})$, respectivamente.

Observamos que, con la suma habitual de matrices, la suma de dos matrices de $MN_{n\times m}(\mathbb{R})$ (respectivamente, de dos matrices de $MP_{n\times m}(\mathbb{R})$) es una matriz que pertenece también a $MN_{n\times m}(\mathbb{R})$ (respectivamente, a $MP_{n\times m}(\mathbb{R})$).
Sin embargo, los conjuntos $MN_{n\times m}(\mathbb{R})$ y $MP_{n\times m}(\mathbb{R})$ no tienen "estructura de grupo", puesto que si $A \in MN_{n\times m}(\mathbb{R})$, entonces su simétrica respecto a la suma, la matriz $-A$, no pertenece a $MN_{n\times m}(\mathbb{R})$ y lo mismo ocurre en el caso del conjunto $MP_{n\times m}(R)$.
Finalmente, observamos que si $A \in MN_{n\times m}(\mathbb{R})$, $B \in MN_{m\times p}(\mathbb{R})$ (respectivamente, $A \in MP_{n\times m}(\mathbb{R})$, $B \in MP_{m\times p}(\mathbb{R})$) la matriz producto AB es una matriz no negativa (respectivamente, positiva). Esto es, $AB \in MN_{n\times p}(\mathbb{R})$ (respectivamente, $MP_{n\times p}(\mathbb{R})$).

Relaciones de orden

En los conjuntos $MN_{n\times m}(R)$ (respectivamente, $MP_{n\times m}(\mathbb{R})$) podemos definir las siguientes relaciones.

Definición 4.1.3. Dadas dos matrices A, B de $MN_{n\times m}(\mathbb{R})$ (respectivamente, de $MP_{n\times m}(\mathbb{R})$) diremos que A es mayor o igual que B (y lo notaremos por $A \geq B$) si y sólo si $a_{ij} \geq b_{ij}$ para todo $i \in \{1, \ldots, n\}$ y para todo $j \in \{1, \ldots, m\}$.
Si $A \geq B$ y $A \neq B$ diremos que A es mayor que B (y lo notaremos por $A > B$).
Finalmente, diremos que A es estrictamente mayor que B (y lo notaremos por $A \gg B$) si y sólo si $a_{ij} > b_{ij}$ para todo $i \in \{1, \ldots, n\}$ y para todo $j \in \{1, \ldots, m\}$.

Observaciones
1.- Estas tres relaciones son relaciones de orden, (es decir verifican las propiedades de 1- reflexividad, 2- antisimetría y transitividad).
2.- En ninguno de los tres casos el orden es total ya que fácilmente podemos encontrar dos matrices que no estén relacionadas. Por ejemplo, basta considerar las matrices

$$A = \begin{pmatrix} 2 & 4 \\ 3 & 5 \end{pmatrix} \quad \text{y} \quad B = \begin{pmatrix} 1 & 5 \\ 4 & 3 \end{pmatrix}$$

que no están relacionadas por ninguna de las tres relaciones descritas anteriormente.

3.- Si A es una matriz no negativa, entonces $A \geq 0$. Si, además, $A \neq 0$, entonces $A > 0$. Si A es una matriz positiva, entonces $A \gg 0$.

4.2. Matrices irreducibles

La noción de matriz reducible se da en general para matrices cuadradas. Aquí analizaremos las propiedades que se pueden deducir en el caso particular en que las matrices sean además no negativas o bien positivas.

Definición 4.2.1. Diremos que una matriz $A \in M_n(\mathbb{R})$ $(n \geq 2)$ es *reducible* cuando existe una matriz P cuyas columnas son una reordenación de las columnas de la matriz identidad (a una tal matriz la llamaremos matriz permutación) de modo que

$$PAP^{-1} = \begin{pmatrix} B & 0 \\ C & D \end{pmatrix}$$

donde las matrices B y D son matrices cuadradas.
En caso contrario se dice que la matriz es *irreducible*.

Observación 1.
Algunos autores definen las matrices reducibles como aquellas matrices A, para las cuales existe una matriz permutación P tal que

$$PAP^{-1} = \begin{pmatrix} D & C \\ 0 & B \end{pmatrix}$$

Nótese que ambas definiciones son equivalentes ya que

$$\begin{pmatrix} 0 & I_1 \\ I_2 & 0 \end{pmatrix} \begin{pmatrix} B & 0 \\ C & D \end{pmatrix} \begin{pmatrix} 0 & I_1 \\ I_2 & 0 \end{pmatrix}^{-1} = \begin{pmatrix} 0 & I_1 \\ I_2 & 0 \end{pmatrix} \begin{pmatrix} B & 0 \\ C & D \end{pmatrix} \begin{pmatrix} 0 & I_2 \\ I_1 & 0 \end{pmatrix} = \begin{pmatrix} D & C \\ 0 & B \end{pmatrix}$$

Observación 2. Puesto que la matriz P es una reordenación de las columnas de la matriz identidad se cumple que $P^{-1} = P^t$ (es decir, P es una matriz ortogonal).

Ejemplo 4.2.1. (a) La matriz

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

es reducible (basta considerar $P = I_n$).
(b) La matriz

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

es reducible. Para ello basta considerar

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

y comprobar que PAP^{-1} es de la forma deseada. En efecto,

$$PAP^{-1} = \begin{pmatrix} 1 & 2 & 0 \\ 1 & 0 & 0 \\ 2 & 1 & 1 \end{pmatrix}$$

Observamos que en este caso $B = \begin{pmatrix} 1 & 2 \\ 1 & 0 \end{pmatrix}$, $C = \begin{pmatrix} 2 & 1 \end{pmatrix}$ y $D = \begin{pmatrix} 1 \end{pmatrix}$.
(c) La matriz

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

es irreducible. En efecto, las posibles matrices permutación son, en este caso,

$$P_1 = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}, \quad P_2 = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{pmatrix}, \quad P_3 = \begin{pmatrix} 0 & 0 & 1 \\ 0 & 1 & 0 \\ 1 & 0 & 0 \end{pmatrix},$$

$$P_4 = \begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix}, \quad P_5 = \begin{pmatrix} 0 & 0 & 1 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix}, \quad P_6 = \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & 0 & 0 \end{pmatrix}$$

Con ellas obtenemos las matrices

$$P_1AP_1^{-1} = A, \quad P_2AP_2^{-1} = A^t, \quad P_3AP_3^{-1} = A^t,$$
$$P_4AP_4^{-1} = A^t, \quad P_5AP_5^{-1} = A, \quad P_6AP_6^{-1} = A$$

Teniendo en cuenta que una matriz permutación es una matriz de cambio de base que consiste en una reordenación de los vectores de la base canónica se tiene la siguiente descripción geométrica de una matriz reducible.

Observación 3. Se dice que dos matrices A y E son *cogredientes* si y sólo si existe una matriz permutación P tal que

$$PAP^{-1} = E$$

En particular se tiene que una matriz A es reducible si y sólo si es cogrediente a una matriz de la forma

$$\begin{pmatrix} B & 0 \\ C & D \end{pmatrix}$$

con B y D matrices cuadradas.
Es fácil probar que la cogrediencia es una relación de equivalencia.

Proposición 4.2.1. *Una matriz $A \in MN_n(\mathbb{R})$ es reducible si y sólo si existe un subespacio invariante de dimensión $r < n$ generado por r vectores de la base canónica de $\mathbb{R}^n$.*

Basándonos en esta proposición es fácil observar que la matriz A del ejemplo (4.2.1) c, anterior, es irreducible ya que el vector propio de valor propio 1 es $(1, 1, 1)$ que no pertenece a la base canónica: y el subespacio invariante de dimensión 2 es ortogonal a este vector, por lo que tampoco contiene ningún vector de la base canónica.

Ejemplo 4.2.2. La matriz, en la base canónica, de la aplicación lineal que deja invariante el subespacio $F = [(1, 0, 0, 0), (0, 0, 1, 0)]$ es reducible. En efecto, la matriz de la aplicación en la base $v_1 = (0, 1, 0, 0), v_2 = (0, 0, 0, 1), v_3 = (1, 0, 0, 0), v_4 = (0, 0, 1, 0)$ es de la forma

$$\begin{pmatrix} B & 0 \\ C & D \end{pmatrix}$$

Observación 4. Considerando el ejemplo anterior, la matriz de la aplicación, en la base canónica de $\mathbb{R}^4$ es de la forma

$$A = \begin{pmatrix} a_{11} & a_{12} & a_{13} & a_{14} \\ 0 & a_{22} & 0 & a_{24} \\ a_{31} & a_{32} & a_{33} & a_{34} \\ 0 & a_{42} & 0 & a_{44} \end{pmatrix}$$

Si permutamos las filas y las correspondientes columnas la matriz A se transforma en

$$PAP^{-1} = \begin{pmatrix} a_{22} & a_{24} & 0 & 0 \\ a_{42} & a_{44} & 0 & 0 \\ a_{12} & a_{14} & a_{11} & a_{13} \\ a_{32} & a_{34} & a_{31} & a_{33} \end{pmatrix}, \quad \text{con} \quad P = \begin{pmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{pmatrix}$$

Vemos que el número de elementos nulos de la matriz se mantiene. Esto ocurre en general, puesto que las transformaciones permitidas son permutaciones de filas y columnas. De aquí se deduce que toda matriz positiva es irreducible, si bien, tal como vemos en el caso de la matriz del ejemplo (4.2.1)-c, el que la matriz inicial tenga ceros no es condición suficiente para la reducibilidad.

Proposición 4.2.2. *Si A es una matriz reducible, cualquier potencia A^p, de dicha matriz, también es reducible.*

Demostración. Sea P la matriz permutación tal que

$$PAP^{-1} = \begin{pmatrix} B & 0 \\ C & D \end{pmatrix}$$

con B y D matrices cuadradas. Entonces

$$PA^pP^{-1} = (PAP^{-1})^p = \begin{pmatrix} B^p & 0 \\ E & D^p \end{pmatrix}$$

donde la matriz E es:
$E = CB + DC$, para $p = 2$
$E = CB^2 + DCB + D^2C$, para $p = 3$
etc.
Así pues, A^p es también reducible. □

Cualquier potencia A^p de una matriz no negativa A es también no negativa. Queremos ver que si la matriz A tiene suficientes elementos no nulos podemos, para un cierto p, conseguir que A^p sea positiva. Para ello necesitamos previamente la siguiente proposición.

Proposición 4.2.3. *Sea A una matriz no negativa irreducible de orden $n > 1$. Entonces*

$$(I + A)^{n-1} > 0.$$

Demostración. Sea $x \in \mathbb{R}^n$, $x > 0$, y consideremos

$$y = (I + A)x = x + Ax$$

Puesto que A es no negativa se cumple $Ax \geq 0$ y, por lo tanto, $y > 0$.
Si x no es positivo veamos que y tiene, por lo menos, un elemento nulo menos que el vector x. Para ello consideremos la matriz de permutación P tal que $Px = \begin{pmatrix} 0 \\ u \end{pmatrix}$, $u > 0$. Entonces

$$Py = Px + PAx = \begin{pmatrix} 0 \\ u \end{pmatrix} + PAP^{-1} \begin{pmatrix} 0 \\ u \end{pmatrix}$$

Partimos ahora la matriz PAP^{-1} según la partición de Py:

$$PAP^{-1} = \begin{pmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{pmatrix}$$

y $A_{12} \neq 0$, ya que A es irreducible.
Tenemos, pues,

$$Py = \begin{pmatrix} 0 \\ u \end{pmatrix} + \begin{pmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{pmatrix} \begin{pmatrix} 0 \\ u \end{pmatrix} = \begin{pmatrix} 0 \\ u \end{pmatrix} + \begin{pmatrix} A_{12}u \\ A_{22}u \end{pmatrix}$$

al ser $A_{12} \neq 0$ y $u > 0$ tenemos que $A_{12}u \neq 0$. Luego Py (y, por lo tanto, y) tiene, como mínimo, un cero menos que x.
Si y no es positivo, repetimos el proceso con la matriz $(I + A)y$. Observamos que a lo sumo tendremos que repetir este proceso $n - 1$ veces. Así pues, si aplicamos esto a los vectores x de la base canónica tenemos que $(I + A)^{n-1} > 0$, como queríamos demostrar. □

Ejemplo 4.2.3. Consideremos la matriz irreducible

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

$$(I + A)^2 = \begin{pmatrix} 1 & 1 & 2 \\ 2 & 1 & 1 \\ 1 & 2 & 1 \end{pmatrix} > 0$$

El siguiente corolario nos permite reducir el número de veces que tendríamos que repetir el proceso para conseguir la positividad de $I + A$.

Corolario 4.2.1. Si A es una matriz no negativa irreducible de orden n, entonces

$$(I + A)^{m-1} > 0$$

siendo m el grado del polinomio mínimo de A.

La siguiente proposición nos proporciona un resultado recíproco de éste.

Proposición 4.2.4. *Si $A \in M_n(\mathbf{R})$ es tal que $(I + A)^j > 0$, para un cierto j, entonces A es irreducible.*

Demostración. Si A fuese reducible existiría una matriz de permutación P tal que

$$PAP^{-1} = \begin{pmatrix} B & 0 \\ C & D \end{pmatrix}$$

Ahora bien,

$$P(I + A)P^{-1} = (I + PAP^{-1}) = \begin{pmatrix} I + B & 0 \\ C & I + D \end{pmatrix}$$

y como hemos visto en la proposición (4.2.2), cualquier potencia de esta matriz es de esta forma, por lo que ninguna potencia de $I + A$ podría ser positiva. □

Veamos ahora algunas propiedades espectrales importantes para matrices no negativas irreducibles.

