UNIVERSIDAD NACIONAL AUTÓNOMA DE MÉXICO

FACULTAD DE CIENCIAS

Introducción a la teoría de matrices oscilatorias

T E S I S

QUE PARA OBTENER EL TÍTULO DE:

Matemático

PRESENTA:

Gerardo Ulises Peña González

DIRECTOR DE TESIS:

Dr. Christof Geiss Hahn

Junio 2010

Act. Mauricio Aguilar González
Jefe de la División de Estudios Profesionales de la
Facultad de Ciencias.
Presente.

Por este medio hacemos de su conocimiento que hemos revisado el trabajo escrito intitulado

Introducción a la teoría de matrices oscilatorias

(138 pp., 2010) realizado por Gerardo Ulises Peña González quien cubrió los créditos de la carrera de Matemáticas en la Facultad de Ciencias, Universidad Nacional Autónoma de México, con número de cuenta 093362483.

Dicho trabajo cuenta con nuestro voto aprobatorio.

Atentamente

Director de Tesis
Propietario Dr. Christof Geiss Hahn

Propietario Dr. Michael Barot Schlatter

Propietario Dra. Edith Corina Sáenz Valadez

Suplente Dra. Diana Avella Alaminos

Suplente Dr. Francisco Marmolejo Rivas

Índice general

Introducción

La teoría de matrices oscilatorias tiene su origen en la teoría de pequeñas oscilaciones y, al mismo tiempo, representa una base matemática natural para la investigación de las llamadas propiedades oscilatorias de vibraciones armónicas pequeñas de *continuos elásticos lineales* (e.g., vibraciones transversales de cuerdas, varillas, oscilaciones de torsión de rayos, etc.)

Durante los años 1930's F. R. Gantmacher y M. G. Krein realizaron diversas investigaciones sobre matrices totalmente no-negativas en conexión con las observaciones que realizaron sobre las propiedades oscilatorias de vibraciones. Como resultado de estas investigaciones una importante clase de matrices hizo su aparición: las matrices cuadradas totalmente no-negativas A con la característica de que cierta potencia A^k es totalmente positiva. El siguiente hecho relevante observado fue que las propiedades espectrales fundamentales de estas matrices (i.e., positividad y simplicidad de los valores propios, las leyes de alternancia de los signos de las coordenadas de los vectores propios, etc.) no están relacionadas de ninguna manera con la simetría de una matriz y son comunes a matrices oscilatorias, sean simétricas o no.

En las aplicaciones a problemas concretos de oscilación, por lo regular se manejan matrices oscilatorias simétricas. No obstante, dada la relevancia de este tipo de matrices desde un punto de vista puramente algebraico, nuestro objetivo a lo largo de este texto será dar una introducción a la teoría de matrices oscilatorias con la generalidad de no ceñir nuestro análisis a matrices simétricas solamente, tal como se expone en el libro intitulado *Oscillation Matrices and Kernels and Small Vibrations of Mechanical Systems* [6] de F. R. Gantmacher y M. G. Krein, con el cual se trabajó bajo la premisa de exponer con mayor claridad y precisión casi la totalidad de los dos primeros capítulos.

Así pues, las matrices oscilatorias aparecieron por vez primera en ciertos manuscritos de Gantmacher y Krein que posteriormente aparecieron recopilados en un libro publicado en Rusia en el año 1941, la primera edición, y en 1950 en una segunda edición. Escrito originalmente en ruso, varias traducciones se han hecho de él primero en alemán y poco después también en inglés. El libro [6] tal como apareció en 2002 está basado en estos tres antecedentes.

El primer capítulo es introductorio al estudio general de las matrices y sus operaciones así como de las formas cuadráticas. Si

$$A = \begin{pmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \dots & \dots & \dots & \dots \\ a_{m1} & a_{m2} & \dots & a_{mn} \end{pmatrix}$$

es una matriz rectangular con m filas y n columnas, cuyos elementos a_{ik} $(i = 1, 2, \dots, m; k = 1, 2, \dots, n)$ son números complejos, definimos el menor de un matriz como el determinante

$$A \begin{pmatrix} i_1 & i_2 & \dots & i_p \\ k_1 & k_2 & \dots & k_p \end{pmatrix} := \begin{vmatrix} a_{i_1k_1} & a_{i_1k_2} & \dots & a_{i_1k_p} \\ a_{i_2k_1} & a_{i_2k_2} & \dots & a_{i_2k_p} \\ \dots & \dots & \dots & \dots \\ a_{i_pk_1} & a_{i_pk_2} & \dots & a_{i_pk_p} \end{vmatrix} \qquad (p \le \text{mín}(m, n)).$$

En la primera sección del primer capítulo se demuestra el teorema de Binet-Cauchy:

Supongamos que la matriz cuadrada C es el producto de dos matrices rectangulares A y B de dimensiones $m \times n$ y $n \times m$, respectivamente. Entonces

$$C \begin{pmatrix} 1 & 2 & \dots & m \\ 1 & 2 & \dots & m \end{pmatrix} = \sum_{1 \le k_1 < k_2 < \dots < k_m \le n} A \begin{pmatrix} 1 & 2 & \dots & m \\ k_1 & k_2 & \dots & k_m \end{pmatrix} B \begin{pmatrix} k_1 & k_2 & \dots & k_m \\ 1 & 2 & \dots & m \end{pmatrix}$$

De donde se deriva el conocido resultado sobre el determinante de un producto de matrices cuadradas. La demostración de este teorema fue tomada del libro [5]. En la sección 1.2 se obtiene una fórmula para los menores de una matriz inversa que será de utilidad para demostrar la identidad de Sylvester, la cual se establece de la siguiente manera:

Si $A = (a_{ik})_1^n$ es una matriz arbitraria y $1 \leq p < n$, definimos

$$b_{ik} := A\begin{pmatrix} 1 & 2 & \dots & p & i \\ 1 & 2 & \dots & p & k \end{pmatrix} \qquad (i, k = 1, 2, \dots, n)$$

y $B := (b_{ik})_{p+1}^n$. Entonces

$$B\begin{pmatrix} p+1 & \dots & n \\ p+1 & \dots & n \end{pmatrix} = A\begin{pmatrix} 1 & 2 & \dots & p \\ 1 & 2 & \dots & p \end{pmatrix}^{n-p-1} A\begin{pmatrix} 1 & 2 & \dots & n \\ 1 & 2 & \dots & n \end{pmatrix}.$$

Esta misma fórmula servirá para probar algunos otros resultados en secciones posteriores. La sección 1.3 se caracteriza por ser un repaso a ciertos resultados fundamentales sobre vectores y valores propios de una matriz, y en la sección 1.4 este repaso se extiende a vectores y valores propios de matrices reales simétricas. Análogamente, en las secciones 1.5 y 1.6 repasaremos el concepto de forma cuadrática y veremos que una forma cuadrática puede reducirse a una suma de cuadrados, así como la relevancia de este hecho y un método para llevar a cabo tal reducción. La sección 1.7 contiene un criterio de positividad y uno de no-negatividad para una forma cuadrática. En la sección 1.8 se establecen tres desigualdades de Hadamard íntimamente relacionadas, de las cuales la más sencilla establece que si la forma $A(x, x)$ es positiva, entonces se satisface la desigualdad siguiente para su discriminante:

$$\begin{vmatrix} a_{11} & \dots & a_{1n} \\ \dots & \dots & \dots \\ a_{n1} & \dots & a_{nn} \end{vmatrix} \leq a_{11}a_{22}\cdots a_{nn},$$

aquí la igualdad se satisface si, y sólo si $a_{ik} = 0$ para $i \neq k$ $(i, k = 1, 2, \dots, n)$. En la sección 1.9, la última del primer capítulo, se estudia el concepto de matriz asociada de orden p de una matriz. Si A es una matriz de orden n, consideremos todas las combinaciones (subconjuntos con p elementos)

$$(i_1, i_2, \ldots, i_p)$$

de n índices $1, 2, \ldots, n$, y escribamos cada combinación en orden creciente: $i_1 < i_2 < \ldots < i_p$. De esta manera podemos ordenar todas las combinaciones al convenir en que una combinación $(i_1, i_2, \ldots, i_p)$ antecede a otra combinación $(k_1, k_2, \ldots, k_p)$ si, y sólo si el primer término distinto de cero en la sucesión

$$k_1 - i_1, k_2 - i_2, \ldots, k_p - i_p$$

es positivo. En consecuencia, podemos asignar un número bien definido s localizado entre 1 y $N := \binom{n}{p}$ a cada combinación. Para los menores de orden p de la matriz A introducimos la siguiente notación:

$$\mathfrak{a}_{st} := A\begin{pmatrix} i_1 & i_2 & \ldots & i_p \\ k_1 & k_2 & \ldots & k_p \end{pmatrix} \qquad (i_1 < i_2 < \ldots < i_p;\ k_1 < k_2 < \ldots < k_p),$$

donde s y t son, respectivamente, los números de las combinaciones $(i_1, i_2, \ldots, i_p)$ y $(k_1, k_2, \ldots, k_p)$. A la matriz

$$\mathfrak{A}_p = (\mathfrak{a}_{st})_1^N$$

la llamaremos *matriz asociada p-ésima* de A.

En esta sección se establecen ciertas propiedades de la matriz asociada que la relacionan con la matriz original.

En el segundo capítulo se entra de lleno al estudio de las matrices oscilatorias. Se dice que una matriz $A = (a_{ik})_1^n$ es *totalmente no-negativa* (respectivamente, *totalmente positiva*) si todos sus menores de cualquier orden son no-negativos (respectivamente, positivos):

$$A\begin{pmatrix} i_1 & i_2 & \dots & i_p \\ k_1 & k_2 & \dots & k_p \end{pmatrix} \geq 0 \quad (\text{resp.} > 0) \quad \text{para} \quad 1 \leq \begin{matrix} i_1 < i_2 < \dots < i_p \\ k_1 < k_2 < \dots < k_p \end{matrix} \leq n$$

$$(p = 1, 2, \dots, n).$$

A una matriz A la llamaremos *oscilatoria* si es totalmente no-negativa, y existe un entero positivo κ tal que A^κ es totalmente positiva. El menor de tales exponentes es el *exponente* de la matriz oscilatoria. A modo de introducción a este capítulo y como una forma de familiarizarse con el tipo de técnicas empleadas a lo largo del capítulo, en la sección 2.1 se desarrollan varias propiedades características de las matrices oscilatorias en las matrices de Jacobi normales, y esto se lleva a cabo de forma independiente al desarrollo general de la teoría de matrices oscilatorias que comprende las siguientes secciones. En la sección 2.2 se establecen las propiedades elementales de una matriz oscilatoria. La siguiente sección, 2.3, está dedicada por completo a la demostración del teorema de Perron dada por Frobenius, teorema cuyo enunciado es el siguiente:

Si todos los elementos de una matriz de orden n, A, son positivos, entonces existe un valor propio simple (i.e., una raíz simple del polinomio característico) y positivo ρ, cuyo módulo es mayor que el de todos los demás valores propios. A este "máximo" valor propio le corresponde un vector propio con coordenadas positivas.

En la sección 2.4 se establecen propiedades fundamentales de los valores y vectores propios de las matrices oscilatorias por medio del teorema fundamental de las propiedades espectrales de una matriz oscilatoria. Si $u = (u_1, u_2, \dots, u_n)$ es un vector, llamaremos *u-línea* a la curva en el plano cartesiano, recta por pedazos, cuyos vértices P_k tienen las siguientes coordenadas:

$$x_k = k, \quad y_k = u_k \quad (k = 1, 2, \dots, n).$$

A los puntos donde la u-línea cruza el eje X los llamaremos *nodos* de la u-línea o u-vector. Entonces en el teorema ya mencionado se establece lo siguiente:

Sea $A = (a_{ik})_1^n$ una matriz oscilatoria. Entonces

(i) La matriz A siempre tiene sólo valores propios simples y todos ellos son positivos: $\lambda_1 > \lambda_2 > \cdots > \lambda_n > 0$.

(ii) Si $u^k = (u_{1k}, u_{2k}, \ldots, u_{nk})$ es un vector propio de A correspondiente al valor propio λ_k, el k-ésimo en magnitud entre todos los valores propios $(k = 1, 2, \ldots, n)$, entonces dada cualquier sucesión de coeficientes $c_p, c_{p+1}, \ldots, c_q$ $(1 \leq p \leq q \leq n,\ \sum_{i=p}^{q} c_i^2 > 0)$, el número de cambios de signo en las coordenadas del vector $u := c_p u^p + c_{p+1} u^{p+1} + \cdots + c_q u^q$ se encuentra entre $p-1$ y $q-1$. En particular, entre las coordenadas del vector u^k $(k = 1, 2, \ldots, n)$ hay exactamente $k-1$ cambios de signo.

(iii) Los nodos de dos vectores propios sucesivos u^k y u^{k+1} $(k = 1, 2, \ldots, n-1)$ se alternan.

La sección 2.5 está dedicada a una desigualdad con determinantes que resulta fundamental para nuestro estudio:

$$A\begin{pmatrix} 1 & 2 & \ldots & n \\ 1 & 2 & \ldots & n \end{pmatrix} \leq A\begin{pmatrix} 1 & \ldots & p \\ 1 & \ldots & p \end{pmatrix} A\begin{pmatrix} p+1 & \ldots & n \\ p+1 & \ldots & n \end{pmatrix} \qquad (p < n).$$

Esta desigualdad se ve anteriormente bajo ciertas condiciones como la desigualdad generalizada de Hadamard, sin embargo, en este capítulo tal desigualdad es demostrada bajo distintas condiciones. Finalmente, en la última sección del segundo capítulo, la sección 2.6, se demuestra un teorema que proporciona una caracterización de las matrices oscilatorias, es decir, un criterio bajo el cual se identifica una matriz oscilatoria:

Para que una matriz totalmente no-negativa $A = (a_{ik})_1^n$ sea oscilatoria es necesario y suficiente que

(i) A sea una matriz no-singular.

(ii) $a_{i,i+1} > 0$ y $a_{i+1,i} > 0 \qquad (i = 1, 2, \ldots, n-1)$.

Adicionalmente se establecen algunos resultados que se derivan de forma inmediata al hacer uso de este teorema esencial.

Capítulo 1

Repaso de matrices y formas cuadráticas

1.1. Matrices y operaciones con matrices

Sea

$$A = \begin{pmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \dots & \dots & \dots & \dots \\ a_{m1} & a_{m2} & \dots & a_{mn} \end{pmatrix}$$

una matriz rectangular con m filas y n columnas, cuyos elementos a_{ik} $(i = 1, 2, \dots, m; k = 1, 2, \dots, n)$ son números complejos.

Introducimos la siguiente notación para los menores de esta matriz:

$$A \begin{pmatrix} i_1 & i_2 & \dots & i_p \\ k_1 & k_2 & \dots & k_p \end{pmatrix} := \begin{vmatrix} a_{i_1k_1} & a_{i_1k_2} & \dots & a_{i_1k_p} \\ a_{i_2k_1} & a_{i_2k_2} & \dots & a_{i_2k_p} \\ \dots & \dots & \dots & \dots \\ a_{i_pk_1} & a_{i_pk_2} & \dots & a_{i_pk_p} \end{vmatrix} \qquad (p \leq \text{mín}(m, n))$$

En la mayor parte del texto sólo consideraremos matrices cuadradas $(m = n)$. En este caso denotaremos la matriz A por $(a_{ik})_1^n$ y su determinante por $|A|$.

Dadas dos matrices $A = (a_{ik})_1^n$ y $B = (b_{ik})_1^n$ haremos corresponder una matriz $C = (c_{ik})_1^n$ llamada suma de A y B $(A + B)$ tal que

$$c_{ik} = a_{ik} + b_{ik} \quad (i, k = 1, 2, \ldots, n);$$

y una matriz $D = (d_{ik})_1^n$ llamada el producto de A y B $(D = AB)$ cuyos elementos están dados por

$$d_{ik} = a_{i1}b_{1k} + a_{i2}b_{2k} + \cdots + a_{in}b_{nk} \quad (i, k = 1, 2, \ldots, n)$$

Además, si $A = (a_{ik})_1^n$ y λ es un número, entonces el producto λA es la matriz $(\lambda a_{ik})_1^n$.

Se cumplen las siguientes propiedades de operaciones con matrices del mismo orden:

(i) $A + B = B + A, (A + B) + C = A + (B + C)$.

(ii) $(AB)C = A(BC), (A + B)C = AC + BC, A(B + C) = AB + AC$.

Observemos que la definición anterior del producto de dos matrices es aplicable también para matrices rectangulares, siempre que el número de columnas del primer factor sea igual al número de renglones del segundo factor.

Recordemos que si $C = AB$, entonces $|C| = |A||B|$. Este resultado es un caso especial del siguiente teorema de Binet-Cauchy.

Teorema 1.1. *Supongamos que la matriz cuadrada C es el producto de dos matrices rectangulares A y B de dimensiones $m \times n$ y $n \times m$, respectivamente. Entonces*

$$
\begin{aligned}
&C\begin{pmatrix} 1 & 2 & \dots & m \\ 1 & 2 & \dots & m \end{pmatrix} \\
&\quad = \sum_{1\le k_1<k_2<\dots<k_m\le n} A\begin{pmatrix} 1 & 2 & \dots & m \\ k_1 & k_2 & \dots & k_m \end{pmatrix} B\begin{pmatrix} k_1 & k_2 & \dots & k_m \\ 1 & 2 & \dots & m \end{pmatrix}
\end{aligned}
\tag{1.1}
$$

De acuerdo con esta fórmula, el determinante de C es la suma de todos los posibles productos de un menor de máximo orden (m-ésimo) de A y el menor de B simétrico con respecto a la diagonal principal correspondiente. En caso de tener $m > n$, no hay menores de m-ésimo orden de A ni de B, con lo cual la suma del lado derecho en 1.1 es igual a cero.

Demostración. El determinante de C puede verse de la forma

$$
\begin{aligned}
\begin{vmatrix} c_{11} & \dots & c_{1m} \\ \dots & \dots & \dots \\ c_{m1} & \dots & c_{mm} \end{vmatrix} &= \begin{vmatrix} \sum_{\alpha_1=1}^{n} a_{1\alpha_1}b_{\alpha_1 1} & \dots & \sum_{\alpha_m=1}^{n} a_{1\alpha_m}b_{\alpha_m m} \\ \dots & \dots & \dots \\ \sum_{\alpha_1=1}^{n} a_{m\alpha_1}b_{\alpha_1 1} & \dots & \sum_{\alpha_m=1}^{n} a_{m\alpha_m}b_{\alpha_m m} \end{vmatrix} \\
&= \sum_{\alpha_1,\dots,\alpha_m=1}^{n} \begin{vmatrix} a_{1\alpha_1}b_{\alpha_1 1} & \dots & a_{1\alpha_m}b_{\alpha_m m} \\ \dots & \dots & \dots \\ a_{m\alpha_1}b_{\alpha_1 1} & \dots & a_{m\alpha_m}b_{\alpha_m m} \end{vmatrix} \\
&= \sum_{\alpha_1,\dots,\alpha_m=1}^{n} A\begin{pmatrix} 1 & 2 & \dots & m \\ \alpha_1 & \alpha_2 & \dots & \alpha_m \end{pmatrix} b_{\alpha_1 1}b_{\alpha_2 2}\cdots b_{\alpha_m m}.
\end{aligned}
\tag{1.2}
$$

Si $m > n$, entre los números $\alpha_1, \alpha_2, \dots, \alpha_m$ siempre hay dos que son iguales, entonces cada sumando en el lado derecho de 1.2 es igual a cero. Luego, en este caso $|C| = 0$.

Ahora supongamos que $m \le n$. Entonces se anulan todos aquellos sumandos para los cuales dos de los números α_i son iguales. El resto de los sumandos puede separarse en grupos de $m!$ sumandos, cada uno de los cuales cuenta con términos que sólo se diferencian por el orden de los $\alpha_1, \alpha_2, \dots, \alpha_m$. De este modo, para cada uno de estos grupos tenemos la siguiente suma de sus elementos correspondientes

$$\sum \epsilon(\alpha_1, \alpha_2, \ldots, \alpha_m) A \begin{pmatrix} 1 & 2 & \ldots & m \\ k_1 & k_2 & \ldots & k_m \end{pmatrix} b_{\alpha_1 1} b_{\alpha_2 2} \cdots b_{\alpha_m m} =$$

$$= A \begin{pmatrix} 1 & 2 & \ldots & m \\ k_1 & k_2 & \ldots & k_m \end{pmatrix} \sum \epsilon(\alpha_1, \alpha_2, \ldots, \alpha_m) b_{\alpha_2 2} \cdots b_{\alpha_m m}$$
$$= A \begin{pmatrix} 1 & 2 & \ldots & m \\ k_1 & k_2 & \ldots & k_m \end{pmatrix} B \begin{pmatrix} k_1 & k_2 & \ldots & k_m \\ 1 & 2 & \ldots & m \end{pmatrix}.$$

Aquí, $k_1 < k_2 < \ldots < k_m$ es el orden de los $\alpha_1, \alpha_2, \ldots, \alpha_m$ y $\epsilon(\alpha_1, \alpha_2, \ldots, \alpha_m) = (-1)^N$, donde N es el número de transposiciones de los índices que es necesario componer para obtener la permutación $\alpha_1, \alpha_2, \ldots, \alpha_m$.

En consecuencia, 1.1 se sigue de 1.2. $\square$

Consideremos, ahora, dos matrices cuadradas $A = (a_{ik})_1^n$ y $B = (b_{ik})_1^n$ del mismo orden y digamos que $C := AB$. Entonces la matriz del menor

$$C \begin{pmatrix} i_1 & i_2 & \ldots & i_p \\ k_1 & k_2 & \ldots & k_p \end{pmatrix}$$

es el producto de dos matrices rectangulares:

$$\begin{pmatrix} a_{i_1 1} & a_{i_1 2} & \ldots & a_{i_1 n} \\ a_{i_2 1} & a_{i_2 2} & \ldots & a_{i_2 n} \\ \ldots & \ldots & \ldots & \ldots \\ a_{i_p 1} & a_{i_p 2} & \ldots & a i_p n \end{pmatrix} \quad \text{y} \quad \begin{pmatrix} b_{1k_1} & b_{1k_2} & \ldots & b_{1k_p} \\ b_{2k_1} & b_{2k_2} & \ldots & b_{2k_p} \\ \ldots & \ldots & \ldots & \ldots \\ b_{nk_1} & b_{nk_2} & \ldots & b_{nk_p} \end{pmatrix}$$

Luego, por el teorema de Binet-Cauchy,

$$
\begin{aligned}
&C\begin{pmatrix} i_1 & i_2 & \dots & i_p \\ k_1 & k_2 & \dots & k_p \end{pmatrix} \\
&\qquad = \sum A\begin{pmatrix} i_1 & i_2 & \dots & i_p \\ \alpha_1 & \alpha_2 & \dots & \alpha_p \end{pmatrix} B\begin{pmatrix} \alpha_1 & \alpha_2 & \dots & \alpha_p \\ k_1 & k_2 & \dots & k_p \end{pmatrix}
\end{aligned} \tag{1.3}
$$

donde la suma se extiende sobre todas las posibles combinaciones $(\alpha_1, \alpha_2, \dots, \alpha_p)$ de los n índices $1, 2, \dots, n$, tomando p cada vez. Esta fórmula nos proporciona una expresión para los menores del producto de dos matrices en términos de los menores de los factores.

Otra conclusión respecto de esta fórmula es la siguiente: Si r_A, r_B y r_C denotan los rangos de las matrices A, B y C, respectivamente, entonces se sigue de 1.3 que

$$r_C \leq r_A, r_B.$$

Si adicionalmente también se tiene que la matriz B no es singular, entonces $A = CB^{-1}$, y esto implica que

$$r_A \leq r_C, r_{B^{-1}},$$

luego,

$$r_C = r_A.$$

Así pues, cuando una matriz es multiplicada por una matriz no singular, su rango no cambia.

1.2. La identidad de Sylvester

En esta sección deduciremos una fórmula para los menores de la matriz inversa, y sobre la base para obtener este resultado hallaremos también la *identidad de Sylvester.*

Si $A = (a_{ik})_1^n$ es una matriz, denotaremos por $\tilde{A} = (\tilde{a}_{ik})_1^n$ a la matriz en la que cada entrada $\tilde{a}_{ik}$ es el menor del elemento a_{ik}, es decir,

$$\tilde{a}_{ik} = A\begin{pmatrix} 1 & 2 & \ldots & i-1 & i+1 & \ldots & n \\ 1 & 2 & \ldots & k-1 & k+1 & \ldots & 0 \end{pmatrix}.$$

Proposición 1.2. *La siguiente fórmula se cumple para los menores de $\tilde{A}$:*

$$\tilde{A}\begin{pmatrix} i_1 & i_2 & \ldots & i_p \\ k_1 & k_2 & \ldots & k_p \end{pmatrix} = |A|^{p-1}\begin{pmatrix} j_1 & j_2 & \cdots & j_{n-p} \\ l_1 & l_2 & \ldots & l_{n-p} \end{pmatrix}, \tag{1.4}$$

donde los sistemas de índices $i_1 < i_2 < \ldots < i_p, j_1 < j_2 < \ldots < j_{n-p}$ y $k_1 < k_2 < \ldots < k_p, l_1 < l_2 < \ldots < l_{n-p}$ coinciden con el sistema total de índices $1, 2, \ldots, n$.

Demostración. Probaremos esta fórmula primero para el menor

$$\tilde{A}\begin{pmatrix} 1 & 2 & \ldots & p \\ 1 & 2 & \ldots & p \end{pmatrix}$$

Si convenimos en la simplificación $A_{ik} := (-1)^{i+k}\tilde{a}_{ik}$ $(i, k = 1, 2, \ldots, n)$, entonces al multiplicar el i-ésimo renglón $(i = 1, 2, \ldots, p)$ en la matriz que subyace en el determinante $\tilde{A}\begin{pmatrix} 1 & 2 & \ldots & p \\ 1 & 2 & \ldots & p \end{pmatrix}$ por $(-1)^i$ y la k-ésima columna $(k = 1, 2, \ldots, p)$ por $(-1)^k$, obtenemos la siguiente igualdad:

$$\tilde{A}\begin{pmatrix} 1 & 2 & \dots & p \\ 1 & 2 & \dots & p \end{pmatrix} = \begin{vmatrix} A_{11} & \dots & A_{1p} & A_{1p+1} & \dots & \dots & A_{1n} \\ \dots & \dots & \dots & \dots & \dots & \dots & \dots \\ A_{p1} & \dots & A_{pp} & A_{pp+1} & \dots & \dots & A_{pn} \\ 0 & \dots & 0 & 1 & 0 & \dots & 0 \\ 0 & \dots & 0 & 0 & 1 & \dots & 0 \\ \dots & \dots & \dots & \dots & \dots & \dots & 0 \\ 0 & \dots & 0 & 0 & 0 & \dots & 1 \end{vmatrix}.$$

Multipliquemos ahora ambos lados de esta igualdad por el determinante $|A|$ y apliquemos la identidad $|X||Y| = |XY|$ en el lado derecho; de este modo, obtenemos

$$\tilde{A}\begin{pmatrix} 1 & 2 & \dots & p \\ 1 & 2 & \dots & p \end{pmatrix}|A| = \begin{vmatrix} |A| & 0 & \dots & 0 & 0 & \dots & 0 \\ 0 & |A| & \dots & 0 & 0 & \dots & 0 \\ \dots & \dots & \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & \dots & |A| & 0 & \dots & 0 \\ a_{1,p+1} & a_{2,p+1} & \dots & a_{p,p+1} & a_{p+1,p+1} & \dots & a_{n,p+1} \\ \dots & \dots & \dots & \dots & \dots & \dots & \dots \\ a_{1n} & a_{2n} & \dots & a_{pn} & a_{p+1,n} & \dots & a_{nn} \end{vmatrix}.$$

Pero el determinante del lado derecho es igual a

$$|A|^p A\begin{pmatrix} p+1 & \dots & n \\ p+1 & \dots & n \end{pmatrix}.$$

Luego, si $|A| \neq 0$, entonces

$$\tilde{A}\begin{pmatrix} 1 & 2 & \dots & p \\ 1 & 2 & \dots & p \end{pmatrix} = |A|^{p-1} A\begin{pmatrix} p+1 & \dots & n \\ p+1 & \dots & n \end{pmatrix}. \tag{1.5}$$

Observemos que la función del espacio de matrices de un mismo orden a la recta real que asigna a cada matriz su determinante (o un menor determinado), evidentemente es continua. De ahí concluimos que la igualdad 1.5 se cumple también cuando $|A| = 0$.

Ahora, para obtener la fórmula 1.4 en el caso general, observemos que haciendo el reordenamiento adecuado de renglones y columnas de la matriz A obtenemos 1.4 a partir de 1.5. □

Cuando $p = n$ la fórmula 1.4 adquiere la forma

$$|\tilde{A}| = |A|^{n-1}. \tag{1.6}$$

De la fórmula para los menores de la matriz $\tilde{A}$ obtenemos la siguiente fórmula para los menores de la matriz inversa.

Proposición 1.3. *Si $A = (a_{ik})_1^n$ y $B = A^{-1}$, entonces*

$$B\begin{pmatrix} i_1 & i_2 & \dots & i_p \\ k_1 & k_2 & \dots & k_p \end{pmatrix} = (-1)^{\sum_1^p i_\nu + \sum_1^p k_\nu} \frac{A\begin{pmatrix} l_1 & l_2 & \dots & l_{n-p} \\ j_1 & j_2 & \dots & j_{n-p} \end{pmatrix}}{A\begin{pmatrix} 1 & 2 & \dots & n \\ 1 & 2 & \dots & n \end{pmatrix}} \tag{1.7}$$

donde cada uno de los sistemas de índices $i_1 < i_2 < \ldots < i_p$; $j_1 < j_2 < \ldots < j_{n-p}$ y $k_1 < k_2 < \ldots < k_p$; $l_1 < l_2 < \ldots < l_{n-p}$ coincide con el sistema de índices $1, 2, \ldots, n$.

