Divulgaciones Matemáticas Vol. 9 No. 1(2001), pp. 85–102

Matrices Circulantes

Circulant Matrices

Teodoro Lara (teodorolara@cantv.net)
Departamento de Física y Matemáticas
NURR - ULA, Trujillo, Venezuela.

Resumen

En este trabajo estudiamos las propiedades fundamentales de las matrices circulantes, algunas conocidas y otras no muy conocidas. Hacemos énfasis en las matrices circulantes por bloques, mencionando su utilidad para la controlabilidad de algunos sistemas de ecuaciones diferenciales.
Palabras y frases clave: matrices circulantes, producto tensorial, matriz de Vandermonde, matriz de Fourier.

Abstract

In this work we study the fundamental properties of circulant matrices, some of them known and others not so well known. We emphasize block circulant matrices, mentioning their usefulness for the controllability of some differential equations systems.
Key words and phrases: circulant matrices, tensor product, Vandermonde's matrix, Fourier's matrix.

1 Introducción

Las matrices circulantes son un tópico fascinante que ha venido siendo estudiado por algún tiempo. Recientemente estas matrices han sido usadas para estudiar la controlabilidad de un sistema de ecuaciones diferenciales con control [5]. Debido a su importancia algunos *software* de matemáticas han incorporado las matrices circulantes dentro de la lista de matrices especiales que contienen. Las matrices circulantes también han sido usadas para estudiar objetos geométricos, por ejemplo en [2, 3, 4].

Recibido 2000/10/10. Revisado 2001/04/25. Aceptado 2001/04/27.
MSC (2000): Primary 15-02; Secondary 15A57.

Este trabajo contiene resultados conocidos y estudiados por otros autores como también aportes, específicamente en el caso de matrices circulantes por bloques en los casos constante y no constante. La estructuración de este artículo es como sigue: en la Sección 1 estudiamos algunos preliminares del álgebra lineal, en particular producto tensorial de matrices y matrices de Fourier; la Sección 2 la dedicamos al estudio de las matrices circulantes en general, su definición y propiedades. Por último en la Sección 3 consideramos la generalización de las matrices circulantes, en particular matrices circulantes por bloques así como también matrices por bloques donde cada bloque es una matriz circulante.

2 Preliminares de Algebra Lineal

Definición 2.1. Sea A y B matrices de tamaño $m \times n$ y $p \times q$, respectivamente. El *producto de Kronecker*, o *producto tensorial*, o *producto directo* de A con B es una nueva matriz de tamaño $mp \times nq$ definida y denotada como

$$A \otimes B = \begin{pmatrix} a_{11}B & a_{12}B & \cdots & a_{1n}B \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1}B & a_{m2}B & \cdots & a_{mn}B \end{pmatrix}.$$

Es decir, el producto tensorial de matrices se define para matrices de cualquier tamaño.

Propiedades del Producto de Kronecker

(i) $(\alpha A) \otimes B = A \otimes (\alpha B) = \alpha(A \otimes B)$, α escalar.

(ii) $(A + B) \otimes C = (A \otimes C) + B \otimes C$ (siempre y cuando A y B sean del mismo tamaño).

(iii) $A \otimes (B + C) = (A \otimes B) + (A \otimes C)$ (siempre y cuando B y C sean del mismo tamaño).

(iv) $A \otimes (B \otimes C) = (A \otimes B) \otimes C$.

(v) $(A \otimes B)(C \otimes D) = (AC) \otimes (BD)$ (asumiendo que los productos que aparecen se pueden realizar).

(vi) $\overline{(A \otimes B)} = \bar{A} \otimes \bar{B}$ (donde la barra significa conjugada).

(vii) $(A \otimes B)^T = A^T \otimes B^T$; $(A \otimes B)^* = A^* \otimes B^*$ (T significa transpuesta y $*$ transpuesta conjugada).

(viii) $\operatorname{tr}(A \otimes B) = \operatorname{tr}(A)\operatorname{tr}(B)$ (A y B cuadradas de orden m y n respectivamente).