Proposición 4.2.5. *Si* $0 \leq A \leq B$*, entonces*

$$\rho(A) \leq \rho(B)$$

Como corolario tenemos

Corolario 4.2.2. Si $0 \leq A$, entonces

$$\rho(A) > 0$$

Demostración. Si $0 < A$ existe $\varepsilon > 0$ tal que $\varepsilon I < A$ por lo que

$$\rho(A) \geq \rho(\varepsilon I) = \varepsilon > 0$$

□

Corolario 4.2.3. Sea B una matriz cuadrada obtenida de A eliminando algunas filas y columnas. Entonces

$$\rho(B) \leq \rho(A)$$

Teorema 4.2.1 (Perron)**.** *Una matriz positiva* A *tiene siempre un valor propio* λ_1 *que es simple, real, positivo e igual al radio espectral* $\lambda_1 = \rho(A)$*. Los demás valores propios tienen módulo estrictamente menor. Para este valor propio existe un vector propio con todas las componentes estrictamente positivas.*

Observación 5. Este teorema falla en general si $A \geq 0$ pero no positiva, así por ejemplo si consideramos la matriz

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

los valores propios son 2 y -2 ambos del mismo módulo e igual al radio espectral.

Este teorema fue generalizado por Frobenius al caso de matrices no negativas.

Teorema 4.2.2 (Frobenius)**.** *Una matriz no negativa e irreducible A tiene siempre un valor propio λ_1 que es simple, real, positivo e igual al radio espectral. Para este valor propio existe un vector propio con todas las componentes estrictamente positivas. Además, si existen otros valores propios $\lambda_2, \ldots, \lambda_h$, de módulo $\rho(A)$, entonces $\lambda_1, \lambda_2, \ldots \lambda_h$ son las h raíces distintas de la ecuación*

$$\lambda^h - \rho(A)^h = 0.$$

Proposición 4.2.6. *Sean $A \geq 0$ y $x > 0$. Si existen $\alpha, \beta \in \mathbb{R}$ tales que $\beta x \leq Ax \leq \alpha x$. Entonces*

$$\beta \leq \rho(A) \leq \beta$$

Ejemplo 4.2.4. Consideremos la matriz

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

Para $x = (1, 1, 1)$, tenemos

$$\begin{pmatrix} 3 & 1 & 0 \\ 1 & 3 & 1 \\ 0 & 1 & 3 \end{pmatrix} \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} = \begin{pmatrix} 4 \\ 5 \\ 4 \end{pmatrix}$$

por lo que

$$4x \leq Ax \leq 5x$$

y

$$4 \leq \rho(A) \leq 5$$

De hecho si calculamos los valores propios tenemos

$$\lambda_1 = 3,\ \lambda_2 = 3 + \sqrt{2},\ \lambda_3 = 3 - \sqrt{2}$$

y

$$4 \leq 3 + \sqrt{2} \leq 5$$

Si tomamos $x = (1{,}4, 2, 1{,}4)$ tenemos

$$4{,}2x \leq Ax \leq 4{,}8x$$

luego

$$4{,}2 \leq \rho(A) \leq 4{,}8$$

De hecho si calculamos los valores propios tenemos

$$\lambda_1 = 3,\ \lambda_2 = 3 + \sqrt{2},\ \lambda_3 = 3 - \sqrt{2}$$

y

$$4 \leq 4{,}2 \leq 3 + \sqrt{2} \leq 4{,}8 \leq 5$$

4.3. Matrices primitivas

En este apartado pretendemos clasificar las matrices no negativas en función del valor propio dominante.

Definición 4.3.1. Sea A una matriz no negativa e irreducible. Diremos que A es una matriz *primitiva* si y sólo si tiene un único valor propio de módulo igual a su radio espectral. En caso contrario se dice que es *imprimitiva*.

Ejemplo 4.3.1. Toda matriz positiva es primitiva.

Ejemplo 4.3.2. La matriz

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

es imprimitiva. En efecto, sus valores propios son $\lambda_1 = -\sqrt{2}$, $\lambda_2 = 0$, $\lambda_3 = \sqrt{2}$, y $\rho(A) = \sqrt{2} = |\lambda_1| = |\lambda_3|$.

Teorema 4.3.1. *Una matriz $A \in MN_n(\mathbb{R})$ es primitiva si y sólo si existe $p \in \mathbb{N}$ tal que $A^p \in MP_n(\mathbb{R})$.*

Demostración. Supongamos primero que $A^p > 0$, en cuyo caso A^p es una matriz irreducible.
Teniendo en cuenta la proposición (4.3.2), la matriz A es irreducible. Si la matriz A fuese imprimitiva existirían $\lambda_1, \ldots, \lambda_r$ con $\lambda_1 = |\lambda_2| = \ldots = |\lambda_r| = \rho(A)$, por lo que $\lambda_1^p, \ldots, \lambda_r^p$ serían valores propios de A^p con $\lambda_1^p = |\lambda_2^p| = \ldots = |\lambda_r^p| = \rho(A^p)$ lo que contradeciría el teorema de Perron que afirma que el valor propio de módulo igual al radio espectral es único para matrices positivas. Así pues, la matriz A es primitiva.
Se deja para el lector la demostración del recíproco. □

Observación. Si consideramos la matriz A del ejemplo (4.3.2), tenemos que

$$A^2 = \begin{pmatrix} 1 & 0 & 1 \\ 0 & 2 & 0 \\ 1 & 0 & 1 \end{pmatrix}, \quad \ldots, \quad A^{2m} = 2^{m-1}A^2, \quad A^{2m+1} = 2^m A, \ \forall m \geq 1$$

por lo que no existe ningún $p \in \mathbb{N}$ para el cual A^p sea positiva.

Corolario 4.3.1. Si A es una matriz primitiva, entonces A^p es también primitiva $\forall p > 0$.

Proposición 4.3.1. *Dada una matriz primitiva A, el número p para el cual $A^p > 0$ está acotado superiormente por $n^2 - 2n + 2$*

$$p \leq n^2 - 2n + 2$$

Ejemplo 4.3.3. Sea

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

En este caso la cota es $n^2 - 2n + 2 = 5$

$$A^2 = \begin{pmatrix} 0 & 0 & 1 \\ 1 & 1 & 0 \\ 0 & 1 & 1 \end{pmatrix} \geq 0, \quad A^3 = \begin{pmatrix} 1 & 1 & 0 \\ 0 & 1 & 1 \\ 1 & 1 & 1 \end{pmatrix} \geq 0$$

$$A^4 = \begin{pmatrix} 0 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 2 & 1 \end{pmatrix} \geq 0, \quad A^5 = \begin{pmatrix} 1 & 1 & 1 \\ 1 & 2 & 1 \\ 1 & 2 & 2 \end{pmatrix} > 0$$

Observemos que para esta matriz se alcanza la cota.
Si A es positiva obviamente $p = 1$.

Proposición 4.3.2. *Sea $A \geq 0$ una matriz primitiva tal que $\rho(A) = 1$. Entonces,*

$$\lim_{n\to\infty} A^n = xy^t$$

con $x, y > 0$ dos vectores tales que $< x, y >= 1$, $Ax = x$ y $A^t y = y$.

4.4. Matrices estocásticas

Definición 4.4.1. Una matriz $A = (a_{ij}) \in MN_n(\mathbb{R})$ se dice que es *estocástica* (por filas) si la suma de los elementos de cada fila de A valen 1. Es decir,

$$a_{ij} \geq 0, \qquad \sum_{j=1}^{n} a_{ij} = 1 \text{ para cada } i = 1, \ldots, n.$$

Ejemplo 4.4.1. La matriz

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

es estocástica.

Teorema 4.4.1. *Una matriz no negativa $A \in MN_n(\mathbb{R})$ es estocástica si y sólo si tiene el valor propio $\lambda = 1$ con vector propio asociado $u = \begin{pmatrix} 1 \\ \vdots \\ 1 \end{pmatrix}$.*

Demostración. Si A es una matriz estocástica se tiene

$$Au = \begin{pmatrix} \sum_{j=1}^{n} a_{1j} \\ \vdots \\ \sum_{j=1}^{n} a_{nj} \end{pmatrix} = u,$$

luego u es vector propio de valor propio $\lambda = 1$.
Recíprocamente. Si $Au = u$, entonces $\sum_{j=1}^{n} a_{ij} = 1$ para cada $i \in \{1, \ldots, n\}$. □

Teorema 4.4.2. *Si una matriz no negativa $A \in MN_n(\mathbb{R})$ es estocástica, entonces su radio espectral es 1.*

Demostración. Por el teorema anterior sabemos que 1 es valor propio de A, por lo que es raíz de su polinomio característico $p(\lambda) = \det(\lambda I - A)$. Basta ahora comprobar que la función $p(\lambda)$ es creciente a partir de 1, esto es, comprobar que $p'(\lambda) > 0$.
También se puede demostrar por reducción al absurdo de la siguiente manera Supongamos por ejemplo que $\lambda > 1$ es valor propio, y $v = (v_1, \ldots, v_n)$ un vector propio asociado ($Av = \lambda v$), que sin pérdida de generalidad podemos suponer con componentes de suma la unidad, entonces

$$1 < \lambda = \sum_{i=1}^{n} (\lambda v)_i = \sum_{i=1}^{n} (Av)_i = \sum_{i=1}^{n} \sum_{j=1}^{n} a_{ij} v_j = \sum_{j=1}^{n} (\sum_{i=1}^{n} aij) v_j = \sum_{j} v_j = 1$$

□

Teorema 4.4.3. *Sea $A \in MN_n(\mathbb{R})$ con valor propio maximal λ. Si para este valor propio, existe un vector propio $x = (x_1, \ldots, x_n) > 0$, entonces llamando*

$$X = \begin{pmatrix} x_1 & & \\ & \ddots & \\ & & x_n \end{pmatrix},$$

se tiene que

$$A = \lambda X P X^{-1}$$

Demostración. Basta probar que

$$\frac{1}{\lambda} X^{-1} A X$$

es estocástica.
Observamos que la condición $x > 0$ excluye la posibilidad de que $\lambda = 0$. □

Consideremos ahora una matriz A estocástica y al sucesión de potencias de dicha matriz

$$A, A^2, A^3, \ldots$$

y nos preguntamos por la existencia de límite de dicha sucesión. Tenemos el siguiente resultado

Teorema 4.4.4. *Si A es una matriz irreducible y estocástica, existe* $\lim_{n\to\infty} A^n$, *si y sólo si, A es primitiva.*

Definición 4.4.2. Una matriz $A = (a_{ij}) \in MN_n(\mathbb{R})$ se dice que es *doblemente estocástica* si la suma de los elementos de cada fila y de cada columna de A valen 1. Es decir,

$$a_{ij} \geq 0, \quad \sum_{j=1}^{n} a_{ij} = 1, \quad \sum_{i=1}^{n} a_{ij} = 1 \text{ para todo } i \in \{1, \ldots, n\}, j \in \{1, \ldots n\}.$$

Ejemplo 4.4.2. La matriz

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

es doblemente estocástica.

Matriz de normalización

Vamos a ver ahora como a partir de una matriz positiva podemos construir una matriz estocástica.
Dada una matriz $A = (a_{ij})$ positiva, llamaremos matriz de *normalización*, que notaremos por ΣA, a la matriz

$$\Sigma A = \begin{pmatrix} a_{11} + \ldots + a_{1n} & 0 & \ldots & 0 \\ 0 & a_{21} + \ldots + a_{2n} & \ldots & 0 \\ & & \ddots & \\ 0 & 0 & \ldots & a_{n1} + \ldots + a_{nn} \end{pmatrix}.$$

Observamos que la matriz ΣA es siempre inversible.

Definición 4.4.3. Dada $A \in M_n(\mathbb{R})$ positiva, llamaremos *matriz normalizada de A*, que notaremos por $N(A)$, a la matriz

$$N(A) = (\Sigma A)^{-1} A.$$

La matriz normalizada es tal que la suma de los elementos de cada fila vale 1. Esto es, $N(A)$ es una matriz *estocástica*.

Observación. Se puede generalizar esta definición a matrices no negativas. En este caso no siempre se puede asegurar que la matriz ΣA sea inversible. Veámoslo con un ejemplo. Consideramos la matriz

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

y su correspondiente "matriz de normalización",

$$\Sigma A = \begin{pmatrix} 3 & 0 \\ 0 & 0 \end{pmatrix}$$

Fijamos un valor no nulo $\ell \in \mathbb{R}^+$,

$$\overline{\Sigma A} = \begin{pmatrix} 3 & 0 \\ 0 & \ell \end{pmatrix}$$

con lo que $\overline{\Sigma A}$ es ahora inversible. La "matriz normalizada"de A sería

$$N(A) = (\overline{\Sigma A})^{-1} A = \begin{pmatrix} \frac{1}{3} & 0 \\ 0 & \frac{1}{\ell} \end{pmatrix} \begin{pmatrix} 1 & 2 \\ 0 & 0 \end{pmatrix} = \begin{pmatrix} \frac{1}{3} & \frac{2}{3} \\ 0 & 0 \end{pmatrix}$$

La matriz $N(A)$ en este caso es tal que la suma de los elementos de todas sus filas no nulas es igual a 1.

4.5. Matrices totalmente no negativas

Consideramos ahora matrices que son, no solamente no negativas, sino tales que todos sus menores (de cualquier orden) son también no negativos. Estas matrices tienen importantes aplicaciones en la teoría de pequeñas oscilaciones de sistemas elásticos.