Demostración. En efecto, $B = (b_{ik})_1^n$ donde $b_{ik} = A_{ki}/|A| = (-1)^{i+k}\tilde{a}_{ki}/|A|$ $(i, k = 1, 2, \ldots, n)$. Entonces

$$B\begin{pmatrix} i_1 & i_2 & \dots & i_p \\ k_1 & k_2 & \dots & k_p \end{pmatrix} = (-1)^{\sum_1^p i_\nu + \sum_1^p k_\nu} \frac{\tilde{A}\begin{pmatrix} k_1 & k_2 & \dots & k_p \\ i_1 & i_2 & \dots & i_p \end{pmatrix}}{|A|^p}$$

y de aquí, al aplicar 1.4, obtenemos 1.7. □

Teorema 1.4. (Sylvester). *Sea $A = (a_{ik})_1^n$ una matriz arbitraria y $1 \leq p < n$. Definimos*

$$b_{ik} := A\begin{pmatrix} 1 & 2 & \dots & p & i \\ 1 & 2 & \dots & p & k \end{pmatrix} \qquad (i, k = 1, 2, \dots, n)$$

y $B = (b_{ik})_{p+1}^{n}$. *Entonces*

$$B\begin{pmatrix} p+1 & \dots & n \\ p+1 & \dots & n \end{pmatrix} = A\begin{pmatrix} 1 & 2 & \dots & p \\ 1 & 2 & \dots & p \end{pmatrix}^{n-p-1} A\begin{pmatrix} 1 & 2 & \dots & n \\ 1 & 2 & \dots & n \end{pmatrix} \tag{1.8}$$

Demostración. Por 1.4 tenemos

$$\begin{aligned} &\tilde{A}\begin{pmatrix} p+1 & \dots & r-1 & r+1 & \dots & n \\ p+1 & \dots & s-1 & s+1 & \dots & n \end{pmatrix} \\ &\qquad = A\begin{pmatrix} 1 & 2 & \dots & p & r \\ 1 & 2 & \dots & p & s \end{pmatrix} |A|^{n-p-2} = b_{rs}|A|^{n-p-2}. \end{aligned} \tag{1.9}$$

Por otro lado, si definimos la matriz $(c_{rs})_{p+1}^{n}$ por

$$c_{rs} = \tilde{A}\begin{pmatrix} p+1 & \dots & r-1 & r+1 & \dots & n \\ p+1 & \dots & s-1 & s+1 & \dots & n \end{pmatrix},$$

de acuerdo con la igualdad 1.6, tenemos

$$C\begin{pmatrix} p+1 & \dots & n \\ p+1 & \dots & n \end{pmatrix} = \tilde{A}\begin{pmatrix} p+1 & \dots & n \\ p+1 & \dots & n \end{pmatrix}^{n-p-1} \tag{1.10}$$

Pero de acuerdo con 1.9,

$$C\begin{pmatrix} p+1 & \dots & n \\ p+1 & \dots & n \end{pmatrix} = B\begin{pmatrix} p+1 & \dots & n \\ p+1 & \dots & n \end{pmatrix} |A|^{(n-p-2)(n-p)},$$

y de acuerdo con 1.4,

$$\tilde{A}\begin{pmatrix} p+1 & \dots & n \\ p+1 & \dots & n \end{pmatrix}^{n-p-1} = A\begin{pmatrix} 1 & 2 & \dots & p \\ 1 & 2 & \dots & p \end{pmatrix}^{n-p-1} |A|^{(n-p-1)^2}.$$

Al igualar los lados derechos de estas dos ecuaciones, tomando en cuenta 1.10, llegamos a

$$B\begin{pmatrix} p+1 & \dots & n \\ p+1 & \dots & n \end{pmatrix} |A|^{(n-p-2)(n-p)} = A\begin{pmatrix} 1 & 2 & \dots & p \\ 1 & 2 & \dots & p \end{pmatrix}^{n-p-1} |A|^{(n-p-1)^2}$$

y, al dividir por $|A|^{(n-p-2)(n-p)}$ ambos lados de esta igualdad (bajo la suposición de que $|A| \neq 0$), llegamos a la identidad de Sylvester 1.8. Nuevamente, sobre un argumento de continuidad, tenemos que también se cumple 1.8 cuando $|A| = 0$. □

De ahora en adelante, podemos hacer uso de la fórmula

$$\begin{aligned} B\begin{pmatrix} i_1 & \dots & i_q \\ k_1 & \dots & k_q \end{pmatrix} = & \\ &= A\begin{pmatrix} 1 & 2 & \dots & p \\ 1 & 2 & \dots & p \end{pmatrix}^{q-1} A\begin{pmatrix} 1 & 2 & \dots & p & i_1 & \dots & i_q \\ 1 & 2 & \dots & p & k_1 & \dots & k_q \end{pmatrix} \end{aligned} \tag{1.11}$$

la cual es, desde luego, la identidad de Sylvester para la matriz del determinante

$$A\begin{pmatrix} 1 & 2 & \dots & p & i_1 & \dots & i_q \\ 1 & 2 & \dots & p & k_1 & \dots & k_q \end{pmatrix}.$$

1.3. Valores y vectores propios de una matriz

Para esta sección puede ser de utilidad el material proporcionado por el libro [1] en sus secciones 5 y 6.

Sea A una matriz de orden n. Consideremos el producto $\operatorname{sgn}(\sigma)(a_{1\sigma(1)} - \lambda\delta_{1\sigma(1)})$ $(a_{2\sigma(2)} - \lambda\delta_{2\sigma(2)}) \cdots (a_{n\sigma(n)} - \lambda\delta_{n\sigma(n)})$ como uno de los sumandos en la definición del determinante $|A - \lambda E|$; así pues, σ es una permutación de $1, 2, \ldots, n$ y $\operatorname{sgn}(\sigma)$ denota su signo. Observemos que en el desarrollo de tal producto aparece la potencia $(-\lambda)^k$ si, y sólo si σ deja fijos por lo menos k elementos. En tal caso, $(-\lambda)^k$ se multiplica por productos de la forma $\operatorname{sgn}(\sigma) a_{i_1\sigma(i_1)} a_{i_2\sigma(i_2)} \ldots a_{i_{n-k}\sigma(i_{n-k})}$. Aquí σ puede considerarse como una permutación de $i_1, i_2, \ldots, i_{n-k}$, y esta restricción de la σ original preserva el signo porque se trata de una restricción que sólo excluye elementos que quedan fijos bajo σ. De tal manera que al simplificar la expansión de $|A - \lambda E|$ y obtener la forma general de un polinomio en términos de $-\lambda$, hallaremos que el coeficiente correspondiente a $(-\lambda)^k$ es igual a la suma de todos los menores principales de orden $n - k$ de la matriz A. Es decir,

$$|A - \lambda E| = (-\lambda)^n + S_1(-\lambda)^{n-1} + S_2(-\lambda)^{n-2} + \cdots + S_n,$$

donde $S_1 := \sum_i a_{ii}$, $S_2 := \sum_{i<k} \begin{vmatrix} a_{ii} & a_{ik} \\ a_{ki} & a_{kk} \end{vmatrix}$, etc., es decir; S_k $(k = 1, 2, \ldots, n)$ es la suma de todos los menores principales de k-ésimo orden de la matriz A.

Emplearemos este resultado más adelante.

A un sistema ordenado de n números complejos $x_1, x_2, \ldots, x_n$ lo llamaremos *vector* x y a los números x_i $(i = 1, 2, \ldots, n)$, coordenadas de este vector. Escribiremos $x = (x_1, x_2, \ldots, x_n)$. Al vector $(0, 0, \ldots, 0)$ lo denotaremos por 0.

Dados cualesquiera k vectores

$$x^1 = (x_{11}, x_{21}, \ldots, x_{n1}), x^2 = (x_{12}, x_{22}, \ldots, x_{n2}), \ldots, x^k = (x_{1k}, x_{2k}, \ldots, x_{nk})$$

y k números $a_1, a_2, \ldots, a_k$, podemos formar un vector $x = a_1x^1 + a_2x^2 + \cdots + a_kx^k$

cuyas coordenadas están dadas por la fórmula

$$x_i = a_1 x_{i1} + a_2 x_{i2} + \cdots + a_k x_{ik} \qquad (i = 1, 2, \ldots, n).$$

Diremos que los vectores $x^1, x^2, \ldots, x^k$ son *linealmente dependientes* (*independientes*) si existe (si no existe) un sistema de números $a_1, a_2, \ldots, a_k$ con alguno de ellos distinto de cero tal que

$$a_1 x^1 + a_2 x^2 + \cdots + a_k x^k = 0.$$

Definimos el producto de una matriz $A = (a_{ik})_1^n$ por un vector x como el vector x' dado por

$$x'_i = \sum_{k=1}^{n} a_{ik} x_k \qquad (i = 1, 2, \ldots, n).$$

Escribiremos

$$x' = Ax.$$

Desde luego,

$$A(a_1 x^1 + a_2 x^2 + \cdots + a_k x^k) = a_1 A x^1 + a_2 A x^2 + \cdots + a_k A x^k.$$

Más aún, si

$$x'' = Ax' \qquad y \qquad x' = Bx$$

donde $A = (a_{ik})_1^n$ y $B = (b_{ik})_1^n$. Entonces

$$x''_i = \sum_{j=1}^{n} a_{ij} x'_j, \quad x'_j = \sum_{k=1}^{n} b_{jk} x_k \qquad (i, j = 1, 2, \ldots, n)$$

luego

$$x''_i = \sum_{k=1}^{n} c_{ik} x_k \qquad (i = 1, 2, \ldots, n)$$

donde

$$c_{ik} = \sum_{j=1}^{n} a_{ij} b_{jk} \qquad (i, k = 1, 2, \ldots, n);$$

es decir,

$$x'' = Cx \qquad (C = AB).$$

Dicho de otra manera:

$$A(Bx) = (AB)x.$$

De esto obtenemos, en particular, que si $|A| \neq 0$ y $x' = Ax$, entonces $x = A^{-1}x'$.

Cuando sucede que

$$Ax = \lambda x \tag{1.12}$$

con $x \neq 0$, diremos que x es un vector propio de la matriz A y λ su correspondiente valor propio.

La ecuación 1.12 puede escribirse en coordenadas como el siguiente sistema de ecuaciones

$$\sum_{k=1}^{n} a_{ik} x_k = \lambda x_i \qquad (i = 1, 2, \ldots, n),$$

o, con la ayuda del símbolo de Kronecker δ_{ik} ($\delta_{ik} = 1$ cuando $i = k$ y $\delta_{ik} = 0$ cuando $i \neq k$), en la forma

$$\sum_{k=1}^{n}(a_{ik} - \lambda\delta_{ik})x_k = 0 \qquad (i = 1, 2, \ldots, n). \tag{1.13}$$

El sistema de ecuaciones anterior tiene una solución no trivial ($x \neq 0$) si, y sólo si el determinante

$$\Delta(\lambda) := |a_{ik} - \lambda\delta_{ik}|_1^n = (-\lambda)^n + S_1(-\lambda)^{n-1} + S_2(-\lambda)^{n-2} + \cdots + S_n$$

es igual a cero. Aquí los coeficientes S_k tienen la definición que se hizo al principio de esta sección. $\Delta(k)$ es llamado usualmente el *determinante característico* de la matriz A y la ecuación $\Delta(k) = 0$ es llamada *ecuación característica* de A.

Así pues, un número λ es un valor propio de la matriz A si, y sólo si coincide con alguna raíz de la ecuación característica de A. Ya que la ecuación característica es una ecuación algebraica de grado n, ésta tiene n raíces complejas (contadas con su multiplicidad). En consecuencia, toda matriz tiene por lo menos un valor propio.

Resulta obvio el hecho de que si un vector u es un vector propio, entonces el vector cu (donde c es un número distinto de 0) también es un vector propio correspondiente al mismo valor propio. En general, si un mismo valor propio corresponde a varios vectores propios $u^1, u^2, \ldots, u^k$, entonces cualquier combinación lineal de estos vectores $a_1u^1 + a_2u^2 + \cdots + a_ku^k$ produce un vector propio para el mismo valor propio, siempre que esta combinación sea distinta de cero.

De un conocido resultado acerca de las soluciones de un sistema homogéneo de ecuaciones lineales (véase [1], pág. 149), se sigue que el número máximo d_0 de vectores propios linealmente independientes que corresponden a un mismo valor propio λ_0 es igual al defecto (i.e., la diferencia entre el orden y el rango) de la matriz del sistema de ecuaciones 1.13 (con $\lambda = \lambda_0$), es decir, d_0 es igual al defecto de la matriz

$$A - \lambda_0 E = (a_{ik} - \lambda_0 \delta_{ik})_1^n.$$

Proposición 1.5. *Si denotamos por m_0 la multiplicidad de λ_0, visto como una raíz de la ecuación característica $\Delta(\lambda) = 0$, entonces tenemos la siguiente desigualdad*

$$d_0 \leq m_0.$$

Demostración. En efecto, si establecemos las identidades

$$B := A - \lambda_0 E \quad y \quad \rho := \lambda - \lambda_0,$$

tenemos

$$\Delta(\lambda) = |A - \lambda E| = |B - \rho E| = T_n - \rho T_{n-1} + \rho^2 T_{n-2} + \cdots + (-\rho)^{n-1} T_1 + (-\rho)^n,$$

donde T_p es la suma de todos los menores principales de orden p de la matriz B $(p = 1, 2, \ldots, n)$. Pero, puesto que el rango de esta matriz es $n - d_0$, tenemos que

$$T_n = T_{n-1} = \cdots = T_{n-d_0+1} = 0$$

y, en consecuencia,

$$\Delta(\lambda) = \pm(\lambda - \lambda_0)^{d_0}(T_{n-d_0} - T_{n-d_0-1}(\lambda - \lambda_0) + \ldots).$$

Luego, la multiplicidad m_0 de λ_0 es por lo menos d_0. □

De ahora en adelante, por lo regular manejaremos matrices para las cuales $d_0 = m_0$. Esto no sucede siempre, como contraejemplo veamos el siguiente:

Consideremos la matriz

$$A = \begin{pmatrix} \lambda_0 & 1 & 0 & \dots & 0 \\ 0 & \lambda_0 & 1 & \dots & 0 \\ 0 & 0 & \lambda_0 & \dots & 0 \\ \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & 0 & \dots & 1 \\ 0 & 0 & 0 & \dots & \lambda_0 \end{pmatrix}.$$

Entonces

$$\Delta(\lambda) = (\lambda_0 - \lambda)^n.$$

Con lo cual, en este caso $m_0 = n$. Por otro lado, la matriz

$$A - \lambda_0 E = \begin{pmatrix} 0 & 1 & 0 & \dots & 0 \\ 0 & 0 & 1 & \dots & 0 \\ \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & 0 & \dots & 1 \\ 0 & 0 & 0 & \dots & 0 \end{pmatrix}$$

tiene rango $n - 1$, luego, $d_0 = 1$.

Proposición 1.6. *Sea A una matriz arbitraria. Ocurre que no importa qué vectores propios $u^1, u^2, \dots, u^k$ correspondientes a valores propios distintos escojamos, éstos siempre serán linealmente independientes.*

Demostración. Supongamos que

$$c_1 u^1 + c_2 u^2 + \dots + c_k u^k = 0 \tag{1.14}$$

Veamos que en esta relación cualquier coeficiente, por ejemplo c_k, es igual a cero.

Para este propósito, multiplicaremos ambos lados de 1.14 por la matriz A. Así, obtenemos

$$\lambda_1 c_1 u^1 + \lambda_2 c_2 u^2 + \cdots + \lambda_k c_k u^k = 0.$$

Al restar esta ecuación a la ecuación 1.14 multiplicada por λ_1 obtenemos

$$c_2' u^2 + \cdots + c_k' u^k = 0 \tag{1.15}$$

donde $c_k' = (\lambda_1 - \lambda_k)c_k$. Si ahora multiplicamos 1.15 por A para obtener

$$\lambda_2 c_2' u^2 + \cdots + \lambda_k c_k' u^k = 0,$$

y, al igual que antes, restamos esta ecuación a la ecuación 1.15 multiplicada por λ_2 obtendremos una ecuación con un sumando menos y con el vector u^k multiplicado por el escalar $(\lambda_2 - \lambda_k)(\lambda_1 - \lambda_k)c_k$.

Siguiendo de esta manera, eventualmente obtendremos la ecuación

$$(\lambda_1 - \lambda_k)(\lambda_2 - \lambda_k)\cdots c_k u^k = 0.$$

De aquí que $(\lambda_1 - \lambda_k)(\lambda_2 - \lambda_k)\cdots c_k = 0$ .Como estamos considerando valores propios distintos, se sigue que $c_k = 0$. □

Si

$$\lambda_1, \lambda_2, \ldots, \lambda_p$$

son todos los valores propios distintos de la matriz A y $d_1, d_2, \ldots, d_p$ los defectos de las matrices $A - \lambda_1 E$, $A - \lambda_2 E$,..., $A - \lambda_p E$, sea

$$u_1^1, u_1^2, \ldots, u_1^{d_1}; u_2^1, u_2^2, \ldots, u_2^{d_2}; \ldots; u_p^1, u_p^2, \ldots, u_p^{d_p} \tag{1.16}$$

un sistema de vectores propios de la matriz A que consista de d_1 vectores propios linealmente independientes correspondientes al número λ_1, d_2 vectores propios linealmente independientes correspondientes al número λ_2, etc. De la proposición 1.6 se sigue que los vectores del sistema 1.16 son linealmente independientes. A un sistema de vectores como el de 1.16 lo llamaremos *sistema completo de vectores propios* de la matriz A.

Diremos que la matriz A es *diagonalizable* si tiene n vectores propios linealmente independientes, con n igual al orden de la matriz. Como $d_j \leq m_j$ $(j = 1, 2, \ldots, p)$, donde m_j es la multiplicidad del valor propio correspondiente, y $\sum_{j=1}^{p} m_j = n$, tenemos que el sistema completo de vectores propios 1.16 consiste de n vectores si, y sólo si

$$d_j = m_j \qquad (j = 1, 2, \ldots, p);$$

es decir, la matriz A es diagonalizable si, y sólo si $d_j = m_j$ para cada j.

Sea A una matriz diagonalizable. Enumeremos los vectores propios del sistema completo 1.16 de esta manera:

$$u^1, u^2, \ldots, u^n.$$

Denotaremos los valores propios correspondientes a estos vectores por

$$\lambda_1, \lambda_2, \ldots, \lambda_n.$$

Obviamente en esta lista cada valor propio estará repetido de acuerdo a su multiplicidad, de modo tal que

$$\Delta(\lambda) = (\lambda_1 - \lambda)(\lambda_2 - \lambda)\dots(\lambda_n - \lambda).$$

Introducimos, ahora, la siguiente notación para las coordenadas de los vectores u^j:

$$u^j := (u_{1j}, u_{2j}, \dots, u_{nj}) \qquad (j = 1, 2, \dots, n).$$

A la matriz

$$U = (u_{ij})_1^n$$

la llamaremos *matriz fundamental* de la matriz A.

Proposición 1.7. *Una matriz A es diagonalizable si, y sólo si tiene una representación de la forma*

$$A = U \begin{pmatrix} \lambda_1 & 0 & 0 & \dots & 0 \\ 0 & \lambda_2 & 0 & \dots & 0 \\ \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & 0 & \dots & \lambda_n \end{pmatrix} U^{-1}, \tag{1.17}$$

donde $|U| \neq 0$.

Demostración. Sea A una matriz diagonalizable con matriz fundamental U. Como la j-ésima columna de la matriz fundamental consiste de las coordenadas del vector u^j $(j = 1, 2, \dots, n)$, y estos vectores son linealmente independientes, entonces

$$|U| \neq 0. \tag{1.18}$$

Consideremos las n ecuaciones vectoriales:

$$Au^j = \lambda_j u^j$$

o, en coordenadas,

$$\sum_{k=1}^{n} a_{ik} u_{kj} = \lambda_j u_{ij} \qquad (i = 1, 2, \ldots, n).$$

Al asignar los valores $1, 2, \ldots, n$ a j obtenemos n^2 ecuaciones que en su conjunto podemos expresar en una simple ecuación matricial:

$$AU = U(\lambda_i \delta_{ik})_1^n.$$

De aquí, en virtud de 1.18, obtenemos 1.17.

Dado que todas las afirmaciones hechas hasta ahora son reversibles, tenemos que si A tiene una representación de la forma 1.17 con una matriz U no singular y $\lambda_1, \lambda_2, \ldots, \lambda_n$ n números complejos cualesquiera, podemos concluir que U es una matriz fundamental para la matriz A con valores propios $\lambda_1, \lambda_2, \ldots, \lambda_n$ ordenados adecuadamente. En consecuencia, la matriz A es diagonalizable. $\square$

Llamaremos *matriz transpuesta* de la matriz $A = (a_{ik})_1^n$ a la matriz $A^t = (b_{ik})_1^n$ tal que

$$b_{ik} := a_{ki} \qquad (i, k = 1, 2, \ldots, n)$$

Una matriz $A = (a_{ik})_1^n$ es *simétrica* si coincide con su matriz transpuesta, i.e., si

$$a_{ik} = a_{ki} \qquad (i, k = 1, 2, \ldots, n)$$

Para cualesquiera dos matrices A y B de orden n, evidentemente se cumple que

$$(A^t)^t = A, \quad (A+B)^t = A^t + B^t, \quad (\lambda A)^t = \lambda A^t.$$

También se cumple que

$$(AB)^t = B^t A^t. \tag{1.19}$$

En particular, si $B = A^{-1}$, entonces $B^t A^t = E$, es decir,

$$(A^{-1})^t = (A^t)^{-1}. \tag{1.20}$$

Definamos ahora la matriz $V = (v_{ik})_1^n$ por

$$V := (U^{-1})^t. \tag{1.21}$$

Al pasar a las matrices transpuestas en 1.17, gracias a las propiedades que se acaban de mencionar, obtenemos

$$A^t = V(\lambda_i \delta_{ik})_1^n V^{-1}.$$

Así pues, la matriz V es fundamental para la matriz transpuesta A^t. En consecuencia, los vectores

$$v^j = (v_{1j}, v_{2j}, \ldots, v_{nj}) \qquad (j = 1, 2, \ldots, n)$$

conforman un sistema completo de vectores propios de la matriz A^t con

$$A^t v^j = \lambda_j v^j \qquad (j = 1, 2, \ldots, n).$$

Como es usual, el producto escalar (x, y) de dos vectores $x = (x_1, x_2, \ldots, x_n)$ y $y = (y_1, y_2, \ldots, y_n)$ está definido por la relación

$$(x, y) := \sum_{i=1}^{n} x_i y_i \,,$$

y la longitud de un vector con coordenadas reales $x = (x_1, x_2, \ldots, x_n)$ se define por la ecuación

$$|x| = \sqrt{(x, x)} = \sqrt{\sum_{i=1}^{n} x_i^2}.$$

Dos vectores x y y son *ortogonales* si $(x, y) = 0$. Dos sistemas de vectores $x^1, x^2, \ldots, x^p$ y $y^1, y^2, \ldots, y^p$ son ortogonales si

$$(x^i, y^k) = \delta_{ik} \qquad (i, k = 1, 2, \ldots, p).$$

En particular, si un sistema de vectores $x^1, x^2, \ldots, x^p$ es ortogonal consigo mismo, es decir, si

$$(x^i, x^k) = \delta_{ik} \qquad (i, k = 1, 2, \ldots, p),$$

diremos que es *ortonormal*.

Proposición 1.8. *Sean $x^1, x^2, \ldots, x^p$ y $y^1, y^2, \ldots, y^p$ dos sistemas de vectores ortogonales. Entonces cada uno de los sistemas consiste de vectores linealmente independientes.*

Demostración. Supongamos que

$$\sum_{i=1}^{p} c_i x^i = 0$$

Luego,

$$0 = (\sum_{i=1}^{p} c_i x^i, y^k) = \sum_{i=1}^{p} c_i (x^i, y^k) = c_k (x^k, y^k) = c_k \qquad (k = 1, 2, \ldots, p).$$

Simétricamente se demuestra que el sistema de vectores $y^1, y^2, \ldots, y^p$ es linealmente independiente también. □

Para finalizar con esta sección, notemos que la igualdad 1.21 implica que

$$U^t V = E$$

o, dicho de otra manera,

$$(u^i, v^k) = \sum_{s=1}^{n} u_{si} v_{sk} = \delta_{ik} \qquad (i, k = 1, 2, \ldots, n).$$

De modo tal que los sistemas de vectores $u^1, u^2, \ldots, u^n$ y $v^1, v^2, \ldots, v^n$ son ortogonales.

Existen dos casos relevantes en los que una matriz A es diagonalizable, a saber,

a) La ecuación característica de la matriz $A = (a_{ik})_1^n$ tiene n raíces distintas.

b) La matriz A es real y simétrica.

En el caso a) la matriz A tiene n valores propios distintos que corresponden a n vectores propios linealmente independientes. El caso b) será discutido en la siguiente sección.

1.4. Matrices reales y simétricas

Sea A una matriz cuadrada de orden n. Para cualesquiera dos vectores

$$x = (x_1, x_2, \ldots, x_n)$$

y

$$y = (y_1, y_2, \ldots, y_n)$$

se cumple la siguiente ecuación

$$(Ax, y) = \sum_{i,k=1}^{n} a_{ik} x_k y_i = (x, A^t y). \tag{1.22}$$

Si A es simétrica, entonces para cualesquiera vectores x y y tenemos

$$(Ax, y) = (x, Ay).$$

Proposición 1.9. *Todos los valores propios de una matriz real simétrica son reales.*

Demostración. Supongamos que

$$Ax = \lambda x \tag{1.23}$$

con $x = (x_1, x_2, \ldots, x_n)$ diferente de cero. Pongamos $\bar{x} := (\bar{x}_1, \bar{x}_2, \ldots, \bar{x}_n)$, donde la barra denota *tomar* el conjugado complejo. De acuerdo con 1.23 tenemos que

$$\lambda \sum_{i=1}^{n} |x_i|^2 = \lambda(x, \bar{x}) = (\lambda x, \bar{x}) = (Ax, \bar{x}) = \sum_{i,k=1}^{n} a_{ik} x_k \bar{x}_i =$$

$$= \sum_{i=1}^{n} a_{ii} |x_i|^2 + \sum_{i<k} a_{ik}(x_i \bar{x}_k + x_k \bar{x}_i) = \sum_{i=1}^{n} a_{ii} |x_i|^2 + \sum_{i<k} a_{ik}(x_i \bar{x}_k + \overline{(x_i \bar{x}_k)}).$$

La última de estas expresiones es real porque A es una matriz real, en consecuencia, λ es real. □

Proposición 1.10. *Los vectores propios de una matriz real y simétrica correspondientes a distintos valores propios son ortogonales.*

Demostración. Si tenemos que

$$Ax = \lambda x \quad (x \neq 0)$$

y

$$Ay = \mu y \quad (y \neq 0),$$

entonces

$$\lambda(x, y) = (Ax, y) = (x, Ay) = \mu(x, y).$$

De modo tal que si $\lambda \neq \mu$,

$$(x.y) = 0.$$

□

Proposición 1.11. *Sean A una matriz real y simétrica y λ_0 uno de sus valores propios. Si denotamos por m la multiplicidad del valor propio λ_0, tenemos que a λ_0 le corresponden tantos vectores propios linealmente independientes como m, es decir, $d = m$, donde d es el defecto de la matriz $B = A - \lambda_0 E$.*

Demostración. En la demostración de la proposición 1.5 obtuvimos

$$\Delta(\lambda) = |A - \lambda E| = |B - (\lambda - \lambda_0)E| =$$
$$\pm(\lambda - \lambda_0)^d (T_r - T_{r-1}(\lambda - \lambda_0) + \dots) \qquad (r + d = n)$$

Entonces bastará probar que $T_r \neq 0$. Recordemos que T_k denota la suma de los menores principales de orden k de la matriz B, que en este caso es real y simétrica, y que $r = n - d$ es el rango de esta matriz.

Sin pérdida de generalidad, podemos suponer que los primeros r renglones de la matriz $B = (b_{ik})_1^n$ son linealmente independientes, porque de otro modo podemos aplicar una permutación en filas y columnas de tal manera que se preserve la simetría de B y que coloque los renglones de B linealmente independientes al principio en la matriz. Entonces cada renglón $b^i = (b_{i1}, b_{i2}, \dots, b_{in})$ puede verse como una combinación lineal de los primeros r renglones, es decir,

$$b^i = \sum_{p=1}^{r} c_{ip} b^p \qquad (i = 1, 2, \dots, n),$$

equivalentemente,

$$b_{ik} = \sum_{p=1}^{r} c_{ip} b_{pk} \qquad (i, k = 1, 2, \dots, n). \tag{1.24}$$

Dado que la matriz B es simétrica, las mismas relaciones se cumplen también para las columnas con los mismos coeficientes c_{ip}:

$$b_{ik} = \sum_{p=1}^{r} c_{kp} b_{ip} \qquad (i, k = 1, 2, \dots, n). \tag{1.25}$$

De la conocida igualdad con producto de determinantes junto con 1.24 y 1.25 se siguen las igualdades:

$$B\begin{pmatrix} i_1 & i_2 & \dots & i_r \\ k_1 & k_2 & \dots & k_r \end{pmatrix} = C\begin{pmatrix} i_1 & i_2 & \dots & i_r \\ 1 & 2 & \dots & r \end{pmatrix} B\begin{pmatrix} 1 & 2 & \dots & r \\ k_1 & k_2 & \dots & k_r \end{pmatrix},$$

$$B\begin{pmatrix} 1 & 2 & \dots & r \\ k_1 & k_2 & \dots & k_r \end{pmatrix} = C\begin{pmatrix} k_1 & k_2 & \dots & k_r \\ 1 & 2 & \dots & r \end{pmatrix} B\begin{pmatrix} 1 & 2 & \dots & r \\ 1 & 2 & \dots & r \end{pmatrix}.$$

Luego,

$$\begin{aligned} & B\begin{pmatrix} i_1 & i_2 & \dots & i_r \\ k_1 & k_2 & \dots & k_r \end{pmatrix} = \\ & C\begin{pmatrix} i_1 & i_2 & \dots & i_r \\ 1 & 2 & \dots & r \end{pmatrix} C\begin{pmatrix} k_1 & k_2 & \dots & k_r \\ 1 & 2 & \dots & r \end{pmatrix} B\begin{pmatrix} 1 & 2 & \dots & r \\ 1 & 2 & \dots & r \end{pmatrix}. \end{aligned} \tag{1.26}$$

Como no es posible que todos los menores

$$B\begin{pmatrix} i_1 & i_2 & \dots & i_r \\ k_1 & k_2 & \dots & k_r \end{pmatrix}$$

se anulen simultáneamente, 1.26 implica que

$$B\begin{pmatrix} 1 & 2 & \dots & r \\ 1 & 2 & \dots & r \end{pmatrix} \neq 0, \tag{1.27}$$

y, por otra parte, que

$$B\begin{pmatrix} i_1 & i_2 & \dots & i_r \\ i_1 & i_2 & \dots & i_r \end{pmatrix} = \left[C\begin{pmatrix} i_1 & i_2 & \dots & i_r \\ 1 & 2 & \dots & r \end{pmatrix}\right]^2 B\begin{pmatrix} 1 & 2 & \dots & r \\ 1 & 2 & \dots & r \end{pmatrix}.$$

Se sigue entonces que la suma T_r de todos los menores principales de orden r no es vacía por 1.27 y que todos aquellos que no se anulan tienen el mismo signo, en consecuencia, $T_r \neq 0$ como queríamos. □

La proposición 1.11 afirma el hecho de que una matriz real y simétrica siempre es diagonalizable o, dicho de otra manera, que su sistema completo de vectores propios consiste de n vectores si n es el orden de la matriz.

Teorema 1.12. *Una matriz real y simétrica $A = (a_{ik})_1^n$ tiene n vectores propios reales que conforman un sistema ortonormal.*

Demostración. Por la proposición 1.10 sabemos que los vectores propios correspondientes a valores propios distintos son ortogonales dos a dos. Bastará probar entonces que para cada valor propio λ de multiplicidad m existe un sistema ortonormal de m vectores propios correspondientes a este valor propio.