(ix) Si A y B son no singulares entonces
$$(A \otimes B)^{-1} = A^{-1} \otimes B^{-1}.$$

(x) Si A y B son como en (viii) entonces
$$\det(A \otimes B) = \det{}^n(A)\det{}^m B.$$

Consideremos ahora $n \in \mathbf{Z}_+$ fijo y definamos $w_n = \exp\left(\frac{2\pi i}{n}\right)$, es decir, w_n es la raíz n-sima "primitiva" de la unidad. Entonces las siguientes propiedades se verifican inmediatamente.

(a) $w_n^n = 1, \quad w_n \bar{w}_n = 1, \quad \bar{w}_n = w_n^{-1}$.

(b) $(\bar{w}_n)^k = w_n^{-k} = w_n^{n-k}, \quad \sum_{j=1}^{n} w_n^{j-1} = 0$.

Definición 2.2. La *matriz de Fourier* de orden n se define como aquella F tal que
$$F^* = \frac{1}{\sqrt{n}} \begin{pmatrix} 1 & 1 & 1 & \cdots & 1 \\ 1 & w_n & w_n^2 & \cdots & w_n^{n-1} \\ 1 & w_n^2 & w_n^4 & \cdots & w_n^{2(n-1)} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & w_n^{n-1} & w_n^{2(n-1)} & \cdots & w_n^{(n-1)(n-1)} \end{pmatrix}.$$

La sucesión $\{w_n^k\}_{k=1}^{+\infty}$ es periódica de período n, es decir, $w_n^{k+n} = w_n^k$ para todo $k \in \mathbf{Z}_+$; de tal manera que hay sólo n elementos diferentes en F. Más aún, para $1 \leq j \leq n-1$ se tiene
$$w_n^{j(n-1)} = w_n^{nj-j} = w_n^{-nj} = w_n^{n-j}.$$

De tal manera que la expresión de F^* dada en la Definición 2.2 puede escribirse también como
$$F^* = \frac{1}{\sqrt{n}} \begin{pmatrix} 1 & 1 & 1 & \cdots & 1 \\ 1 & w_n & w_n^2 & \cdots & w_n^{n-1} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & w_n^{n-1} & w_n^{n-2} & \cdots & w_n \end{pmatrix}.$$

Algunas propiedades de la matriz de Fourier F son

(1) F y F^* son simétricas, es decir, $F = F^T$, $F^* = \bar{F}$, $F = \bar{F}^*$.

(2) F es unitaria. En realidad, $FF^* = F^*F = I_n$, o mejor dicho, $F^* = F^{-1}$. En efecto, si $(FF^*)_{k\ell}$ es el $k\ell$-elemento del producto FF^* entonces se puede verificar que

$$(FF^*)_{k\ell} = \sum_{r=0}^{n-1} (w_n^{(\ell-k)})^r = \begin{cases} n \text{ si } \ell = k, \\ 0 \text{ si } \ell \neq k. \end{cases}$$

(3) Los autovalores de F son ± 1, $\pm i$ con sus apropiadas multiplicidades. Esto se deduce del hecho que F es una matriz unitaria y todo autovalor λ de una matriz unitaria satisface $|\lambda| = 1$.

Matrices de uso común en la teoría de matrices circulantes son

$$\Pi = \begin{pmatrix} 0 & 1 & \cdots & 0 \\ \vdots & \ddots & \ddots & \vdots \\ \vdots & \cdots & 0 & 1 \\ 1 & \cdots & \cdots & 0 \end{pmatrix} \quad \text{y} \quad \Gamma = \begin{pmatrix} 1 & 0 & \cdots & 0 & 0 & 0 \\ 0 & 0 & \cdots & 0 & 0 & 1 \\ 0 & 0 & \cdots & 0 & 1 & 0 \\ 0 & 0 & \cdots & 1 & 0 & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots & \vdots \\ 0 & 1 & \cdots & 0 & 0 & 0 \end{pmatrix},$$

ambas de orden n. Claramente estas matrices satisfacen

$$\Pi^n = I_n, \quad \Pi^T = \Pi^* = \Pi^{-1} = \Pi^{n-1}$$

y

$$\Gamma^2 = I, \quad \Gamma^* = \Gamma^T = \Gamma = \Gamma^{-1}.$$

Definición 2.3. Por *matriz de Vandermonde* $V(z_1, z_2, \ldots, z_n)$ entendemos la matriz de orden n

$$V(z_1, z_2, \ldots, z_n) = \begin{pmatrix} 1 & 1 & \cdots & 1 \\ z_1 & z_2 & \cdots & z_n \\ z_1^2 & z_2^2 & \cdots & z_n^2 \\ \vdots & \vdots & \ddots & \vdots \\ z_1^{n-1} & z_2^{n-1} & \cdots & z_n^{n-1} \end{pmatrix}.$$