Definición 4.5.1. Una matriz $A \in M_{n\times m}(\mathbb{R})$ se dice que es *totalmente no negativa* si todos los menores de cualquier orden son no negativos.

Una consecuencia inmediata de la definición es que toda matriz totalmente no negativa es no negativa (los menores de orden 1 son no negativos).

Ejemplo 4.5.1. Dadas las matrices

$$A_1 = \begin{pmatrix} 1 & 2 \\ 1 & 2 \end{pmatrix}, \qquad A_2 = \begin{pmatrix} 1 & 2 \\ 2 & 1 \end{pmatrix}$$

observamos que A_1 es una matriz totalmente no negativa, pero A_2 no es totalmente no negativa puesto que $\det A_2 = -3 < 0$.

Definición 4.5.2. Una matriz $A \in M_{n\times m}(\mathbb{R})$ se dice que es *totalmente positiva* si todos los menores de cualquier orden son positivos.

Una consecuencia inmediata de la definición es que toda matriz totalmente positiva es positiva (los menores de orden 1 son positivos).

Ejemplo 4.5.2. Las matrices

$$A_1 = \begin{pmatrix} 1 & 2 \\ 1 & 3 \end{pmatrix}, \qquad A_2 = \begin{pmatrix} 1 & 1 & 1 \\ 2 & 3 & 4 \\ 4 & 9 & 16 \end{pmatrix}$$

son totalmente positivas. Sin embargo las matrices positivas del ejemplo anterior no son totalmente positivas.

4.6. Ejercicios Resueltos

1.- Estudiar si son o no reducibles las siguientes matrices.

$$A = \begin{pmatrix} 1 & 0 & 2 \\ 0 & 3 & 0 \\ 2 & 1 & 3 \end{pmatrix}, \quad B = \begin{pmatrix} 2 & 1 & 1 \\ 1 & 2 & 1 \\ 1 & 1 & 2 \end{pmatrix}$$

Solución: Analicemos la matriz A. Observamos que el subespacio $F = [e_1, e_3]$ es invariante por la matriz A ya que $A(e_1 = e_1 + 2e_3 \in [e_1, e_3]$, $A(e_3) = 2e_1 + 3e_3 \in [e_1, e_3]$.

De hecho tenemos

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

por lo que A es reducible.
Pasemos a analizar la matriz B.
Es fácil observar que ninguno de los subespacios $[e_1]$, $[e_2]$, $[e_3]$,$[e_1, e_2]$,$[e_1, e_3]$, $[e_2, e_3]$ son invariantes, por lo que la matriz B es irreducible.

2.- Estudiar si son o no primitivas las matrices

$$C = \begin{pmatrix} 4 & 2 & 2 \\ 2 & 4 & 2 \\ 2 & 2 & 4 \end{pmatrix}, \quad D = \begin{pmatrix} 0 & 1 & 1 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{pmatrix}$$

Solución
Estudiemos la matriz C.
Primero observamos que la matriz es no negativa ($a_{ij} > 0$, $\forall i, j$) y es irreducible ($C = 2B$ siendo B la matriz del ejercicio anterior). Los valores propios de la matriz son 2 doble y 8 simple, claramente el radio espectral es 8 por lo que la matriz es primitiva.
Pasemos ahora a analizar la matriz D
Observamos que

$$(I_3 + D)^2 = \begin{pmatrix} 1 & 2 & 3 \\ 0 & 1 & 2 \\ 0 & 0 & 1 \end{pmatrix}$$

es triangular, por tanto $(I_3 + D)^{n-1}$ no es positiva. Por lo que D no puede ser primitiva.
3.- Comprobar el Teorema de Frobenius para la matriz

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

Solución
Observamos que la matriz es no negativa y primitiva. Sus valores propios son las tres raíces de la unidad, por lo tanto el radio espectral es 1 e igual a la raíz real (simple).

4.- ¿Para qué valores de a, b, c, d la siguiente matriz es estocástica?

$$A = \begin{pmatrix} a & b & c & d \\ d & b & a & c \\ c & a & d & b \\ b & c & d & a \end{pmatrix}$$

¿Para qué valors es doblemente estocástica?
Solución:
La matriz es estocástica si y sólo si

$$a + b + c + d = 1, \quad a, b, c, d \geq 0$$

Para que sea doblemente estocástica tiene además que verificarse que

$$2b + a + c = 1, \quad \text{y} \quad a + c + 2d = 1$$

por lo que $b = d$.

4.7. Ejercicios Propuestos

1.- Sea A una matriz cuadrada no negativa de $MN_n(\mathbb{R})$. Probar que si x, y son dos vectores de $\mathbb{R}^n$ tales que $x \geq y$, entonces se cumple $Ax \geq Ay$.
2.- Sean $A, B \in MN_n(\mathbb{R})$ tales que $A \geq B$. Probar que $\rho(A) \geq \rho(B)$.
3.- Sea A una matriz cuadrada no negativa. Demostrar que si existe un vector x tal que $Ax \leq \lambda x$ entonces $\lambda \geq \rho(A)$.
4.- Sea A una matriz cuadrada no negativa. Demostrar que $(\lambda I - A)^{-1}$ es no negativa si y sólo si $\lambda > \rho(A)$.
5.- Sea $A \in MN_n(\mathbb{R})$ irreducible. Probar que para todo $\varepsilon > 0$ la matriz $\varepsilon I_n + A$ es primitiva.
6.- Estudiar si son, o no, primitivas las siguientes matrices

$$A_1 = \begin{pmatrix} \frac{9}{10} & \frac{1}{10} & 0 & 0 \\ 0 & \frac{7}{10} & \frac{2}{10} & \frac{1}{10} \\ \frac{1}{10} & 0 & \frac{8}{10} & \frac{1}{10} \\ \frac{1}{10} & \frac{1}{10} & 0 & \frac{8}{10} \end{pmatrix}, \qquad A_2 = \begin{pmatrix} 0 & \frac{3}{4} & \frac{1}{4} \\ \frac{1}{2} & 0 & \frac{1}{2} \\ \frac{2}{3} & \frac{1}{3} & 0 \end{pmatrix}$$

7.- Sean $A, B \in MN_n(\mathbb{R})$ dos matrices estocásticas. Probar que la matriz

$$\lambda A + (1 - \lambda)B, \quad \text{con } 0 \leq \lambda \leq 1$$

es también estocástica.

8.- Sean $A, B \in MN_n(\mathbb{R})$ dos matrices doblemente estocásticas. ¿Es AB doblemente estocástica?

9.- Probar que la matriz

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

es doblemente estocástica.

10.- Decir si son o no totalmente no negativas las siguientes matrices.

$$A = \begin{pmatrix} 2 & 1 & 0 \\ 1 & 2 & 1 \\ 0 & 1 & 2 \end{pmatrix}, \qquad B = \begin{pmatrix} 3 & 1 & 3 \\ 1 & 2 & 1 \\ 3 & 5 & 7 \end{pmatrix}$$

Capítulo 5

Cadenas de Markov finitas

5.1. Cadenas de Markov a variable discreta

Consideremos un sistema físico que puede existir solamente en uno de un número finito de estados y tal que su estado puede evolucionar sólo en puntos discretos del tiempo. Sea $\{X_0, X_1, \ldots, X_k, \ldots\}$ el conjunto de estados. Utilizamos la expresión $X_k = j$ para indicar que el proceso está en estado j en el tiempo k.
Suponemos que hay una probabilidad (que puede o no depender del tiempo k), de que un sistema que está en el estado j en el tiempo k evolucione hacia el estado i en el tiempo $k+1$. Esta restricción permite utilizar un modelo matemático (matricial) que estudia la evolución a lo largo del tiempo y que vamos a tratar a continuación.

Definición 5.1.1. Una *cadena de Markov* es un proceso aleatorio sin memoria.

Esto es, un sistema que evoluciona en el tiempo ($k = 0, 1, \ldots$), está en el estado i_k en el instante k si en el instante $k-1$ está en el estado i_{k-1}:

$$P(X_k = i_k \mid X_0 = i_0, \ldots, X_{k-1} = i_{k-1}) = \\ P(X_k = i_k \mid X_{k-1} = i_{k-1}) = p_{i_{k-1}, i_k}$$

Intuitivamente, esta ecuación implica que conocido el estado "presente"de un sistema, el estado "futuro.es independiente de su estado "pasado".
Tal y como decíamos al principio vamos a describir matricialmente este proceso. Sean $p_1(k)$, ..., $p_n(k)$ las probabilidades de los n estados en el tiempo k. Sean $p_{ij}(k)$ la probabilidad de la evolución del estado i al j en el tiempo k, $1 \leq i, j \leq n$. Entonces

$$\left.\begin{aligned} p_1(k+1) &= p_{11}(k)p_1(k) + \ldots + p_{1n}(k)p_n(k) \\ &\vdots \\ p_n(k+1) &= p_{n1}(k)p_1(k) + \ldots + p_{nn}(k)p_n(k) \end{aligned}\right\}$$

lo que puede expresarse matricialmente de la forma:

$$\begin{pmatrix} p_1(k+1) \\ \vdots \\ p_n(k+1) \end{pmatrix} = \begin{pmatrix} p_{11}(k) & \ldots & p_{1n}(k) \\ \vdots & & \vdots \\ p_{n1}(k) & \ldots & p_{nn}(k) \end{pmatrix} \begin{pmatrix} p_1(k) \\ \vdots \\ p_n(k) \end{pmatrix}$$

Definición 5.1.2. Llamaremos *matriz de transición de estados* a la matriz:

$$A(k) = \begin{pmatrix} p_{11}(k) & \ldots & p_{1n}(k) \\ \vdots & & \vdots \\ p_{n1}(k) & \ldots & p_{nn}(k) \end{pmatrix}$$

Observación 1.

$$p(k+1) = A(k)p(k)$$

Definición 5.1.3. Si $p_{ij}(k)$ no depende de k para todos los pares (i, j), $1 \leq i, j \leq n$, diremos que la cadena es *homogénea*.

Ejemplo 5.1.1. Consideremos una red de comunicación que consiste en una sucesión de estados de canales de comunicación binarios, donde $\begin{pmatrix} x_n \\ y_n \end{pmatrix}$ representa el dígito saliente del estado n-ésimo del sistema. y $\begin{pmatrix} x_0 \\ y_0 \end{pmatrix}$ el dígito

entrante del primer estado. Supongamos que los canales de comunicación binarios son estocásticamente idénticos. La red de comunicación se representa mediante la siguiente relación:

$$x_n = (1-a)x_{n-1} + by_{n-1}$$
$$y_n = ax_{n-1} + (1-b)y_{n-1}$$

por lo que la matriz de transición de estados es

$$A = \begin{pmatrix} 1-a & b \\ a & 1-b \end{pmatrix}$$

De ahora en adelante supondremos que las cadenas son homogéneas. En dicho caso tenemos la siguiente proposición.

Proposición 5.1.1. *Sea* $p(0) = (p_1(0), \ldots, p_n(0))$ *una probabilidad inicial. Entonces*

$$p(k) = A^k p(0).$$

Demostración. Vamos a hacer una demostración por inducción. Comprobemos, en primer lugar, que el enunciado es cierto para $k = 1$ y $k = 2$:

$$p(1) = Ap(0),\ p(2) = Ap(1) = AAp(0) = A^2p(0)$$

Supongamos ahora que $p(k-1) = A^{k-1}p(0)$. Veamos que el enunciado es cierto para k:

$$p(k) = Ap(k-1) = AA^{k-1}p(0) = A^kp(0)$$

□

Observación 2. Puesto que

$$p_{11} + \ldots + p_{n1} = 1$$
$$\vdots$$
$$p_{1n} + \ldots + p_{nn} = 1$$

y la matriz A es no negativa vemos que A^t es una matriz estocástica. La matriz A recibe el nombre de *matriz de Markov*.

Comportamiento asintótico

Consideremos una cascada de comunicación binaria libre de error. Esto es, la matriz de transición de estados es

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

por lo que

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

En otras palabras, la cadena nunca cambia de estado y un dígito transmitido es recibido correctamente después de un número arbitrario de estados. Obviamente, esto no siempre sucede así. Nos interesa conocer la evolución a lo largo del tiempo de una cadena de Markov.

Definición 5.1.4. Llamaremos *distribución estacionaria* de una cadena de Markov en caso de existir a

$$\pi = \lim_{k\to\infty} p(k)$$

Observación. Si la cadena de Markov es homogénea (como los casos que estamos considerando) la existencia de π equivale a la existencia de $\lim_{k\to\infty} A^k$.

$$\lim_{k\to\infty} p(k) = \lim_{k\to\infty} A^k p(0) = (\lim_{k\to\infty} A^k)p(0)$$

Proposición 5.1.2. *Si A es irreducible entonces exite π. Además, si A es primitiva entonces π es independiente de la distribución de probabilidad $p(0)$ inicial. De hecho,*

$$\pi = \frac{1}{\sum v_i} \begin{pmatrix} v_1 \\ \vdots \\ v_n \end{pmatrix}$$

siendo $\begin{pmatrix} v_1 \\ \vdots \\ v_n \end{pmatrix}$ *un vector propio de valor propio 1.*

Demostración. El teorema de Perron-Frobenius nos asegura que 1 es el valor propio igual al radio espectral y es el único con dicho módulo en caso de ser A primitiva. También nos asegura la existencia de un vector propio no negativo para dicho valor propio. □

El hecho de que la matriz A sea irreducible se traduce, en términos de cadenas de Markov, en que cada estado puede ser alcanzado por otro estado mediante un número finito de pasos. Esto es, existe $n \geq 0$ para el cual $p_{ij}(n) > 0$. Las cadenas que verifican esta condición reciben el nombre de *cadenas de Markov irreducibles o ergódicas*. Las cadenas de Markov en que la matriz de transición de estados es primitiva reciben el nombre de *regulares*.
El hecho de que la cadena de Markov no sea primitiva no impide que exista distribución estacionaria, si bien ésta dependerá de la distribución inicial.