Puesto que la matriz A es real y todos sus valores propios también lo son, podemos escoger un vector propio real u correspondiente a un valor propio λ, pues la ecuación $(\lambda I - A)u = 0$ tiene solución *sobre* los números reales. Denotemos por u^1 al vector normalizado

$$\frac{u}{\sqrt{(u,u)}}.$$

Ahora, observemos que si $k < m$ y $u^1, u^2, \ldots, u^k$ es un sistema ortonormal de vectores propios reales correspondientes al valor propio λ, y x es un vector propio real correspondiente al mismo λ y linealmente independiente de los vectores $u^1, u^2, \ldots, u^k$, entonces el vector

$$x' = x - \sum_{i=1}^{k} c_i u^i \quad \text{con} \quad c_i = (x, u^i) \quad (i = 1, 2, \ldots, k)$$

es real, diferente de cero y ortogonal a cada uno de los vectores $u^1, u^2, \ldots, u^k$. De esta manera, si hacemos

$$u^{k+1} := \frac{x'}{\sqrt{(x', x')}},$$

obtenemos un sistema ortonormal $u^1, u^2, \ldots, u^{k+1}$ de $k+1$ vectores propios correspondientes a λ.

De esta forma podemos construir un sistema ortonormal conformado por m vectores propios reales $u^1, u^2, \ldots, u^m$ correspondientes a un valor propio dado λ de multiplicidad m. Esto termina la demostración. □

1.5. Reducción de una forma cuadrática a una suma de cuadrados

Llamaremos *forma cuadrática* de n variables $x_1, x_2, \ldots, x_n$ a un polinomio homogéneo de segundo grado en estas variables. Siempre podremos expresar una forma cuadrática de $x_1, x_2, \ldots, x_n$ como

$$\sum_{i,k=1}^{n} a_{ik} x_i x_k,$$

donde $A = (a_{ik})_1^n$ es una matriz simétrica.

Si x es el vector cuyas coordenadas son $x_1, x_2, \ldots, x_n$, establecemos la siguiente notación abreviada para una ecuación cuadrática:

$$A(x, x) := \sum_{i,k=1}^{n} a_{ik} x_i x_k. \tag{1.28}$$

Consideraremos siempre formas reales, i.e., supondremos siempre que los coeficientes a_{ik} $(i, k = 1, 2, \ldots, n)$ son reales. En conjunción con 1.28 consideraremos también la correspondiente forma bilineal

$$A(x, y) := \sum_{i,k=1}^{n} a_{ik} x_i y_k. \tag{1.29}$$

Haciendo uso de la notación vectorial y del producto escalar, podemos observar que

$$A(x, y) = (Ax, y) = (x, Ay). \tag{1.30}$$

El determinante $|A| = |a_{ik}|_1^n$ es usualmente llamado el *discriminante* de la forma cuadrática $A(x, x)$. Si el discriminante $|A|$ es igual a cero, diremos que la forma cuadrática $A(x, x)$ es *singular*.

Observemos también que por efecto de la bilinealidad de la forma $A(x, y)$, para cualesquiera vectores $x^1, x^2, \ldots, x^l; y^1, y^2, \ldots, y^m$ y escalares $c_1, c_2, \ldots, c_l, d_1, d_2, \ldots, d_m$, se satisface la siguiente igualdad

$$A\left(\textstyle\sum_{i=1}^{l} c_i x^i, \sum_{j=1}^{m} d_j y^j\right) = \sum_{i=1}^{l} \sum_{j=1}^{m} c_i d_j A(x^i, y^j). \tag{1.31}$$

Proposición 1.13. *Supongamos que las variables de una forma cuadrática $A(x, x)$ están sujetas a una transformación lineal*

$$x_i = \sum_{k=1}^{n} g_{ik} \xi_k \qquad (i = 1, 2, \ldots, n).$$

Definamos los vectores $g^k := (g_{1k}, g_{2k}, \ldots, g_{nk}) \quad (k = 1, 2, \ldots, n)$, de modo tal que la transformación descrita antes se puede expresar en una ecuación vectorial simple

$$x = \sum_{k=1}^{n} \xi_k g^k.$$

Entonces cuando la forma $A(x,x)$ es transformada por $G = (g_{ik})_1^n$, el discriminante de la forma se multiplica por el cuadrado del discriminante de la transformación.

Demostración. Por 1.31 tenemos

$$A(x,x) = A\left(\textstyle\sum_{p=1}^n \xi_p g^p, \sum_{q=1}^n \xi_q g^q\right) = \sum_{p,q=1}^n \tilde{a}_{pq}\xi_p\xi_q,$$

donde

$$\tilde{a}_{pq} = A(g^p, g^q) = \sum_{i,k=1}^n a_{ik} g_{ip} g_{kq} \qquad (p,q = 1,2,\ldots,n).$$

Estas expresiones para los coeficientes $\tilde{a}_{pq}$ nos dicen que las matrices A, $\tilde{A} = (\tilde{a}_{ik})_1^n$ y G están relacionadas de la siguiente manera

$$\tilde{A} = G^t A G. \tag{1.32}$$

En consecuencia,

$$|\tilde{A}| = |A||G|^2$$

tal como queríamos ver. □

De 1.32 junto con el último resultado visto en la sección 1.1 también se sigue que si la transformación G es no singular, entonces el rango de la matriz de los coeficientes no cambia. El rango de una matriz simétrica A es el *rango* de la forma $A(x,x)$.

Una matriz G es *ortogonal* si se satisface la siguiente igualdad

$$GG^t = E. \tag{1.33}$$

Se sigue entonces que los renglones de una matriz ortogonal son, a su vez, ortogonales dos a dos y que la suma de los cuadrados de los elementos de cada renglón es igual a uno:

$$\sum_{k=1}^{n} g_{ik}g_{jk} = \delta_{ij} \qquad (i, j = 1, 2, \ldots, n).$$

Relaciones similares tenemos también para las columnas puesto que de 1.33 se deduce fácilmente que $G^t = G^{-1}$ y, por consiguiente,

$$G^tG = E. \tag{1.34}$$

Teorema 1.14. *Para cada forma cuadrática real*

$$A(x, x) = \sum_{i,k=1}^{n} a_{ik}x_ix_k, \quad (a_{ik} = a_{ki})$$

existe una transformación ortogonal

$$x_i = \sum_{k=1}^{n} u_{ik}\xi_k \qquad (i = 1, 2, \ldots, n)$$

que la reduce a una suma de cuadrados

$$A(x, x) = \sum_{i=1}^{n} \lambda_i {\xi_i}^2.$$

Aquí $U = (u_{ik})_1^n$ debe ser una matriz ortogonal fundamental de la matriz A y los λ_i $(i = 1, 2, \ldots, n)$ son los valores propios de la matriz A.

Demostración. Sea

$$u^1, u^2, \ldots, u^n$$

un sistema ortonormal completo de vectores propios de la matriz simétrica real A y $\lambda_1, \lambda_2, \ldots, \lambda_n$ los valores propios correspondientes:

$$Au^j = \lambda_j u^j \qquad (j = 1, 2, \ldots, n), \tag{1.35}$$

$$(u^j, u^k) = \delta_{jk} \qquad (j, k = 1, 2, \ldots, n). \tag{1.36}$$

Formamos la matriz fundamental

$$U = (u_{ij})_1^n$$

con las coordenadas de los vectores

$$u^j = (u_{1j}, u_{2j}, \ldots, u_{nj}) \qquad (j = 1, 2, \ldots, n).$$

$UU^t = E$ por 1.36, i.e., la matriz U es ortogonal. Por otro lado, de acuerdo con 1.17,

$$A = U(\lambda_i \delta_{ik})_1^n U^{-1},$$

es decir,

$$(\lambda_i \delta_{ik})_1^n = U^{-1} A U$$

y, como U es una matriz ortogonal, tenemos que

$$(\lambda_i \delta_{ik})_1^n = U^t A U. \tag{1.37}$$

Pero por 1.32 sabemos que $U^t A U$ es la matriz de la forma transformada si las variables $x_i \quad (i = 1, 2, \ldots, n)$ están sujetas a la transformación

$$x_i = \sum_{k=1}^{n} u_{ik} \xi_k \qquad (i = 1, 2, \ldots, n).$$

Se sigue que bajo tal transformación la forma cuadrática se reduce a

$$A(x, x) = \sum_{i=1}^{n} \lambda_i {\xi_k}^2.$$

La segunda parte del teorema se sigue porque cada argumento expuesto hasta ahora es reversible. □

Sea A una matriz real y simétrica.

Si no nos restringimos a considerar sólo transformaciones ortogonales de las variables, entonces una forma cuadrática $A(x, x)$ puede verse como

$$A(x, x) = \sum_{i=1}^{r} a_i {X_i}^2 \tag{1.38}$$

en una infinidad de maneras distintas, donde $a_i \neq 0$ $(i = 1, 2, \ldots, r)$ y los

$$X_i := \sum_{k=1}^{n} b_{ik} x_k \qquad (i = 1, 2, \ldots, r)$$

son formas lineales independientes de $x_1, x_2, \ldots, x_n$.

Teorema 1.15. (Ley de inercia de formas cuadráticas). *Cuando una forma cuadrática $A(x,x)$ es representada como una suma de cuadrados linealmente independientes*

$$A(x,x) = \sum_{i=1}^{r} a_i {X_i}^2, \tag{1.39}$$

el número de a_i positivos y el número de a_i negativos (y entonces también r) son independientes de la representación.

Demostración. Supongamos que en la representación 1.39,

$$a_1 > 0, a_2 > 0, \ldots, a_s > 0, a_{s+1} < 0, \ldots, a_r < 0.$$

Sea

$$A(x,x) = \sum_{i=1}^{\rho} b_i Y_i^2 \tag{1.40}$$

otra representación como suma de cuadrados linealmente independientes de la forma $A(x,x)$ en la cual

$$b_1 > 0, \ldots, b_\sigma > 0, b_{\sigma+1} < 0, \ldots, b_\rho < 0.$$

Supongamos lo contrario de lo que queremos demostrar; por ejemplo, que $\sigma < s$. Entonces $\nu := \sigma + (r - s) < r \leq n$, por consiguiente, el sistema de $\nu < n$ ecuaciones con n incógnitas $x_1, x_2, \ldots, x_n$:

$$Y_1 = 0, Y_2 = 0, \ldots, Y_\sigma = 0, X_{s+1} = 0, \ldots, X_r = 0, \tag{1.41}$$

tienes soluciones no triviales. Por otro lado, el sistema de ecuaciones

$$X_1 = 0, X_2 = 0, \ldots, X_r = 0 \tag{1.42}$$

consiste de r ecuaciones independientes, en consecuencia, no puede ser equivalente al sistema de $\nu < r$ ecuaciones 1.41. Así pues, existen valores x_i^0 tales que cuando $x_i = x_i^0$ $(i = 1, 2, \ldots, n)$, simultáneamente ocurre que se satisface el sistema 1.41 y que alguna de las ecuaciones X_j $(j = 1, 2, \ldots, s)$ es diferente de cero. Pero también se sigue que si $x^0 := (x_1^0, x_2^0, \ldots, x_n^0)$, entonces $A(x^0, x^0) > 0$ por 1.39 y $A(x^0, x^0) \leq 0$ por 1.40. Esta contradicción nos lleva a concluir que $\sigma = s$.

Simétricamente también tenemos que $\rho - \sigma = r - s$. □

A la cantidad de números positivos (negativos) a_i en la representación 1.39 la llamaremos *número de cuadrados positivos (negativos)* de la forma $A(x, x)$.

De acuerdo con el teorema 1.14 tenemos la siguiente

Regla: El número de valores propios positivos (negativos) de la matriz simétrica real A es igual al número de cuadrados positivos (negativos) de la forma $A(x, x)$

Observemos también que, como el número de valores propios distintos de cero de una matriz diagonalizable es igual al rango de la matriz, en la representación 1.39 el número de cuadrados r es igual al rango de la matriz A.

1.6. Las fórmulas de Sylvester y Jacobi

Para determinar el número de cuadrados positivos o el de negativos de la forma $A(x, x)$ se han empleado varios recursos. De estos, los más importantes se derivan de la fórmula de Sylvester descubierta en 1851. Para una mayor documentación acerca de este y otros resultados relacionados puede consultarse [11].

En esta fórmula emplearemos la notación siguiente

$$A_i(x) := \sum_{k=1}^{n} a_{ik} x_k \qquad (i = 1, 2, \ldots, n).$$

Teorema 1.16. (La fórmula de Sylvester). *Sea* $A = (a_{ik})_1^n$ *una matriz simétrica tal que*

$$A_p := A \begin{pmatrix} 1 & 2 & \ldots & p \\ 1 & 2 & \ldots & p \end{pmatrix} \neq 0$$

para algún p *entre* 1 *y* n*. Entonces se cumple la siguiente igualdad*

$$A(x,x) = -\frac{1}{A_p} \begin{vmatrix} a_{11} & \ldots & a_{1p} & A_1(x) \\ \ldots & \ldots & \ldots & \ldots \\ a_{p1} & \ldots & a_{pp} & A_p(x) \\ A_1(x) & \ldots & A_p(x) & 0 \end{vmatrix} + \frac{1}{A_p} \sum_{i,k=1}^{n} A \begin{pmatrix} 1 & 2 & \ldots & p & i \\ 1 & 2 & \ldots & p & k \end{pmatrix} x_i x_k. \tag{1.43}$$

Demostración. Evidentemente tenemos la siguiente igualdad

$$\begin{vmatrix} a_{11} & \ldots & a_{1p} & A_1(x) \\ \ldots & \ldots & \ldots & \ldots \\ a_{p1} & \ldots & a_{pp} & A_p(x) \\ A_1(x) & \ldots & A_p(x) & A(x,x) \end{vmatrix} = \begin{vmatrix} a_{11} & \ldots & a_{1p} & \sum_{k=1}^n a_{1k} x_k \\ \ldots & \ldots & \ldots & \ldots \\ a_{p1} & \ldots & a_{pp} & \sum_{k=1}^n a_{pk} x_k \\ \sum_{i=1}^n a_{i1} x_i & \ldots & \sum_{i=1}^n a_{ip} x_i & \sum_{i,k=1}^n a_{ik} x_i x_k \end{vmatrix}. \tag{1.44}$$

En este último determinante podemos ver el último renglón y la última columna como una suma de n renglones y una suma de n columnas, respectivamente, así que tal determinante en 1.44 puede verse como una suma de n^2 determinantes

$$\begin{vmatrix} a_{11} & \dots & a_{1p} & a_{1k}x_k \\ \dots & \dots & \dots & \dots \\ a_{p1} & \dots & a_{pp} & a_{pk}x_k \\ a_{i1}x_i & \dots & a_{ip}x_i & a_{ik}x_ix_k \end{vmatrix} \qquad (i, k = 1, 2, \dots, n)$$

que es igual a la suma de los términos

$$A\begin{pmatrix} 1 & 2 & \dots & p & i \\ 1 & 2 & \dots & p & k \end{pmatrix} x_ix_k \qquad (i, k = 1, 2, \dots, n).$$

En consecuencia, llegamos a la igualdad

$$\begin{vmatrix} a_{11} & \dots & a_{1p} & A_1(x) \\ \dots & \dots & \dots & \dots \\ a_{p1} & \dots & a_{pp} & A_p(x) \\ A_1(x) & \dots & A_p(x) & A(x,x) \end{vmatrix} = \sum_{i,k=1}^{n} A\begin{pmatrix} 1 & 2 & \dots & p & i \\ 1 & 2 & \dots & p & k \end{pmatrix} x_ix_k. \tag{1.45}$$

Pero también tenemos que

$$\begin{vmatrix} a_{11} & \dots & a_{1p} & A_1(x) \\ \dots & \dots & \dots & \dots \\ a_{p1} & \dots & a_{pp} & A_p(x) \\ A_1(x) & \dots & A_p(x) & A(x,x) \end{vmatrix} = \\ A(x,x)A_p + \begin{vmatrix} a_{11} & \dots & a_{1p} & A_1(x) \\ \dots & \dots & \dots & \dots \\ a_{p1} & \dots & a_{pp} & A_p(x) \\ A_1(x) & \dots & A_p(x) & 0 \end{vmatrix}, \tag{1.46}$$

entonces al combinar 1.45 y 1.46 teniendo en cuenta que $A_p \neq 0$, se obtiene la fórmula de Sylvester 1.43. □

Notemos que en el lado derecho de la fórmula de Sylvester la suma se extiende sólo sobre los valores de i y k que son mayores que p, pues si i o k es menor o igual que p, entonces $A\begin{pmatrix} 1 & 2 & \dots & p & i \\ 1 & 2 & \dots & p & k \end{pmatrix} = 0$.

Veamos, ahora, un método específico para reducir una forma cuadrática a una suma de cuadrados.

Convengamos en la siguiente notación:

$$\Delta_0 = 1,\ \Delta_1 = a_{11},\ \Delta_2 = \begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix}, \dots, \Delta_n = \begin{vmatrix} a_{11} & \dots & a_{1n} \\ \dots & \dots & \dots \\ a_{n1} & \dots & a_{nn} \end{vmatrix}.$$

Jacobi observó y posteriormente empleó, tal como puede verse en [8], el hecho siguiente: Si cada uno de los menores principales consecutivos Δ_k $(k = 1, 2, \dots, n)$ es distinto de cero, entonces

$$A(x, x) = \sum_{k=1}^{n} \frac{{X_k}^2}{\Delta_k \Delta_{k-1}}, \tag{1.47}$$

donde

$$X_1 := A_1(x),\ \ X_k := \begin{vmatrix} a_{11} & \dots & a_{1,k-1} & A_1(x) \\ \dots & \dots & \dots & \dots \\ \dots & \dots & \dots & \dots \\ a_{k1} & \dots & a_{k,k-1} & A_k(x) \end{vmatrix} \quad (k = 2, 3, \dots, n). \tag{1.48}$$

Probaremos un resultado un tanto más general:

Teorema 1.17. (La fórmula de Jacobi). *Si una forma $A(x, x)$ tiene rango r y*

$$\Delta_k \neq 0 \qquad (k = 1, 2, \ldots, r),$$

entonces

$$A(x,x) = \sum_{k=1}^{r} \frac{{X_k}^2}{\Delta_k \Delta_{k-1}}. \tag{1.49}$$

Demostración. Si hacemos $p = r$ en la fórmula de Sylvester 1.43, obtenemos la identidad conocida como identidad de Kronecker:

$$A(x,x) = -\frac{1}{\Delta_r} \begin{vmatrix} a_{11} & \ldots & a_{1r} & A_1(x) \\ \ldots & \ldots & \ldots & \ldots \\ a_{r1} & \ldots & a_{rr} & A_r(x) \\ A_1(x) & \ldots & A_r(x) & 0 \end{vmatrix}.$$

Definamos

$$P_k(x,x) = \begin{vmatrix} a_{11} & \ldots & a_{1k} & A_1(x) \\ \ldots & \ldots & \ldots & \ldots \\ a_{k1} & \ldots & a_{kk} & A_k(x) \\ A_1(x) & \ldots & A_k(x) & 0 \end{vmatrix} \qquad (k = 1, 2, \ldots, r).$$

Si $C := (c_{ij})_1^{k+1}$ es una matriz de orden $k+1$ y ponemos $b_{jl} := C \begin{pmatrix} 1 & 2 & \ldots & k-1 & j \\ 1 & 2 & \ldots & k-1 & l \end{pmatrix}$ $(j, l = k, k+1)$, tenemos que

$$B \begin{pmatrix} k & k+1 \\ k & k+1 \end{pmatrix} = C \begin{pmatrix} 1 & 2 & \ldots & k-1 \\ 1 & 2 & \ldots & k-1 \end{pmatrix} C \begin{pmatrix} 1 & 2 & \ldots & k & k+1 \\ 1 & 2 & \ldots & k & k+1 \end{pmatrix} \tag{1.50}$$

al aplicar el teorema 1.4 (Sylvester). Como

$$B\begin{pmatrix} k & k+1 \\ k & k+1 \end{pmatrix} = C\begin{pmatrix} 1 & 2 & \dots & k \\ 1 & 2 & \dots & k \end{pmatrix} C\begin{pmatrix} 1 & 2 & \dots & k-1 & k+1 \\ 1 & 2 & \dots & k-1 & k+1 \end{pmatrix}$$
$$-C\begin{pmatrix} 1 & 2 & \dots & k-1 & k \\ 1 & 2 & \dots & k-1 & k+1 \end{pmatrix} C\begin{pmatrix} 1 & 2 & \dots & k-1 & k+1 \\ 1 & 2 & \dots & k-1 & k \end{pmatrix},$$

de 1.50 se sigue que

$$\begin{aligned} &C\begin{pmatrix} 1 & 2 & \dots & k-1 \\ 1 & 2 & \dots & k-1 \end{pmatrix} C\begin{pmatrix} 1 & 2 & \dots & k & k+1 \\ 1 & 2 & \dots & k & k+1 \end{pmatrix} = \\ &\quad C\begin{pmatrix} 1 & 2 & \dots & k \\ 1 & 2 & \dots & k \end{pmatrix} C\begin{pmatrix} 1 & 2 & \dots & k-1 & k+1 \\ 1 & 2 & \dots & k-1 & k+1 \end{pmatrix} \\ &\qquad -C\begin{pmatrix} 1 & 2 & \dots & k-1 & k \\ 1 & 2 & \dots & k-1 & k+1 \end{pmatrix} C\begin{pmatrix} 1 & 2 & \dots & k-1 & k+1 \\ 1 & 2 & \dots & k-1 & k \end{pmatrix}. \end{aligned} \tag{1.51}$$

En particular, para un determinante simétrico $P_k(x,x)$, la identidad 1.51 nos proporciona la igualdad

$$\Delta_{k-1}P_k(x,x) = P_{k-1}(x,x)\Delta_k - X_k^2$$

o, equivalentemente,

$$-\frac{P_k(x,x)}{\Delta_k} = \frac{X_k^2}{\Delta_k\Delta_{k-1}} - \frac{P_{k-1}(x,x)}{\Delta_{k-1}} \qquad (k = 1,2,\dots,n; P_0 := 0).$$

Luego, por la identidad de Kronecker,

$$A(x,x) = -\frac{P_r(x,x)}{\Delta_r} = \sum_{k=1}^{r} \frac{X_k^2}{\Delta_k\Delta_{k-1}}.$$

$\square$

Si expandimos en 1.48 cada $A_i(x)$ $(i = 1,2,\dots,n)$ en términos de $x_1, x_2, \dots, x_n$, y aplicamos el teorema sobre la suma de determinantes, encontramos que

$$X_k = \sum_{j=k}^{n} \begin{vmatrix} a_{11} & \dots & a_{1,k-1} & a_{1j} \\ \dots & \dots & \dots & \dots \\ a_{k1} & \dots & a_{k,k-1} & a_{kj} \end{vmatrix} x_j = \Delta_k x_k + \dots \qquad (k = 1, 2, \dots, r).$$

Dado que $\Delta_k \neq 0$, observemos que *yendo hacia atrás* cada forma X_k tiene una nueva variable con respecto a X_{k+1}, en consecuencia, las formas X_k $(k = 1, 2, \dots, r)$ son linealmente independientes.

Así, la fórmula de Jacobi 1.49 nos da una representación de la forma $A(x, x)$ como una suma de cuadrados independientes. Adicionalmente, por el teorema 1.15 y la forma de los coeficientes en la representación 1.49 de la forma $A(x, x)$, tenemos el siguiente

Teorema 1.18. *Si una matriz simétrica $A = (a_{ik})_1^n$ tiene rango r y los menores principales consecutivos*

$$\Delta_0 = 1, \Delta_1, \Delta_2, \dots, \Delta_r \tag{1.52}$$

de esta matriz son todos distintos de cero, entonces el número de cuadrados positivos de la forma $A(x, x)$ es igual a r menos el número de cambios de signo en la serie 1.52, y el número de cuadrados negativos es igual al número de cambios de signo en la misma serie.

1.7. Formas cuadráticas positivas

A una forma cuadrática real

$$A(x, x) = \sum_{i,k=1}^{n} a_{ik} x_i x_k$$

la llamaremos *no-negativa* si para cada vector x se tiene que

$$A(x,x) \geq 0, \tag{1.53}$$

y *positiva* si para cada vector $x = (x_1, x_2, \ldots, x_n) \neq 0$, tenemos que

$$A(x,x) > 0. \tag{1.54}$$

La representación de la forma

$$A(x,x) = \sum_{i=1}^{n} \lambda_i \xi_i^2,$$

donde λ_i $(i = 1, 2, \ldots, n)$ son los valores propios de la matriz A (Teorema 1.14) nos lleva al siguiente

Teorema 1.19. *Para que una forma $A(x,x)$ sea no-negativa (respectivamente positiva) es necesario y suficiente que todos los valores propios $\lambda_1, \lambda_2, \ldots, \lambda_n$ de la matriz A sean no-negativos (resp. positivos).*

De modo tal que una forma cuadrática positiva $A(x,x)$ puede representarse siempre como una suma de n cuadrados positivos independientes y una forma no-negativa puede representarse siempre como una suma de r cuadrados positivos independientes, donde r es el rango de dicha forma.

Teorema 1.20. (Jacobi) *Para que una forma cuadrática $A(x,x)$ sea positiva, una condición necesaria y suficiente es que todos los menores principales consecutivos de su matriz sean positivos, es decir, tener que*

$$a_{11} > 0, \begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix} > 0, \ldots, \begin{vmatrix} a_{11} & \ldots & a_{1n} \\ \ldots & \ldots & \ldots \\ a_{n1} & \ldots & a_{nn} \end{vmatrix} > 0.$$

Demostración. Primero establezcamos la siguiente notación:

$$\Delta_0 := 1, \ldots, \Delta_k := \begin{vmatrix} a_{11} & \ldots & a_{1k} \\ \ldots & \ldots & \ldots \\ a_{k1} & \ldots & a_{kk} \end{vmatrix} \qquad (k = 1, 2, \ldots, n).$$

Si la forma cuadrática $A(x, x)$ es positiva, entonces $\lambda_i > 0$ para cada $i = 1, 2, \ldots, n$, y así tenemos que el discriminante tiene la forma $\Delta_n = \lambda_1\lambda_2 \ldots \lambda_n > 0$ (Proposición 1.7). Por otro lado, si por lo menos uno de los $x_1, x_2, \ldots, x_p$ es distinto de cero y $x_{p+1} = \cdots = x_n = 0$, de 1.54 se sigue que

$$\sum_{i,k=1}^{p} a_{ik}x_ix_k > 0 \qquad (p = 1, 2, \ldots, n-1),$$

i.e., las formas *recortadas* $\sum_{i,k=1}^{p} a_{ik}x_ix_k$ $(p = 1, 2, \ldots, n-1)$ también son positivas. En consecuencia, sus respectivos discriminantes satisfacen $\Delta_p > 0$ $(p = 1, 2, \ldots, n-1)$.

Para la suficiencia, supongamos que

$$\Delta_1 > 0, \Delta_2 > 0, \ldots, \Delta_n > 0. \tag{1.55}$$

De la fórmula de Jacobi 1.47 obtenemos:

$$A(x.x) = \sum_{i=1}^{n} \frac{X_i^2}{\Delta_i\Delta_{i-1}} > 0,$$

siempre que no todos los X_i $(i = 1, 2, \ldots, n)$ sean iguales a cero. Pero de la independencia lineal de las formas X_i $(i = 1, 2, \ldots, n)$, se sigue que éstas se anulan simultáneamente sólo cuando $x_i = 0$ con $i = 1, 2, \ldots, n$. □

Corolario 1.21. *Si los menores principales consecutivos 1.55 de una matriz simétrica $A = (a_{ik})_1^n$ son positivos, entonces todos sus menores principales son positivos.*

Demostración. Si se tiene 1.55, entonces la forma $A(x, x)$ es positiva. Por otro lado, dado un menor principal cualquiera de A, sea B la matriz que se obtiene al hacer el reordenamiento de renglones y columnas de A necesario para poder incluir tal menor principal entre los menores principales consecutivos. De modo tal que si x es un vector distinto de cero y x' es el vector que se obtiene al reordenar las coordenadas de x en el mismo orden en que fueron intercambiadas las columnas (o los renglones) de A para obtener B; tenemos entonces que

$$B(x', x') = A(x, x) > 0.$$

Por consiguiente, la forma $B(x, x)$ es positiva y los menores principales consecutivos de B son positivos. En consecuencia, cualquier menor principal de A es positivo. □

Teorema 1.22. (Criterio de no-negatividad de una forma cuadrática). *Con respecto a tener una forma cuadrática no-negativa $A(x, x)$ es necesario y suficiente que todos los menores principales de la matriz A sean no-negativos.*

Demostración. Para cada $\epsilon > 0$ definamos la matriz

$$A_\epsilon = A + \epsilon E = (a_{ik} + \epsilon\delta_{ik})_1^n.$$

Entonces

$$A_\epsilon(x, x) = \sum_{i,k=1}^{n} a_{ik} x_i x_k + \epsilon \sum_{i=1}^{n} {x_1}^2.$$

En consecuencia, si la forma $A(x, x)$ es no-negativa, entonces la forma $A_\epsilon(x, x)$ es positiva y de aquí se sigue que todos sus menores principales son positivos. Como $A_\epsilon \to A$ cuando $\epsilon \to 0$, tenemos que los menores principales de la matriz A son no-negativos.

En la dirección opuesta, si ningún menor principal de la matriz A es negativo, entonces para $\epsilon > 0$ tenemos que

$$|a_{ik} + \epsilon\delta_{ik}|_1^p = \epsilon^p + \sum a_{ii}\epsilon^{p-1} + \sum \begin{vmatrix} a_{ii} & a_{ik} \\ a_{ki} & a_{kk} \end{vmatrix} \epsilon^{p-2} + \ldots > 0 \qquad (p = 1, 2, \ldots, n),$$

i.e., los menores principales consecutivos de la matriz A_ϵ son positivos. Entonces, por el corolario 1.21, la forma $A_\epsilon(x, x)$ es positiva, luego, la forma $A(x, x) = \lim_{\epsilon \to 0} A_\epsilon(x, x)$ es no-negativa. □

Observemos que del hecho de que los menores principales consecutivos de una matriz simétrica A sean no-negativos no se sigue que la forma $A(x, x)$ sea no-negativa y, en consecuencia, que los demás menores principales sean no-negativos.

Consideremos, por ejemplo, una forma arbitraria $\sum_{i,k=2}^n a_{ik}x_i x_k$ que no sea no-negativa y veámosla como una forma cuadrática de n variables $x_1, x_2, \ldots, x_n$ al tomar $a_{i1} = a_{1i} = 0$ $(i = 1, 2, \ldots, n)$. Entonces todos los menores consecutivos

$$A \begin{pmatrix} 1 & 2 & \ldots & p \\ 1 & 2 & \ldots & p \end{pmatrix} \qquad (p = 1, 2, \ldots, n)$$

son iguales a cero (son no-negativos), sin embargo, la forma cuadrática $\sum_{i,k=1}^n a_{ik}x_i x_k$ no es no-negativa.