Entonces $V(1, w_n, w_n^2, \ldots, w_n^{n-1}) = \sqrt{n}F^*$ y $V(1, \bar{w}_n, (\bar{w}_n)^2, \ldots, (\bar{w}_n)^{n-1}) = \sqrt{n}F$. Es posible probar por inducción que $\det V(z_1, \ldots, z_n) = \prod_{k \neq l}(z_k - z_l)$, luego si $z_k \neq z_l$ para $k \neq l$ la matriz de Vandermonde es no singular ([4]).

3 Matrices Circulantes

Definición 3.1. Dados $n \in \mathbf{Z}_+$ y $(c_1, c_2, \ldots, c_n)$, la matriz $n \times n$ definida como

$$\operatorname{circ}(c_1, c_2, \ldots, c_n) = \begin{pmatrix} c_1 & c_2 & c_3 & \cdots & c_n \\ c_n & c_1 & c_2 & \cdots & c_{n-1} \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ c_3 & \cdots & c_n & c_1 & c_2 \\ c_2 & \cdots & c_3 & c_n & c_1 \end{pmatrix}$$

es llamada *matriz circulante* de orden n.

Note que los elementos de cada fila de C son idénticos a los de la fila anterior, pero han sido movidos una posición hacia la derecha y luego enrollados. Es fácil ver que la suma de matrices circulantes es de nuevo una matriz circulante; lo mismo ocurre cuando se multiplica por un escalar. Note también que la matriz Π definida en la sección anterior es una matriz circulante, de hecho $\Pi = \operatorname{circ}(0, 1, 0, \ldots, 0)$.

Proposición 3.2. *Sea A una matriz de tamaño $n \times n$ y Π la matriz de orden n anteriormente definida. Entonces A es circulante si y sólo si $\Pi A = A\Pi$.*

Prueba. Sea $A = (a_{ij})$, entonces $\Pi A \Pi^* = \Pi A \Pi^{-1} = (a_{(i+1)(j+1)})$ (donde los índices son tomados módulo n) y esta matriz es igual a $A = (a_{ij})$ si y sólo si A es circulante. □

Es claro de esta proposición que A es circulante siempre y cuando A^* lo sea.

Notando que las potencias Π^k $(k = 0, \ldots, n-1)$ consisten en permutaciones de las filas de Π, se puede ver que

$$C = \operatorname{circ}(c_1, \ldots, c_n) = c_1 I + c_2 \Pi + \cdots + c_n \Pi^{n-1}$$

y si

$$P(Z) = c_1 + c_2 Z + \cdots + c_n Z^{n-1}$$

entonces $\operatorname{circ}(c_1, \ldots, c_n) = P(\Pi)$. El polinomio P dado arriba es llamado *representante* de la circulante C y algunas veces se denota como P_C.

En algunos casos la función $\phi(\theta) = c_1 + c_2 e^{i\theta} + \cdots + c_n e^{i(n-1)\theta}$ también es útil como representante de $\operatorname{circ}(c_1, c_2, \ldots, c_n)$.

Propiedades de las matrices circulantes

(1) Como polinomios en la misma matriz conmutan (respecto al producto), entonces el producto de matrices circulantes conmuta. Más aún, el producto de circulantes es de nuevo una circulante. En efecto, si $C = P_C(\Pi)$ y $D = P_D(\Pi)$, entonces claramente $CD = P_C(\Pi)P_D(\Pi) = P_{CD}(\Pi)$.

(2) Como C y C^* conmutan entonces toda matriz circulante es normal.

(3) De (1) y (2) se obtiene que si A es circulante y $k \in \mathbf{Z}_+$ entonces A^k también es circulante.

Definición 3.3. La matriz cuadrada $T = (t_{ij})$, de orden n, se llama *matriz de Toeplitz* si $t_{ij} = t_{(i+1)\,(j+1)}$ para $i, j = 1, \ldots, n-1$.