Ejemplo 5.1.2. Consideremos una cascada de comunicación binaria tan ruidosa que el dígito transmitido siempre es erróneamente interpretado. Esto es, la matriz de transición de estados es

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

por lo que

$$A^k = \begin{cases} \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \text{ si } k \text{ es par} \\ \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \text{ si } k \text{ es impar} \end{cases}$$

Observamos que no existe $\lim_{k\to\infty} A^k$. Sin embargo, para la distribución de probabilidad inicial $p(0) = \begin{pmatrix} 0{,}5 \\ 0{,}5 \end{pmatrix}$ se tiene que $\lim_{k\to\infty} A^k p(0) = p(0)$. (Nótese que $p(0)$ es el único vector propio de valor propio 1 de A con suma de componentes igual a 1).

Definición 5.1.5. Diremos que un estado i en una cadena de Markov es *absorbente* si, cuando ocurre en una repetición del experimento entonces ocurre en cada repetición posterior. Es decir, si $p_{ii} = 1$ y $p_{ij} = 0$ para $i \neq j$. Una cadena de Markov se dice que es *absorbente* si existe algún estado absorbente. En una cadena de Markov absorbente, los estados que no lo son se denominan *transitorios*.

Ejemplo 5.1.3. Consideremos un sistema aleatorio cuya matriz de transición de estados viene dada por

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

En este caso los estados 4 y 5 son absorbentes y los estados 1,2,3 son transitorios.

Dada una cadena de Markov absorbente siempre podemos renombrar los estados de forma que la matriz de transición de estados sea de la forma

$$A = \begin{pmatrix} T & 0 \\ Q & I \end{pmatrix}$$

siendo I la matriz identidad de orden el número de estados absorbentes. En el ejemplo anterior,

$$A = \begin{pmatrix} T & 0 \\ Q & I_2 \end{pmatrix} \quad \text{con} \quad T = \begin{pmatrix} 0 & \frac{1}{2} & 0 \\ \frac{1}{2} & 0 & \frac{1}{2} \\ 0 & \frac{1}{2} & 0 \end{pmatrix},\ Q = \begin{pmatrix} \frac{1}{2} & 0 & 0 \\ 0 & 0 & \frac{1}{2} \end{pmatrix}.$$

Si partimos el vector de probabilidad $p(k) = \begin{pmatrix} p_1(k) \\ \vdots \\ p_n(k) \end{pmatrix}$ en dos vectores según los órdenes de las matrices T e I, respectivamante, tenemos que $p(k) = \begin{pmatrix} u(k) \\ v(k) \end{pmatrix}$ con lo que

$$\begin{pmatrix} u(k+1) \\ v(k+1) \end{pmatrix} = \begin{pmatrix} T & 0 \\ Q & I \end{pmatrix} \begin{pmatrix} u(k) \\ v(k) \end{pmatrix}$$

y

$$\begin{aligned} u(k+1) &= Tu(k) \\ v(k+1) &= v(k) + Qu(k) \end{aligned}$$

Por consiguiente,

$$u(k) = T^k u(0)$$
$$v(k) = v(k-1) + QT^{k-1}u(0) = v(0) + \sum_{r=0}^{k-1} QT^r u(0)$$

En el caso particular en que $\rho(T) < 1$ entonces

$$\lim_{k\to\infty} T^k = 0$$
$$\sum_{r=0}^{\infty} T^r = \lim_{k\to\infty} \sum_{r=0}^{k} T^r = (I-T)^{-1}$$

y

$$\lim_{k\to\infty} v(k) = v(0) + Q(I-T)^{-1}u(0).$$

Así en el ejemplo 5.1.3, tenemos que

$$\lim_{k\to\infty} v(k) = \begin{pmatrix} \frac{1}{2} & 0 & 0 \\ 0 & 0 & \frac{1}{2} \end{pmatrix} \begin{pmatrix} \frac{3}{2} & 1 & \frac{1}{2} \\ 1 & 2 & 1 \\ \frac{1}{2} & 1 & \frac{3}{2} \end{pmatrix} \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix} = \begin{pmatrix} \frac{1}{2} \\ \frac{1}{2} \end{pmatrix}$$

Proposición 5.1.3. *Si A es la matriz de transición de estados de una cadena de Markov absorbente tal que la matriz $T \in M_r(\mathbb{R})$ es irreducible y Q es no nula, entonces $\lim_{k\to\infty} T^k = 0$ y, por lo tanto, existe $\lim_{k\to\infty} v(k)$.*

Demostración. Por ser $Q \neq 0$, se tiene $\sum_{i=1}^{r} t_{ij} \leq 1$. La desigualdad es estricta al menos en el caso de una columna.
Por otra parte, al ser la matriz T irreducible se tiene que $\rho(T)$ es tal que

$$s = \min_j \left(\sum_{i=1}^{r} t_{ij} \right) \leq \rho(T) \leq \max_j \left(\sum_{i=1}^{r} t_{ij} \right) = S \leq 1$$

En particular, $\rho(T) \leq 1$. Además $s = \rho(T)$ o $S = \rho(T)$ si y sólo si $s = S$ en cuyo caso $\rho(T) < 1$ ya que $s < 1$. Luego siempre, se cumple que $\rho(T) < 1$. □

Análisis moderno de las cadenas de Markov

Recientes estudios acerca de las cadenas de Markov homogéneas muestran cómo podemos obtener información acerca de dichas cadenas mediante la inversa generalizada de una cierta matriz asociada a la matriz de transición que define la cadena.
Dada una matriz de transición de estados A de una cadena de Markov, observamos que la matriz $B = I - A$ es tal que $b_{ii} \geq 0$ y $b_{ij} \leq 0$. Las matrices B tales que pueden expresarse como $B = sI - A$ con A no negativa y $b_{ii} \geq 0$, $b_{ij} \leq 0$ para un cierto $s \leq \rho(A)$, reciben el nombre de *M-matrices*.

Proposición 5.1.4. *Sea A una matriz de transición de estados de una cadena de Markov. Entonces, existe* $\lim_{k\to\infty} A^k$ *y dicho límite vale*

$$L = I - BB^{+}.$$

(Una M-matriz B que verifica la condición de convergencia "existe $\lim_{k\to\infty} \left(\frac{1}{s}A\right)^k$*", recibe el nombre de M-matriz con propiedad c).*

Demostración. Por ser $A = (a_{ij})$ una matriz de transición de estados se tiene que $0 \leq a_{ij} \leq 1$. Luego $b_{ii} \geq 1$ y $b_{ij} \leq 0$ si $i \neq j$. □

Observación. No todas las M-matrices verifican la condición c. Veámoslo con un ejemplo. Sea B la matriz $\begin{pmatrix} 0 & -1 \\ 0 & 0 \end{pmatrix}$. Entonces B puede representarse como $B = sI - A$ con $s > 0$ y $A = \begin{pmatrix} s & 1 \\ 0 & s \end{pmatrix}$. Puesto que A es no negativa, B es una M-matriz. Por otra parte, dada una matriz T existe $\lim_{k\to\infty} T^k$ si y sólo si $\rho(T) < 1$. En nuestro caso, para todo $s > 0$, $\rho\left(\frac{1}{s}A\right) = 1$, por lo que no puede verificarse la condición c.
También cabe destacar que si B es una M-matriz con $s > \max_i b_{ii}$ entonces verifica la condición c. Así por ejemplo, sea $B = \begin{pmatrix} 1 & -1 \\ -1 & 1 \end{pmatrix}$. Entonces la sucesión de término general

$$\left(\frac{1}{s}A\right)^k = \left(\frac{1}{s}\begin{pmatrix} s-1 & 1 \\ 1 & s-1 \end{pmatrix}\right)^k$$

no es convergente para $s = 1$. Sin embargo, sí lo es para cualquier $s > 1$.

5.2. Ejercicios Resueltos

1.- Sea una cadena de Markov a dos variables cuya matriz de transición de estados es

$$P = \begin{pmatrix} 0{,}5 & 0{,}3 \\ 0{,}5 & 0{,}7 \end{pmatrix}$$

Calcular $\lim_{k\to\infty} p(k)$.
Solución:
La cadena de Markov es homogénea, por lo que hemos de calcular

$$\lim k \to \infty P^k$$

Tenemos que

$$P = SDS^{-1}$$

con

$$S = \begin{pmatrix} 3 & 1 \\ 5 & -1 \end{pmatrix}, \quad S^{-1} = \frac{1}{8}\begin{pmatrix} 1 & 1 \\ 5 & -3 \end{pmatrix}, \quad D = \begin{pmatrix} 1 & \\ & 0{,}2 \end{pmatrix}$$

Por lo que

$$\lim_{k\to\infty} P^k = S(\lim_{k\to\infty} D^k)S^{-1}p(0)$$

Calculemos

$$\lim_{k\to\infty} D^k = \lim_{k\to\infty} \begin{pmatrix} 1^k & \\ & 0{,}2^k \end{pmatrix} = \begin{pmatrix} 1 & \\ & 0 \end{pmatrix}$$

Finalmente

$$\lim_{k\to\infty} P^k = \frac{1}{8}\begin{pmatrix} 3 & 1 \\ 5 & -1 \end{pmatrix}\begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}\begin{pmatrix} 1 & 1 \\ 5 & -3 \end{pmatrix}p(0) = \frac{1}{8}\begin{pmatrix} 3 & 3 \\ 5 & 5 \end{pmatrix}p(0)$$

2.- Una ciudad tiene tres supermercados X, Y, Z. Considerando un determinado periodo de tiempo observamos que por diferentes razones como precio, calidad,..., algunos habitantes deciden cambiar de cadena. Deseamos expresar en un modelo matemático y analizar el movimiento de clientes de una cadena a otra suponiendo que la proporción de clientes que cambian al día de supermercado se mantiene constante durante un mes.
Aplicarlo al caso en que la proporción de clientes de X, Y, Z el 31 de Diciembre es $(0{,}2, 0{,}3, 0{,}5) = u(0)$ respecto da la población total y que la proporción de clientes que permanecen en el supermercado o que cambian de uno a otro es:

que se mantiene en X 0.8, que pasa de Y, Z a X es 0.2, 0.1 resp.
que se mantiene en Y 0.7, que pasa de X, Z a Y es 0.1, 0.3 resp.
que se mantiene en Z 0.6, que pasa de X, Y a Z es 0.1, 0.1 resp.

Solución:

Supongamos que la proporción de clientes de X, Y, Z el 31 de Diciembre es $(x_0, y_0, z_0) = u(0)$ respecto da la población total y que el 31 de Enero es $(x_1, y_1, z_1) = u(1)$.

Puesto que la totalidad de la población compra en estos tres únicos establecimientos tenemos que

$$\begin{aligned} x_0 + y_0 + z_0 &= 1 \\ x_1 + y_1 + z_1 &= 1 \end{aligned}$$

Ahora bien, la proporción de clientes en Enero en cada uno de los supermercados es

$$\left.\begin{aligned} x_1 &= a_{11}x_0 + a_{12}y_0 + a_{13}z_0 \\ y_1 &= a_{21}x_0 + a_{22}y_0 + a_{23}z_0 \\ z_1 &= a_{31}x_0 + a_{32}y_0 + a_{33}z_0 \end{aligned}\right\},$$

donde a_{ij} con $j \neq i$ es la proporción de clientes del supermercado j que absorbe el supermercado i, y a_{ii} es la proporción de clientes del supermercado i que se mantiene en i.

Puesto que la proporción de clientes que cambia de supermercado se matiene constante durante un mes tenemos

$$\left.\begin{aligned} x_{k+1} &= a_{11}x_k + a_{12}y_k + a_{13}z_k \\ y_{k+1} &= a_{21}x_k + a_{22}y_k + a_{23}z_k \\ z_{k+1} &= a_{31}x_k + a_{32}y_k + a_{33}z_k \end{aligned}\right\},$$

por lo que, en lenguaje matricial tenemos

$$\begin{pmatrix} x(k+1) \\ y(k+1) \\ z(k+1) \end{pmatrix} = \begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix} \begin{pmatrix} x(k) \\ y(k) \\ z(k) \end{pmatrix}.$$

Apliquémoslo a nuestro caso concreto y tenemos

$$\begin{pmatrix} x(k) \\ y(k) \\ z(k) \end{pmatrix} = \begin{pmatrix} 0{,}8 & 0{,}2 & 0{,}1 \\ 0{,}1 & 0{,}7 & 0{,}3 \\ 0{,}1 & 0{,}1 & 0{,}6 \end{pmatrix}^k \begin{pmatrix} 0{,}2 \\ 0{,}3 \\ 0{,}5 \end{pmatrix}$$

Los valores propios de A son 1, 0.6, 0.5 y los vectores propios respectivos son

$$v_1 = (0{,}45, 0{,}35, 0{,}20),\ v_2 = (1, -1, 0),\ v_3 = (1, -2, 1).$$

Por lo tanto y teniendo en cuenta que

$$u(0) = 1v_1 + (-0{,}55)v_2 + 0{,}3v_3$$

tenemos que

$$A^k u(0) = 1{,}1^k v_1 + (-0{,}55)(0{,}6)^k v_2 + 0{,}3(0{,}5)^k v_3$$

Finalmente, observamos que

$$\lim_k A^k u(0) = v_1 = (0{,}45, 0{,}35, 0{,}20)$$

3.- Revisando la cantidad de socios de un club de ua ciudad, se observa que el 80 % de las personas que son socias un año lo son también el siguiente y que el 30 % de los que no son socias lo son al siguiente. ¿Cuál será la situación a largo plazo de dicho club?