1.8. La desigualdad de Hadamard

La estimación más simple y, al mismo tiempo, la más precisa que puede hacerse de un determinante la propuso Hadamard por vez primera en 1893 (véase [9], pág. 383).

Teorema 1.23. (Desigualdad de Hadamard). *Si la forma $A(x, x)$ es positiva, entonces se satisface la desigualdad siguiente para su discriminante*

$$\begin{vmatrix} a_{11} & \ldots & a_{1n} \\ \ldots & \ldots & \ldots \\ a_{n1} & \ldots & a_{nn} \end{vmatrix} \leq a_{11}a_{22}\cdots a_{nn}. \tag{1.56}$$

Aquí la igualdad se satisface si, y sólo si $a_{ik} = 0$ para $i \neq k$ $(i, k = 1, 2, \ldots, n)$.

Demostración. Para probar la desigualdad 1.56 será suficiente mostrar que

$$A\begin{pmatrix} 1 & 2 & \ldots & n \\ 1 & 2 & \ldots & n \end{pmatrix} \leq A\begin{pmatrix} 1 & 2 & \ldots & n-1 \\ 1 & 2 & \ldots & n-1 \end{pmatrix} a_{nn} \tag{1.57}$$

con la igualdad si, y sólo si $a_{1n} = a_{2n} = \cdots = a_{n-1,n} = 0$. Pues si la forma $A(x, x)$ es positiva, entonces las formas *recortadas*

$$\sum_{i,k=1}^{p} a_{ik} x_i x_k \qquad (p = 1, 2, \ldots, n)$$

son todas positivas. Luego, al aplicar 1.57 a cada una de estas formas obtenemos:

$$\begin{gathered} A\begin{pmatrix} 1 & 2 & \ldots & n \\ 1 & 2 & \ldots & n \end{pmatrix} \leq A\begin{pmatrix} 1 & 2 & \ldots & n-1 \\ 1 & 2 & \ldots & n-1 \end{pmatrix} a_{nn} \leq \\ A\begin{pmatrix} 1 & 2 & \ldots & n-2 \\ 1 & 2 & \ldots & n-2 \end{pmatrix} a_{n-1,n-1} a_{nn} \leq \cdots \leq a_{11} a_{22} \cdots a_{nn}. \end{gathered} \tag{1.58}$$

Como la desigualdad se satisface en 1.56 si, y sólo si ocurre en cada relación en 1.58, llegamos a la segunda parte del teorema.

Para probar 1.57 consideremos todos aquellos $\delta > 0$ para los cuales la forma

$$A_\delta(x, x) := A(x, x) - \delta x_n^2$$

es positiva. Según el criterio de positividad para una forma cuadrática, la forma $A_\delta(x, x)$ es positiva si los menores principales consecutivos de la matriz A_δ son positivos. Pero todos estos menores, con excepción del último, coinciden con los menores correspondientes de la matriz A, pues es claro que los elementos de la matriz A coinciden con los de la matriz A_δ excepto el elemento a_{nn}. De tal manera que todos estos menores son positivos por el mismo criterio aplicado, ahora, a la forma $A(x, x)$ que es positiva por hipótesis. Luego, los δ considerados simplemente deben satisfacer la condición

$$\begin{vmatrix} a_{11} & \dots & a_{1n} \\ \dots & \dots & \dots \\ a_{n1} & \dots & a_{nn} - \delta \end{vmatrix} = \Delta_n - \delta\Delta_{n-1} > 0.$$

Así pues, el número $\frac{\Delta_n}{\Delta_{n-1}}$ es la cota superior mínima para los valores de δ en consideración. Por otro lado, si la forma $A_\delta(x, x)$ es positiva, entonces todos los menores principales de A_δ son positivos (corolario 1.21) y, en particular, el elemento $a_{nn} - \delta > 0$. Luego, el número a_{nn} es una cota superior para los números δ, entonces

$$\frac{\Delta_n}{\Delta_{n-1}} \leq a_{nn} \quad \text{o} \quad \Delta_n \leq \Delta_{n-1} a_{nn}, \tag{1.59}$$

de donde se sigue la desigualdad 1.57.

Veamos, ahora, cuándo es posible tener la igualdad en 1.57. Si se da la igualdad en 1.57, entonces a_{nn} es la cota superior mínima de los valores de δ para los cuales la forma $A_\delta(x, x)$ es positiva. En consecuencia, la forma

$$B(x, x) := A(x, x) - a_{nn} x_n^2$$

que se obtiene al sustituir δ por a_{nn} en la forma $A_\delta(x, x)$ es no-negativa. Luego, los menores principales de la matriz B,

$$\begin{vmatrix} b_{kk} & b_{kn} \\ b_{nk} & b_{nn} \end{vmatrix} = \begin{vmatrix} a_{kk} & a_{kn} \\ a_{nk} & 0 \end{vmatrix} = -a_{kn}^2 \qquad (k = 1, 2, \dots, n-1)$$

son no-negativos, lo cual implica que

$$a_{kn} = 0 \quad \text{para} \quad k = 1, 2, \ldots, n-1.$$

Esto termina la demostración del teorema. $\square$

Teorema 1.24. (Desigualdad de Hadamard generalizada). *Si una forma* $A(x,x) = \sum_{i,k=1}^{n} a_{ik} x_i x_k$ *es positiva, entonces para cada* $p < n$ *se satisface la desigualdad:*

$$A\begin{pmatrix} 1 & 2 & \ldots & n \\ 1 & 2 & \ldots & n \end{pmatrix} \leq A\begin{pmatrix} 1 & 2 & \ldots & p \\ 1 & 2 & \ldots & p \end{pmatrix} A\begin{pmatrix} p+1 & \ldots & n \\ p+1 & \ldots & n \end{pmatrix}, \tag{1.60}$$

donde la igualdad se da si, y sólo si

$$a_{ik} = a_{ki} = 0 \quad \text{para} \quad i = 1, 2, \ldots, p; \; k = p+1, \ldots, n. \tag{1.61}$$

Demostración. Haremos la prueba del teorema por inducción sobre n. La validez del teorema para $n = 2$ se tiene porque en este caso $p = 1$ y

$$A\begin{pmatrix} 1 & 2 \\ 1 & 2 \end{pmatrix} = a_{11}a_{22} - a_{12}^2 \leq a_{11}a_{22}$$

con la igualdad si, y sólo si $a_{12} = 0$.

Si $n > 2$, entonces alguno de los números p o $n - p$ es mayor que 1. Sin pérdida de generalidad, podemos suponer que $p > 1$ siempre, porque en caso contrario, al invertir el orden de las filas y las columnas de A simultáneamente y aplicar el teorema para esta nueva matriz y para $n - p > 1$, se tiene el mismo resultado que para A y p.

Escribamos

$$b_{ik} := A\begin{pmatrix} 1 & 2 & \ldots & p-1 & i \\ 1 & 2 & \ldots & p-1 & k \end{pmatrix} \qquad (i, k = p, p+1, \ldots, n).$$

Por la identidad de Sylvester (Teorema 1.4) para $p \leq q \leq n$, tenemos que

$$B\begin{pmatrix} p & p+1 & \dots & q \\ p & p+1 & \dots & q \end{pmatrix} = \left[A\begin{pmatrix} 1 & 2 & \dots & p-1 \\ 1 & 2 & \dots & p-1 \end{pmatrix}\right]^{q-p} A\begin{pmatrix} 1 & 2 & \dots & q \\ 1 & 2 & \dots & q \end{pmatrix} > 0,$$

de modo tal que la forma

$$B(x,x) = \sum_{i,k=p}^{n} b_{ik}x_i x_k$$

es positiva. Luego, por la hipótesis de inducción,

$$\begin{aligned} B\begin{pmatrix} p & p+1 & \dots & n \\ p & p+1 & \dots & n \end{pmatrix} \leq b_{pp} B\begin{pmatrix} p+1 & \dots & n \\ p+1 & \dots & n \end{pmatrix} = \\ A\begin{pmatrix} 1 & 2 & \dots & p \\ 1 & 2 & \dots & p \end{pmatrix} \left[A\begin{pmatrix} 1 & 2 & \dots & p-1 \\ 1 & 2 & \dots & p-1 \end{pmatrix}\right]^{n-p-1} \\ \times A\begin{pmatrix} 1 & 2 & \dots & p-1 & p+1 & \dots & n \\ 1 & 2 & \dots & p-1 & p+1 & \dots & n \end{pmatrix}. \end{aligned} \tag{1.62}$$

Por otra parte, si hacemos $x_p = 0$ en la forma $A(x,x)$, obtenemos otra forma positiva de $n-1$ variables $x_1, x_2, \dots, x_{p-1}, x_{p+1}, \dots, x_n$ con discriminante

$$A\begin{pmatrix} 1 & 2 & \dots & p-1 & p+1 & \dots & n \\ 1 & 2 & \dots & p-1 & p+1 & \dots & n \end{pmatrix}.$$

De este modo, podemos aplicar la desigualdad 1.60 a este determinante (por la hipótesis de inducción):

$$\begin{aligned} A\begin{pmatrix} 1 & 2 & \dots & p-1 & p+1 & \dots & n \\ 1 & 2 & \dots & p-1 & p+1 & \dots & n \end{pmatrix} \\ \leq A\begin{pmatrix} 1 & 2 & \dots & p-1 \\ 1 & 2 & \dots & p-1 \end{pmatrix} A\begin{pmatrix} p+1 & \dots & n \\ p+1 & \dots & n \end{pmatrix}. \end{aligned} \tag{1.63}$$

En consecuencia,

$$B\begin{pmatrix} p & p+1 & \dots & n \\ p & p+1 & \dots & n \end{pmatrix}$$
$$\leq \left[A\begin{pmatrix} 1 & 2 & \dots & p-1 \\ 1 & 2 & \dots & p-1 \end{pmatrix}\right]^{n-p} A\begin{pmatrix} 1 & 2 & \dots & p \\ 1 & 2 & \dots & p \end{pmatrix} A\begin{pmatrix} p+1 & \dots & n \\ p+1 & \dots & n \end{pmatrix},$$

y, como la identidad de Sylvester nos proporciona la igualdad

$$B\begin{pmatrix} p & p+1 & \dots & n \\ p & p+1 & \dots & n \end{pmatrix} = \left[A\begin{pmatrix} 1 & 2 & \dots & p-1 \\ 1 & 2 & \dots & p-1 \end{pmatrix}\right]^{n-p} A\begin{pmatrix} 1 & 2 & \dots & n \\ 1 & 2 & \dots & n \end{pmatrix},$$

finalmente llegamos a la desigualdad deseada

$$A\begin{pmatrix} 1 & 2 & \dots & n \\ 1 & 2 & \dots & n \end{pmatrix} \leq A\begin{pmatrix} 1 & 2 & \dots & p \\ 1 & 2 & \dots & p \end{pmatrix} A\begin{pmatrix} p+1 & \dots & n \\ p+1 & \dots & n \end{pmatrix}.$$

La igualdad es posible aquí sólo cuando ocurre en 1.62 y en 1.63. Si en 1.63 se tiene la igualdad, por la hipótesis de inducción tenemos que

$$a_{ik} = a_{ki} = 0 \quad \text{para} \quad k = p+1, \dots, n. \tag{1.64}$$

Por otro lado, si la igualdad ocurre en 1.62, nuevamente por la hipótesis de inducción tenemos que

$$b_{pk} = b_{kp} = A\begin{pmatrix} 1 & 2 & \dots & p-1 & k \\ 1 & 2 & \dots & p-1 & p \end{pmatrix} = 0 \quad \text{para} \quad k = p+1, p+2, \dots, n.$$

Pero al expandir el determinante b_{kp} con respecto al último renglón obtenemos:

$$A\begin{pmatrix} 1 & 2 & \dots & p-1 & k \\ 1 & 2 & \dots & p-1 & p \end{pmatrix} = a_{kp}A\begin{pmatrix} 1 & 2 & \dots & p-1 \\ 1 & 2 & \dots & p-1 \end{pmatrix} \qquad (k = p+1, \dots, n),$$

de donde se sigue que ($A(x, x)$ es positiva)

$$a_{kp} = a_{pk} = 0 \quad \text{para} \quad k = p+1, \dots, n. \tag{1.65}$$

Las afirmaciones 1.64 y 1.65 juntas nos dan 1.61. Aquí termina la demostración. □

No obstante, la usualmente llamada desigualdad de Hadamard no es la desigualdad 1.56 sino la desigualdad del siguiente teorema:

Teorema 1.25. *Para cada matriz real y no singular $C = (c_{ik})_1^n$ se satisface la siguiente desigualdad:*

$$\begin{vmatrix} c_{11} & \dots & c_{1n} \\ \dots & \dots & \dots \\ c_{n1} & \dots & c_{nn} \end{vmatrix}^2 \leq \sum_{k=1}^{n} c_{1k}^2 \cdot \sum_{k=1}^{n} c_{2k}^2 \cdots \sum_{k=1}^{n} c_{nk}^2, \tag{1.66}$$

aquí la igualdad ocurre si, y sólo si las filas de la matriz C son ortogonales dos a dos, i.e.,

$$\sum_{k=1}^{n} c_{ik}c_{jk} = 0 \quad \text{para} \quad i \neq j \qquad (i, j = 1, 2, \dots, n). \tag{1.67}$$

Demostración. Veremos que la desigualdad 1.56 implica la desigualdad 1.66 y viceversa. Si $C = (c_{ik})_1^n$ es una matriz real no singular, pongamos

$$a_{ij} := \sum_{k=1}^{n} c_{ik}c_{jk} \qquad (i, j = 1, 2, \dots, n). \tag{1.68}$$

Tal definición de los números a_{ij} nos proporciona una forma

$$\sum_{i,j=1}^{n} a_{ij}x_i x_j = \sum_{k=1}^{n}(c_{1k}x_1 + c_{2k}x_2 + \cdots + c_{nk}x_n)^2 \tag{1.69}$$

positiva, luego, la desigualdad 1.56 se cumple para la matriz $A = (a_{ij})_1^n$. Por 1.68 tenemos:

$$A\begin{pmatrix} 1 & 2 & \dots & n \\ 1 & 2 & \dots & n \end{pmatrix} = \left[C\begin{pmatrix} 1 & 2 & \dots & n \\ 1 & 2 & \dots & n \end{pmatrix}\right]^2, \tag{1.70}$$

de donde 1.56 coincide con 1.66, y, por 1.68, la condición $a_{ij} = 0$ para $i \neq j$ $(i, j = 1, 2, \dots, n)$ coincide con 1.67.

Por otra parte, puesto que cualquier forma positiva $A(x, x)$ puede representarse como una suma de n cuadrados linealmente independientes con coeficientes positivos (i.e., en la forma 1.69 con $|C| \neq 0$), obtenemos 1.56 de 1.66. $\square$

1.9. Matrices asociadas y el teorema de Kronecker

Sea $A = (a_{ik})_1^n$ una matriz de orden n. Comenzaremos por definir la matriz asociada p-ésima $\mathfrak{A}_p$ de la matriz A.

Consideremos todas las combinaciones (subconjuntos con p elementos)

$$(i_1, i_2, \dots, i_p)$$

de n índices $1, 2, \ldots, n$ y escribamos cada combinación en orden creciente: $i_1 < i_2 < \ldots < i_p$. De esta manera podemos ordenar todas las combinaciones al convenir en que una combinación $(i_1, i_2, \ldots, i_p)$ antecede a otra combinación $(k_1, k_2, \ldots, k_p)$ si, y sólo si el primer término distinto de cero en la sucesión

$$k_1 - i_1, k_2 - i_2, \ldots, k_p - i_p$$

es positivo. En consecuencia, podemos asignar un número bien definido s localizado entre 1 y $N := \binom{n}{p}$ a cada combinación. Por ejemplo, si $n = 5$ y $p = 3$, obtenemos la siguiente sucesión de combinaciones:

$$(123), (124), (125), (134), (135), (145), (234), (235), (245), (345).$$

De modo tal que la combinación (125) tiene el número 3 y la combinación (235) el número 8.

Para los menores de orden p de la matriz A introducimos la siguiente notación:

$$\mathfrak{a}_{st} := A\begin{pmatrix} i_1 & i_2 & \ldots & i_p \\ k_1 & k_2 & \ldots & k_p \end{pmatrix} \qquad (i_1 < i_2 < \ldots < i_p;\ k_1 < k_2 < \ldots < k_p),$$

donde s y t son los números de las combinaciones $(i_1, i_2, \ldots, i_p)$ y $(k_1, k_2, \ldots, k_p)$, respectivamente. A la matriz

$$\mathfrak{A}_p = (\mathfrak{a}_{st})_1^N$$

la llamaremos *matriz asociada p-ésima* de A.

Proposición 1.26. *La matriz asociada a un producto es igual al producto de las matrices asociadas de los factores.*

Demostración. Sea $C = AB$, como ya sabemos de la primera sección, se cumple que

$$C\begin{pmatrix} i_1 & \dots & i_p \\ k_1 & \dots & k_p \end{pmatrix} = \sum A \begin{pmatrix} i_1 & \dots & i_p \\ \alpha_1 & \dots & \alpha_p \end{pmatrix} B \begin{pmatrix} \alpha_1 & \dots & \alpha_p \\ k_1 & \dots & k_p \end{pmatrix}. \tag{1.71}$$

Consideremos las matrices asociadas $\mathfrak{A}_p, \mathfrak{B}_p$ y $\mathfrak{C}_p$ de las matrices A, B y C, respectivamente. Entonces la relación 1.71 puede escribirse de la forma

$$\mathfrak{c}_{st} = \sum_{r=1}^{N} \mathfrak{a}_{sr}\mathfrak{b}_{rt},$$

donde s, t y r son los números de las combinaciones

$$(i_1, \dots, i_p), \quad (k_1, \dots, k_p) \quad y \quad (\alpha_1, \dots, \alpha_p),$$

respectivamente. Así pues, $\mathfrak{C}_p = \mathfrak{A}_p\mathfrak{B}_p$ por el teorema de Binet-Cauchy. □

Corolario 1.27. *La matriz asociada a la matriz inversa es igual a la matriz inversa de la matriz asociada.*

Demostración. Al aplicar la propiedad anterior al producto $AB = E$, obtenemos que $\mathfrak{A}_p\mathfrak{B}_p = \mathfrak{E}_p$. Como la matriz p-ésima asociada a E, $\mathfrak{E}_p$, es igual a la matriz $(\delta_{st})_1^N$ evidentemente, se sigue el resultado deseado. □

Teorema 1.28. (Kronecker)*. Sean $\lambda_1, \lambda_2, \dots, \lambda_n$ el sistema completo de valores propios de la matriz A. Entonces el sistema completo de valores propios de la matriz asociada $\mathfrak{A}_p$ consiste de todos los posibles productos con p factores de los números $\lambda_1, \lambda_2, \dots, \lambda_n$.*

Demostración. Escribamos la matriz A mediante la representación

$$A = PTP^{-1}, \tag{1.72}$$

donde

$$T = \begin{vmatrix} \lambda_1 & * & \dots & * \\ 0 & \lambda_2 & \dots & * \\ \dots & \dots & \dots & \dots \\ 0 & 0 & \dots & \lambda_n \end{vmatrix}.$$

Tomando en cuenta la proposición 1.26 y el corolario 1.27, de 1.72 obtenemos

$$\mathfrak{A}_p = \mathfrak{P}_p \mathfrak{T}_p \mathfrak{P}_p^{-1}.$$

Ahora, si

$$\mathfrak{t}_{st} = T \begin{pmatrix} i_1 & \dots & i_p \\ k_1 & \dots & k_p \end{pmatrix}$$

es un elemento de la matriz $\mathfrak{T}_p$ con $s > t$, entonces para algún $r < p$ se tiene que

$$i_1 = k_1, \dots, i_r = k_r, \quad i_{r+1} > k_{r+1}.$$

En este caso

$$T \begin{pmatrix} i_1 & i_2 & \dots & i_p \\ k_1 & k_2 & \dots & k_p \end{pmatrix} = \begin{vmatrix} \lambda_{i_1} & * & \dots & * & * & \dots & * \\ 0 & \lambda_{i_2} & \dots & * & * & \dots & * \\ \dots & \dots & \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & \dots & \lambda_{i_r} & * & \dots & * \\ 0 & 0 & \dots & 0 & 0 & \dots & * \\ \dots & \dots & \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & \dots & 0 & 0 & \dots & 0 \end{vmatrix} = 0.$$

Luego, la matriz $\mathfrak{T}_p$ es triangular. Como cada elemento $\mathfrak{t}_{ss}$ de la diagonal principal de esta matriz es igual a

$$T\begin{pmatrix} i_1 & \dots & i_p \\ i_1 & \dots & i_p \end{pmatrix} = \lambda_{i_1}\lambda_{i_2}\dots\lambda_{i_p}$$

y los elementos de la diagonal principal de la matriz triangular $\mathfrak{T}_p$ nos dan un sistema completo de valores propios de la matriz $\mathfrak{A}_p$, el teorema de Kronecker está probado. $\square$

Proposición 1.29. *La matriz $\mathfrak{U}_p$ asociada a una matriz fundamental para una matriz diagonalizable A (véase la pág. 33) es una matriz fundamental para la matriz asociada $\mathfrak{A}_p$.*

Demostración. Cualquier matriz A diagonalizable admite una representación

$$A = U(\lambda_i \delta_{ik})_1^n U^{-1},$$

donde U es una matriz fundamental para A. Al pasar en esta relación a matrices asociadas, obtenemos

$$\mathfrak{A}_p = \mathfrak{U}_p(\Lambda_s \delta_{st})_1^N \mathfrak{U}_p^{-1} ,$$

donde

$$\Lambda_s = \lambda_{i_1} \cdots \lambda_{i_p}$$

y s es el número de la combinación $(i_1, i_2, \dots, i_p)$. $\square$

De esta proposición es inmediato el siguiente resultado:

Corolario 1.30. *Dada una combinación $(k_1, k_2, \dots, k_p)$, los menores*

$$U\begin{pmatrix} \alpha_1 & \alpha_2 & \dots & \alpha_p \\ k_1 & k_2 & \dots & k_p \end{pmatrix}$$

que se obtienen al hacer variar la combinación $(\alpha_1, \alpha_2, \ldots, \alpha_p)$ *nos proporcionan todas las coordenadas del vector propio de la matriz* $\mathfrak{A}_p$ *correspondiente al valor propio* $\lambda_{k_1} \ldots \lambda_{k_p}$.

Capítulo 2

Matrices Oscilatorias

De aquí en adelante, consideraremos sólo matrices con entradas reales.

2.1. Matrices de Jacobi

Antes de hacer un estudio general de las matrices oscilatorias, veremos algunos resultados en relación a las matrices de Jacobi normales. Independientemente de la teoría general uno puede deducir para este tipo de matrices, mediante resultados elementales, cada una de las propiedades básicas que caracterizan a las matrices oscilatorias.

A una matriz $J = (a_{ik})_1^n$ tal que $a_{ik} = 0$ si $|i-k| > 1$ la llamaremos *matriz de Jacobi*. Si convenimos en la notación

$$a_i = a_{ii} \qquad (i = 1, 2, \ldots, n),$$
$$b_i = -a_{i,i+1}\ , \quad c_i = -a_{i+1,i} \quad (i = 1, 2, \ldots, n-1),$$

podemos escribir la matriz J en la forma

$$J = \begin{pmatrix} a_1 & -b_1 & 0 & 0 & \dots & 0 \\ -c_1 & a_2 & -b_2 & 0 & \dots & 0 \\ 0 & -c_2 & a_3 & -b_3 & \dots & 0 \\ \dots & \dots & \dots & \dots & \dots & \dots \\ \dots & \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & \dots & \dots & a_{n-1} & -b_{n-1} \\ 0 & 0 & \dots & \dots & -c_{n-1} & a_n \end{pmatrix} \tag{2.1}$$

Establezcamos la notación siguiente:

$$D_0(x) \equiv 1, \; D_k(x) := |a_{ij} - \delta_{ij} x|_1^k \qquad (k = 1, 2, \dots, n).$$

Luego, $D_n(x)$ es el polinomio característico de la matriz J.

La siguiente fórmula recursiva para los polinomios $D_k(x)$ se obtiene fácilmente al hacer una expansión por menores del determinante $|a_{ij} - \delta_{ij} x|_1^k$:

$$D_k(x) = (a_k - x) D_{k-1}(x) - b_{k-1} c_{k-1} D_{k-2}(x) \qquad (k = 2, 3, \dots, n) \tag{2.2}$$

Dado que $D_0 \equiv 1$ y $D_1(x) = a_1 - x$, haciendo uso de esta fórmula se puede calcular de forma sucesiva cada polinomio $D_k(x)$ $(k = 2, 3, \dots, n)$. Esto implica que en la expresión $D_k(x)$ $(k = 0, 1, \dots, n)$ los números b_k y c_k $(k = 1, 2, \dots, n-1)$ aparecen sólo en forma de productos $b_k c_k$ $(k = 1, 2, \dots, n-1)$.

Confinaremos nuestros argumentos sólo al caso en el que

$$b_k c_k > 0 \qquad (k = 1, 2, \dots, n-1).$$

En este caso, por la fórmula recursiva 2.2, la sucesión de polinomios en $\mathbb{R}[x]$,

$$D_m, D_{m-1}, \dots, D_0 \qquad (m \leq n),$$

tiene las dos primeras propiedades de una *sucesión de Sturm*, a saber,

(i) $D_0(x)$ tiene signo constante ($D_0 \equiv 1$).

(ii) Si $D_k(x)$ con $1 \leq k < m$ se anula en $x = \lambda$, entonces $D_{k-1}(\lambda)$ y $D_{k+1}(\lambda)$ son distintos de cero y tienen signos opuestos.

Teorema 2.1. (Regla de Sturm). *Si las condiciones* (i) *y* (ii) *se satisfacen, entonces la diferencia entre el número de cambios de signo en la sucesión*

$$D_m(x), D_{m-1}(x), \dots, D_0(x) \tag{2.3}$$

cuando $x = \alpha$ y el número de cambios de signo cuando $x = \beta$ ($\alpha < \beta$, $D_m(\alpha) \neq 0$, $D_m(\beta) \neq 0$) es igual a la diferencia entre el número de raíces del polinomio $D_m(x)$ en (α, β) con las que el producto $D_m(x)D_{m-1}(x)$ cambia su signo de $+$ a $-$ (al ir la variable x de α a β), y el número de raíces de $D_m(x)$ en el mismo intervalo con las que el producto $D_m(x)D_{m-1}(x)$ cambia su signo de $-$ a $+$.

Demostración. Como x varía entre α y β, el número de cambios de signo en la sucesión 2.3 puede modificarse sólo cuando x *pasa por* una raíz de los polinomios $D_k(x)$ ($k = 1, \dots, m$). Sin embargo, cuando algún polinomio intermedio $D_k(x)$ ($1 \leq k < m$) se anula en algún valor x_0, los polinomios adyacentes toman valores con signos opuestos en $x = x_0$, de modo tal que para valores en una vecindad de x_0 con radio lo suficientemente pequeño, la sucesión

$$D_{k+1}(x),\ D_k(x),\ D_{k-1}(x)$$

contiene exactamente un cambio de signo.

Luego, el número de cambios de signo en la sucesión 2.3 se ve afectado sólo por las raíces del polinomio $D_m(x)$, de modo tal que si el producto $D_m(x)D_{m-1}(x)$ cambia su signo de $+$ a $-$ cuando x pasa por una de estas raíces, entonces se agrega un cambio de signo a la sucesión 2.3 (debido a los dos primeros términos $D_m(x)$ y $D_{m-1}(x)$); si este producto cambia su signo de - a $+$, entonces se pierde un cambio de signo en la sucesión 2.3; finalmente, si este producto no cambia de signo, entonces el número de cambios de signo en 2.3 no se modifica. □

La sucesión 2.3 también cuenta con las propiedades que se enuncian en el siguiente teorema y que enumeraremos como continuación de las dos propiedades anteriores.

Teorema 2.2. *Las afirmaciones siguientes son válidas:*

(iii) *Todas las raíces del polinomio $D_m(x)$ son reales y distintas.*

(iv) *Cuando x pasa por una raíz de $D_m(x)$, el producto $D_m(x)D_{m-1}(x)$ cambia su signo de $+$ a $-$.*

(v) *El número de raíces del polinomio $D_m(x)$ en el intervalo (α, β) $(\alpha < \beta)$ es igual al incremento en el número de cambios de signo en la sucesión 2.3 al variar x de α a β.*

(vi) *Entre dos raíces adyacentes cualesquiera del polinomio $D_m(x)$ existe exactamente una raíz del polinomio $D_{m-1}(x)$ $(m = 2, 3, \ldots, n)$.*

(vii) *Si denotamos por $\lambda_1 < \lambda_2 < \ldots < \lambda_n$ los distintos valores propios de la matriz J, es decir, todas las raíces del polinomio $D_n(x)$, entonces la sucesión*

$$D_{n-1}(x),\ D_{n-2}(x), \ldots, D_0(x) \tag{2.4}$$

tiene $j-1$ cambios de signo cuando $x = \lambda_j$.

Demostración. Como $D_k(x) = (-x)^k + \ldots$, tenemos que cuando x varía de $-\infty$ a ∞, la sucesión 2.3 acumula exactamente m cambios de signo porque para algún x lo suficientemente pequeño el polinomio $D_k(x)$ es mayor que cero y para algún x lo suficientemente grande $D_k(x) < 0$ si k es impar y $D_k(x) > 0$ si k es par. Luego, por la regla de Sturm aplicada al intervalo $(-\infty, \infty)$ obtenemos las propiedades (iii) y (iv). Al combinar la propiedad (iv) con la regla de Sturm obtenemos la propiedad (v). De las propiedades (iii) y (iv) se sigue la propiedad (vi).

Finalmente, para probar la propiedad (vii) observemos que como las raíces del polinomio $D_n(x)$ separan las raíces del polinomio $D_{n-1}(x)$, el intervalo $(-\infty, \lambda_j)$ contiene exactamente $j-1$ raíces del polinomio $D_{n-1}(x)$, ubicadas en los intervalos $(\lambda_1, \lambda_2), \ldots, (\lambda_{j-1}, \lambda_j)$, respectivamente. De modo tal que el incremento en el número de cambios de signo en la sucesión 2.4 al variar x de $-\infty$ a λ_j es igual a $j-1$, pero para valores de x lo suficientemente pequeños no hay cambios de signo en la sucesión 2.4. De aquí se sigue la propiedad (vii). □

Veamos ahora cómo calcular las coordenadas de un vector propio $u = (u_1, u_2, \ldots, u_n)$ correspondiente a un valor propio λ ($D_n(\lambda) = 0$). La ecuación vectorial $Ju - \lambda u = 0$ en términos de sus coordenadas se ve así:

$$\left.\begin{array}{ll} (a_1-\lambda)u_1 - b_1u_2 & = 0, \\ -c_1u_1 + (a_2-\lambda)u_2 - b_2u_3 & = 0, \\ \cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots & \\ -c_{n-1}u_{n-1} + (a_n-\lambda)u_n & = 0 \end{array}\right\}. \tag{2.5}$$

Como el determinante $D_n(\lambda) = 0$ y $D_{n-1}(\lambda) \neq 0$ (por (vi)), se sigue que las primeras $n-1$ ecuaciones de 2.5 son linealmente independientes y la última de estas ecuaciones se obtiene de las primeras $n-1$.