Entonces una matriz de Toeplitz es aquella que es constante a lo largo de las diagonales paralelas a la diagonal principal.

Ejemplo 3.4.

$$T = \begin{pmatrix} a & b & c \\ d & a & b \\ e & d & a \end{pmatrix}.$$

Observación 3.5. Toda circulante es Toeplitz pero el recíproco no es cierto.

Una matriz circulante C de orden $n = pq$ es automáticamente circulante por bloques, y cada bloque es una matriz de Toeplitz. Los bloques son de orden q en un arreglo de $p \times q$ bloques.

Ejemplo 3.6. Consideremos una circulante de orden 6; la podemos dividir en 3×3 bloques de orden 2

$a\ b$ $f\ a$	$c\ d$ $b\ c$	$e\ f$ $d\ e$
$e\ f$ $d\ e$	$a\ b$ $f\ a$	$c\ d$ $b\ c$
$c\ d$ $b\ c$	$e\ f$ $d\ e$	$a\ b$ $f\ a$

ó

$$\begin{pmatrix} A & B & C \\ C & A & B \\ B & C & A \end{pmatrix}$$

circulante por bloques.

Note que un bloque circulante no necesariamente es una matriz circulante.

Proposición 3.7. *Si* $C = \mathrm{circ}(c_1, c_2, \ldots, c_n)$ *entonces* $\det C = \prod_{j=1}^{n} h(w_k)$, *donde* $h(x) = \sum_{j=1}^{n} c_j x^{j-1}$ *y* w_k *(*$k = 1, \ldots, n$*) son las diferentes raíces n-simas de la unidad.*

Prueba. Sea

$$V_n \equiv V(w_1, \ldots, w_n) = \begin{pmatrix} 1 & 1 & \cdots & 1 \\ w_1 & w_2 & \cdots & w_n \\ \vdots & \vdots & \ddots & \vdots \\ w_1^{n-1} & w_2^{n-1} & \cdots & w_n^{n-1} \end{pmatrix}$$

la matriz de Vandermonde de orden n en las n-raíces distintas de la unidad. Primero notamos que $CV_n = V_n \,\mathrm{diag}(h(w_k))_{k=1}^{n}$. Pero entonces

$$\det(CV_n) = \det(V_n) \det(\mathrm{diag}(h(w_k)_{k=1}^{n}))$$

y como $w_k \neq w_l$ si $l \neq k$, $\det V_n \neq 0$; por tanto $\det C = \prod_{k=1}^{n} h\,(w_k)$. □

Corolario 3.8. *Los autovalores y autovectores de* $C = \mathrm{circ}(c_1, \ldots, c_n)$ *están dados respectivamente por*

$$\lambda_k = h(w_k),\ V_{\lambda_k} = (1, w_k, \ldots, w_k^{n-1})^T \quad k = 1, \ldots, n.$$

Prueba. Lo que hay que observar es que

$$C = V_n \,\mathrm{diag}(h\,(w_k))_{k=1}^{n} V_n^{-1}.$$

□

Pero entonces hemos probado un resultado más general, a saber:

Teorema 3.9. *Toda matriz circulante es diagonalizable.*

Sean $C = \mathrm{circ}(c_1, c_2, \ldots, c_n)$ y $D = \mathrm{circ}(d_1, d_2, \ldots, d_n)$ matrices circulantes, $w_1, \ldots, w_n$ las distintas raíces n-simas de la unidad, V_n la matriz de Vandermonde dada más arriba y

$$h(w_k) = \sum_{j=1}^{n} c_j w_k^{j-1}, \quad q(w_k) = \sum_{j=1}^{n} d_j w_k^{j-1}.$$

Entonces

$$C = V_n \,\mathrm{diag}(h(w_k))_{k=1}^{n} V_n^{-1}, \quad D = V_n \,\mathrm{diag}(q(w_k))_{k=1}^{n} V_n^{-1}$$

y por tanto

$$CD = V_n \operatorname{diag}(h(w_k)q(w_k))_{k=1}^n V_n^{-1}.$$

Sean C y h como anteriormente, entonces $C = h(\Pi)$. Si $C \neq 0$ y asumiendo que los últimos coeficientes de h pudieran ser nulos, por ejemplo