Solución:

Si denominamos por $s(k)$ a la cantidad de socios del año k y $n(k)$ a la cantidad de personas que no lo son tenemos la siguiente relación entre los socios $p(k) = \begin{pmatrix} s(k) \\ n(k) \end{pmatrix}$ de un año y el siguiente

$$p(k+1) = \begin{pmatrix} s(k+1) \\ n(k+1) \end{pmatrix} = \begin{pmatrix} 0{,}8 & 0{,}3 \\ 0{,}2 & 0{,}7 \end{pmatrix} \begin{pmatrix} s(k) \\ n(k) \end{pmatrix} = Ap(k)$$

Hemos de calcular $\lim_{k\to\infty} p(k) = (\lim_{k\to\infty} A^k)p(0)$

$$A = SDS^{-1}, \quad \lim_{k\to\infty} A^k = S(\lim_{k\to\infty} D^k)S^{-1}$$

con

$$S = \begin{pmatrix} 3 & 1 \\ 2 & -1 \end{pmatrix}, \quad D = \begin{pmatrix} 1 & \\ & 0{,}5 \end{pmatrix}, \quad \lim_{k\to\infty} D^k = \begin{pmatrix} 1 & \\ & 0 \end{pmatrix}$$

5.3. Ejercicios Propuestos

1.- Sea A una matriz de Markov positiva. Determinar el límite de A^n cuando n crece indefinidamente.

2.- Sea A una matriz de Markov tal que A^2 es positiva. Estudiar el límite de A^{2n} y de A^{2n+1} cuando n crece indefinidamente. ¿Qué puede decirse de la convergencia de A^n?

3.- Consideremos una cadena de Markov cuya matriz de transición de estados es

$$A = \begin{pmatrix} 0{,}6 & 0{,}1 & 0{,}6 \\ 0{,}2 & 0{,}8 & 0 \\ 0{,}2 & 0{,}1 & 0{,}4 \end{pmatrix}$$

Probar que la cadena es irreducible.

4.- Sea A la matriz

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

que es la matriz de transición de estados de una cadena de Markov absorbente. Calcular $\lim_{k\to\infty} p(k)$.

5.- Calcular π para el caso en que la matriz de transición de estados A de una cadena de Markov sea primitiva y doblemente estocástica.

6.- Sea A la matriz de transición de estados de una cadena de Markov siguiente.

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

Probar que tiene una única distribución de probabilidad estacionaria.

7.- Sea $A(p)$ la matriz

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

que para cada p es la matriz de transición de estados de una cadena de Markov. Determinar para que valores de p admite distribución estacionaria. Para los valores que exista ¿es única?

8.- Sea

$$P = \begin{pmatrix} 0{,}4 & 0{,}2 \\ 0{,}6 & 0{,}8 \end{pmatrix}$$

la matriz de transición de estados de una cadena de Markov a variable continua. Determinar P_t.

9. Se trata de estudiar como una noticia es susceptible de ser cambiada al ser propagada de una persona a otra.

Sea $p(k)$ la probabilidad de que la k-ésima persona escucha la noticia correctamente y $q(k)$ la probabilidad de que la k-ésima persona escucha la noticia erroneamente.

Supongamos que la probabilidad de que una persona comunique la noticia correctamente es 0,95 y que la probabilidad de que lo haga erroneamente es 0.05.

a) Describir dicha propagación, mediante un sistema de ecuaciones de la forma $x(k+1) = Ax(k)$. Dar los puntos de equilibrio. Dar los valores propios de A

b) Comprobar que

$$p(k+1) + q(k+1) = p(k) + q(k)$$

y por tanto $p(k)+q(k) = p(0)+q(0) = constante$. ¿Cual es dicha constante?

c) Deducir del apartado anterior, una ecuación en diferencias de primer orden, que describa la probabilidad de transmitir correctamente la noticia.

d) Resolver la ecuación hallada en c), suponiendo que la primera persona transmite la noticia correctamente.

Determinar, en caso de existir, el punto de equilibrio. ¿Es estable?

e) Calcular $q(k)$ y $\lim_k q(k)$

f) Interpretación del resultado.

10.- Un señor está paseando a lo largo de un misma calle en la que, en un extremo hay un Bar y en el otro extremo está su casa, entre medio hay tres travesías.

Bar — 1 — 2 — 3 — Casa

Llamemos $p_1(n)$, $p_2(n)$, $p_3(n)$ las probabilidades de que el señor al llegar a una esquina ésta sea 1, 2, 3 respectivamente, y llamemos $p_4(n)$ y $p_5(n)$ la probabilidad de que la esquina sea el bar o su casa.

Este señor se encuentre en una determinada esquina s, supongamos que la probalilidad de que camine hacia $s-1$ o $s+1$ es $\frac{1}{2}$ y que cuando llega a la nueva esquina sigue adelante salvo que ésta sea el Bar o su casa en cuyo caso se queda. ¿Que probabilidad hay de que este señor acabe en el bar? ¿y en su casa?, suponiendo que inicialmente, esta a mitad de camino de ambos (esquina 2).

Capítulo 6

Modelos de Leontief en economía

En este capítulo presentamos una aplicación de la teoría de matrices no negativas al campo de las ciencias económicas, concretamente al análisis de Leontief de entradas y salidas.
El análisis de Leontief de entradas y salidas trata de la siguiente cuestión: ¿qué nivel de salidas debería producir cada una de n industrias en una situación económica particular de manera que sea suficiente para poder satisfacer la demanda total del mercado para este producto?

Ejemplo 6.0.1. Consideremos una simple e hipotética economía que consiste en tres sectores que pueden ser por ejemplo: (1) pesca, (2) manufacturados y (3) servicios. Cada uno de los sectores produce una clase de salida, a saber: (1) pescado, (2) manufacturados y (3) servicios. Estos tres sectores son independientes. Todos los productos acabados y servicios que revierten en el proceso de producción son utilizados por el sector externo como mercancías, etc.
Con todas estas premisas damos una tabla hipotética de movimiento de capital tanto de productos como de servicios, calculados en miles de euros.

salidas / entradas	pesca	manufacturados	servicios	demanda externa	total
pesca	20	15	35	30	100
manufacturados	25	15	50	110	200
servicios	30	50	–	80	150

Las filas denotan la distribución de entradas de los distintos sectores y usuarios y las columnas indican las entradas necesarias para la producción. Así leyendo la primera fila (pesca) vemos que del total de 100 euros de productos de pesca 20 se utilizan para una próxima producción, 15 son ventas a manufacturas, 35 son ventas a servicios y finalmente 30 satisfacen la demanda externa. Análogamente leyendo la segunda columna tenemos que en orden a producir 200 euros de salidas totales de manufacturados entran 15 de pesca, 15 de productos propios y 50 de servicios.
Para analizar la tabla, necesitamos la siguiente notación. Denotamos por 1 la pesca, 2 los productos manufacturados, 3 los servicios y

x_i salidas totales del sector i, para $i = 1, 2, 3$;
x_{ij} ventas del sector i al sector j;
d_i demanda total en el sector i.

Entonces la relación básica de las filas de la tabla es

$$x_i = x_{i1} + x_{i2} + x_{i3} + d_i, \quad \text{para } i = 1, 2, 3$$

lo que nos dice que la salida total de un sector consiste en las ventas de productos intermedios a varios sectores de producción y la salida final que tiene en cuenta a los consumidores y al sector abierto.
Agrupamos las salidas totales de todos los sectores en el vector

$$x = \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix}$$

y la demanda final en el vector

$$d = \begin{pmatrix} d_1 \\ d_2 \\ d_3 \end{pmatrix}.$$

Para el análisis de entradas y salidas es mucho más conveniente que la tabla indique las entradas necesarias para la producción de una unidad de salidas para cada sector. Para construir la nueva tabla hacemos la división de cada entrada por el total en el sector. Esto es, llamando t_{ij} al coeficiente de entrada que indica la cantidad de producto i necesario para producir una unidad de salida del producto j, tenemos

$$t_{ij} = \frac{x_{ij}}{x_j}, \quad 1 \leq i, j \leq 3$$

que en nuestro caso particular es

$$
\begin{aligned}
x_1 &= 100, \\
x_{11} = 20, \quad x_{12} &= 15, \quad x_{13} = 35, \quad d_1 = 30 \\
x_2 &= 200, \\
x_{21} = 25, \quad x_{22} &= 15, \quad x_{23} = 50, \quad d_2 = 110 \\
x_3 &= 150, \\
x_{31} = 30, \quad x_{32} &= 50, \quad x_{33} = 0, \qquad d_3 = 80 \\
t_{11} = \frac{x_{11}}{x_1} = 0{,}20, \quad t_{12} &= \frac{x_{12}}{x_2} = 0{,}075, \quad t_{13} = \frac{x_{13}}{x_3} = 0{,}233 \\
t_{21} = \frac{x_{21}}{x_1} = 0{,}25, \quad t_{22} &= \frac{x_{22}}{x_2} = 0{,}075, \quad t_{23} = \frac{x_{23}}{x_3} = 0{,}333 \\
t_{31} = \frac{x_{31}}{x_1} = 0{,}30, \quad t_{32} &= \frac{x_{32}}{x_2} = 0{,}25, \quad t_{33} = \frac{x_{33}}{x_3} = 0
\end{aligned}
$$

y la tabla queda de la siguiente manera

salidas / entradas	pesca	manufacturados	servicios
pesca	0.20	0.075	0.233
manufacturados	0.25	0.075	0.333
servicios	0.30	0.25	0

Observamos ahora que si llamamos T a la matriz (t_{ij}) se tiene que

$$x = Tx + d$$

por lo que llamando $A = I - T$ nos queda el sistema de ecuaciones lineales siguiente

$$Ax = d. \tag{6.1}$$

Suponemos que los datos de la tabla son fijos aunque la demanda total y la salida total puedan cambiar. Entonces la pregunta hecha al inicio, de qué nivel de salidas totales x_i para cada uno de los tres sectores se han de tener para satisfacer las entradas requeridas y las demandas externas, se traduce en la resolución del sistema (6.1).

6.1. Modelo abierto

El análisis interindustrial del tipo de Leontief se refiere a sistemas en los que los factores tales como máquinas, material, etc. son usados como factores de producción de otros productos. El modelo abierto de Leontief es el más simple de todos.

Supongamos que la economía está dividida en n sectores, cada uno produciendo una mercancía de manera que la producción pueda ser consumida por el propio sector, por otras industrias o fuera del sector.

Identificando al sector i con la mercancía i tenemos

x_i salidas totales del sector i,

x_{ij} ventas del sector i al j,

d_i demanda final en el sector i,

t_{ij} coeficiente de entradas (número de unidades de la mercancía i necesarias para producir una unidad de la mercancía j).

El *balance* global de entradas y salidas de toda la economía puede ser expresado de la forma

$$\left.\begin{aligned} x_1 &= x_{11} + x_{12} + \ldots + x_{1n} + d_1 \\ &\vdots \\ x_n &= x_{n1} + x_{n2} + \ldots + x_{nn} + d_n \end{aligned}\right\}$$

Suponemos ahora que si cambia el nivel de salidas, entonces también cambian proporcionalmente los totales de todas las entradas necesarias. Esto es, suponiendo una proporción fija del factor de entradas, los coeficientes de entradas t_{ij} son constantes y satisfacen la relación

$$t_{ij} = \frac{x_{ij}}{x_j}, \quad 1 \leq i, j \leq n.$$

Entonces el sistema de ecuaciones anterior se transforma en

$$\left.\begin{aligned} x_1 &= t_{11}x_1 + t_{12}x_2 + \ldots + t_{1n}x_n + d_1 \\ &\vdots \\ x_n &= t_{n1}x_1 + t_{n2}x_2 + \ldots + t_{nn}x_n + d_n \end{aligned}\right\}$$

Llamemos T a la matriz (t_{ij}), $A = I - T$. El balance total de entradas y salidas se expresa como un sistema de ecuaciones lineales con n incógnitas:

$$Ax = d \tag{6.2}$$

donde $x = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}$ es el *vector de salidas* y $d = \begin{pmatrix} d_1 \\ \vdots \\ d_n \end{pmatrix}$ es el *vector de demandas final*.

El modelo aquí descrito recibe el nombre de *modelo de Leontief abierto* y la matriz T el de *matriz de entradas del modelo*. La matriz A es tal que $a_{ij} \leq 0$ si $i \neq j$. Estas matrices reciben el nombre de *matrices de Leontief* o *matrices esencialmente no positivas*.

La existencia de solución del sistema (6.2) para salidas no negativas para cada $1 \leq i \leq n$, significa la viabilidad del modelo. Esto nos lleva a la siguiente definición.

Definición 6.1.1. Un modelo abierto de Leontief con matriz de entradas T se dice *viable* si el sistema (6.2) tiene una solución no negativa para cada vector de demanda d.

Teniendo en cuenta que una M-matriz (ver definición en página 98), es no singular si y sólo si $A = sI - B$ con $s > 0$, $B \geq 0$ y $s > \rho(B)$, tenemos el siguiente teorema.

Teorema 6.1.1. *Consideremos un modelo abierto de Leontief con matriz de entradas T y sea $A = I - T$. El modelo es viable si y sólo si, A es una M-matriz no singular.*

Ejemplo 6.1.1. Como ejemplo estudiemos la viabilidad del sistema presentado en el ejemplo 6.0.1.