Consideremos las primeras $n-1$ ecuaciones para un valor cualquiera λ:

$$-c_{k-1}u_{k-1} + (a_k - \lambda)u_k - b_k u_{k+1} = 0$$
$$(k = 1, 2, \ldots, n-1; c_0 = u_0 = 0).$$

Al hacer el cambio de variables:

$$v_1 = u_1, \;\; v_k = b_1 b_2 \ldots b_{k-1} u_k \qquad (k = 2, 3, \ldots, n),$$

obtenemos

$$v_{k+1} = (a_k - \lambda)v_k - b_{k-1}c_{k-1}v_{k-1} \quad (k = 1, 2, \ldots, n-1; v_0 = 0). \tag{2.6}$$

Observemos que la fórmula recursiva 2.6 que determina los v_k coincide con la fórmula 2.2 con la cual se definen los $D_k(x)$ y, como evidentemente $u_1 \neq 0$, tenemos que

$$v_1 = CD_0(\lambda),\ v_2 = CD_1(\lambda) \quad (C = \text{const.} \neq 0).$$

Luego,

$$v_k = CD_{k-1}(\lambda) \qquad (k = 1, 2, \ldots, n).$$

Así pues,

$$u_k = Cb_1^{-1}b_2^{-1}\ldots b_{k-1}^{-1}D_{k-1}(\lambda) \quad (k = 1, 2, \ldots, n). \tag{2.7}$$

Hagamos ahora $\lambda = \lambda_j$ $(j = 1, 2, \ldots, n)$, de ese modo obtenemos para la k-ésima coordenada u_{kj} del j-ésimo vector propio u^j la siguiente expresión:

$$u_{kj} = C_j b_1^{-1}\ldots b_{k-1}^{-1}D_{k-1}(\lambda_j) \quad (k = 1, 2, \ldots, n;\ j = 1, 2, \ldots, n). \tag{2.8}$$

Definición 2.3. *Diremos que una matriz de Jacobi J es* normal *si*

$$b_k > 0,\ c_k > 0 \quad (k = 1, 2, \ldots, n-1).$$

Teorema 2.4. *Las matrices normales de Jacobi cuentan con las siguientes propie-*

dades:

(i) Todos sus valores propios son reales y simples.

(ii) La sucesión de las coordenadas del j-ésimo vector tiene exactamente $j-1$ cambios de signo.

Demostración. Se sigue de 2.8 y de (iii) y (vii) del teorema 2.2. □

Definición 2.5. *Sea $u = (u_1, u_2, \ldots, u_n)$ un vector. Llamaremos* u-línea *a la curva en el plano cartesiano, recta por pedazos, cuyos vértices P_k tienen las siguientes coordenadas:*

$$x_k = k,\ y_k = u_k \quad (k = 1, 2, \ldots, n).$$

Definición 2.6. *Sea $u = (u_1, u_2, \ldots, u_n)$ un vector. A los puntos donde la u-línea cruza el eje X los llamaremos* nodos *de la u-línea o u-vector.*

Si hacemos $C = 1$ en la ecuación 2.7, obtenemos un vector $u(\lambda) = (u_1(\lambda), u_2(\lambda), \ldots, u_n(\lambda))$ cuyas coordenadas son funciones de λ:

$$u_1(\lambda) \equiv 1,\ u_k(\lambda) = b_1^{-1} b_2^{-1} \ldots\ b_{k-1}^{-1} D_{k-1}(\lambda) \quad (k = 2, 3, \ldots, n).$$

De acuerdo con la definición 2.5, a este vector le corresponde una $u(\lambda)$-línea cuya ecuación puede escribirse en la forma

$$y(x;\lambda) = (k-x)u_{k-1}(\lambda) + (x-k+1)u_k(\lambda)$$
$$\text{para } k-1 \leq x \leq k \quad (k = 2, 3, \ldots, n).$$

Estudiaremos la forma de esta $u(\lambda)$-línea y el resultado que tiene sobre los nodos el hacer variar λ.

Observación 2.7. Observemos que un punto de intersección de una $u(\lambda)$-línea con el eje X en el intervalo $(1, n)$ es un nodo de esta $u(\lambda)$-línea si, y sólo si se tiene que cuando una coordenada $u_k(\lambda)$ $(1 < k < n)$ de $u(\lambda)$ es cero, las dos coordenadas adyacentes $u_{k-1}(\lambda)$ y $u_{k+1}(\lambda)$ tienen signos opuestos. Luego, por la propiedad (ii) de los polinomios $D_k(\lambda)$ $(k = 1, \dots, n-1)$, cualquier punto que tengan en común la $u(\lambda)$-línea y el eje X en el intervalo $(1, n)$ es un nodo de la $u(\lambda)$-línea.

Por el teorema 2.4, una $u(\lambda)$-línea con $\lambda = \lambda_j$ tiene exactamente $j-1$ nodos. En general, por la propiedad (v) del teorema 2.2, el número de nodos de la $u(\lambda)$-línea para un λ arbitrario es igual al número de raíces del polinomio $D_{n-1}(\lambda)$ en el intervalo $(-\infty, \lambda)$. Si *simetrizamos* la matriz J al reemplazar los números b_k y c_k por $\sqrt{b_k c_k} > 0$ $(k = 1, 2, \dots, n-1)$ y llamamos matriz *simetrizada* J_s a esta nueva matriz, encontramos que J_s tiene su propia $\bar{u}(\lambda)$-línea $\bar{P}_1, \dots, \bar{P}_n$, la cual difiere de la $u(\lambda)$-línea $P_1, \dots, P_n$, y se tiene que

$$\begin{aligned} \bar{u}_k(\lambda) &= C\sqrt{b_1^{-1} b_2^{-1} \dots b_{k-1}^{-1} c_1^{-1} c_2^{-1} \dots c_{k-1}^{-1}}\, D_{k-1}(\lambda) \\ &= \sqrt{\frac{b_1 \dots b_{k-1}}{c_1 \dots c_{k-1}}}\, u_k(\lambda) \qquad (k = 1, 2, \dots, n). \end{aligned} \tag{2.9}$$

De 2.9 se siguen las afirmaciones siguientes:

a) Estas líneas siempre coinciden en el número de nodos.

b) Si una de estas líneas tiene un nodo entre los puntos $x = k-1$ y $x = k$, entonces la otra línea también tiene un nodo entre los mismos puntos

c) Si una variación de λ ocasiona que los nodos de una línea se corran a la izquierda, lo mismo ocurre con la segunda línea.

De tal manera que para probar los teoremas y el lema siguientes, podemos suponer sin pérdida de generalidad que la matriz J es simétrica:

$$b_k = c_k \qquad (k = 1, 2, \dots, n-1).$$

De este modo, las cantidades $u_k(\lambda)$ están relacionadas por las ecuaciones

$$-b_{k-1}u_{k-1}(\lambda) + a_k u_k(\lambda) - b_k u_{k+1}(\lambda) = \lambda u_k(\lambda) \quad (k = 1, 2, \ldots, n-1). \tag{2.10}$$

Lema 2.8. *Para dos números arbitrarios λ y μ se satisface la siguiente fórmula:*

$$\begin{aligned} &b_{p-1}[u_{p-1}(\lambda)u_p(\mu) - u_{p-1}(\mu)u_p(\lambda)] \\ &-b_q[u_q(\lambda)u_{q+1}(\mu) - u_q(\mu)u_{q+1}(\lambda)] = \sum_{k=p}^{q}(\mu-\lambda)u_k(\lambda)u_k(\mu) \quad (1 \leq p \leq q < n-1). \end{aligned} \tag{2.11}$$

Demostración. Para un k fijo consideremos dos ecuaciones: la ecuación que aparece en 2.10 con λ y la misma ecuación pero con μ en lugar de λ. Entonces al eliminar el coeficiente a_k de ambas ecuaciones obtenemos

$$\begin{aligned} &b_{k-1}[u_{k-1}(\lambda)u_k(\mu) - u_{k-1}(\mu)u_k(\lambda)] \\ &-b_k[u_k(\lambda)u_{k+1}(\mu) - u_k(\mu)u_{k+1}(\lambda)] = (\mu-\lambda)u_k(\lambda)u_k(\mu). \end{aligned} \tag{2.12}$$

Si hacemos k igual a $p, p+1, ..., q$ $(1 \leq p \leq q < n-1)$ y sumamos las ecuaciones resultantes obtenemos 2.11. □

En particular, cuando $p = 1$ la fórmula 2.11 se simplifica notablemente:

$$\begin{aligned} -b_q[u_q(\lambda)u_{q+1}(\mu) - u_q(\mu)u_{q+1}(\lambda)]& \\ = \sum_{k=1}^{q}(\mu-\lambda)u_k(\lambda)u_k(\mu) &\quad (1 \leq q < n-1). \end{aligned} \tag{2.13}$$

Teorema 2.9. *Si $\lambda < \mu$, entonces entre dos nodos cualesquiera de la $u(\lambda)$-línea existe al menos un nodo de la $u(\mu)$-línea.*

Demostración. Sean α y β $(\alpha < \beta)$ dos nodos adyacentes de la $u(\lambda)$-línea, y sean p y q los enteros tales que

$$p-1 \leq \alpha < p, \;\; q < \beta \leq q+1 \quad (p \leq q).$$

De modo tal que $y(x;\lambda) \neq 0$ para $\alpha < x < \beta$, y que

$$\left.\begin{aligned} (p-\alpha)u_{p-1}(\lambda) + (\alpha-p+1)u_p(\lambda) &= 0, \\ (q+1-\beta)u_q(\lambda) + (\beta-q)u_{q+1}(\lambda) &= 0 \end{aligned}\right\}. \tag{2.14}$$

Supongamos lo contrario de lo que queremos demostrar, es decir, que

$$y(x;\mu) \neq 0 \quad \text{para} \quad \alpha < x < \beta.$$

Sin pérdida de generalidad, podemos suponer que

$$y(x;\lambda) > 0 \quad \text{para} \quad \alpha < x < \beta, \tag{2.15}$$

y que

$$y(x;\mu) > 0 \quad \text{para} \quad \alpha < x < \beta, \tag{2.16}$$

porque en caso contrario, consideraríamos $-y(x;\lambda)$ (o $-y(x;\mu)$) en lugar de $y(x;\lambda)$ (de $y(x;\mu)$), pero la fórmula 2.11 que vamos a emplear se mantiene igual si reemplazamos $u_i(\lambda)$ (o $u_i(\mu)$) por $-u_i(\lambda)$ (por $-u_i(\mu)$). De esta manera tenemos que

$$u_p(\lambda) > 0, \; u_{p+1}(\lambda) > 0, \ldots, u_q(\lambda) > 0, \tag{2.17}$$

y que

$$u_p(\mu) > 0, \; u_{p+1}(\mu) > 0, \ldots, u_q(\mu) > 0. \tag{2.18}$$

Además, de 2.16 también se sigue que

$$\left.\begin{array}{ll}(p-\alpha)u_{p-1}(\mu)+(\alpha-p+1)u_p(\mu) & \geq 0,\\ (q+1-\beta)u_q(\mu)+(\beta-q)u_{q+1}(\mu) & \geq 0\end{array}\right\}. \tag{2.19}$$

Al eliminar α de las dos primeras ecuaciones de 2.14 y 2.19, respectivamente, obtenemos:

$$\begin{aligned}&u_p(\mu)(u_{p-1}(\lambda)-u_p(\lambda))-u_p(\lambda)(u_{p-1}(\mu)-u_p(\mu))\\ &\quad=u_{p-1}(\lambda)u_p(\mu)-u_{p-1}(\mu)u_p(\lambda)\leq 0,\end{aligned}$$

y al eliminar β de las dos segundas ecuaciones de 2.14 y 2.19, respectivamente, obtenemos:

$$\begin{aligned}&u_q(\mu)(u_{q+1}(\lambda)-u_q(\lambda))+u_q(\lambda)(u_{q+1}(\mu)-u_q(\mu))\\ &\quad=u_q(\lambda)u_{q+1}(\mu)-u_q(\mu)u_{q+1}(\lambda)\geq 0.\end{aligned}$$

En consecuencia, el lado izquierdo de 2.11 en este caso no es positivo. Sin embargo, de 2.17 y 2.18 se sigue que el lado derecho de la misma igualdad 2.11 sí es positivo. Esta contradicción nos lleva a la afirmación del teorema. □

Teorema 2.10. *Conforme λ crece, los nodos de la $u(\lambda)$-línea se mueven a la izquierda.*

Demostración. Denotaremos por $\alpha_1(\lambda) < \alpha_2(\lambda) < \dots.$ los nodos sucesivos de la $u(\lambda)$-línea. De acuerdo con el teorema anterior, para $\mu > \lambda$ al menos uno de los nodos $\alpha_1(\mu) < \alpha_2(\mu) < \dots.$ se encuentra entre dos nodos adyacentes $\alpha_k(\lambda)$ y $\alpha_{k+1}(\lambda)$, en consecuencia, para probar las desigualdades

$$\alpha_k(\mu) < \alpha_k(\lambda) \quad \text{para} \quad \lambda < \mu \qquad (k = 1, 2, \dots)$$

bastará con probar la primera.

Para llevar a cabo un argumento por reducción al absurdo, supongamos que

$$\alpha_1(\mu) \geq \alpha_1(\lambda) \qquad (\lambda < \mu).$$

Sea

$$q < \alpha_1(\lambda) \leq q+1 \qquad (1 \leq q < n-1).$$

Como $u_1(\lambda) \equiv 1$, tenemos que

$$u_1(\lambda) > 0, u_2(\lambda) > 0, \ldots, u_{q-1}(\lambda) > 0, u_q(\lambda) > 0,$$

$$(q+1-\alpha_1(\lambda))u_q(\lambda) + (\alpha_1(\lambda)-q)u_{q+1}(\lambda) = 0,$$

$$u_1(\mu) > 0, u_2(\mu) > 0, \ldots, u_{q-1}(\mu) > 0, u_q(\mu) > 0,$$

$$(q+1-\alpha_1(\lambda))u_q(\mu) + (\alpha_1(\lambda)-q)u_{q+1}(\mu) = 0.$$

Al eliminar el número $\alpha_1(\lambda)$ de las relaciones que lo contienen obtenemos

$$u_q(\lambda)u_{q+1}(\mu) - u_q(\mu)u_{q+1}(\lambda) \geq 0.$$

En consecuencia, el lado izquierdo de 2.13 no es positivo mientras que el lado derecho sí lo es.

Esta contradicción nos lleva al resultado deseado. □

Teorema 2.11. *Los nodos de dos vectores propios sucesivos se alternan.*

Demostración. Por el teorema 2.4, el vector propio u^j, o dicho de otra forma, la $u(\lambda_j)$-línea tiene $j-1$ nodos:

$$\alpha_1(\lambda_j) < \alpha_2(\lambda_j) < \ldots < \alpha_{j-1}(\lambda_j) \quad (j = 2, 3, \ldots, n)$$

Por el teorema 2.10,

$$\alpha_1(\lambda_{j+1}) < \alpha_1(\lambda_j),$$

de modo tal que si también se cumple la desigualdad

$$\alpha_{j-1}(\lambda_j) < \alpha_j(\lambda_{j+1}), \tag{2.20}$$

por el teorema 2.9 se sigue que

$$\alpha_1(\lambda_{j+1}) < \alpha_1(\lambda_j) < \alpha_2(\lambda_{j+1}) < \alpha_2(\lambda_j) < \ldots < \alpha_{j-1}(\lambda_j) < \alpha_j(\lambda_{j+1})$$

tal como deseábamos probar.

Para probar la desigualdad 2.20 escojamos en número positivo arbitrario b_n y definamos

$$u_{n+1}(\lambda) := b_1^{-1} b_2^{-1} \cdots b_n^{-1} D_n(\lambda). \tag{2.21}$$

De esta manera podemos extender nuestra $u(\lambda)$-línea con un segmento de recta más $P_n P_{n+1}$, donde P_{n+1} tiene las coordenadas

$$x_{n+1} = n+1, \;\; y_{n+1} = u_{n+1}(\lambda).$$

Gracias a 2.21 la ecuación 2.10 se satisface ahora también para $k = n$. En consecuencia, todos los resultados anteriores son aplicables a la $u(\lambda)$-línea extendida $P_1 P_2 \ldots P_n P_{n+1}$. En particular, como en la demostración del teorema 2.9 vimos que entre dos ceros de la $u(\lambda)$-línea existe por lo menos un nodo de la $u(\mu)$-línea (¡no necesariamente entre dos nodos!), para la $u(\lambda_j)$-línea extendida tenemos que entre el nodo $\alpha_{j-1}(\lambda_j)$ y el punto $n+1$ existe por lo menos un nodo de la $u(\lambda_{j+1})$-línea extendida porque $u_{n+1}(\lambda_j) = 0$. Pero como $u_{n+1}(\lambda_{j+1}) = 0$, la $u(\lambda_{j+1})$-línea extendida tiene los mismos nodos

$$\alpha_1(\lambda_{j+1}) < \alpha_2(\lambda_{j+1}) < \ldots < \alpha_j(\lambda_{j+1})$$

que ya tenía la $u(\lambda_{j+1})$-línea. Luego,

$$\alpha_{j-1}(\lambda_j) < \alpha_j(\lambda_{j+1}).$$

Esto termina la demostración. □

En la derivación de los teoremas anteriores hicimos uso de la *simetrizabilidad* de la matriz J. Sin embargo, un análisis más refinado de la naturaleza de las propiedades oscilatorias muestra que la presencia de estas propiedades no está conectada con la *simetrizabilidad* ni con la simetría de las matrices como veremos en la sección 2.4.

2.2. Matrices Oscilatorias

En las proposiciones siguientes 2.13, 2.15 y 2.18 junto con la observación 2.16 enumeraremos en conjunto las propiedades más relevantes que atañen al concepto de matriz oscilatoria.

Definición 2.12. *Diremos que una matriz $A = (a_{ik})_1^n$ es* totalmente no-negativa *(*totalmente positiva*, respectivamente) cuando todos sus menores de cualquier orden son no-negativos (*positivos*, respectivamente):*

$$A\begin{pmatrix} i_1 & i_2 & \ldots & i_p \\ k_1 & k_2 & \ldots & k_p \end{pmatrix} \geq 0 \quad (resp. > 0) \quad para \quad 1 \leq \begin{matrix} i_1 < i_2 < \ldots < i_p \\ k_1 < k_2 < \ldots < k_p \end{matrix} \leq n$$

$$(p = 1, 2, \ldots, n).$$

Proposición 2.13. *Se cumplen las siguientes afirmaciones:*

(i) El producto de dos matrices totalmente no-negativas es una matriz totalmente no-negativa.

(ii) El producto de una matriz totalmente positiva con una matriz no-singular totalmente no-negativa es una matriz totalmente positiva.

Demostración. Recordemos la igualdad 1.3 derivada del teorema de Binet-Cauchy (teorema 1.1):

$$C\begin{pmatrix} i_1 & i_2 & \ldots & i_p \\ k_1 & k_2 & \ldots & k_p \end{pmatrix} = \sum_{\alpha_1<\alpha_2<\ldots<\alpha_p} A\begin{pmatrix} i_1 & i_2 & \ldots & i_p \\ \alpha_1 & \alpha_2 & \ldots & \alpha_p \end{pmatrix} B\begin{pmatrix} \alpha_1 & \alpha_2 & \ldots & \alpha_p \\ k_1 & k_2 & \ldots & k_p \end{pmatrix}$$

$$\left(1 \leq \begin{array}{c} i_1 < i_2 < \ldots < i_p \\ k_1 < k_2 < \ldots < k_p \end{array} \leq n\right).$$

La primera afirmación (i) es inmediata de esta relación. Para probar (ii) bastará con observar que la *no-singularidad* de B, digamos, implica que por lo menos uno de los menores

$$B\begin{pmatrix} \alpha_1 & \alpha_2 & \ldots & \alpha_p \\ k_1 & k_2 & \ldots & k_p \end{pmatrix}$$

es diferente de cero, pues de otro modo tendríamos que las columas de B correspondientes a los números $k_1, k_2, \ldots, k_p$ de acuerdo a su posición de izquierda a derecha son linealmente dependientes y, en consecuencia, que $|B| = 0$. □

Sea $A = (a_{ik})_1^n$. Denotaremos por $A^* = (a_{ik}^*)_1^n$ a la matriz con elementos

$$a_{ik}^* := (-1)^{i+k} a_{ik} \qquad (i, k = 1, 2, \ldots, n).$$

Resulta claro que:

a) Si $C = A \pm B$, entonces $C^* = A^* \pm B^*$.

b) Si $C = AB$, entonces $C^* = A^* B^*$.

c) Si $B = A^{-1}$, entonces $B^* = (A^*)^{-1}$.

Definición 2.14. *A una matriz $A = (a_{ik})_1^n$ la llamaremos* signo-regular *si la matriz A^* es totalmente no-negativa o, equivalentemente, si*

$$(-1)^{\sum_1^p i_\nu + \sum_1^p k_\nu} A\begin{pmatrix} i_1 & i_2 & \dots & i_p \\ k_1 & k_2 & \dots & k_p \end{pmatrix} \geq 0 \quad \text{para} \quad 1 \leq \begin{matrix} i_1 < i_2 < \dots < i_p \\ k_1 < k_2 < \dots < k_p \end{matrix} \leq n$$

$$(p = 1, 2, \dots, n).$$

Si todos los menores de cualquier orden de una matriz signo-regular A son distintos de cero, a esta matriz la llamaremos estrictamente signo-regular.

Proposición 2.15. *Las afirmaciones siguientes son válidas:*

(iii) Si una matriz no-singular A es totalmente no-negativa, entonces la matriz inversa $B = A^{-1}$ es signo-regular, y, en sentido opuesto, si una matriz no-singular A es signo-regular, entonces la matriz inversa $B = A^{-1}$ es totalmente no-negativa.

(iv) Si una matriz A es totalmente positiva, entonces la matriz inversa $B = A^{-1}$ es estrictamente signo-regular, e inversamente, si una matriz A es estrictamente signo-regular, entonces la matriz inversa $B = A^{-1}$ es totalmente positiva.

Demostración. Ambas afirmaciones se siguen en forma evidente de la fórmula para los menores de una matriz inversa probada en la proposición 1.3 del capítulo anterior:

$$(-1)^{\sum_1^p i_\nu + \sum_1^p k_\nu} B\begin{pmatrix} i_1 & i_2 & \dots & i_p \\ k_1 & k_2 & \dots & k_p \end{pmatrix} = \frac{A\begin{pmatrix} l_1 & l_2 & \dots & l_{n-p} \\ j_1 & j_2 & \dots & j_{n-p} \end{pmatrix}}{A\begin{pmatrix} 1 & 2 & \dots & n \\ 1 & 2 & \dots & n \end{pmatrix}}$$

donde el sistema de índices $i_1 < i_2 < \dots < i_p$; $j_1 < j_2 < \dots < j_{n-p}$ así como el sistema de índices $k_1 < k_2 < \ldots < k_p$; $l_1 < l_2 < \ldots < l_{n-p}$ a su vez, coinciden con el sistema de índices $1, 2, \ldots, n$. $\square$

Observación 2.16. Como $(A^*)^* = A$, los incisos (iii) y (iv) pueden reescribirse de la siguiente forma:

(v) Si una matriz no-singular A es totalmente no-negativa, entonces la matriz $C = (A^*)^{-1}$ es totalmente no-negativa. Si una matriz A es totalmente positiva, entonces la matriz $C = (A^*)^{-1}$ es totalmente positiva.

Más adelante veremos que los valores y vectores propios de una matriz totalmente positiva tienen varias propiedades notables. Sin embargo, la clase de las matrices totalmente positivas no es lo suficientemente amplia desde el punto de vista de las aplicaciones. Este hecho motivó el estudio de otro tipo de matrices y, eventualmente, llevó al estudio de la clase de las matrices oscilatorias, la cual se puede considerar intermedia entre la clase de las matrices totalmente no-negativas y la clase de las matrices totalmente positivas.

Definición 2.17. *A una matriz $A = (a_{ik})_1^n$ la llamaremos* oscilatoria *si A es totalmente no-negativa y existe un entero positivo κ tal que A^κ es totalmente positiva. Al menor de tales exponentes lo denominaremos el* exponente *de la matriz oscilatoria.*

De acuerdo con esta definición, las matrices totalmente positivas son matrices oscilatorias con exponente 1.

Proposición 2.18. *Las afirmaciones siguientes son válidas:*

(vi) Una matriz oscilatoria es invertible.

(vii) Una potencia A^p $(p = 1, 2, \ldots)$ de una matriz oscilatoria A también es una matriz oscilatoria.

(viii) Si A es una matriz oscilatoria con exponente κ, entonces para cualquier entero $k \geq \kappa$ la matriz A^k es totalmente positiva.

(ix) Si A es una matriz oscilatoria, entonces la matriz $(A^*)^{-1}$ es también una matriz oscilatoria.

Demostración. La afirmación (vi) es inmediata de la igualdad $|A^\kappa| = |A|^\kappa$, donde A es una matriz oscilatoria con exponente κ. La afirmación (vii) se sigue de la igualdad $(A^p)^\kappa = (A^\kappa)^p$ y de la afirmación (ii). De (ii) y de (vi) es inmediata la afirmación (viii). Finalmente, la afirmación (ix) se sigue de la afirmación (v) y de la propiedad b) de la operación $*$. □

Como ejemplos de matrices totalmente positivas veamos los siguientes:

Ejemplo 2.19. Una matriz de Vandermonde generalizada

$$A = (a_i^{\alpha_k})_1^n \qquad (0 < a_1 < a_2 < \ldots < a_n; \alpha_1 < \alpha_2 < \ldots < \alpha_n)$$

es totalmente positiva.

Para ver esto probaremos primero que una función f dada por

$$f(x) = c_1 x^{\alpha_1} + c_2 x^{\alpha_2} + \cdots + c_n x^{\alpha_n} \qquad (\sum c_i^2 > 0)$$

no tiene más de $n-1$ raíces positivas. Procederemos por inducción sobre n. Para $n = 1$ la afirmación es obvia. Supongamos que la afirmación es válida cuando el número de términos es menor que n. Si f tiene n raíces positivas, entonces

$$f_1(x) = (x^{-\alpha_1} f(x))' = (\alpha_2 - \alpha_1) c_2 x^{\alpha_2 - \alpha_1 - 1} + \cdots + (\alpha_n - \alpha_1) c_n x^{\alpha_n - \alpha_1 - 1}$$

tiene $n-1$ raíces positivas según el teorema de Rolle. Pero esto es imposible por la hipótesis de inducción, pues $f_1(x)$ tiene la misma forma que $f(x)$ pero con $n-1$ términos solamente. De esta manera, se sigue el resultado deseado.

Como los números $a_1, a_2, \ldots, a_n$ no pueden ser simultáneamente raíces de f, el sistema homogéneo de ecuaciones con respecto a c_k,

$$\sum_{k=1}^{n} c_k a_i^{\alpha_k} = 0 \qquad (i = 1, 2, \ldots, n).$$

tiene únicamente la solución trivial $c_1 = c_2 = \cdots = c_n = 0$. Entonces

$$|A| = |a_i^{\alpha_k}|_1^n \neq 0.$$

Por otra parte, cuando $\alpha_k = k - 1$ $(k = 1, 2, \ldots, n)$ el determinante $|A|$ es un determinante de Vandermonde. Haciendo uso de la fórmula para el determinante de Vandermonde (una demostración de esta fórmula puede verse en [7]) tenemos que $|a_i^{k-1}|_1^n = \prod_{1 \leq i < j \leq n}(a_j - a_i) > 0$ porque $0 < a_1 < a_2 < \ldots < a_n$. Puesto que hemos probado que el determinante de $|A|$ siempre es distinto de cero, de lo anterior y por la continuidad del determinante se sigue que $|A| > 0$ para $\alpha_1 < \alpha_2 < \ldots < \alpha_n$ arbitrarios. Dado que cualquier menor de la matriz A tiene la estructura de A misma, tenemos que todos los menores de la matriz A también son positivos, luego, A es una matriz totalmente positiva.

Ejemplo 2.20. La matriz

$$A = (e^{-\sigma(\alpha_i - \beta_k)^2})_1^n \qquad (\alpha_1 < \alpha_2 < \ldots < \alpha_n;\ \beta_1 < \beta_2 < \ldots < \beta_n;\ \sigma > 0)$$

es totalmente positiva.

Observemos que así como en el ejemplo anterior bastará con probar que $|A| > 0$.

El determinante $|e^{2\sigma\alpha_i\beta_k}|_1^n$ es positivo porque se obtiene del determinante de la matriz del ejemplo anterior al reemplazar a_i por $e^{2\sigma\alpha_i}$ y α_k por β_k. Como

$$|A| = e^{-\sigma(\alpha_1^2 + \cdots + \alpha_n^2 + \beta_1^2 + \cdots + \beta_n^2)}|e^{2\sigma\alpha_i\beta_k}|_1^n,$$

obtenemos el resultado que queríamos.

Pongamos $\alpha_i = \beta_i = i$ $(i = 1, 2, \ldots, n)$ en la matriz A del ejemplo 2.20 y denotemos la matriz resultante por F_σ:

$$F_\sigma = (e^{-\sigma(i-k)^2})_1^n.$$

Observemos una característica de la matriz F_σ que emplearemos en una sección posterior: los elementos en la diagonal de la matriz F_σ son todos iguales a uno y los elementos restantes tienden a cero cuando σ tiende a $+\infty$. Luego,

$$\lim_{\sigma\to+\infty} F_\sigma = E,$$

donde E es la matriz unitaria.

Un ejemplo de una matriz oscilatoria que no es totalmente positiva lo veremos en la última sección de esta tesis.

2.3. El teorema de Perron

En esta sección estableceremos el teorema de Perron referente al valor propio máximo de una matriz con entradas positivas y su vector propio correspondiente. Este teorema será una base para el estudio de valores y vectores propios de matrices oscilatorias. La prueba del teorema de Perron que daremos se debe a Frobenius. Los artículos originales referentes a esta sección que sugerimos son [10], [4], [2] y [3].

Teorema 2.21. (Perron) *Si todos los elementos de una matriz de orden n, $A = (a_{ik})_1^n$, son positivos, entonces existe un valor propio simple (i.e., una raíz simple del polinomio característico) y positivo ρ cuyo módulo es mayor que el de todos los demás valores propios. A este "máximo" valor propio le corresponde un vector propio con coordenadas positivas.*

Demostración. Denotaremos por $A_{ik}(\lambda)$ al cofactor del elemento $\lambda\delta_{ik} - a_{ik}$ en la matriz $\lambda E - A = (\lambda\delta_{ik} - a_{ik})_1^n$, es decir, $A_{ik}(\lambda)$ es el número que se obtiene al multiplicar $(-1)^{i+k}$ por el determinante de la submatriz de $\lambda E - A$ que se obtiene al suprimir el i-ésimo renglón y la k-ésima columna.