$$h(x) = \sum_{j=1}^{r} c_j x^{j-1},\ 1 \leq r \leq n,\ c_r \neq 0$$

y si

$$h(x) = c_r \prod_{j=1}^{r-1} (x - \mu_j)$$

(es decir que $\mu_1, \mu_2, \ldots, \mu_{r-1}$ son los ceros de $h(x)$) entonces

$$C = h(\Pi) = c_r(\Pi - \mu_1 I)(\Pi - \mu_2 I)\cdots(\Pi - \mu_{r-1} I) = c_r \prod_{j=1}^{r-1} (\Pi - \mu_j I)$$

y tenemos así una factorización de cualquier circulante en un producto de circulantes $\Pi - \mu_k I$, que son de un tipo muy particular.

Supongamos que C es no singular, entonces por el Corolario 3.8, para todo $\lambda_k \in \sigma(C)$ (espectro de C)

$$\lambda_k = h(w_k) \neq 0, \quad k = 1, \ldots, n,$$

o equivalentemente

$$c_r \prod_{j=1}^{r-1} (w_k - \mu_j) \neq 0, \quad k = 1, \ldots, n.$$

Por tanto μ_j no es una raíz n-sima de la unidad y $\mu_j^n \neq 1$, para $j = 1, \ldots, r-1$. Como

$$C = h(\Pi) = c_r \prod_{j=1}^{r-1} (\Pi - \mu_j I)$$

se sigue

$$C^{-1} = c_r^{-1} \prod_{j=1}^{r-1} (\Pi - \mu_j I)^{-1}.$$

Es un hecho fácil de verificar que si $\mu \in \mathbf{C}$ y $\mu^n \neq 1$ entonces

$$(\Pi - \mu I)^{-1} = \frac{1}{1-\mu^n} \sum_{j=1}^{n} \mu^{j-1} \Pi^{n-j}.$$

Esto se deduce multiplicando el lado derecho por $\Pi - \mu I$ y usando que $\Pi^n = I$. Por cierto que la matriz $(\Pi - \mu I)^{-1}$ es llamada la *resolvente* de Π. Hay muchas propiedades de la resolvente que pueden ser obtenidas (ver [4] por ejemplo).

El determinante de $C = \operatorname{circ}(c_1, c_2, \ldots, c_n)$ no es, en general, trivial de calcular, pero por supuesto que usando la forma de C dada en la Proposición 3.7 es muy sencillo obtenerlo. Para dar una idea de cuán complicado es el cálculo de $\det C$ hagámoslo en algunos casos (por supuesto sin usar la forma diagonal).

$$\begin{aligned} n = 3; \quad \det(\operatorname{circ}(c_1, c_2, c_3)) &= c_1^3 + c_2^3 + c_3^3 - 3c_1c_2c_3. \\ n = 4; \quad \det(\operatorname{circ}(c_1, c_2, c_3, c_4)) &= c_1^4 - c_2^4 + c_3^4 - c_4^4 - 2c_1^2(c_3^2 + 2c_2c_4) \\ &\quad + 4c_1(c_2^2c_3 + c_3c_4^2) + 2c_2^2c_4^2 - 4c_2c_3^2c_4. \end{aligned}$$

Volviendo de nuevo a la expresión $C = V_n \operatorname{diag}(h(w_k))_{k=1}^n V_n^{-1}$ dada en el Corolario 3.8, si $\lambda_k = h(w_k)$, para $k = 1, \ldots, n$, y

$$\Lambda_k = \operatorname{diag}(0, 0, \ldots, 0, 1, 0, \ldots, 0),$$

donde el 1 aparece en la k-ésima posición ($k = 1, \ldots, n$), y si Λ es la matriz dada como

$$\Lambda = \operatorname{diag}(\lambda_k)_{k=1}^n = \sum_{k=1}^{n} \lambda_k \Lambda_k,$$

entonces

$$C = V_n \sum_{k=1}^{n} \lambda_k \Lambda_k V_n^{-1} = \sum_{k=1}^{n} \lambda_k V_n \Lambda_k V_n^{-1},$$

y si

$$D_k = V_n \Lambda_k V_n^{-1}, \quad k = 1, \ldots, n,$$

obtenemos

$$C = \sum_{k=1}^{n} \lambda_k D_k.$$

Las matrices D_k son llamadas *componentes* o *matrices idempotentes principales* de la matriz circulante C. Es inmediato ver que D_k, para $k = 1, \ldots, n$, es una matriz circulante y además

$$D_j D_k = (V_n \Lambda_j V_n^{-1})(V_n \Lambda_k V_n^{-1}) = V_n \Lambda_j \Lambda_k V_n^{-1} = 0 \text{ si } j \neq k.$$