La matriz T es la siguiente

$$T = \begin{pmatrix} 0{,}20 & 0{,}075 & 0{,}233 \\ 0{,}25 & 0{,}075 & 0{,}333 \\ 0{,}30 & 0{,}25 & 0 \end{pmatrix}$$

por lo que la matriz $A = I - T$ es

$$A = \begin{pmatrix} 0{,}80 & -0{,}075 & -0{,}233 \\ -0{,}25 & 0{,}975 & -0{,}333 \\ -0{,}30 & -0{,}25 & 1 \end{pmatrix}$$

es una M-matriz ($s = 1 > \rho(T)$).
Calculamos $\det A \neq 0$ por lo que la matriz A es regular y en consecuencia el modelo es viable.

Corolario 6.1.1. (*Condición de Hawkins y Simon*) Un modelo abierto de Leontief con matriz de entradas T es viable si y sólo si, la matriz $A = I - T$ tiene todos sus menores principales positivos.

Observación. Si la matriz de entradas T de un modelo abierto de Leontief es irreducible, entonces el modelo es viable si y sólo si

$$A^{-1} \gg 0$$

6.2. Modelo cerrado

Si el sector externo del modelo abierto es absorbido en un sistema como otra industria entonces el sistema se convierte en cerrado.
En dichos modelos no aparecen ni la demanda final ni las entradas primarias, en su lugar aparecen las entradas requeridas y la salida de la nueva industria considerada.
Todas las mercancías son intermedias puesto que cada mercancía es producida solamente por las entradas requeridas en los sectores o industrias del propio sistema.
En el modelo cerrado de entradas y salidas de Leontief, el consumidor o sector abierto es visto como un sector de producción. Puesto que no hay variable determinada para las salidas, nos preguntamos lo siguiente: dado un sistema de producción ¿cuál es el equilibrio de salida y nivel de precios para que no quede demanda por satisfacer?
A distinción del caso abierto, ahora las demandas finales son consideradas entradas del sector de consumo y cada componente d_i (demanda externa del sector) está relacionada con el *nivel de empleo* $\mathcal{E}$. Definiendo unos coeficientes técnicos c_i de la forma

$$d_i = c_i \mathcal{E},$$

la relación entradas-salidas del modelo cerrado queda descrita por el sistema

$$\left.\begin{array}{r} x_1 = t_{11}x_1 + \ldots + t_{1n-1}x_{n-1} + c_1\mathcal{E} \\ \vdots \\ x_{n-1} = t_{n-11}x_1 + \ldots + t_{n-1n-1}x_{n-1} + c_{n-1}\mathcal{E} \end{array}\right\} \tag{6.3}$$

el nivel de empleo es la suma del trabajo realizado en cada sector

$$\mathcal{E} = L_1 + \ldots + L_n$$

siendo L_i el trabajo realizado en el sector i.
Definiendo los coeficientes (fijos) de trabajo l_i por

$$\begin{array}{l} l_i = L_i/x_i, \quad i = 1, \ldots, n-1 \\ l_n = L_n/\mathcal{E} \end{array}$$

tenemos que el sistema (6.3) queda

$$\left.\begin{array}{r} x_1 = t_{11}x_1 + \ldots + t_{1n-1}x_{n-1} + c_1\mathcal{E} \\ \vdots \\ x_{n-1} = t_{n-11}x_1 + \ldots + t_{n-1n-1}x_{n-1} + c_{n-1}\mathcal{E} \\ \mathcal{E} = l_1x_1 + \ldots + l_{n-1}x_{n-1} + l_n\mathcal{E} \end{array}\right\}$$

llamando

$$T = \begin{pmatrix} t_{11} & \ldots & t_{1n-1} & c_1 \\ \vdots & & & \vdots \\ t_{n-11} & \ldots & t_{n-1n-1} & c_{n-1} \\ l_1 & \ldots & l_{n-1} & l_n \end{pmatrix}$$

y si renombramos

$$\begin{array}{l} t_{in} = c_i \quad i = 1, \ldots, n-1 \\ t_{nj} = l_i \quad j = 1, \ldots, n \end{array}$$

$$x = \begin{pmatrix} x_1 \\ \vdots \\ x_{n-1} \\ \mathcal{E} \end{pmatrix}$$

la matriz T es (t_{ij}) la relación de entradas-salidas para el modelo cerrado queda descrito a través del siguiente sistema lineal de ecuaciones

$$x = Tx \tag{6.4}$$

T recibe el nombre de *matriz de entradas* del modelo.
Llamando ahora $A = I - T$ el sistema (6.4) puede describirse de la forma

$$Ax = 0$$

Este sistema es homogéneo, luego siempre es compatible. Nos interesarán aquellos sistemas que tengan soluciones positivas (en el modelo abierto nos interesaban las soluciones no negativas).

Definición 6.2.1. Un modelo cerrado de Leontief con matriz de entradas T se dice *viable* si y sólo si, existe algún vector x tal que

$$Tx \leq x \qquad x \gg 0$$

Si el tal vector x existe recibe el nombre de *solución de salida viable* para el modelo.

Ejemplo 6.2.1. Si la matriz de entradas y salidas de un modelo cerrado es

$$T = \begin{pmatrix} 0{,}6 & 0{,}5 & 0{,}2 \\ 0{,}1 & 0{,}5 & 0{,}1 \\ 0{,}3 & 0{,}2 & 0{,}1 \end{pmatrix}$$

observamos que

$$x = \begin{pmatrix} 2 \\ 1 \\ 1 \end{pmatrix} \gg 0$$

es tal que

$$Tx \leq x$$

Definición 6.2.2. Un vector x recibe el nombre de *vector de salida de equilibrio* para el modelo cerrado de Leontief con matriz de entradas T si

$$Tx = x \qquad x \gg 0$$

esto es, x es un vector propio de valor propio 1 de la matriz T y positivo.

Como consecuencia inmediata de esta definición tenemos el siguente corolario.

Corolario 6.2.1. Un modelo de entradas y salidas cerrado tiene un vector de salida de equilibrio si y sólo si el sistema homogéneo

$$Ax = 0$$

tiene una solución positiva.

Al igual que en el modelo abierto podemos caracterizar los modelos viables mediante M-matrices.

Teorema 6.2.1. *Un modelo cerrado de Leontief con matriz de entradas T es viable si y sólo si la matriz $A = I - T$ es una M-matriz con la propiedad c.*

Corolario 6.2.2. Supongamos que T es irreducible, entonces el modelo tiene un vector de salida de equilibrio x que es único salvo múltiplos escalares positivos.

6.3. Ejercicio Resueltos

1.- Consideremos el modelo abierto de economía con tres sectores a saber energía, manufacturados, y servicios donde la matriz T de entradas y salidas viene dada por

$$T = \begin{pmatrix} 0{,}1 & 0{,}2 & 0{,}1 \\ 0{,}4 & 0{,}2 & 0{,}2 \\ 0{,}2 & 0{,}3 & 0{,}5 \end{pmatrix}$$

y el vector demanda d por

$$d = \begin{pmatrix} 92 \\ 69 \\ 115 \end{pmatrix}$$

Encontrar el vector de producción.
Solución: El balance total de entradas y salidas es

$$Ax = d, \quad \text{con } A = I - T$$

por lo que tenemos que resolver el sistema

$$\begin{pmatrix} 0{,}9 & -0{,}2 & -0{,}1 \\ -0{,}4 & 0{,}8 & -0{,}2 \\ -0{,}2 & -0{,}3 & 0{,}5 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} 92 \\ 69 \\ 115 \end{pmatrix}$$

Cuya solución es

$$x = A^{-1}d = \begin{pmatrix} 235 \\ 335 \\ 525 \end{pmatrix}$$

2.- Consideremos una economía abierta, con tres sectores: minería del carbón, generadores de electricidad y una planta de montaje de automóviles. Para producir 1 euro de carbón, la explotación minera debe adquirir 0.1 euros de su propia producción, 0.30 euros de electricidad y 0.1 euro en automóviles para el transporte. Para producir 1 euro de eléctricidad, se necesitan 0.25 euros de carbón, 0.4 euros de electricidad y 0.15 euros de automóvil. Por último, para producir 1 euro de valor de automóvil, la planta de montaje debe comprar 0.2 euros de carbón, 0.5 euros de electricidad y el consumo de 0.1 euros de automóvil. Supongamos también que, durante un período de una semana, la economoía tiene un demanda por valor de 50000 euros de carbón, 75000 euros de electricidad, y 125000 euros de automóviles. Encontrar el nivel de producción de cada una de las tres industrias en ese período de tiempo con el fin de satisfacer exactamente la demanda.
Solución: La matriz de entradas y salidas de esta economía es

$$A = \begin{pmatrix} 0{,}1 & 0{,}25 & 0{,}2 \\ 0{,}3 & 0{,}4 & 0{,}5 \\ 0{,}1 & 0{,}15 & 0{,}1 \end{pmatrix}$$

y el vector demanda es

$$d = \begin{pmatrix} 50000 \\ 75000 \\ 125000 \end{pmatrix}$$

La ecuación de entradas en función de la demanda

$$Ax = d$$

con

$$A = (I - T)^{-1}$$

y

$$I - T = \begin{pmatrix} 0{,}9 & -0{,}25 & -0{,}2 \\ -0{,}3 & 0{,}6 & -0{,}5 \\ -0{,}1 & -0{,}15 & 0{,}9 \end{pmatrix}$$

$$(I - T)^{-1} = \begin{pmatrix} 1{,}464 & 0{,}803 & 0{,}771 \\ 1{,}007 & 2{,}488 & 1{,}606 \\ 0{,}330 & 0{,}503 & 1{,}464 \end{pmatrix}$$

Obteniendo:

$$x = \begin{pmatrix} 1{,}464 & 0{,}803 & 0{,}771 \\ 1{,}007 & 2{,}488 & 1{,}606 \\ 0{,}330 & 0{,}503 & 1{,}464 \end{pmatrix} \begin{pmatrix} 50000 \\ 75000 \\ 125000 \end{pmatrix} = \begin{pmatrix} 229921{,}59 \\ 437795{,}25 \\ 237401{,}57 \end{pmatrix}$$

Por lo tanto, la producción total de carbón de la explotación minera debe ser 229921,59 euros, la producción total de electricidad de la planta generadora es de 437795,27 euros y la producción total para el montaje de automóviles es de 237401,57 euros.

3.- Supongamos que la economía de una determinada región depende de tres sectores: servicios, electricidad y producción de petróleo. Supervisando las operaciones de estas tres industrias durante un período de un año, hemos podido llegar a las siguientes observaciones:

1. Para producir 1 unidad de valor de servicio, la industria de servicios debe consumir 0.3 unidades de su propia producción, 0.3 unidades de electricidad y 0.3 unidades de aceite para realizar sus operaciones.

2. Para producir 1 unidad de electricidad, el poder de generación de la planta debe comprar 0.4 unidades de servicio, 0.1 unidades de su propia producción, y 0.5 unidades de petróleo.

3. Por último, la compañía de producción de petróleo requiere 0.3 unidades de servicio, 0.6 unidades de electricidad y 0.2 unidades de su propia producción para producir 1 unidad de petróleo.

Encuentar el nivel de producción de cada una de estas industrias con el fin de satisfacer la demanda, en el supuesto de que el modelo es cerrado.
Solución: Consideremos las siguientes variables:
p_1 nivel de producción servicio
p_2 nivel de producción electricidad
p_3 nivel de producción petróleo
Dado que el modelo es cerrado, el consumo total de cada industria debe ser igual a su producción total. Esto nos proporciona el siguiente sistema lineal

$$\left.\begin{array}{ll} 0{,}3p_1 + 0{,}3p_2 + 0{,}3p_3 & = p_1 \\ 0{,}4p_1 + 0{,}1p_2 + 0{,}5p_3 & = p_2 \\ 0{,}3p_1 + 0{,}6p_2 + 0{,}2p_3 & = p_3 \end{array}\right\}$$

La matriz de entradas es

$$A = \begin{pmatrix} 0{,}3 & 0{,}3 & 0{,}3 \\ 0{,}4 & 0{,}1 & 0{,}5 \\ 0{,}3 & 0{,}6 & 0{,}2 \end{pmatrix}$$

Este sistema se puede escribir como $(A-I)P = 0$. Este sistema es homogéneo por lo que tiene siempre solución, de hecho tiene infinitas soluciones ya que la matriz A^t es estocástica por lo que 1 es valor propio,
Resolvemos el sistema obteniendo

$$\left.\begin{array}{ll} p_1 & = 0{,}82t \\ p_2 & = 0{,}92t \\ p_3 & = t \end{array}\right\}$$

Los valores de las variables en este sistema debe ser no negativas a fin de que el modelo tenga sentido, en otras palabras, $t \geq 0$. Tomando $t = 100$, por ejemplo, tendríamos la solución

$$\left.\begin{array}{ll} p_1 & = 82 \text{ unidades} \\ p_2 & = 92 \text{ unidades} \\ p_3 & = 100 \text{ unidades} \end{array}\right\}$$

6.4. Ejercicios Propuestos

1.- Analizar la viabilidad del modelo abierto de Leontief cuya matriz de entradas viene dada por

$$T = \begin{pmatrix} 0{,}1 & 0{,}2 & 0{,}4 \\ 0{,}3 & 0{,}1 & 0{,}5 \\ 0{,}2 & 0 & 0{,}3 \end{pmatrix}$$

2.- Consideremos el modelo de Leontief del ejercicio anterior. Analizar el crecimiento de las salidas suponiendo que aumenta la demanda de un solo prodructo.