Probaremos por inducción sobre n el teorema de Perron junto con la afirmación:

$$A_{ik}(\lambda) > 0 \quad \text{para} \quad \lambda \geq \rho \quad (i, k = 1, 2, \dots, n). \tag{2.22}$$

Para $n = 1$ ciertamente es válido el teorema.

Ahora supondremos el enunciado del teorema y la afirmación 2.22 para matrices de orden menor que n y probaremos que en tal caso tanto el teorema como la afirmación 2.22 son ciertos para matrices de orden n.

Definamos $D_m(\lambda) := |\lambda\delta_{ik} - a_{ik}|_1^m$ $(m = 1, 2, \dots, n)$. Al expandir $D_n(\lambda)$ con respecto al último renglón y a la última columna obtenemos

$$\begin{aligned} D_n(\lambda) &= (\lambda - a_{nn})D_{n-1}(\lambda) + \sum_{i=1}^{n-1} -a_{in}(-1)^{i+n} \sum_{k=1}^{n-1} -a_{nk}(-1)^{n+k-1} A_{ik}^{(n-1)}(\lambda) \\ &= (\lambda - a_{nn})D_{n-1}(\lambda) - \sum_{i,k=1}^{n-1} A_{ik}^{(n-1)}(\lambda) a_{in} a_{nk} \,, \end{aligned} \tag{2.23}$$

donde $A_{ik}^{(n-1)}(\lambda)$ denota el cofactor del elemento $\lambda\delta_{ik} - a_{ik}$ en la matriz que subyace en $D_{n-1}(\lambda)$. Si denotamos por ρ_m el valor propio máximo de la matriz truncada $A_m = (a_{ik})_1^m$ $(m = 1, 2, \dots, n-1)$ que es positivo por hipótesis de inducción, entonces por 2.23 con $\lambda = \rho_{n-1}$ obtenemos:

$$D_n(\rho_{n-1}) = -\sum_{i,k=1}^{n-1} A_{ik}^{(n-1)}(\rho_{n-1}) a_{in} a_{nk} < 0.$$

Por otro lado, tenemos que

$$\lim_{\lambda \to \infty} D_n(\lambda) = \lim_{\lambda \to \infty} (\lambda^n + \dots) = +\infty \,.$$

En consecuencia, la ecuación $D_n(\lambda) = 0$ tiene una raíz positiva en el intervalo (ρ_{n-1}, ∞).

Denotaremos por ρ_n la solución más grande de la ecuación $D_n(\lambda) = 0$. Entonces $\rho_n > \rho_{n-1}$. Análogamente, $\rho_{n-1} > \rho_{n-2} > \dots$, de modo tal que $\rho_n > \rho_m$ si $m < n$. El número ρ_m es la raíz "máxima" de $D_m(\lambda)$ que puede verse como un menor principal de la matriz que subyace en $D_n(\lambda)$. Al reordenar los renglones y las columnas precisos podemos hacer que cualquiera de los menores principales de orden $m < n$ de esta matriz juegue el papel del menor $D_m(\lambda)$. Así pues, ρ_n es mayor que la más grande de las soluciones de cualquier menor principal de la matriz subyacente de $D_n(\lambda)$. Adicionalmente, como el coeficiente en el término de mayor grado en la expansión en potencias de λ de cualquier menor principal de la matriz subyacente de $D_n(\lambda)$ es igual a uno, tenemos que todos los menores principales en esta matriz son positivos cuando $\lambda \geq \rho_n$. En particular,

$$A_{ii}(\lambda) > 0 \quad \text{para} \quad \lambda \geq \rho_n \qquad (i = 1, 2, \dots, n).$$

Luego,

$$D_n^{'}(\rho_n) = \sum_{k=1}^{n} A_{ii}(\rho_n) > 0.$$

De donde ρ_n es una raíz simple de la ecuación característica $D_n(\lambda) = 0$.

Consideremos ahora el cofactor $A_{ik}(\lambda)$ para $i \neq k$. Entonces al expandir $A_{ik}(\lambda)$ con respecto al k-ésimo renglón y a la i-ésima columna en $D_n(\lambda)$ obtenemos

$$A_{ik}(\lambda) = a_{ki}C(\lambda) + \sum_{p \neq i,k} a_{pi} \sum_{q \neq i,k} a_{kq} C_{pq}(\lambda) = a_{ki}C(\lambda) + \sum_{p,q \neq i,k} C_{pq}(\lambda) a_{pi} a_{kq}, \quad (2.24)$$

aquí $C(\lambda)$ es el menor principal de orden $n-2$ que se obtiene de la matriz que subyace en $D_n(\lambda)$ al quitar el renglón y la columna i-ésimos así como el renglón y la columna k-ésimos, y $C_{pq}(\lambda)$ $(p, q \neq i, k)$ denota el cofactor del elemento $\lambda\delta_{pq} - a_{pq}$ en la matriz subyacente de $C(\lambda)$. A esta matriz podemos aplicarle el teorema de Perron así como la afirmación 2.22 por la hipótesis de inducción. Como ya vimos, la raíz máxima de la ecuación $C(\lambda) = 0$ es menor que ρ_n. En consecuencia, tenemos que

$$C(\lambda) > 0 \quad y \quad C_{pq}(\lambda) > 0 \quad \text{para} \quad \lambda \geq \rho_n.$$

Luego, por 2.24,

$$A_{ik}(\lambda) > 0 \quad \text{para} \quad \lambda \geq \rho_n.$$

De modo tal que si

$$u_1 = A_{11}(\rho_n), u_2 = A_{12}(\rho_n), \ldots, u_n = A_{1n}(\rho_n),$$

tenemos que

$$\sum_{k=1}^{n} (\rho\delta_{ik} - a_{ik})u_k = 0 \qquad (i = 1, 2, \ldots, n).$$

Consecuentemente, $u = (u_1, u_2, \ldots, u_n)$ es un vector propio con coordenadas positivas correspondiente al valor propio ρ_n de la matriz A.

Para completar la demostración, sólo nos resta probar que

$$\rho_n > |\lambda_0|,$$

donde λ_0 es un valor propio cualquiera de A distinto de ρ_n.
Si escribimos $u_i = A_{1i}(\rho_n)$ $(i = 1, 2, \ldots, n)$ como antes, tenemos que

$$\sum_{k=1}^{n} a_{ik}u_k = \rho_n u_i \qquad (i = 1, 2, \ldots, n). \tag{2.25}$$

Sea $v = (v_1, v_2, \ldots, v_n)$ un vector propio cualquiera de la matriz transpuesta A^t correspondiente al valor propio λ_0, es decir,

$$\sum_{i=1}^{n} a_{ik} v_i = \lambda_0 v_k \qquad (k = 1, 2, \ldots, n). \tag{2.26}$$

Entonces

$$\sum_{i=1}^{n} a_{ik} |v_i| \geq |\lambda_0||v_k| \qquad (k = 1, 2, \ldots, n). \tag{2.27}$$

De 2.25 y 2.27 hallamos que

$$\rho_n \sum_{i=1}^{n} u_i |v_i| = \sum_{i,k=1}^{n} a_{ik} u_k |v_i| \geq |\lambda_0| \sum_{k=1}^{n} u_k |v_k|.$$

Como $\sum_{i=1}^{n} u_i |v_i| \neq 0$, tenemos que $\rho_n \geq |\lambda_0|$. La igualdad se da si, y sólo si ocurre en cada una de las relaciones 2.27, i.e., si y sólo si

$$\sum_{i=1}^{n} a_{ik} |v_i| = |\sum_{i=1}^{n} a_{ik} v_i| \qquad (k = 1, 2, \ldots, n),$$

es decir, todos los números complejos v_i distintos de cero $(i = 1, 2, \ldots, n)$ tienen el mismo argumento. Pero entonces, por 2.26, tendríamos que $\lambda_0 > 0$. Luego, $\rho_n = |\lambda_0| = \lambda_0$, lo cual contradice la suposición hecha al principio. Entonces $\rho_n > |\lambda_0|$. Con esto queda terminada la prueba del teorema de Perron. □

2.4. Valores y vectores propios de una matriz oscilatoria.

Sea

$$u_1, u_2, \ldots, u_n \tag{2.28}$$

una sucesión de números reales. Si a cada término igual a cero de esta sucesión le asignamos un signo arbitrario, podemos calcular el número de cambios de signo en la sucesión 2.28. Este número depende de la elección de los signos para los términos cero de la sucesión 2.28 que hayamos hecho. Al mayor valor que pueda tomar este número lo llamaremos el *máximo número de cambios de signo* de la sucesión, y al menor de los valores para este número lo llamaremos el *mínimo número de cambios de signo* de la sucesión. Estos valores los denotaremos por S_u^+ y S_u^-, respectivamente. En caso de que $S_u^+ = S_u^-$ diremos que hay un *número exacto de cambios de signo* en la sucesión 2.28, que denotaremos simplemente por S_u.

Proposición 2.22. *Se satisfacen las siguientes afirmaciones:*

(i) En la sucesión 2.28 hay un número exacto de cambios de signo si, y sólo si se satisfacen las dos condiciones siguientes:

a) $u_1 u_n \neq 0$.

b) si algún u_i *es cero* $(1 < i < n)$, *entonces* $u_{i-1}u_{i+1} < 0$.

(ii) Si entre los términos u_i $(i = 1, 2, \ldots, n)$ *hay* k $(0 \leq k < n)$ *términos cero, entonces* $S_u^- \leq n - k - 1$ *y* $S_u^+ \geq k$.

Demostración. Para que no haya variaciones en el número de cambios de signo en la sucesión 2.28, desde luego es necesario que si algún término intermedio u_i $(1 < i < n)$ es cero, los términos adyacentes sean distintos de cero y tengan signos opuestos; es decir, la condición b) es necesaria. Si alguno de los extremos de la sucesión 2.28, por ejemplo u_1, es cero, resulta obvio que el número de cambios de signo en 2.28 cambiará de acuerdo a la elección de signo que hagamos para u_1. Luego, la condición a) también es necesaria para que haya un número exacto de cambios de signo en 2.28. Si, por otro lado, se satisfacen las condiciones a) y b), entonces no habrá ninguna variación en el número de cambios de signo en la sucesión 2.28 al asignar de forma arbitraria un signo a cada término cero, pues los únicos términos de la sucesión iguales a cero se encuentran en medio de dos términos con signos opuestos. Así pues, hemos probado (i).

Para probar (ii) observemos que hay $n-1$ posibilidades para un cambio de signo en 2.28 y por cada cero podemos *evitar* por lo menos un cambio de signo. Luego, $S_u^- \leq n-k-1$. Por otra parte, observemos también que por cada cero podemos generar por lo menos un cambio de signo en 2.28. En consecuencia, $S_u^+ \geq k$. □

Observación 2.23. Antes de continuar con el teorema fundamental de las propiedades espectrales de una matriz oscilatoria, hagamos un par de observaciones más sobre los números S_u^- y S_u^+. Resulta claro de la proposición anterior que bajo la suposición (i), S_u es igual al número de cambios de signo en la sucesión 2.28 después de remover todos los términos iguales a cero. También tenemos que si $u_i = 0$ para cada $i = 1, 2, \ldots, n$, entonces $S_u^- = 0$ y $S_u^+ = n-1$.

Teorema 2.24. *Sea $A = (a_{ik})_1^n$ una matriz oscilatoria. Entonces*

(i) La matriz A sólo tiene valores propios simples y todos ellos son positivos:

$$\lambda_1 > \lambda_2 > \cdots > \lambda_n > 0. \tag{2.29}$$

(ii) Si $u^k = (u_{1k}, u_{2k}, \ldots, u_{nk})$ es un vector propio de A correspondiente al valor propio λ_k, el k-ésimo en magnitud entre todos los valores propios ($k = 1, 2, \ldots, n$), entonces dada cualquier sucesión de coeficientes reales $c_p, c_{p+1}, \ldots, c_q$ ($1 \leq p \leq q \leq n,\ \ \sum_{i=p}^q c_i^2 > 0$), el número de cambios de signo en las coordenadas del vector

$$u := c_p u^p + c_{p+1} u^{p+1} + \cdots + c_q u^q$$

se encuentra entre $p-1$ y $q-1$. Es decir,

$$p-1 \leq S_u^- \leq S_u^+ \leq q-1. \tag{2.30}$$

En particular, entre las coordenadas del vector u^k ($k = 1, 2, \ldots, n$) hay exactamente $k-1$ cambios de signo, i.e.,

$$S_{u^k} = k-1 \qquad (k = 1, 2, \ldots, n). \tag{2.31}$$

(iii) Los nodos de dos vectores propios sucesivos u^k y u^{k+1} ($k = 1, 2, \ldots, n-1$) se alternan.

Demostración. Probaremos el teorema primero para el caso en el que A es totalmente positiva.

Prueba de (i): Si la matriz A es totalmente positiva, entonces para cualquier $q \leq n$ los elementos de la matriz asociada q-ésima $\mathfrak{A}_q$ son positivos porque todos los menores de q-ésimo orden de la matriz A son positivos. Entonces el teorema de Perron (teo. 2.21)es aplicable a la matriz $\mathfrak{A}_q$. Por otro lado, según el teorema de Kronecker (teorema 1.28), los valores propios de $\mathfrak{A}_q$ son todos los posibles productos de q valores propios de A. Si enumeramos los valores propios de A en orden decreciente de acuerdo a su valor absoluto: $|\lambda_1| \geq |\lambda_2| \geq \cdots \geq |\lambda_n|$, tenemos que el más grande en valor absoluto de los valores propios de $\mathfrak{A}_q$ es el producto $\lambda_1\lambda_2\cdots\lambda_q$. Entonces al aplicar el teorema de Perron a la matriz asociada $\mathfrak{A}_q$ obtenemos lo siguiente:

$$\lambda_1\lambda_2\cdots\lambda_q > 0 \qquad (q = 1, 2, \ldots, n),$$
$$\lambda_1\lambda_2\cdots\lambda_q > |\lambda_1\lambda_2\cdots\lambda_{q-1}\lambda_{q+1}| \quad (q = 1, 2, \ldots, n-1).$$

De la primera desigualdad se sigue que

$$\lambda_q > 0 \qquad (q = 1, 2, \ldots, n)$$

y de la segunda desigualdad que

$$\lambda_q > \lambda_{q+1} \qquad (q = 1, 2, \ldots, n-1).$$

Estas son las desigualdades 2.29.

Prueba de (ii): Hemos probado que una matriz totalmente positiva no puede tener valores propios múltiples. De modo tal que un vector propio $u^k = (u_{1k}, u_{2k}, \ldots, u_{nk})$ correspondiente al k-ésimo valor propio λ_k está determinado en forma única salvo un factor. Por el corolario 1.30 del capítulo anterior tenemos que para una q fija $(q = 1, 2, \ldots, n)$, los menores

$$U\begin{pmatrix} i_1 & i_2 & \dots & i_q \\ 1 & 2 & \dots & q \end{pmatrix} \qquad (1 \leq i_1 < i_2 < \cdots < i_q \leq n) \tag{2.32}$$

de la matriz fundamental $U = (u_{ik})_1^n$ son las coordenadas del vector propio de la matriz asociada $\mathfrak{A}_q$ correspondiente al mayor de los valores propios, $\lambda_1\lambda_2\cdots\lambda_q$. Luego, por el teorema 2.21, todos los menores 2.32 son distintos de cero y todos tienen el mismo signo ϵ_q para cualquier $q \leq n$. De tal manera que si multiplicamos los vectores $u^1, u^2, \dots, u^n$ por los factores $\epsilon_1, \epsilon_2/\epsilon_1, \dots, \epsilon_n/\epsilon_{n-1}$, respectivamente, obtendremos las desigualdades:

$$U\begin{pmatrix} i_1 & i_2 & \dots & i_q \\ 1 & 2 & \dots & q \end{pmatrix} > 0 \quad (1 \leq i_1 < i_2 < \cdots < i_q \leq n, \;\; q = 1, 2, \dots, n). \tag{2.33}$$

Supongamos que para un entero q,

$$u = \sum_{k=1}^{q} c_k u^k \qquad \left(\sum_{k=1}^{q} c_k^2 > 0\right), \tag{2.34}$$

y probemos que 2.34 implica la desigualdad

$$S_u^+ \leq q - 1. \tag{2.35}$$

Si por el contrario, $S_u^+ \geq q$, entonces podemos escoger $q+1$ coordenadas $u_{i_1}, u_{i_2}, \dots, u_{i_{q+1}}$ del vector u tales que

$$u_{i_\alpha} u_{i_{\alpha+1}} \leq 0 \qquad (\alpha = 1, 2, \dots, q). \tag{2.36}$$

Aquí las coordenadas $u_{i_1}, u_{i_2}, \ldots, u_{i_q}$ no son cero simultáneamente porque en caso contrario los números c_k ($k = 1, 2, \ldots, q,\ \sum_{k=1}^{q} c_k^2 > 0$) satisfacen el sistema de ecuaciones homogéneo

$$\sum_{k=1}^{q} c_k u_{i_\alpha k} = 0 \qquad (\alpha = 1, 2, \ldots, q)$$

con determinante 2.32 distinto de cero.
Consideremos el determinante

$$\begin{vmatrix} u_{i_1 1} & \ldots & u_{i_1 q} & u_{i_1} \\ \ldots & \ldots & \ldots & \ldots \\ \ldots & \ldots & \ldots & \ldots \\ u_{i_{q+1} 1} & \ldots & u_{i_{q+1} q} & u_{i_{q+1}} \end{vmatrix} = 0.$$

Al expandir este determinante con respecto a la última columna obtenemos

$$\sum_{\alpha=1}^{q+1} (-1)^{q+\alpha+1} u_{i_\alpha} U \begin{pmatrix} i_1 & \ldots & i_{\alpha-1} & i_{\alpha+1} & \ldots & i_{q+1} \\ 1 & \ldots & \ldots & \ldots & \ldots & q \end{pmatrix} = 0.$$

Pero de 2.33 y 2.36 se sigue que todos los términos distintos de cero del lado izquierdo de esta ecuación (que sabemos que existen) tienen el mismo signo, de modo tal que la ecuación misma resulta imposible. Así, hemos probado que 2.35 se sigue de 2.33 y 2.34.

Para terminar con la prueba de (ii) consideremos un vector

$$u = c_p u^p + c_{p+1} u^{p+1} + \cdots + c_q u^q$$

como se definió en el enunciado del teorema. Ya probamos que

$$S_u^+ \leq q - 1.$$

Si

$$B = (b_{ik})_1^n := A^{-1} \quad \text{y} \quad C := B^* := (b_{ik}^*)_1^n := ((-1)^{i+k} b_{ik})_1^n,$$

entonces C es una matriz totalmente positiva, como ya habíamos observado. Por otro lado, como

$$Au^k = \lambda_k u^k \qquad (k = 1, 2, \dots, n),$$

tenemos que

$$Bu^k = \lambda_k^{-1} u^k \qquad (k = 1, 2, \dots, n)$$

o, dicho de otra manera,

$$\sum_{j=1}^n b_{ij} u_{jk} = \lambda_k^{-1} u_{ik} \qquad (i, k = 1, 2, \dots, n). \tag{2.37}$$

A cada vector $u = (u_1, u_2, \dots, u_n)$ le asignamos un vector $u^* = (u_1^*, u_2^*, \dots, u_n^*)$, donde $u_i^* = (-1)^i u_i$ $(i = 1, 2, \dots, n)$. Luego, de 2.35 se sigue que

$$\sum_{j=1}^n b_{ij}^* u_{jk}^* = \lambda_k^{-1} u_{ik}^* \quad (u_{ik}^* = (-1)^{i+k} u_{ik};\ i,\ k = 1, 2, \dots, n),$$

i.e.,

$$Cu^{k*} = \lambda_k^{-1} u^{k*} \qquad (k = 1, 2, \dots, n).$$

Entonces $u^{n*}, u^{(n-1)*}, \dots, u^{1*}$ es un sistema completo de vectores propios de una matriz totalmente positiva C. Los valores propios correspondientes están dispuestos en orden decreciente: $\lambda_n^{-1} > \lambda_{n-1}^{-1} > \dots > \lambda_1^{-1}$. Como

$$u^* = c_q u^{q*} + \dots + c_p u^{p*},$$

podemos deducir una relación análoga a 2.35 para u^* simplemente notando que en esta nueva relación el papel del número q lo desempeña ahora el número $n-p+1$:

$$S^+_{u^*} \leq n - p. \tag{2.38}$$

Ahora, si asignamos a cada coordenada de u igual a cero un signo, de modo tal que el número de cambios de signo en las coordenadas de u sea igual a S^-_u, y bajo esta condición después multiplicamos cada coordenada u_i por $(-1)^i$, habremos hecho desaparecer los cambios de signo que había entre las coordenadas de u y habremos hecho aparecer cambios de signo donde antes no los había. Luego, el número S^-_u más el número de cambios de signo en esta nueva sucesión de coordenadas es igual a $n-1$. Pero resulta evidente que el número de cambios de signo en tal sucesión es igual $S^+_{u^*}$. En consecuencia,

$$S^-_u + S^+_{u^*} = n - 1. \tag{2.39}$$

De 2.38 y 2.39 se sigue que

$$S^-_u \geq p - 1. \tag{2.40}$$

Al combinar 2.35 con 2.40 obtenemos 2.30. Esto termina la prueba de (ii).

Prueba de (iii): Recordemos que los nodos de un vector u son los nodos de la correspondiente u-línea. Como

$$S^-_{u^k} = S^+_{u^k} \quad y \quad S^-_{u^{k+1}} = S^+_{u^{k+1}}\,, \tag{2.41}$$

por la proposición (ii) tenemos que todos los ceros de la u^k-línea y los de la u^{k+1}-línea son nodos.

Emplearemos los siguientes resultados derivados de la proposición (ii):

$$S_{u^k} = k - 1, \quad S_{u^{k+1}} = k, \tag{2.42}$$

y, para un vector

$$u = cu^k + du^{k+1} \tag{2.43}$$

con c y d arbitrarios ($c^2 + d^2 > 0$) tenemos que

$$k - 1 \leq S_u^- \leq S_u^+ \leq k. \tag{2.44}$$

Supongamos que entre cierto par de nodos sucesivos α y β de la u^{k+1}-línea no hay ningún nodo de la u^k-línea. Entonces las funciones $u^k(x)$ y $u^{k+1}(x)$ (las ordenadas de las u^k y u^{k+1} líneas) son diferentes de cero y tienen signo constante en el intervalo (α, β). Sin pérdida de generalidad podemos suponer que

$$u^k(x) > 0, \;\; u^{k+1}(x) > 0 \qquad (\alpha < x < \beta). \tag{2.45}$$

Pongamos $d = -1$ en 2.43 y consideremos la función

$$u(x) = cu^k - u^{k+1}(x).$$

Veamos primero que $u^k(\alpha) \neq 0$ y $u^k(\beta) \neq 0$. Por ejemplo, supongamos que $u^k(\alpha) = 0$. Entonces $u(\alpha) = 0$ para cualquier c. Elijamos γ tal que

$$\alpha < \gamma < \text{mín}\{\beta, [\alpha + 1]\}$$

y hagamos $c = u^{k+1}(\gamma)/u^k(\gamma)$. Entonces $u(\gamma) = 0$ también. Luego, un segmento de la u-línea yace sobre el eje x. De aquí que dos coordenadas sucesivas del vector u son iguales a cero. Luego, $S_u^+ - S_u^- \geq 2$, lo cual contradice 2.44. De modo tal que la u^k-línea y la u^{k+1}-línea no tienen nodos en común. Observemos ahora que existe un número ζ en $[\alpha, \beta]$ tal que $u^{k+1}(\zeta) \geq u^{k+1}(x)$ para cada x en $[\alpha, \beta]$. Si $c_0 := u^{k+1}(\zeta)/u^k(\zeta)$, entonces

$$u_0(x) := c_0 u^k(x) - u^{k+1}(x) \geq 0 \quad (\alpha \leq x \leq \beta)$$

y

$$u_0(\zeta) = 0.$$

Adicionalmente, tenemos que la función $u_0(x)$ no se anula en $x = \alpha$ ni en $x = \beta$ porque $c_0 > 0$. Entonces la línea quebrada $y = u_0(x)$ se encuentra de un lado del eje x y tiene un punto en común $P(\zeta, 0)$ $(\alpha < \zeta < \beta)$ con este eje. Si el segmento de la línea $y = u_0(x)$ que contiene al punto $P(\zeta, 0)$ yace por completo en el eje x, entonces tendríamos que $S_{u_0}^+ - S_{u_0}^- \geq 2$ que, como ya vimos, contradice 2.44. Por otra parte, si la línea quebrada tiene sólo un vértice $P(\zeta, 0)$ en el eje x, entonces los dos vértices adyacentes a este vértice se encuentran de un mismo lado del eje x, de modo tal que entre las coordenadas del vector u_0 hay una que es igual a cero y que es adyacente a dos coordenadas con el mismo signo. En este caso también tenemos la desigualdad $S_{u_0}^+ - S_{u_0}^- \geq 2$, y nuevamente la contradicción con 2.44.

Consecuentemente, tenemos que entre dos nodos sucesivos de la u^{k+1}-línea existe por lo menos un nodo de la u^k-línea. Pero de acuerdo con 2.42, la u^k-línea tiene $k-1$ nodos y la u^{k+1}-línea tiene k nodos. Por tal motivo, entre dos nodos sucesivos de la u^{k+1}-línea hay exactamente un nodo de la u^k-línea. Luego, los nodos de la u^{k+1}-línea y los de la u^k-línea se alternan.

Así pues, hemos probado el teorema para el caso en el que la matriz A es totalmente positiva.

Si A es cualquier matriz oscilatoria con exponente κ, entonces las matrices A^κ y $A^{\kappa+1}$ son totalmente positivas (proposición 2.18). Los valores propios de estas matrices son, respectivamente:

$$\lambda_1^\kappa > \lambda_2^\kappa > \cdots > \lambda_n^\kappa > 0 \quad \text{y} \quad \lambda_1^{\kappa+1} > \lambda_2^{\kappa+1} > \cdots > \lambda_n^{\kappa+1} > 0.$$

De aquí se sigue 2.29.

Las matrices A y A^κ tienen los mismos vectores propios, y el vector propio u^k corresponde a los k-ésimos valores propios λ_k y λ_k^κ de las matrices A y A^κ, respectivamente. De modo tal que las proposiciones (ii) y (iii) son válidas también para los vectores propios de la matriz oscilatoria A.

Así, hemos demostrado el teorema por completo. $\square$

2.5. Una desigualdad fundamental

Definición 2.25. *Sea $A = (a_{ik})_1^n$ una matriz. A un menor de la matriz A lo llamaremos* casi principal *cuando dado $p > 1$ el menor tenga la forma*

$$A\begin{pmatrix} i_1 & \dots & i_{\alpha-1} & i & i_\alpha & \dots & i_{\beta-1} & i_\beta & i_{\beta+1} & \dots & i_p \\ i_1 & \dots & i_{\alpha-1} & i_\alpha & i_{\alpha+1} & \dots & i_\beta & k & i_{\beta+1} & \dots & i_p \end{pmatrix} \quad \text{con} \quad i \neq k. \tag{2.46}$$

Los elementos a_{ik} con $i \neq k$ serán menores casi principales de orden 1.

Entonces los menores casi principales de orden p se obtienen de un menor principal de orden $p-1$ al agregar un renglón y una columna con distintos índices.

Teorema 2.26. *Si todos los menores principales y casi principales de una matriz $A = (a_{ik})_1^n$ son no-negativos, entonces para cada $p < n$ se cumple la desigualdad* *siguiente:*

$$A\begin{pmatrix} 1 & 2 & \ldots & n \\ 1 & 2 & \ldots & n \end{pmatrix} \leq A\begin{pmatrix} 1 & \ldots & p \\ 1 & \ldots & p \end{pmatrix} A\begin{pmatrix} p+1 & \ldots & n \\ p+1 & \ldots & n \end{pmatrix}. \tag{2.47}$$

Demostración. Probaremos el teorema por inducción sobre n. Para $n = 2$ tenemos que

$$\begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix} = a_{11}a_{22} - a_{12}a_{21} \leq a_{11}a_{22} \tag{2.48}$$

porque $a_{12}a_{21} \geq 0$ por hipótesis.

Supongamos que $n > 2$ y que el teorema es válido si el orden de A es menor que n. Entonces uno de los números p o $n-p$ es mayor que 1. Sin pérdida de generalidad podemos suponer que $p > 1$, pues en caso contrario podemos reordenar los renglones y columnas de la matriz A en orden inverso para obtener la desigualdad 2.47 con esta nueva matriz que, en realidad, se trata de la misma desigualdad con la matriz original A.

Consideraremos dos casos.

1. Supongamos primero que

$$A\begin{pmatrix} 1 & 2 & \ldots & p-1 \\ 1 & 2 & \ldots & p-1 \end{pmatrix} \neq 0.$$

Sea $D = (d_{ik})_p^n$ la matriz definida por

$$d_{ik} := A\begin{pmatrix} 1 & \ldots & p-1 & i \\ 1 & \ldots & p-1 & k \end{pmatrix} \qquad (i, k = p, \ldots, n). \tag{2.49}$$

Por la identidad de Sylvester (teorema 1.4) tenemos que

$$D\begin{pmatrix} i_1 & i_2 & \dots & i_q \\ k_1 & k_2 & \dots & k_q \end{pmatrix}$$

$$= A\begin{pmatrix} 1 & 2 & \dots & p-1 \\ 1 & 2 & \dots & p-1 \end{pmatrix}^{q-1} A\begin{pmatrix} 1 & 2 & \dots & p-1 & i_1 & i_2 & \dots & i_q \\ 1 & 2 & \dots & p-1 & k_1 & k_2 & \dots & k_q \end{pmatrix},$$

de donde los menores principales de D son positivos y los casi principales son no-negativos.