3.- Estudiar la viabilidad del modelo abierto de Leontief cuya matriz de entradas es

$$T = \begin{pmatrix} 0{,}2 & 0{,}3 & 0{,}2 \\ 0{,}4 & 0{,}1 & 0{,}2 \\ 0{,}1 & 0{,}3 & 0{,}2 \end{pmatrix}$$

Obtener las salidas de dicho sistema para las entradas $d_1 = 10$, $d_2 = 5$, $d_3 = 6$.

4.- Probar que si un modelo de Leontief abierto es viable, entonces alguna de las columnas de la matriz T es tal que la suma de sus componentes es menor que 1. Concluir la no viabilidad del sistema cuya matriz de entradas es

$$T = \begin{pmatrix} 0{,}1 & 0{,}2 & 0{,}1 \\ 0{,}8 & 0{,}7 & 0{,}4 \\ 0{,}3 & 0{,}6 & 0{,}7 \end{pmatrix}$$

5.- Estudiar la viabilidad del sistema de Leontief cerrado cuya relación de entradas y salidas viene dado mediante

$$\left.\begin{aligned} x_1 &= 0{,}3x_1 + 0{,}3\mathcal{E} \\ x_2 &= 0{,}2x_1 + 0{,}2x_2 \\ \mathcal{E} &= 0{,}5x_1 + 0{,}8x_2 + 0{,}7\mathcal{E} \end{aligned}\right\}$$

donde x_1 y x_2 son las salidas de dos sectores internos y $\mathcal{E}$ es el nivel de empleo. ¿Tiene puntos de equilibrio? En caso afirmativo ¿existe alguno cuya suma de componentes sea 1?

6.- Una hipotética economía formada por tres sectores (1), (2) y (3) tales que las ventas x_{ij} del sector (i) al sector (j) son

$$\begin{array}{lll} x_{11} = 15 & x_{21} = 10 & x_{31} = 18 \\ x_{12} = 10 & x_{22} = 13 & x_{32} = 21 \\ x_{13} = 20 & x_{23} = 22 & x_{33} = 10 \end{array}$$

y la demanda final d_i en el sector i es

$$d_1 = 13, \quad d_2 = 10, \quad d_3 = 18.$$

Estudiar la viabilidad del sistema.

Capítulo 7

Métodos iterativos para sistemas lineales

En este capítulo presentamos una aplicación de la teoría de matrices no negativas al estudio de métodos iterativos para resolver un sistema de ecuaciones de la forma

$$Ax = b \tag{7.1}$$

donde A es una matriz cuadrada de orden n y b es un vector columna de orden $n \times 1$.
Así para $n = 3$ tenemos

$$\begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} b_1 \\ b_2 \\ b_3 \end{pmatrix}$$

que también se puede escribir de la siguiente forma:

$$\left.\begin{aligned} a_{11}x_1 + a_{12}x_2 + a_{13}x_3 &= b_1 \\ a_{21}x_1 + a_{22}x_2 + a_{23}x_3 &= b_2 \\ a_{31}x_1 + a_{32}x_2 + a_{33}x_3 &= b_3 \end{aligned}\right\}$$

Los métodos iterativos consisten básicamente en asociar al sistema (7.1) un sistema lineal $x = Bx+c$ que resolvemos de forma iterativa. Esto es, partiendo de un x^0 obtenemos $x^1 = Bx^0 + c$ logrando así una sucesión $x^0, x^1, \ldots, x^k, \ldots$ mediante la ley de recurrencia

$$x^{k+1} = Bx^k + c, \quad k \geq 0$$

Si la sucesión $\{x^k\}$ converge hacia el vector $\bar{x}$, entonces este vector es la solución del sistema $x = Bx + c$ y por tanto del sistema $Ax = b$, que es el sistema que pretendíamos resolver.

7.1. Métodos iterativos básicos

En esta sección vamos a describir algunos métodos iterativos básicos tales como el método de Jacobi o el método de Gauss-Seidel.
Supongamos que en el sistema (7.1) se ha realizado la siguiente partición

$$\begin{pmatrix} A_{1,1} & \dots & A_{1,q} \\ \vdots & & \vdots \\ A_{q,1} & \dots & A_{q,q} \end{pmatrix} \begin{pmatrix} X_1 \\ \vdots \\ X_q \end{pmatrix} = \begin{pmatrix} b_1 \\ \vdots \\ b_q \end{pmatrix} \tag{7.2}$$

donde $A_{i,j} \in M_{n_i \times n_j}(\mathbb{R})$ con $n_1 + \ldots + n_q = n$ y X_i, b_i son vectores columna de orden n_i para $i = 1, \ldots, q$.
La matriz A puede expresarse como

$$A = D - C_L - C_U \tag{7.3}$$

donde D es la matriz diagonal por bloques

$$D = \begin{pmatrix} A_{1,1} & & \\ & \ddots & \\ & & A_{q,q} \end{pmatrix}$$

C_L la matriz triangular inferior

$$C_L = - \begin{pmatrix} 0 & 0 & 0 & \dots & 0 \\ A_{2,1} & 0 & 0 & \dots & 0 \\ A_{3,1} & A_{3,2} & 0 & \dots & 0 \\ \vdots & \vdots & \vdots & \dots & 0 \\ A_{q,1} & A_{q,2} & A_{q,3} & \dots & 0 \end{pmatrix}$$

y C_U la matriz triangular superior

$$C_U = - \begin{pmatrix} 0 & A_{1,2} & A_{1,3} & \dots & A_{1,q} \\ 0 & 0 & A_{2,3} & \dots & A_{2,q} \\ 0 & 0 & 0 & \dots & A_{3,q} \\ \vdots & \vdots & \vdots & \dots & \vdots \\ 0 & 0 & 0 & \dots & 0 \end{pmatrix}$$

Observación 1. Para cada $i = 1, \ldots, q$, n_i puede ser igual a 1, en cuyo caso si la matriz A es

$$\begin{pmatrix} a_{11} & \ldots & a_{1n} \\ \vdots & & \vdots \\ a_{n1} & \ldots & a_{nn} \end{pmatrix}$$

la partición queda del siguiente modo

$$D = \begin{pmatrix} a_{11} & & \\ & \ddots & \\ & & a_{nn} \end{pmatrix}$$

y

$$C_L = -\begin{pmatrix} 0 & 0 & \ldots & 0 \\ a_{21} & 0 & \ldots & 0 \\ \vdots & \vdots & \ldots & \vdots \\ a_{n1} & a_{n2} & \ldots & 0 \end{pmatrix}, \qquad C_U = -\begin{pmatrix} 0 & a_{12} & \ldots & a_{1n} \\ 0 & 0 & \ldots & a_{2n} \\ \vdots & \vdots & \ldots & \vdots \\ 0 & 0 & \ldots & 0 \end{pmatrix}$$

Método de Jacobi

El método de Jacobi relativo a la partición (7.2) viene dado por

$$A_{i,i}X_i^{n+1} = -\sum_{\substack{j=1 \\ j \neq i}}^{q} A_{i,i}X_j^n + b_i, \quad i = 1, 2, \ldots, q \tag{7.4}$$

Para que el método de Jacobi sea operativo es necesario que dado un y_i, los subsistemas

$$A_{i,i}X_i^{n+1} = y_i$$

puedan resolverse fácilmente.
Si tenemos en cuenta (7.3) el método de Jacobi se expresa de la forma

$$x^{n+1} = B_J x^n + k_J$$

donde

$$B_J = D^{-1}(C_L + C_U) = I - D^{-1}A$$

y

$$x^n = \begin{pmatrix} X_1^n \\ \vdots \\ X_q^n \end{pmatrix}, \qquad k_J = D^{-1} \begin{pmatrix} b_1 \\ \vdots \\ b_q \end{pmatrix}$$

La matriz B recibe el nombre de *matriz de iteración de Jacobi.*

Ejemplo 7.1.1. Consideremos el sistema

$$\begin{pmatrix} 4 & -1 \\ -1 & 2 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 3 \\ 1 \end{pmatrix}$$

La matriz de iteración es

$$B_J = \begin{pmatrix} 0 & 1/4 \\ 1/2 & 0 \end{pmatrix},$$

y

$$k_J = \begin{pmatrix} 1/4 \\ 3/2 \end{pmatrix}$$

Los dos primeros pasos de la iteración a partir de $x^0 = \begin{pmatrix} 1 \\ 0 \end{pmatrix}$ son

$$x^1 = \begin{pmatrix} 1/4 \\ 2 \end{pmatrix}, \qquad x^2 = \begin{pmatrix} 3/4 \\ 13/8 \end{pmatrix}.$$

Método de Gauss-Seidel

Este método aplicado a la partición (7.2) queda definido por

$$A_{ii} X_i^{n+1} = -\sum_{j=1}^{i-1} A_{i,j} X_j^{(n+1)} - \sum_{j=i+1}^{q} A_{i,j} X_j^n + b_i \quad i = 1, 2, \ldots, q. \tag{7.5}$$

Como en el caso anterior, cada paso de la iteración de Gauss-Seidel requiere la resolución de un subsistema de la forma

$$A_{i,i} X_i^{(n+1)} = y_i$$

Utilizando la notación de (7.3) la ecuación (7.5) queda

$$(D - C_L) x^{n+1} = C_U x^n + b$$

por lo que el método de Gauss-Seidel se expresa de la forma

$$x^{n+1} = B_{GS}x^n + k_{GS}$$

donde

$$B_{GS} = (I - L)^{-1}U, \qquad k_{GS} = (I - L)^{-1}D^{-1}b$$

y

$$L = D^{-1}C_L, \qquad U = D^{-1}C_U$$

La matriz B_{GS} recibe el nombre de *matriz de iteración de Gauss-Seidel.*

Ejemplo 7.1.2. Consideremos el sistema

$$\begin{pmatrix} 4 & -1 \\ -1 & 2 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 3 \\ 1 \end{pmatrix}$$

La matriz de iteración es

$$B_{GS} = \begin{pmatrix} 0 & 1/4 \\ 0 & 1/8 \end{pmatrix}$$

y

$$k_{GS} = \begin{pmatrix} 25/28 \\ 4/7 \end{pmatrix}$$

Los dos primeros pasos de la iteración a partir de $x^0 = \begin{pmatrix} 1 \\ 0 \end{pmatrix}$ son

$$x^1 = \begin{pmatrix} 25/28 \\ 4/7 \end{pmatrix}, \qquad x^2 = \begin{pmatrix} 29/28 \\ 9/14 \end{pmatrix}$$

Métodos de sobrerelajación

Estos métodos generalizan el de Gauss-Seidel y se basan en la equivalencia de la ecuación $Ax = b$ con

$$x = x + \omega((D^{-1}C_L - I + D^{-1}C_U)x - D^{-1}b)$$

con ω un parámetro real llamado *factor de relajación*, y que a veces son utilizados para acelerar la convergencia del método de Gauss-Seidel.

Se define el método iterativo de sobrerelajación con factor ω mediante la ley de recurrencia

$$x^{n+1} = B_\omega x^n + k_\omega$$

donde

$$B_\omega = (I - \omega L)^{-1}(\omega U + (1-\omega)I), \qquad \text{y} \qquad k_\omega = (I - \omega L)^{-1}\omega D^{-1} b$$

Observamos que si $\omega = 1$ el método coincide con el de Gauss-Seidel.

7.2. No negatividad y convergencia

Vamos a analizar en esta sección la convergencia de los métodos iterativos. En todo lo que sigue supondremos que la matriz A es no singular.

Definición 7.2.1. Un método iterativo se dice convergente si para x^0, la sucesión $x^1, x^2, \ldots$, converge hacia $\bar{x}$ solución de $Ax = b$.

Proposición 7.2.1. *Una condición necesaria y suficiente para la convergencia es que el radio espectral de la matriz de iteración B, sea menor que 1.*

$$\rho(B) < 1.$$

Demostración. Restando $\bar{x} = B\bar{x} + k$ a la ecuación $x^{k+1} = Bx^k + k$ tenemos

$$\begin{aligned} x^{k+1} - \bar{x} &= (Bx^k + k) - (B\bar{x} + k) = B(x^k - \bar{x}) \\ &= B^{k+1}(x^0 - \bar{x}) \end{aligned}$$

Entonces, la sucesión $x^0, x^1, \ldots$ converge hacia $\bar{x}$ para cada x^0 si y sólo si

$$\lim_{k\to\infty} B^k = 0.$$

Equivalentemente, si y sólo si

$$\rho(B) < 1.$$

□

Observación 1. El vector $x^k - \bar{x}$ recibe el nombre de *vector error* y permite medir la velocidad de la convergencia.
Observación 2. La proposición anterior reduce el problema de la convergencia a demostrar que $\rho(B) < 1$.
Veremos como la no negatividad de B juega un papel importante en el análisis de dicho problema.

Definición 7.2.2. Consideremos el método de iteración $x^{k+1} = Bx^k + k$ cuya matriz de iteración $B \in M_n(\mathbb{C})$ es tal que $\rho(B) < 1$ y sea $\bar{x} = B\bar{x} + k$. Sea

$$\alpha = \sup\{\lim_{k\to\infty} \|x^k - \bar{x}\|^{1/k} \mid x^0 \in \mathbb{C}^n\},$$

para una norma de matriz cualquiera El número

$$R_\infty(B) = -\ln \alpha$$

recibe el nombre de *coeficiente asintótico de convergencia* de la iteración.

Observación 3. El coeficiente asintótico de convergencia no depende de la norma escogida.

Proposición 7.2.2. *Sea $B \in M_n(\mathbb{R})$ la matriz de iteración de un sistema y supongamos que $\rho(B) < 1$. Entonces*

$$\alpha = \rho(B)$$

y por lo tanto

$$R_\infty(B) = -\ln \rho(B).$$

Vamos ahora a dar una condición necesaria y suficiente para garantizar la cota exigida al radio espectral de la matriz de iteración que nos asegura la convergencia.

Proposición 7.2.3. *Sea $A = D - (C_L + C_U)$. Supongamos que la matriz D es no singular y que la matriz de iteración de Jacobi B_J es no negativa. Entonces,*

$$\rho(B_J) < 1$$

si y sólo si

$$A^{-1}(C_L + C_U) \geq 0.$$

En cuyo caso

$$\rho(B_J) = \frac{\rho(A^{-1}(C_L + CU))}{1 + \rho(A^{-1}(C_L + C_U))}$$

Demostración. Supongamos que $\rho(B_J) \geq 1$

$$A^{-1}(C_L + C_U) = (D(I - D^{-1}(C_L + C_U)))^{-1}(C_L + C_U) =$$
$$(I - D^{-1}(C_L + C_U))^{-1}D^{-1}(C_L + C_U) =$$
$$(I - B_J)^{-1}B_J = \left(\sum_{k=0}^{\infty} B_J^k\right) B_J = \sum_{k=1}^{\infty} B_J^k$$

La penúltima igualdad se verifica debido a que, que $\rho(B_J) < 1$, la serie $\sum_{k=0}^{\infty} B_J^k$ converge y su suma es $(I - B_J)^{-1}$.
Ahora bien, puesto que la matriz B_J es no negativa la serie también es no negativa por lo que $A^{-1}(C_L + C_U)$ es no negativa.
Recíprocamente, supongamos ahora que $A^{-1}(C_L+C_U)$ es no negativa. Puesto que B_J es no negativa, por el teorema de Perron-Frobenius, existe un vector positivo $v > 0$ tal que

$$B_J v = \rho(B_J)v.$$

En cuyo caso

$$A^{-1}(C_L + C_U)v = (I - B_J)^{-1}B_J v = (I - B_J)^{-1}(\rho(B_J)v) =$$
$$= \frac{\rho(B_J)}{1 - \rho(B_J)}v$$

Puesto que $A^{-1}(C_L + C_U)v \geq 0$ y $v \geq 0$ se tiene $\dfrac{\rho(B_J)}{1 - \rho(B_J)} \geq 0$ por lo que

$$\rho(B_J) < 1.$$

Además, puesto que

$$\rho(A^{-1}(C_L + C_U)) \geq \frac{\rho(B_J)}{1 - \rho(B_J)}$$

se sigue que

$$\rho(B_J) \leq \frac{\rho(A^{-1}(C_L + C_U))}{1 + \rho(A^{-1}(C_L + C_U))} \tag{7.6}$$

Análogamente, puesto que $A^{-1}(C_L + C_U) \geq 0$ existe $w > 0$ tal que $A^{-1}(C_L + C_U)y = \rho(A^{-1}(C_L + C_U))y$, por lo que

$$B_J y = \frac{\rho(A^{-1}(C_L + C_U))}{1 + \rho(A^{-1}(C_L + C_U))}y$$

y

$$\rho(B_J \geq \frac{\rho(A^{-1}(C_L + C_U))}{1 + \rho(A^{-1}(C_L + C_U))} \tag{7.7}$$

De (7.6) y (7.7) se concluye el resultado. □

De forma análoga se puede demostrar el siguiente resultado para la matriz de iteración de Gauss-Seidel.

Proposición 7.2.4. *Sea $A = (D - C_L) - C_U$. Supongamos que la matriz $D - C_L$ es no singular y que la matriz de iteración de Gauss-Seidel B_{GS} es no negativa. Entonces,*

$$\rho(B_{GS}) < 1$$

si y sólo si

$$A^{-1}C_U \geq 0$$

En cuyo caso

$$\rho(B_{GS}) = \frac{\rho(A^{-1}C_U)}{1 + \rho(A^{-1}C_U)}$$

Con respecto al método iterativo de sobrerelajación, tenemos la siguiente condición para la convergencia.

Proposición 7.2.5. *Sea $A \in M_n(\mathbb{C})$ con todos los elementos de la diagonal no nulos. Entonces el método iterativo de sobrerelajación converge sólo si*

$$0 < \omega < 2$$

Demostración. Es fácil ver que

$$\det B_\omega = (1 - \omega)^n = \prod_{i=1}^{n} \lambda_i$$

□

Finalizamos dando condiciones a los vectores iniciales de los métodos iterativos de manera que, no sólo nos aproximen la solución del sistema (7.1) sino que proporcionan cotas para la solución.

Proposición 7.2.6. *Sea $A = D - (C_L + C_U)$ con A y D no singulares, supongamos que B_J es no negativa. Consideremos el sistema $Ax = b$ y el método iterativo de Jacobi.*

(i) *Si los vectores v^0 y w^0 son tales que $v^0 \leq v^1$, $v^0 \leq w^0$ y $w^1 \leq w^0$ siendo v^1 y w^1 los vectores obtenidos por el método iterativo de Jacobi a partir de v^0 y w^0, respectivamente. Entonces*

$$v^0 \leq v^1 \leq \ldots \leq v^i \leq \ldots \leq A^{-1}b \leq \ldots \leq w^i \leq \ldots \leq w^1 \leq w^0$$

y para algún escalar λ se tiene

$$A^{-1}b = \lambda \lim_{i\to\infty} v^i + (1-\lambda) \lim_{i\to\infty} w^i$$

(ii) *Si $\rho(B_J) < 1$ entonces existen los vectores v^0 y w^0*

Demostración. Veamos (i). Por hipótesis $v^0 \leq v^1$. Supongamos que $v^{i-1} \leq v^i$. Entonces, puesto que B_J es no negativa se tiene,

$$B_J v^{i-1} \leq B_J v^i$$

por lo que

$$v^i = B_J v^{i-1} + D^{-1}b \leq B_J v^i + D^{-1}b = v^{i+1}$$

Análogamente se prueba que

$$v^i \leq w^i \qquad \text{y} \qquad w^{i+1} \leq w^i$$

Luego ambas sucesiones convergen por lo que convergen hacia $A^{-1}b$.
Para (ii) supongamos que $\rho(B_J) < 1$. Entonces por el teorema de Perron-Frobenius, existe un vector $x > 0$ tal que $B_J x = \rho(B_J)x < x$.
Consideremos

$$v^0 = A^{-1}b - x$$

entonces

$$\begin{aligned} v^1 &= B_J v^0 + D^{-1}b = B_J A^{-1}b - B_J x + D^{-1}b = \\ &= (I - D^{-1}A)A^{-1}b - \rho(B_J)x + D^{-1}b = A^{-1}b - \rho(B_J)x \geq A^{-1}b - x = v^0 \end{aligned}$$

Sea $w^0 = A^{-1}b + x$, entonces

$$\begin{aligned} w^1 &= B_J w^0 + D^{-1}b = B_J A^{-1}b + B_J x + D^{-1}b = \\ &= (I - D^{-1}A)A^{-1}b + \rho(B_J)x + D^{-1}b = A^{-1}b + \rho(B_J)x \leq A^{-1}b + x = w^0 \end{aligned}$$

Además

$$w^0 - v^0 = 2x \geq 0$$

□

De forma análoga se demuestra la siguiente proposición, para el método iterativo de Gauss-Seidel.

Proposición 7.2.7. *Sea $A = D - (C_L + C_U)$ con A y D no singulares, y supongamos que B_{GS} es no negativa. Consideremos el sistema $Ax = b$ y el método iterativo de Gauss-Seidel.*

(i) *Si los vectores v^0 y w^0 son tales que $v^0 \leq v^1$, $v^0 \leq w^0$ y $w^1 \leq w^0$ siendo v^1 y w^1 los vectores obtenidos por el método iterativo de Gauss-Seidel a partir de v^0 y w^0, respectivamente. Entonces*

$$v^0 \leq v^1 \leq \ldots \leq v^i \leq \ldots \leq A^{-1}b \leq \ldots \leq w^i \leq \ldots \leq w^1 \leq w^0$$

y para algún escalar λ se tiene

$$A^{-1}b = \lambda \lim_{i\to\infty} v^i + (1-\lambda) \lim_{i\to\infty} w^i.$$

(ii) *Si $\rho(B_{GS}) < 1$ entonces existen los vectores v^0 y w^0.*

7.3. Ejercicios

1.- Consideremos el sistema lineal

$$\begin{pmatrix} 2 & -1 & 0 & 0 \\ -1 & 2 & -1 & 0 \\ 0 & -1 & 2 & -1 \\ 0 & 0 & -1 & 2 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \\ t \end{pmatrix} = \begin{pmatrix} 2 \\ 1 \\ 2 \\ 1 \end{pmatrix}.$$

Obtener tres iteraciones para los métodos de Jacobi y Gauss-Seidel a partir del valor inicial $x^0 = (1, 1, 1, 1)$.

2.- Sea

$$A = \begin{pmatrix} 1 & a & a \\ a & 1 & a \\ a & a & 1 \end{pmatrix}.$$

Estudiar para que valores de a el método de Jacobi converge.

3.- Sea

$$\left.\begin{aligned} 3x + 2y &= 1 \\ x + 3y &= 0 \end{aligned}\right\}$$

Calcular el coeficiente asintótico de convergencia para los métodos de Jacobi y de Gauss-Seidel.

4.- Sea la matriz

$$A = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 \\ 5 & 1 & 2 & 3 & 4 \\ 4 & 5 & 1 & 2 & 3 \\ 3 & 4 & 5 & 1 & 2 \\ 2 & 3 & 4 & 5 & 1 \end{pmatrix}$$

Estudiar la convergencia de los métodos de Jacobi y Gauss-Seidel para la resolución del sistema $(A + \lambda I)x = b$, en función del parámetro λ.

5.- Sea

$$A_1 = \begin{pmatrix} 1 & 2 & -2 \\ 1 & 1 & 1 \\ 2 & 2 & 1 \end{pmatrix}, \qquad A_2 = \begin{pmatrix} 2 & -1 & 1 \\ 2 & 2 & 2 \\ -1 & -1 & 2 \end{pmatrix}$$

Estudiar la convergencia de los métodos de Jacobi y de Gauss-Seidel para estas matrices.

Bibliografía

[1] A. Berman, R.J. Plemmons, *Nonnegative Matrices in Mathematical Sciences*, SIAM. Philadelphia (1994).

[2] F. Chatelin, *Valeurs Propres de Matrices*, Ed. Masson. París (1988).

[3] J. W. Demmel, *Applied Numerical Linear Algebra*, SIAM. Philadelphia (1997).

[4] S. N. Elaydi, *An Introduction to Difference Equations, Springer*, New York (1996).

[5] Gantmacher, *Matrix Theory*, vol. I, II. Chelsea (1977).

[6] M$^{\underline{a}}$ I. García, M. D. Magret, *Matrices no negativas y aplicaciones*. Editado por las autoras, (2002).

[7] L.A. Hageman, D.M. Young, *Applied Iterative methods*, Academic Press (1981).

[8] P. Lancaster, M. Tismenetsky, *The Theory of Matrices with Applications*, Academic Press (1985).

[9] C.C. MacDuffee, *The Theory of Matrices*, Chelsea (1946).

[10] K. Sh. Trivedi, *Probability and Statistics with Reliability, Queuing, and Computer Science Applications*, Prentice-Hall, Inc. Englewwod Cliffs, N.J. (1982).

Índice alfabético

adjunta de una matriz, 20
anillo, 13
aplicación, 37
aplicación lineal, 38

base, 32

cadena de Markov, 91
    absorbente, 95
    irreducible, 95
    regular, 95
cadena homgénea, 92
combinación lineal, 28
coordenadas, 34

defectivo, 57
determinante, 19
dimensión, 33
doblemente estocástica, 84

escalar, matriz escalar, 16
espacio vectorial, 15, 27
espectro, 53
estocástica, 82

Frobenius, 78

grupo abeliano, 12

imagen, 37, 40
invariante
    subespacio invariante, 43
inversa de Moore-Penrose, 22
inversa de una matriz, 21

Leontief, 105
linealmente dependiente, 29
linealmente independientes, 29

M-matriz, 98
método Jacobi, 121
métodos iterativos, 120
Matriz, 11
matriz
    ortogonal, 72
    antisimtrica, 18
    de cambio de base, 35
    de ietración, 122
    de iteración, 123
    de la aplicación lineal, 39
    de Leontief, 109
    de Markov, 93
    de permutación, 35
    diagonal, 15
    diagonalizable, 57
    irreducible, 72
    no negativa, 69
    normalizada, 84
    positiva, 70
    primitiva, 80
    reducible, 72
    regular, 20

simétrica, 17
singular, 20
totalmente no negativa, 85
totalmente positiva, 86
transición de estados, 92
traspuesta, 16
triangular, 15
matriz invers
inversa por la derecha, 21
matriz inversa
inversa por la izquierda, 21
menor de una matriz, 20
modelo abierto de Leontief, 108
modelo de Leontief cerrado, 110
morfismo, 41
multiplicidad
algebraica, 56
geomtrica, 56
mtodo de Gauss-Seidel, 122

núcleo, 40
normalización, 84

orden, relación de orden, 71

Perron, 78
polinomio característico, 51
polinomio mínimo, 58
probabilidad, 91

radio espectral, 58
rango, 20
rango de f, 41

semisimple, 57
sistema de generadores, 30
subespacio vectorial, 28
suma de subespacios, 36
suma directa, 36
teorema de Steinitz, 32

valor propio, 51, 53
valores singulares, 59
vector propio, 51, 54
viable, 109, 112