Entonces al hacer uso de la hipótesis de inducción y aplicar la identidad de Sylvester dos veces hallamos que

$$\begin{aligned} A\begin{pmatrix} 1 & \dots & n \\ 1 & \dots & n \end{pmatrix} &= \frac{D\begin{pmatrix} p & p+1 & \dots & n \\ p & p+1 & \dots & n \end{pmatrix}}{A\begin{pmatrix} 1 & \dots & p-1 \\ 1 & \dots & p-1 \end{pmatrix}^{n-p}} \leq \frac{d_{pp}D\begin{pmatrix} p+1 & \dots & n \\ p+1 & \dots & n \end{pmatrix}}{A\begin{pmatrix} 1 & \dots & p-1 \\ 1 & \dots & p-1 \end{pmatrix}^{n-p}} \\ &= \frac{A\begin{pmatrix} 1 & \dots & p \\ 1 & \dots & p \end{pmatrix} A\begin{pmatrix} 1 & \dots & p-1 & p+1 & \dots & n \\ 1 & \dots & p-1 & p+1 & \dots & n \end{pmatrix}}{A\begin{pmatrix} 1 & \dots & p-1 \\ 1 & \dots & p-1 \end{pmatrix}} \\ &\leq A\begin{pmatrix} 1 & \dots & p \\ 1 & \dots & p \end{pmatrix} A\begin{pmatrix} p+1 & \dots & n \\ p+1 & \dots & n \end{pmatrix}. \end{aligned} \tag{2.50}$$

2. Supongamos ahora que

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

Entonces $a_{11} = 0$ o existe $q < p-1$ tal que

$$A\begin{pmatrix} 1 & 2 & \dots & q \\ 1 & 2 & \dots & q \end{pmatrix} \neq 0 \qquad \text{y} \qquad A\begin{pmatrix} 1 & 2 & \dots & q+1 \\ 1 & 2 & \dots & q+1 \end{pmatrix} = 0. \tag{2.51}$$

Supongamos primero que $a_{11} = 0$. Como

$$-a_{i1}a_{1k} = \begin{vmatrix} a_{11} & a_{1k} \\ a_{i1} & a_{ik} \end{vmatrix} = A\begin{pmatrix} 1 & i \\ 1 & k \end{pmatrix} \geq 0 \quad \text{y} \quad a_{i1}a_{1k} \geq 0 \qquad (i,k = 2,3,\dots,n),$$

se sigue que $a_{i1}a_{1k} = 0 \ \ (i,k = 2,3,\dots,n)$. Entonces

$$a_{i1} = 0 \qquad (i = 2,\dots,n)$$

o

$$a_{1k} = 0 \qquad (k = 2,\dots,n).$$

Como $a_{11} = 0$, en cualquier caso tenemos que

$$A\begin{pmatrix} 1 & 2 & \dots & n \\ 1 & 2 & \dots & n \end{pmatrix} = 0. \tag{2.52}$$

Si $a_{11} \neq 0$, tenemos 2.51. Definamos la matriz $B = (b_{ik})_{q+1}^{p}$ por

$$b_{ik} := A\begin{pmatrix} 1 & \dots & q & i \\ 1 & \dots & q & k \end{pmatrix} \qquad (i,k = q+1,\dots,n). \tag{2.53}$$

Por la identidad de Sylvester

$$B\begin{pmatrix} i_1 & i_2 & \dots & i_g \\ k_1 & k_2 & \dots & k_g \end{pmatrix} = A\begin{pmatrix} 1 & \dots & q \\ 1 & \dots & q \end{pmatrix}^{g-1} A\begin{pmatrix} 1 & \dots & q & i_1 & i_2 & \dots & i_g \\ 1 & \dots & q & k_1 & k_2 & \dots & k_g \end{pmatrix} \tag{2.54}$$

y la hipótesis del teorema, se sigue que todos los menores principales y casi principales de la matriz B son no-negativos.

Como el orden de la matriz B es menor que n, por hipótesis de inducción y porque $b_{q+1,q+1} = 0$, se sigue que

$$B\begin{pmatrix} q+1 & \dots & n \\ q+1 & \dots & n \end{pmatrix} \leq b_{q+1,q+1} B\begin{pmatrix} q+2 & \dots & n \\ q+2 & \dots & n \end{pmatrix} = 0.$$

Luego,

$$B\begin{pmatrix} q+1 & \dots & n \\ q+1 & \dots & n \end{pmatrix} = 0.$$

Al aplicar nuevamente la identidad de Sylvester 2.54 y la afirmación 2.51 obtenemos 2.52 por segunda ocasión.

De tal manera que en este segundo caso el lado izquierdo de 2.47 es igual a cero. Como su lado derecho es no-negativo, en este caso también se tiene la desigualdad 2.47.

Esto termina la prueba. $\square$

Lema 2.27. *Si $A = (a_{ik})_1^n$ es una matriz totalmente no-negativa y*

$$A\begin{pmatrix} 1 & 2 & \dots & p-1 & q \\ 1 & 2 & \dots & p-1 & p \end{pmatrix} = 0 \qquad (1 < p < q \leq n), \tag{2.55}$$

entonces

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

o

$$a_{qk} = 0 \qquad (k = 1, 2, \dots, p).$$

Demostración. Supongamos que

$$A\begin{pmatrix} 1 & 2 & \dots & p-1 & p \\ 1 & 2 & \dots & p-1 & p \end{pmatrix} \neq 0.$$

Por el teorema 2.26 tenemos que

$$A\begin{pmatrix} 1 & 2 & \dots & p-1 & p \\ 1 & 2 & \dots & p-1 & p \end{pmatrix} \leq A\begin{pmatrix} 1 & 2 & \dots & p-1 \\ 1 & 2 & \dots & p-1 \end{pmatrix} a_{pp},$$

en consecuencia,

$$A\begin{pmatrix} 1 & 2 & \dots & p-1 \\ 1 & 2 & \dots & p-1 \end{pmatrix} > 0. \tag{2.56}$$

De 2.55 y 2.56 se sigue que en la matriz

$$\begin{pmatrix} a_{11} & \dots & a_{1p} \\ \dots & \dots & \dots \\ a_{p-1,1} & \dots & a_{p-1,p} \\ \dots & \dots & \dots \\ a_{q1} & \dots & a_{qp} \\ \dots & \dots & \dots \\ a_{n1} & \dots & a_{np} \end{pmatrix}$$

los primeros $p-1$ renglones son linealmente independientes y que el q-ésimo renglón es una combinación lineal de estos $p-1$ renglones:

$$a_{qk} = \sum_{\nu=1}^{p-1} \lambda_\nu a_{\nu k} \qquad (k = 1, 2, \ldots, p).$$

Mostremos que todos los números λ_ν son iguales a cero. Para cualquier ν se tiene lo siguiente:

$$\begin{aligned} & A\begin{pmatrix} 1 & 2 & \ldots & \nu-1 & \nu+1 & \ldots & p-1 & q \\ 1 & 2 & \ldots & \nu-1 & \nu & \ldots & p-2 & p-1 \end{pmatrix} \\ = & \textstyle\sum_{\mu=1}^{p-1} \lambda_\mu A\begin{pmatrix} 1 & 2 & \ldots & \nu-1 & \nu+1 & \ldots & p-1 & \mu \\ 1 & 2 & \ldots & \nu-1 & \nu & \ldots & p-2 & p-1 \end{pmatrix} \\ = & (-1)^{p-\nu-1}\lambda_\nu A\begin{pmatrix} 1 & 2 & \ldots & p-1 \\ 1 & 2 & \ldots & p-1 \end{pmatrix}, \end{aligned} \tag{2.57}$$

$$\begin{aligned} & A\begin{pmatrix} 1 & 2 & \ldots & \nu-1 & \nu+1 & \ldots & p & q \\ 1 & 2 & \ldots & \nu-1 & \nu & \ldots & p-1 & p \end{pmatrix} \\ = & (-1)^{p-\nu}\lambda_\nu A\begin{pmatrix} 1 & 2 & \ldots & p \\ 1 & 2 & \ldots & p \end{pmatrix}. \end{aligned} \tag{2.58}$$

Puesto que todos los menores de A son no-negativos y los menores en los lados derechos respectivos de 2.57 y 2.58 son positivos, simultáneamente tenemos las desigualdades

$$(-1)^{p-\nu-1}\lambda_\nu \geq 0 \qquad y \qquad (-1)^{p-\nu}\lambda_\nu \geq 0;$$

es decir,

$$\lambda_\nu = 0 \qquad (\nu = 1, 2, \ldots, p-1).$$

Esto prueba el lema. □

Teorema 2.28. *Si $A = (a_{ik})_1^n$ es una matriz totalmente no-negativa, entonces para cualquier $p < n$ la desigualdad 2.47 se satisface:*

$$A\begin{pmatrix} 1 & 2 & \dots & n \\ 1 & 2 & \dots & n \end{pmatrix} \leq A\begin{pmatrix} 1 & \dots & p \\ 1 & \dots & p \end{pmatrix} A\begin{pmatrix} p+1 & \dots & n \\ p+1 & \dots & n \end{pmatrix},$$

y la igualdad en esta expresión se tiene sólo en los siguientes casos:

(i) Uno de los factores en el lado derecho es igual a cero.

(ii) Los elementos a_{ik} $(i = 1, 2, \dots, p;\ k = p+1, \dots, n)$ *o los elementos* a_{ik} $(i = p+1, \dots, n;\ k = 1, \dots, p)$ *son iguales a cero.*

Demostración. Puesto que cualquier matriz totalmente no-negativa satisface las condiciones del teorema 2.23, tenemos que se cumple 2.47.

Probaremos, ahora, la segunda parte del teorema por inducción sobre n.

Para $n = 2$ este resultado es obvio, ya que en este caso:

$$\begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix} = a_{11}a_{22},$$

de donde $a_{12}a_{21} = 0$.

Supongamos que la afirmación es cierta para matrices de orden menor que n. Así como en la demostración del teorema 2.26, podemos suponer que $p > 1$ sin pérdida de generalidad.

Dada cualquier n, mostraremos que si

$$A\begin{pmatrix} 1 & \dots & n \\ 1 & \dots & n \end{pmatrix} = A\begin{pmatrix} 1 & \dots & p \\ 1 & \dots & p \end{pmatrix} A\begin{pmatrix} p+1 & \dots & n \\ p+1 & \dots & n \end{pmatrix} \neq 0, \tag{2.59}$$

entonces nos encontramos en el caso (ii). Al aplicar la desigualdad 2.47 a la matriz totalmente no-negativa $(a_{ik})_1^p$ tenemos la desigualdad

$$A\begin{pmatrix} 1 & \dots & p \\ 1 & \dots & p \end{pmatrix} \leq A\begin{pmatrix} 1 & \dots & p-1 \\ 1 & \dots & p-1 \end{pmatrix} a_{pp},$$

la cual junto con 2.59 implica que

$$A\begin{pmatrix} 1 & \dots & p-1 \\ 1 & \dots & p-1 \end{pmatrix} \neq 0. \tag{2.60}$$

Entonces podemos escribir la cadena de desigualdades 2.50 de la prueba del teorema 2.23. Por 2.59 tenemos igualdades a lo largo de esta cadena, en consecuencia,

$$D\begin{pmatrix} p & p+1 & \dots & n \\ p & p+1 & \dots & n \end{pmatrix} = d_{pp}D\begin{pmatrix} p+1 & \dots & n \\ p+1 & \dots & n \end{pmatrix} \neq 0. \tag{2.61}$$

Al aplicar la segunda parte del teorema al determinante $D\begin{pmatrix} p & p+1 & \dots & n \\ p & p+1 & \dots & n \end{pmatrix}$ obtenemos alguno de los dos sistemas de ecuaciones siguientes:

$$d_{ip} = 0 \qquad (i = p+1, \dots, n), \tag{2.62}$$

$$d_{pk} = 0 \qquad (k = p+1, \dots, n). \tag{2.63}$$

Si, por ejemplo,

$$d_{ip} = A\begin{pmatrix} 1 & \dots & p-1 & i \\ 1 & \dots & p-1 & p \end{pmatrix} = 0 \qquad (i = p+1, \dots, n).$$

Entonces por el lema 2.27 y 2.59 se sigue que

$$a_{ik} = 0 \qquad (i = p+1, \dots, n;\ k = 1, 2, \dots, p).$$

Del mismo modo, de 2.63 obtenemos:

$$a_{ik} = 0 \qquad (i = 1, 2, \dots, p;\ k = p+1, \dots, n)$$

Así, terminamos la demostración de este teorema. $\square$

Observación 2.29. En el teorema 2.28 la desigualdad 2.47 se estableció bajo la hipótesis de que los menores de la matriz A son no-negativos y en el teorema 1.24 de la sección 1.8 se supone que la matriz A es simétrica, pues ésta se conforma de los coeficientes de una forma cuadrática positiva $\sum_{i,k=1}^{n} a_{ik}x_i x_k$ (la desigualdad generalizada de Hadamard). Este par de condiciones no se siguen una de la otra. Sin embargo, veremos cómo la desigualdad 2.47, bajo cualquiera de estas condiciones, se sigue del siguiente teorema cuyas condiciones son más generales.

Teorema 2.30. *Si en una matriz $A = (a_{ik})_1^n$ todos los menores principales son positivos y el producto de dos menores casi-principales simétricos con respecto a la diagonal principal es no-negativo, entonces la desigualdad 2.47 ocurre para cualquier $p < n$.*

Demostración. Haremos la prueba del teorema por inducción sobre n.

Para $n = 2$ tenemos que

$$\begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix} = a_{11}a_{22} - a_{12}a_{21} \leq a_{11}a_{22}.$$

Supongamos, ahora, que $n > 2$ y que el teorema es válido para matrices de orden menor que n. Nuevamente podemos suponer sin pérdida de generalidad que $p > 1$. Consideremos el caso 1 de la demostración por inducción que realizamos en el teorema 2.26.

Ahí definimos la matriz $D = (d_{ik})_p^n$ que, por la identidad de Sylvester, satisface la igualdad

$$D\begin{pmatrix} i_1 & i_2 & \dots & i_q \\ k_1 & k_2 & \dots & k_q \end{pmatrix}$$

$$= A\begin{pmatrix} 1 & 2 & \dots & p-1 \\ 1 & 2 & \dots & p-1 \end{pmatrix}^{q-1} A\begin{pmatrix} 1 & 2 & \dots & p-1 & i_1 & i_2 & \dots & i_q \\ 1 & 2 & \dots & p-1 & k_1 & k_2 & \dots & k_q \end{pmatrix},$$

de donde se sigue que la matriz D satisface las condiciones del teorema 2.30. En consecuencia, tenemos la serie de desigualdades 2.50 de la prueba del teorema 2.26. Así, llegamos a la desigualdad 2.47. □

Veamos ahora que, en efecto, la validez de la desigualdad 2.47 en los casos que comprenden el teorema 1.24 y el teorema 2.28, respectivamente, se sigue del teorema anterior.

Si A es una matriz simétrica y la forma cuadrática $\sum_{i,k=1}^{n} a_{ik}x_i x_k$ es positiva, entonces por el teorema 1.20 y el corolario 1.21 tenemos que todos los menores principales de A son positivos. Como A es simétrica, tenemos que dos submatrices de A simétricas con respecto a la diagonal principal resultan ser transpuestas una de la otra, de aquí se sigue que también se satisface la segunda hipótesis del teorema 2.30 para esta matriz A. Luego, tenemos la desigualdad 2.47 en este caso. Si la forma $\sum_{i,k=1}^{n} a_{ik}x_i x_k$ es no-negativa simplemente, podemos considerarla como el límite de una sucesión de formas positivas:

$$\sum_{\substack{i,k=1\\ i\neq k}}^{n} a_{ik}x_i x_k + \sum_{i=1}^{n}\left(a_{ii} + \frac{1}{\eta}\right) x_i^2.$$

De este modo, al pasar al límite obtenemos 2.47 también en este caso.

Si la matriz A es totalmente no-negativa con determinante cero, entonces la desigualdad 2.47 se cumple porque en este caso su lado izquierdo es cero mientras que su lado derecho es no-negativo.

Si A es una matriz totalmente no-negativa y no es singular, uno puede aproximarse a ella mediante matrices totalmente positivas haciendo uso del inciso (ii) de la proposición 2.13:

$$A = \lim_{\sigma\to\infty} F_\sigma A,$$

donde F_σ es la matriz totalmente positiva definida en la página 91 de esta tesis. Las matrices totalmente positivas satisfacen las condiciones del teorema 2.30. Luego, para estas matrices también se cumple la desigualdad 2.47. En consecuencia, para las matrices totalmente no-negativas y no-singulares también se tiene 2.47.

Corolario 2.31. *Si las condiciones de alguno de los teoremas 2.26, 2.28 o 2.30 se satisfacen, entonces se tiene la siguiente desigualdad:*

$$\begin{aligned} &A\begin{pmatrix} 1 & \dots & p & p+1 & \dots & q & q+1 & \dots & n \\ 1 & \dots & p & p+1 & \dots & q & q+1 & \dots & n \end{pmatrix} A\begin{pmatrix} p+1 & \dots & q \\ p+1 & \dots & q \end{pmatrix} \\ &\leq A\begin{pmatrix} 1 & \dots & p & p+1 & \dots & q \\ 1 & \dots & p & p+1 & \dots & q \end{pmatrix} A\begin{pmatrix} p+1 & \dots & q & q+1 & \dots & n \\ p+1 & \dots & q & q+1 & \dots & n \end{pmatrix}. \end{aligned} \tag{2.64}$$

Demostración. Si $|A| = 0$, la desigualdad es obvia. Supongamos, entonces, que $|A| \neq 0$.

Consideremos la matriz $\tilde{A} = (A_{ik})_1^n$ dada por

$$A_{ik} = A\begin{pmatrix} 1 & \dots & i-1 & i+1 & \dots & n \\ 1 & \dots & k-1 & k+1 & \dots & n \end{pmatrix} \qquad (i, k = 1, 2, \dots, n). \tag{2.65}$$

Al aplicar la fórmula para los menores de esta matriz (Proposición 1.2) obtenemos:

$$\tilde{A}\begin{pmatrix} h_1 & \dots & h_p \\ l_1 & \dots & l_p \end{pmatrix} = |A|^{p-1} A\begin{pmatrix} h_1^{'} & \dots & h_{n-p}^{'} \\ l_1^{'} & \dots & l_{n-p}^{'} \end{pmatrix}, \tag{2.66}$$

donde $h_1^{'}, \dots, h_{n-p}^{'}$ y $l_1^{'}, \dots, l_{n-p}^{'}$ son los sistemas de índices complementarios para los sistemas de índices $h_1, \dots, h_p$ y $l_1, \dots, l_p$, respectivamente. De esta manera podemos concluir que para esta matriz $\tilde{A}$, del mismo modo que para A, se satisfacen las condiciones de alguno de los teoremas 2.26, 2.28 o 2.30. Estas condiciones se satisfacen también para la matriz con entradas A_{ik} $(i, k = 1, \dots, p; q+1, \dots, n)$. En consecuencia,

$$\tilde{A}\begin{pmatrix} 1 & \dots & p & q+1 & \dots & n \\ 1 & \dots & p & q+1 & \dots & n \end{pmatrix} \leq \tilde{A}\begin{pmatrix} 1 & \dots & p \\ 1 & \dots & p \end{pmatrix} \tilde{A}\begin{pmatrix} q+1 & \dots & n \\ q+1 & \dots & n \end{pmatrix}. \quad (2.67)$$

Al expresar los menores de $\tilde{A}$ en términos de los menores de A con 2.66, de 2.67 obtenemos

$$|A|^{n-(q-p)-1} A\begin{pmatrix} p+1 & \dots & q \\ p+1 & \dots & q \end{pmatrix} \leq |A|^{p-1}\begin{pmatrix} p+1 & \dots & n \\ p+1 & \dots & n \end{pmatrix} |A|^{n-q-1}\begin{pmatrix} 1 & \dots & q \\ 1 & \dots & q \end{pmatrix}.$$

Al dividir ambos lados de esta desigualdad por $|A|^{n+p-q-2}$ obtenemos 2.64. $\square$

2.6. Criterio para una matriz respecto a ser oscilatoria

En esta sección estableceremos un criterio con el cual es posible determinar si una matriz totalmente no-negativa es oscilatoria o no.

Lema 2.32. *Si $A = (a_{ik})_1^n$ es una matriz oscilatoria, entonces cualquier matriz truncada $(a_{ik})_p^q$ $(1 \leq p \leq q \leq n)$ también es oscilatoria.*

Demostración. Si probamos la afirmación para la matriz truncada $(a_{ik})_2^n$, obviamente la afirmación será válida para cualquier matriz truncada de la forma $(a_{ik})_p^n$. Puesto que al invertir el orden de los renglones y columnas de una matriz oscilatoria se tiene nuevamente una matriz oscilatoria, tenemos que la afirmación es también válida para matrices truncadas de la forma $(a_{ik})_1^q$ y, consecuentemente, para matrices truncadas de la forma general $(a_{ik})_p^q$.

Así pues, bastará con probar el lema para la matriz $B := (a_{ik})_2^n$.

Denotemos por χ el exponente de A, entonces la matriz A^χ es totalmente positiva. Veremos que la matriz B^χ también es totalmente positiva, es decir, veremos que para cualquier par de sistemas de índices

$$2 \leq \begin{matrix} i_1 & < & i_2 & < & \dots & < & i_p \\ k_1 & < & k_2 & < & \dots & < & k_p \end{matrix} \leq n,$$

el menor

$$\begin{aligned} B^{\chi} & \begin{pmatrix} i_1 & i_2 & \dots & i_p \\ k_1 & k_2 & \dots & k_p \end{pmatrix} \\ & = \sum_{\substack{2 \leq \alpha_1^{'} < \dots < \alpha_p^{'} \\ \dots \\ 2 \leq \alpha_1^{(\chi-1)} < \dots < \alpha_p^{(\chi-1)}}} A \begin{pmatrix} i_1 & i_2 & \dots & i_p \\ \alpha_1^{'} & \alpha_2^{'} & \dots & \alpha_p^{'} \end{pmatrix} A \begin{pmatrix} \alpha_1^{'} & \alpha_2^{'} & \dots & \alpha_p^{'} \\ \alpha_1^{''} & \alpha_2^{''} & \dots & \alpha_p^{''} \end{pmatrix} \\ & \qquad \dots A \begin{pmatrix} \alpha_1^{(\chi-1)} & \alpha_2^{(\chi-1)} & \dots & \alpha_p^{(\chi-1)} \\ k_1 & k_2 & \dots & k_p \end{pmatrix} \end{aligned} \tag{2.68}$$

es positivo. Con este propósito, consideremos el "correspondiente" menor de la matriz A^{χ}:

$$\begin{aligned} A^{\chi} & \begin{pmatrix} 1 & i_1 & i_2 & \dots & i_p \\ 1 & k_1 & k_2 & \dots & k_p \end{pmatrix} \\ & = \sum_{\substack{\alpha_0^{'} < \alpha_1^{'} < \dots < \alpha_p^{'} \\ \dots \\ \alpha_0^{(\chi-1)} < \alpha_1^{(\chi-1)} < \dots < \alpha_p^{(\chi-1)}}} A \begin{pmatrix} 1 & i_1 & \dots & i_p \\ \alpha_0^{'} & \alpha_1^{'} & \dots & \alpha_p^{'} \end{pmatrix} A \begin{pmatrix} \alpha_0^{'} & \alpha_1^{'} & \dots & \alpha_p^{'} \\ \alpha_0^{''} & \alpha_1^{''} & \dots & \alpha_p^{''} \end{pmatrix} \\ & \qquad \dots A \begin{pmatrix} \alpha_0^{(\chi-1)} & \alpha_1^{(\chi-1)} & \dots & \alpha_p^{(\chi-1)} \\ 1 & k_1 & \dots & k_p \end{pmatrix}. \end{aligned} \tag{2.69}$$

Como este menor es positivo, al menos un término en el lado derecho de 2.69 es positivo. Digamos que este término es el producto

$$A \begin{pmatrix} 1 & i_1 & \dots & i_p \\ \xi_0^{'} & \xi_1^{'} & \dots & \xi_p^{'} \end{pmatrix} A \begin{pmatrix} \xi_0^{'} & \xi_1^{'} & \dots & \xi_p^{'} \\ \xi_0^{''} & \xi_1^{''} & \dots & \xi_p^{''} \end{pmatrix} \dots A \begin{pmatrix} \xi_0^{(\chi-1)} & \xi_1^{(\chi-1)} & \dots & \xi_p^{(\chi-1)} \\ 1 & k_1 & \dots & k_p \end{pmatrix}.$$

Luego, cada uno de los menores que aparecen como factores en este producto son positivos, entonces por la desigualdad 2.47 de la sección 2.5, el producto

$$A\begin{pmatrix} i_1 & \dots & i_p \\ \xi_1' & \dots & \xi_p' \end{pmatrix} A\begin{pmatrix} \xi_1' & \dots & \xi_p' \\ \xi_1'' & \dots & \xi_p'' \end{pmatrix} \dots A\begin{pmatrix} \xi_1^{(\chi-1)} & \dots & \xi_p^{(\chi-1)} \\ k_1 & \dots & k_p \end{pmatrix}$$

es positivo. Pero este producto aparece como uno de los sumandos de la suma 2.68. En consecuencia, esta suma es positiva, tal como queríamos probar. □

Lema 2.33. *Si una matriz rectangular*

$$A = \begin{pmatrix} a_{11} & \dots & a_{1n} \\ \dots & \dots & \dots \\ a_{m1} & \dots & a_{mn} \end{pmatrix}$$

es totalmente no-negativa (la definición de este tipo de matriz es análoga al de una matriz cuadrada totalmente no-negativa) con p renglones $1 = i_1 < i_2 < \ldots < i_p = m$ linealmente dependientes tales que los primeros y los últimos $p-1$ son linealmente independientes (cada i_k está dado por la posición del renglón en A); entonces el rango de la matriz A es igual a $p-1$.

Demostración. Desde luego, uno puede restringirse sólo al caso $p \leq n+1$. De las condiciones del lema se sigue que existen números $\lambda_1, \lambda_2, \ldots, \lambda_{p-1}$ tales que

$$a_{i_p k} = \sum_{\nu=1}^{p-1} \lambda_\nu a_{i_\nu k} \qquad (k = 1, 2, \ldots, n). \tag{2.70}$$

Aquí $\lambda_1 \neq 0$, pues de otro modo, el sistema de igualdades 2.70 nos daría una dependencia lineal entre los renglones $i_2, \ldots, i_p$.

Supongamos que $i_h < j < i_{h+1}$. Luego, si $k_1 < k_2 < \cdots < k_p$ es un sistema arbitrario de índices y escogemos ciertos índices $k_1^0 < k_2^0 < \cdots < k_{p-1}^0$ de manera tal que

$$A\begin{pmatrix} i_1 & \dots & i_{p-1} \\ k_1^0 & \dots & k_{p-1}^0 \end{pmatrix} \neq 0 \tag{2.71}$$

(esto es posible debido a que los renglones $i_1, \dots, i_{p-1}$ son linealmente independientes por hipótesis), entonces

$$\begin{aligned} A\begin{pmatrix} i_2 & \dots & i_h & j & i_{h+1} & \dots & i_p \\ k_1 & \dots & \dots & \dots & \dots & \dots & k_p \end{pmatrix} &= \\ \sum_{\nu=1}^{p-1} \lambda_\nu A\begin{pmatrix} i_2 & \dots & i_h & j & i_{h+1} & \dots & i_\nu \\ k_1 & \dots & \dots & \dots & \dots & \dots & k_p \end{pmatrix} &= \\ (-1)^{p-1}\lambda_1 A\begin{pmatrix} i_1 & \dots & i_h & j & i_{h+1} & \dots & i_{p-1} \\ k_1 & \dots & \dots & \dots & \dots & \dots & k_p \end{pmatrix} & \end{aligned} \tag{2.72}$$

y

$$A\begin{pmatrix} i_2 & \dots & i_h & i_{h+1} & \dots & i_p \\ k_1^0 & \dots & \dots & \dots & \dots & k_{p-1}^0 \end{pmatrix} = (-1)^p \lambda_1 A\begin{pmatrix} i_1 & \dots & i_{p-1} \\ k_1^0 & \dots & k_{p-1}^0 \end{pmatrix}. \tag{2.73}$$

Puesto que todos los menores de A que aparecen tanto en 2.72 como en 2.73 son no-negativos, de 2.73 y 2.71 se sigue que

$$(-1)^p \lambda_1 > 0.$$

Entonces, por 2.72, podemos concluir que

$$A\begin{pmatrix} i_1 & \dots & i_h & j & i_{h+1} & \dots & i_{p-1} \\ k_1 & \dots & \dots & \dots & \dots & \dots & k_p \end{pmatrix} = 0 \quad \text{con} \quad 1 \leq k_1 < k_2 < \dots < k_p \leq n.$$

De aquí se sigue que el j-ésimo renglón es una combinación lineal de los renglones $i_1, i_2, \ldots, i_{p-1}$. Pero j es un índice arbitrario diferente de $i_1, i_2, \ldots, i_p$. Como el i_p-ésimo renglón también es una combinación lineal de los renglones $i_1, i_2, \ldots, i_{p-1}$ por hipótesis del lema, finalmente hallamos que el rango de la matriz A es $p-1$. Esto prueba el lema. □

Lema 2.34. *Si una matriz rectangular*

$$A\begin{pmatrix} a_{11} & \ldots & a_{1n} \\ \ldots & \ldots & \ldots \\ a_{m1} & \ldots & a_{mn} \end{pmatrix}$$

es totalmente no-negativa y alguno de sus menores

$$A\begin{pmatrix} i_1 & i_2 & \ldots & i_p \\ k_1 & k_2 & \ldots & k_p \end{pmatrix} = 0 \qquad \begin{pmatrix} i_1 = 1 < i_2 < \cdots < i_p = m \\ k_1 = 1 < k_2 < \cdots < k_p = n \end{pmatrix}$$

pero

$$A\begin{pmatrix} i_1 & i_2 & \ldots & i_{p-1} \\ k_1 & k_2 & \ldots & k_{p-1} \end{pmatrix} \neq 0 \qquad \text{y} \qquad A\begin{pmatrix} i_2 & i_3 & \ldots & i_p \\ k_2 & k_3 & \ldots & k_p \end{pmatrix} \neq 0,$$

entonces el rango de A es igual a $p-1$.

Demostración. Apliquemos dos veces el lema anterior. Primero a la matriz con entradas a_{ik} con $i = 1, 2, \ldots, m$ y $k = k_1, k_2, \ldots, k_p$, para obtener que las columnas de A con índices $k_1, k_2, \ldots, k_p$ son linealmente dependientes. Luego, al aplicar el mismo lema 2.33 a A^t, la matriz transpuesta de A, obtenemos que el rango de la matriz A^t es $p-1$. Esto termina la prueba del lema. □

Definición 2.35. *A un menor* $A\begin{pmatrix} i_1 & i_2 & \ldots & i_p \\ k_1 & k_2 & \ldots & k_p \end{pmatrix}$ *de la matriz* $A = (a_{ik})_1^n$ *lo llamaremos* cuasi-principal *si*

$$\left.\begin{array}{c} 1 \leq i_1, k_1 < i_2, k_2 < \cdots < i_p, k_p \leq n, \\ |i_1 - k_1| \leq 1, |i_2 - k_2| \leq 1, \ldots, |i_p - k_p| \leq 1. \end{array}\right\} \tag{2.74}$$

Nótese que dada de la similitud en las denominaciones dadas en esta definición y en la definición 2.25, se debe poner especial cuidado en no confundir los tipos de menores que se definen en cada caso.

Teorema 2.36. (Criterio para una matriz con respecto a ser oscilatoria). *Para que una matriz totalmente no-negativa $A = (a_{ik})_1^n$ sea oscilatoria es necesario y suficiente que se satisfagan dos condiciones, a saber:*

(i) A sea una matriz no-singular.

(ii) $a_{i,i+1} > 0$ *y* $a_{i+1,i} > 0 \qquad (i = 1, 2, \ldots, n-1)$.

Demostración. Si A es una matriz oscilatoria, la condición (i) se sigue de la proposición 2.18 (vi). La condición (ii) se sigue del lema 2.32. En efecto, por el lema 2.32 sabemos que la matriz

$$A_i = \begin{pmatrix} a_{ii} & a_{i,i+1} \\ a_{i+1,i} & a_{i+1,i+1} \end{pmatrix}$$

es oscilatoria. También sabemos que $a_{i+1,i} \geq 0$ y $a_{i,i+1} \geq 0$. Pero si $a_{i+1,i} = 0$, por ejemplo, entonces la entrada con los mismos subíndices en la potencia A_i^κ será igual a cero para cualquier κ. Esto contradice el hecho de que A_i sea una matriz oscilatoria.

Veamos ahora que las condiciones (i) y (ii) son suficientes.

Estableceremos primero que las condiciones (i) y (ii) implican que todos los menores cuasi-principales de A son positivos.

Los menores cuasi-principales de primer orden son los elementos de la forma $a_{i+1,i}$ y $a_{i,i+1}$ $(i = 1, 2, \ldots, n-1)$ junto con los elementos a_{ii} $(i = 1, 2, \ldots, n)$ de la diagonal principal. Por hipótesis tenemos que $a_{i+1,i} > 0$ y $a_{i,i+1} > 0$ $(i = 1, 2, \ldots, n-1)$. También los elementos de la diagonal principal son positivos, pues a estos se les puede aplicar el lema 2.32 al verlos como matrices truncadas $(a_{ik})_p^q$ con $p = q$.

Ahora supongamos que los menores cuasi-principales de orden $< p$ son positivos y probemos que los menores cuasi-principales de orden p también son positivos. Supongamos que sucede lo contrario, es decir, que se satisfacen las relaciones 2.74 pero

$$A\begin{pmatrix} i_1 & i_2 & \dots & i_p \\ k_1 & k_2 & \dots & k_p \end{pmatrix} = 0. \tag{2.75}$$

Por hipótesis de inducción:

$$A\begin{pmatrix} i_1 & \dots & i_{p-1} \\ k_1 & \dots & k_{p-1} \end{pmatrix} A\begin{pmatrix} i_2 & \dots & i_p \\ k_2 & \dots & k_p \end{pmatrix} \neq 0. \tag{2.76}$$

Así pues, podemos aplicar el lema 2.34 a la matriz rectangular de dimensiones $i_p \times k_p$ dada por

$$\text{entradas } a_{ik} \text{ con } i = i_1, i_1 + 1, \dots, i_p \text{ y } k = k_1, k_1 + 1, \dots, k_p. \tag{2.77}$$

Luego, el rango de esta matriz es $p - 1$.

Si $h := \text{máx}(i_1, k_1)$, por 2.74 tenemos que

$$i_1, k_1 \leq h < i_2, k_2 < \dots < i_p, k_p;$$

luego,

$$i_1, k_1 \leq h, h + p - 1 \leq i_p, k_p. \tag{2.78}$$

De modo tal que el determinante

$$A\begin{pmatrix} h & h+1 & \dots & h+p-1 \\ h & h+1 & \dots & h+p-1 \end{pmatrix} \tag{2.79}$$

es un menor de orden p de la matriz dada por 2.77; en consecuencia, es igual a cero. Luego, por la desigualdad 2.47 establecida en la sección 2.5, tenemos que

$$0 \leq A\begin{pmatrix} 1 & \dots & n \\ 1 & \dots & n \end{pmatrix} \leq$$
$$A\begin{pmatrix} 1 & \dots & h-1 \\ 1 & \dots & h-1 \end{pmatrix} A\begin{pmatrix} h & \dots & h+p-1 \\ h & \dots & h+p-1 \end{pmatrix} A\begin{pmatrix} h+p & \dots & n \\ h+p & \dots & n \end{pmatrix} = 0,$$

lo cual contradice la condición (i) del teorema.

Así pues, hemos demostrado que todos los menores cuasi-principales de A son positivos si se tienen (i) y (ii).

Demostremos, ahora, que la matriz $B = A^{n-1}$ es totalmente positiva. La positividad de los menores cuasi-principales y la ya conocida identidad derivada del teorema de Binet-Cauchy:

$$\begin{aligned} B\begin{pmatrix} i_1 & \dots & i_p \\ k_1 & \dots & k_p \end{pmatrix} = \sum A\begin{pmatrix} i_1 & \dots & i_p \\ \alpha_1^{'} & \dots & \alpha_p^{'} \end{pmatrix} \\ \times A\begin{pmatrix} \alpha_1^{'} & \dots & \alpha_p^{'} \\ \alpha_1^{''} & \dots & \alpha_p^{''} \end{pmatrix} \dots A\begin{pmatrix} \alpha_1^{(n-2)} & \dots & \alpha_p^{(n-2)} \\ k_1 & \dots & k_p \end{pmatrix}, \end{aligned} \tag{2.80}$$

implican que será suficiente con mostrar que dado cualquier par de sistemas de índices, $i_1 < i_2 \dots < i_p$ y $k_1 < k_2 \dots < k_p$, uno siempre puede hallar sistemas intermedios de índices: $\alpha_1^{'} < \alpha_2^{'} < \dots < \alpha_p^{'}$; $\alpha_1^{''} < \alpha_2^{''} < \dots < \alpha_p^{''}$; $\dots$; $\alpha_1^{(n-2)} < \alpha_2^{(n-2)} < \dots < \alpha_p^{(n-2)}$, tales que el sumando de 2.80 correspondiente a éstos se conforma sólo de menores cuasi-principales.

A continuación, haremos una construcción de estos sistemas intermedios de índices.

A cada índice i_ν pongámosle una etiqueta: el símbolo $+$ en caso de que $i_\nu < k_\nu$, el símbolo $-$ en caso de que $i_\nu > k_\nu$, y el símbolo 0 en caso de que $i_\nu = k_\nu$. De este modo, habremos dividido el sistema de índices $i_1, i_2, \dots, i_p$ en varios grupos sucesivos, cada uno de los cuales incluye sólo índices con la misma etiqueta y no hay dos grupos adyacentes con la misma etiqueta. De acuerdo con la etiqueta, cada uno de estos grupos será positivo, negativo o cero. En cada grupo positivo incrementaremos el último de los índices en una unidad, y en cada grupo negativo disminuiremos el primero de los índices en una unidad; al resto de los índices lo dejaremos sin modificación alguna. Al sistema resultante lo denotaremos por $\alpha_1', \alpha_2', \ldots, \alpha_p'$.

En general, para obtener el sistema de índices $\alpha_1^{(s)}, \alpha_2^{(s)}, \ldots, \alpha_p^{(s)}$ primero etiquetamos cada uno de los índices del sistema anterior $\alpha_1^{(s-1)}, \alpha_2^{(s-1)}, \ldots, \alpha_p^{(s-1)}$ ($s = 1, 2, \ldots; \alpha_1^{(0)} := i_1, \alpha_2^{(0)} := i_2, \ldots, \alpha_p^{(0)} := i_p$) de acuerdo con el procedimiento ya descrito, pero los agrupamos en base a la agrupación previamente establecida para los índices $i_1, i_2, \ldots, i_p$. Por conveniencia, seguiremos llamando positivo, negativo o cero a cada uno de estos grupos (obsérvese que entonces un grupo positivo o negativo aquí puede contener índices con etiqueta 0). Acto seguido, en cada grupo positivo agregaremos una unidad a cada índice que todavía conserve la etiqueta $+$ y que se encuentre entre los últimos s índices de este grupo; en cada grupo negativo agregaremos una unidad a cada índice que todavía conserve la etiqueta $-$ y que se encuentre entre los primeros s índices de este grupo; y dejaremos sin cambio alguno el resto de los índices.

Verificaremos ahora que los menores

$$A\begin{pmatrix} \alpha_1^{(s-1)} & \alpha_2^{(s-1)} & \ldots & \alpha_p^{(s-1)} \\ \alpha_1^{(s)} & \alpha_2^{(s)} & \ldots & \alpha_p^{(s)} \end{pmatrix} \qquad (s = 1, 2, \ldots) \tag{2.81}$$

son cuasi-principales.

El segundo renglón de las relaciones en 2.74 siempre se tiene para estos menores porque en cada paso de la construcción que acabamos de establecer sólo se agrega una unidad a ciertos índices. De modo tal que para mostrar que los menores 2.81 son cuasi-principales será suficiente verificar las relaciones

$$1 \leq \alpha_1^{(s-1)}, \alpha_1^{(s)} < \alpha_2^{(s-1)}, \alpha_2^{(s)} < \cdots < \alpha_p^{(s-1)}, \alpha_p^{(s)} \leq n \quad (s = 1, 2, \ldots). \tag{2.82}$$

Supongamos que $s = 1$. Consideraremos varios casos para formar la cadena de relaciones 2.82 con $s = 1$. Si $i_\nu < k_\nu$ e i_ν es el último de los índices en un grupo positivo, entonces

$$\alpha'_{\nu-1} \leq i_{\nu-1} < i_\nu,$$
$$\alpha'_\nu = i_\nu + 1 \leq k_\nu < k_{\nu+1} \leq i_{\nu+1};$$

y en caso de que $i_{\nu+1} = k_{\nu+1}$, tenemos que $\alpha'_{\nu+1} = i_{\nu+1}$. Luego,

$$i_{\nu-1}, \alpha'_{\nu-1} < i_\nu, \alpha'_\nu < i_{\nu+1}, \alpha'_{\nu+1}.$$

En caso de que $i_{\nu+1} > k_{\nu+1}$, se sigue que $\alpha'_\nu = i_\nu + 1 < i_{\nu+1} - 1 = \alpha'_{\nu+1}$, entonces nuevamente tenemos que

$$i_{\nu-1}, \alpha'_{\nu-1} < i_\nu, \alpha'_\nu < i_{\nu+1}, \alpha'_{\nu+1}.$$

De forma análoga puede verse que si $i_\nu > k_\nu$ e i_ν es el primero de los índices en un grupo negativo, entonces

$$i_{\nu-1}, \alpha'_{\nu-1} < i_\nu, \alpha'_\nu < i_{\nu+1}, \alpha'_{\nu+1}.$$

Observemos que para obtener las relaciones 2.82 cuando $s = 1$ sólo hace falta notar que si i_ν no es el último de los índices en un grupo positivo ni el primero de los índices en un grupo negativo y tampoco es adyacente a algún índice de estos tipos, entonces $\alpha'_{\nu-1} = i_{\nu-1}$, $\alpha'_\nu = i_\nu$ y $\alpha'_{\nu+1} = i_{\nu+1}$; con lo cual obtenemos

$$i_{\nu-1}, \alpha'_{\nu-1} < i_\nu, \alpha'_\nu < i_{\nu+1}, \alpha'_{\nu+1}.$$

Ahora supongamos que todos los menores 2.81 con $s < t$ ($t > 1$) son cuasi-principales, es decir, que se satisfacen las relaciones 2.82 para $s < t$, y mostremos que entonces los menores 2.81 con $s = t$ también son cuasi-principales.

Para esto consideremos un menor de la forma 2.81 con $s = t$. Observemos que para nuestro propósito bastará con mostrar lo siguiente:

a) Si los índices $\alpha^{(t-1)}_{\mu+1}, \alpha^{(t-1)}_{\mu+2}, \ldots, \alpha^{(t-1)}_{\mu+t}$ son los t últimos en un grupo positivo, entonces

$$\alpha^{(t-1)}_{\mu}, \alpha^{(t)}_{\mu} < \alpha^{(t-1)}_{\mu+1}, \alpha^{(t)}_{\mu+1} < \cdots < \alpha^{(t-1)}_{\mu+t}, \alpha^{(t)}_{\mu+t} < \alpha^{(t-1)}_{\mu+t+1}, \alpha^{(t)}_{\mu+t+1}. \tag{2.83}$$

b) Si los índices $\alpha^{(t-1)}_{\mu+1}, \alpha^{(t-1)}_{\mu+2}, \ldots, \alpha^{(t-1)}_{\mu+t}$ son los t primeros en un grupo negativo, entonces nuevamente tenemos las relaciones 2.83.

c) Si el índice $\alpha^{(t-1)}_{\nu}$ no es ninguno de los que aparecen en 2.83 ya sea en el caso a) o en el caso b), entonces

$$\alpha^{(t-1)}_{\nu-1}, \alpha^{(t)}_{\nu-1} < \alpha^{(t-1)}_{\nu}, \alpha^{(t)}_{\nu} < \alpha^{(t-1)}_{\nu+1}, \alpha^{(t)}_{\nu+1}. \tag{2.84}$$

Prueba de a). Si el índice $\alpha^{(t-1)}_{\mu+\lambda}$ $(\lambda = 1, 2, \ldots, t)$ tiene la etiqueta $+$ y el índice $\alpha^{(t-1)}_{\mu+\lambda-1}$ tiene la etiqueta $-$ o la etiqueta 0, entonces por un argumento similar al ya expuesto para el caso $s = 1$ tenemos:

$$\alpha^{(t-1)}_{\mu+\lambda-1}, \alpha^{(t)}_{\mu+\lambda-1} < \alpha^{(t-1)}_{\mu+\lambda}, \alpha^{(t)}_{\mu+\lambda}.$$

Si el índice $\alpha^{(t-1)}_{\mu+\lambda}$ $(\lambda = 1, 2, \ldots, t)$ tiene la etiqueta $+$ y el índice $\alpha^{(t-1)}_{\mu+\lambda+1}$ tiene la etiqueta $-$ o la etiqueta 0, nuevamente en base al argumento dado para el caso $s = 1$ podemos concluir que

$$\alpha^{(t-1)}_{\mu+\lambda}, \alpha^{(t)}_{\mu+\lambda} < \alpha^{(t-1)}_{\mu+\lambda+1}, \alpha^{(t)}_{\mu+\lambda+1}.$$

Si el índice $\alpha^{(t-1)}_{\mu+\lambda}$ $(\lambda = 1, 2, \ldots, t-1)$ tiene la etiqueta $+$ y el índice $\alpha^{(t-1)}_{\mu+\lambda+1}$ también tiene la etiqueta $+$ hay que considerar dos subcasos: cuando $\lambda = 1$ y cuando $\lambda \neq 1$. Supongamos primero que $\lambda = 1$. Entonces $\alpha^{(t-1)}_{\mu+\lambda} = \alpha^{(t-2)}_{\mu+\lambda}$ y $\alpha^{(t-1)}_{\mu+\lambda+1} = \alpha^{(t-2)}_{\mu+\lambda+1} + 1$, por construcción. También por construcción tenemos que $\alpha^{(t)}_{\mu+\lambda} = \alpha^{(t-1)}_{\mu+\lambda} + 1$ y que

$\alpha_{\mu+\lambda+1}^{(t)} = \alpha_{\mu+\lambda+1}^{(t-1)} + 1$. Por hipótesis de inducción también tenemos que $\alpha_{\mu+\lambda}^{(t-2)} < \alpha_{\mu+\lambda+1}^{(t-2)}$. Luego,

$$\alpha_{\mu+\lambda}^{(t-1)} < \alpha_{\mu+\lambda+1}^{(t-1)}, \alpha_{\mu+\lambda+1}^{(t)},$$

$$\alpha_{\mu+\lambda}^{(t)} = \alpha_{\mu+\lambda}^{(t-2)} + 1 < \alpha_{\mu+\lambda+1}^{(t-2)} + 1 = \alpha_{\mu+\lambda+1}^{(t-1)}$$

y

$$\alpha_{\mu+\lambda}^{(t)} = \alpha_{\mu+\lambda}^{(t-2)} + 1 < \alpha_{\mu+\lambda+1}^{(t-2)} + 2 = \alpha_{\mu+\lambda+1}^{(t)};$$

es decir,

$$\alpha_{\mu+\lambda}^{(t-1)}, \alpha_{\mu+\lambda}^{(t)} < \alpha_{\mu+\lambda+1}^{(t-1)}, \alpha_{\mu+\lambda+1}^{(t)}.$$

Supongamos ahora que $\lambda \neq 1$. Por construcción tenemos las siguientes igualdades: $\alpha_{\mu+\lambda}^{(t-1)} = \alpha_{\mu+\lambda}^{(t-2)} + 1$, $\alpha_{\mu+\lambda+1}^{(t-1)} = \alpha_{\mu+\lambda+1}^{(t-2)} + 1$, $\alpha_{\mu+\lambda}^{(t)} = \alpha_{\mu+\lambda}^{(t-1)} + 1$ y $\alpha_{\mu+\lambda+1}^{(t)} = \alpha_{\mu+\lambda+1}^{(t-1)} + 1$. Como también tenemos que $\alpha_{\mu+\lambda}^{(t-2)} < \alpha_{\mu+\lambda+1}^{(t-2)}$ por hipótesis de inducción, se sigue que

$$\alpha_{\mu+\lambda}^{(t-1)}, \alpha_{\mu+\lambda}^{(t)} < \alpha_{\mu+\lambda+1}^{(t-1)}, \alpha_{\mu+\lambda+1}^{(t)}.$$

Finalmente, si el índice $\alpha_{\mu+\lambda}^{(t-1)}$ $(\lambda = 1, 2, \ldots, t-1)$ tiene la etiqueta 0 y el índice $\alpha_{\mu+\lambda+1}^{(t-1)}$ también tiene la etiqueta 0, entonces $\alpha_{\mu+\lambda}^{(t-1)} = \alpha_{\mu+\lambda}^{(t)}$ y $\alpha_{\mu+\lambda+1}^{(t-1)} = \alpha_{\mu+\lambda+1}^{(t)}$. Dado que $\alpha_{\mu+\lambda}^{(t-1)} < \alpha_{\mu+\lambda+1}^{(t-1)}$ según la hipótesis de inducción, se sigue que

$$\alpha_{\mu+\lambda}^{(t-1)}, \alpha_{\mu+\lambda}^{(t)} < \alpha_{\mu+\lambda+1}^{(t-1)}, \alpha_{\mu+\lambda+1}^{(t)}.$$

Así pues, al recoger todos los subcasos considerados podemos concluir que se satisfacen las relaciones 2.83. Esto termina la prueba de a).

Para probar b) se sigue un procedimiento similar al de la prueba de a).

Prueba de c). Supongamos que el índice $\alpha_{\nu}^{(t-1)}$ no es del tipo de índices que aparecen en 2.83 en cualquiera de los casos a) o b). Entonces se satisfacen las igualdades $\alpha_{\nu-1}^{(t-1)} = \alpha_{\nu-1}^{(t)}$, $\alpha_{\nu}^{(t-1)} = \alpha_{\nu}^{(t)}$ y $\alpha_{\nu+1}^{(t-1)} = \alpha_{\nu+1}^{(t)}$. De modo tal que al hacer uso nuevamente de la hipótesis de inducción llegamos a las relaciones 2.84.

Con esto termina la prueba de c) y, consecuentemente, la prueba por inducción de que todos los menores 2.81 son cuasi-principales.

Para concluir la demostración del teorema, por último mostraremos que el $n-1$-ésimo de los menores 2.81 (i.e., el menor con $s = n-1$) tiene en su segunda fila un sistema de índices que coincide con $k_1, k_2, \ldots, k_p$, es decir,

$$\alpha_{\nu}^{(n-1)} = k_{\nu} \qquad (\nu = 1, 2, \ldots, p). \tag{2.85}$$

Para esto observemos primero que

$$\nu \leq i_{\nu} \quad \text{y} \quad k_{\nu} \leq n - p + \nu \qquad (\nu = 1, 2, \ldots, p), \tag{2.86}$$

de donde se sigue que

$$|i_{\nu} - k_{\nu}| \leq n - p \qquad (\nu = 1, 2, \ldots, p). \tag{2.87}$$

De nuestra regla para la construcción de los sistemas de índices que requerimos se sigue que si $i_{\nu} \neq k_{\nu}$, al comienzo los índices $\alpha_{\nu}^{(0)} := i_{\nu}, \alpha_{\nu}^{'}, \alpha_{\nu}^{''}, \ldots$ se mantienen con el mismo valor i_{ν} hasta cierto paso que no rebasa al $p-1$-ésimo, y a partir de ahí, en cada transición, se acercan cada vez más a k_{ν} en una unidad. Como hay un total de $n-1$ pasos en la construcción de sistemas de índices que realizamos, por 2.87 tenemos que $a_{\nu}^{(t)} = k_{\nu}$ para algún $t \leq n-1$, lo cual implica que $\alpha_{\nu}^{(n-1)} = k_{\nu}$.

Esto termina la prueba del teorema. □

En la demostración del teorema anterior se encuentran implícitos los dos corolarios siguientes.

Corolario 2.37. *El exponente de una matriz oscilatoria siempre es menor o igual que* $n-1$.

Corolario 2.38. *En una matriz oscilatoria todos los menores cuasi-principales (en particular, todos los principales) son positivos.*

Adicionalmente, tenemos los corolarios siguientes también del teorema 2.36.

Corolario 2.39. *El producto* $A = A_1 A_2 \cdots A_m$ $(m \geq n-1)$ *de matrices oscilatorias* A_i $(i = 1, 2, \ldots, m)$ *es totalmente positivo.*

Demostración. Basta con aplicar a la matriz A los argumentos que se dieron en la demostración del teorema 2.36 para la matriz B. □

Corolario 2.40. *El producto de dos matrices oscilatorias también es una matriz oscilatoria y, más aún, el exponente de este producto es* $\leq \left[\dfrac{n}{2}\right]$, *donde* n *es el orden de las matrices.*

Demostración. Sea $C = AB$. Luego,

$$C^{[n/2]} = \underbrace{ABAB\cdots AB}_{m}$$

con $m = 2\left[\dfrac{n}{2}\right] \geq n-1$. Entonces por el corolario 2.39 se sigue el resultado deseado. □

Corolario 2.41. *El producto de una matriz oscilatoria con una matriz no-singular totalmente no-negativa es una matriz oscilatoria.*

Demostración. Sea $C = AB$ con A una matriz oscilatoria y B una matriz totalmente no-negativa y no-singular. Entonces

$$|C| = |A||B| > 0. \tag{2.88}$$

Por otro lado,

$$c_{i,i+1} = \sum_{\nu=1}^{n} a_{i\nu} b_{\nu,i+1} \qquad (i = 1, 2, \ldots, n-1). \tag{2.89}$$

Aquí la suma del lado derecho consiste de términos no-negativos con uno de ellos positivo, a saber,

$$a_{i,i+1} b_{i+1,i+1} > 0;$$

porque $a_{i,i+1} > 0$ por la caracterización de las matrices oscilatorias (teorema 2.36), y $b_{i+1,i+1} > 0$ por las desigualdades

$$0 < |B| \leq b_{11} b_{22} \cdots b_{nn}.$$

En consecuencia,

$$c_{i,i+1} > 0 \qquad (i = 1, 2, \ldots, n-1). \tag{2.90}$$

De forma análoga, podemos concluir también que

$$c_{i+1,i} > 0 \qquad (i = 1, 2, \ldots, n-1). \tag{2.91}$$

Luego, haciendo uso del teorema 2.36 y de 2.88, 2.90 y 2.91, se sigue que C es una matriz oscilatoria. $\square$

Consideremos una matriz de Jacobi

$$J = \begin{pmatrix} a_1 & b_1 & 0 & 0 & \dots & 0 & 0 \\ c_1 & a_2 & b_2 & 0 & \dots & 0 & 0 \\ 0 & c_2 & a_3 & b_3 & \dots & 0 & 0 \\ \dots & \dots & \dots & \dots & \dots & \dots & \dots \\ \dots & \dots & \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & 0 & 0 & \dots & c_{n-1} & a_n \end{pmatrix}. \tag{2.92}$$

Verificaremos el siguiente

Teorema 2.42. *Las afirmaciones siguientes son válidas.*

(i) Para que una matriz de Jacobi 2.92 sea oscilatoria es necesario y suficiente que todos los números b y c sean positivos y que los menores principales sucesivos también sean positivos:

$$a_1 > 0, \quad \begin{vmatrix} a_1 & b_1 \\ c_1 & a_2 \end{vmatrix} > 0, \quad \begin{vmatrix} a_1 & b_1 & 0 \\ c_1 & a_2 & b_2 \\ 0 & c_2 & a_3 \end{vmatrix} > 0, \dots \tag{2.93}$$

(ii) El exponente de una matriz de Jacobi oscilatoria siempre es igual a $n-1$, donde n es el orden de la matriz.

Para ello primero estableceremos la siguiente fórmula para los menores de la matriz J:

a) Si

$$1 \leq \begin{matrix} i_1 < i_2 < \dots < i_p \\ k_1 < k_2 < \dots < k_p \end{matrix} \leq n \tag{2.94}$$

y

$$i_1 = k_1, i_2 = k_2, \ldots, i_{\nu_1} = k_{\nu_1}; i_{\nu_1+1} \neq k_{\nu 1+1}, \ldots, i_{\nu_2} \neq k_{\nu_2};$$

$$; i_{\nu_2+1} = k_{\nu_2+1}, \ldots, i_{\nu_3} = k_{\nu_3}; \ldots$$

, entonces

$$\begin{aligned} &J\begin{pmatrix} i_1 & i_2 & \ldots & i_p \\ k_1 & k_2 & \ldots & k_p \end{pmatrix} = \\ &\quad J\begin{pmatrix} i_1 & \ldots & i_{\nu_1} \\ k_1 & \ldots & k_{\nu_1} \end{pmatrix} J\begin{pmatrix} i_{\nu_1+1} \\ k_{\nu_1+1} \end{pmatrix} \ldots J\begin{pmatrix} i_{\nu_2} \\ k_{\nu_2} \end{pmatrix} J\begin{pmatrix} i_{\nu_2+1} & \ldots & i_{\nu_3} \\ k_{\nu_2+1} & \ldots & k_{\nu_3} \end{pmatrix} \ldots \end{aligned} \tag{2.95}$$

Para probar esta fórmula basta con mostrar que si $i_\nu \neq k_\nu$ y se satisface 2.94, tenemos que

$$J\begin{pmatrix} i_1 & i_2 & \ldots & i_p \\ k_1 & k_2 & \ldots & k_p \end{pmatrix} = J\begin{pmatrix} i_1 & \ldots & i_\nu \\ k_1 & \ldots & k_\nu \end{pmatrix} J\begin{pmatrix} i_{\nu+1} & \ldots & i_p \\ k_{\nu+1} & \ldots & k_p \end{pmatrix} \tag{2.96}$$

y que

$$J\begin{pmatrix} i_1 & i_2 & \ldots & i_p \\ k_1 & k_2 & \ldots & k_p \end{pmatrix} = J\begin{pmatrix} i_1 & \ldots & i_{\nu-1} \\ k_1 & \ldots & k_{\nu-1} \end{pmatrix} J\begin{pmatrix} i_\nu & \ldots & i_p \\ k_\nu & \ldots & k_p \end{pmatrix}. \tag{2.97}$$

Notemos que la segunda igualdad se sigue de la primera al invertir de forma simultánea el orden de los renglones y columnas en la matriz que subyace en el menor $J\begin{pmatrix} i_1 & i_2 & \ldots & i_p \\ k_1 & k_2 & \ldots & k_p \end{pmatrix}$. Observemos que si $i_\nu < k_\nu$, entonces tenemos que

$$a_{i_\lambda k_\mu} = 0 \qquad (\lambda = 1, 2, \ldots, \nu; \mu = \nu + 1, \ldots, p)$$

en la matriz $J = (a_{ik})_1^n$; y si $i_\nu > k_\nu$, tenemos que

$$a_{i_\lambda k_\mu} = 0 \qquad (\lambda = \nu + 1, \ldots, p; \mu = 1, 2, \ldots, \nu).$$

En cualquier caso, la expansión de Laplace del determinate

$$J \begin{pmatrix} i_1 & i_2 & \ldots & i_p \\ k_1 & k_2 & \ldots & k_p \end{pmatrix}$$

con respecto a las primeras ν columnas origina la igualdad 2.96. Así, hemos probado 2.95.

De la fórmula 2.95 se desprende la siguiente afirmación.

Proposición 2.43. *Para que una matriz de Jacobi 2.92 sea totalmente no-negativa es necesario y suficiente que todos sus menores principales y todos los números b y c sean no-negativos.*

En caso de que la matriz J sea no-singular y no-negativa, tenemos que la forma cuadrática $J_s(x, x)$ (ver página 80) es no-negativa, dado que los menores principales de la matriz simetrizada J_s y los de J son los mismos. Pero si adicionalmente el rango de J y por consiguiente el de J_s es igual a n, tenemos que la forma $J_s(x, x)$ es positiva. Luego, todos los menores principales sucesivos de J_s (los de J) son positivos. En sentido opuesto, si todos los menores principales sucesivos de la matriz J son positivos, entonces la forma $J_s(x, x)$ es positiva, lo cual implica que todos los menores principales de J_s son positivos, es decir, todos los menores principales de J. Así pues, tenemos la siguiente

Proposición 2.44. *Para que una matriz de Jacobi no-singular sea totalmente no-negativa es necesario y suficiente que todas las entradas b y c sean no-negativas y que todos los menores principales sucesivos 2.93 sean positivos.*

Demostración del teorema 2.42. La primera afirmación es inmediata de la caracterización de las matrices oscilatorias (teorema 2.36) y de la proposición anterior.

Para probar la afirmación (ii) observemos que si $J^\kappa = (h_{ik})_1^n$ y $\kappa < n-1$, entonces $h_{1n} = 0$. En consecuencia, J^κ no es totalmente positiva. Pero sabemos que el exponente de cualquier matriz oscilatoria es $\leq n-1$ por el corolario 2.37. En consecuencia, el exponente de J es $n-1$. $\square$

Bibliografía

[1] S. H. Friedberg, A. J. Insel, y L. E. Spence. *Linear Algebra.* Prentice Hall, 1989.

[2] G. Frobenius. über matrizen aus positiven elementen, 1. *Sitzungsber. Königl. Preuss. Akad. Wiss.*, págs. 471–476, 1908.

[3] G. Frobenius. über matrizen aus positiven elementen, 2. *Sitzungsber. Königl. Preuss. Akad. Wiss.*, págs. 514–518, 1909.

[4] G. Frobenius. Zur theorie der matrices. *Sitzungsber. Königl. Preuss. Akad. Wiss.*, págs. 456–477, 1912.

[5] F. R. Gantmacher. *Matrix Theory, vol. I.* Chelsea, 1977.

[6] F. R. Gantmacher y M. G. Krein. *Oscillation Matrices and Kernels and Small Vibrations of Mechanical Systems: Revised Edition.* AMS Chelsea Publishing, 2002.

[7] I. Gessel. Tournaments and vandermonde's determinant. *J. Graph Theory*, 3:305–307, 1979.

[8] C. G. J. Jacobi. Ueber eine neue auflösungsart der bei der methode der kleinsten quadraten vorkommenden lineare gleichungen. *Astr. Nachr.*, 22:297–306, 1845.

[9] V. Maz'ya y T. O. Shaposhnikova. *Jacques Hadamard: A Universal Mathematician.* AMS, 1999.

[10] O. Perron. Zur theorie der matrices. *Mathematische Annalen*, 64 (2):248–263, 1907.

[11] J. J. Sylvester y H. F. Baker. *The collected mathematical papers of James Joseph Sylvester, vol. I.* Cambridge, University press, 1904.