Por otro lado

$$D_k^2 = V_n \Lambda_n^2 V_n^{-1} = V_n \Lambda_k V_n^{-1} = D_k.$$

Sea A una matriz y $m(\lambda)$ su polinomio mínimo, digamos

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

Se sabe que A es diagonalizable si y sólo si su polinomio mínimo tiene sólo ceros simples. Entonces si A es circulante

$$m(\lambda) = \prod_{i=1}^{s} (\lambda - \lambda_i), \quad m(A) = 0,$$

donde $m(\lambda)$ es el polinomio mónico de grado mínimo que tiene como ceros los autovalores de A.

Vamos ahora a generalizar la idea de matriz circulante constante considerando

$$C \equiv C(t) = \operatorname{circ}(c_1(t), c_2(t), \ldots, c_n(t))$$

donde $c_j(t)$ para $j = 1, \ldots, n)$ son funciones de t. Entonces siguiendo con la notación de la Proposición 3.7 y el Corolario 3.8 tenemos la siguiente:

Proposición 3.10. *Para $C(t)$ como antes y $h(x)(t) = \sum_{j=1}^{n} c_j(t) x^{j-1}$ se tiene que los autovalores de $C(t)$ están dados por*

$$\lambda_k(t) = h(w_k(t)) = \sum_{j=1}^{n} c_j(t) w_k^{j-1}(t).$$

Más aún, $C(t)$ puede ser representada en la forma

$$C(t) = V_n \operatorname{diag}(\lambda_k(t))_{k=1}^{n} V_n^{-1}$$

y $V_n = V_n(w_1, \ldots, w_n)$. Si $C_j(t)$ es función diferenciable en t entonces

$$\frac{dC(t)}{dt} = V_n \left(\frac{d\lambda_k(t)}{dt} \right)_{k=1}^{n} V_n^{-1}.$$

La prueba sigue esencialmente los mismos lineamientos de las pruebas de la Proposición 3.7 y el Corolario 3.8.

4 Generalización de Circulantes

Consideremos la forma cuadrática $Q(z) = Z^*QZ$ donde Q es una matriz circulante y $Z = (z_1, \dots, z_n)^T$. Este tipo de formas cuadráticas envolviendo matrices circulantes son usadas en geometría ([2]); damos aquí algunas de ellas.

(a) Si $Q = I$; entonces $Q(z) = z^*Iz = |z_1|^2 + \cdots + |z_n|^2$ y representa el momento polar de inercia alrededor de $z = 0$ de figuras geométricas.

(b) Si $Q = (I - \Pi)^*(I - \Pi)$ entonces

$$Q(z) = \sum_{k=1}^{n} |z_{k+1} - z_k|^2$$

y representa la suma de los cuadrados de los lados de un n-gono Z ([2, 1]).

(c) Si $Q = \frac{1}{4i}(\Pi - \Pi^*)$ entonces sus autovalores son

$$\frac{1}{2}\,\mathrm{sen}\left(\frac{(j-1)2\Pi}{n}\right), \quad j = 1, \dots, n,$$

$$Q(z) = \frac{1}{2}\sum_{j=1}^{n} \mathrm{sen}\left(\frac{(j-1)2\Pi}{n}\right)|\hat{Z}_j|^2,$$

donde $\hat{Z} = (\hat{Z}_1, \dots, \hat{Z}_n)$ y $\hat{Z} = FZ$, siendo F la matriz de Fourier de orden n.

Observación 4.1. Hay una generalización de las matrices circulantes llamadas g-circulantes; el lector interesado puede consultar [2] al respecto.

Definición 4.2. Sean $A_1, A_2, \dots, A_m$ matrices cuadradas, cada una de orden n. Por *bloque circulante* de tipo (m, n) y orden mn entendemos la matriz de tamaño $mn \times mn$

$$\mathrm{circ}(A_1, A_2, \dots, A_m) = \begin{pmatrix} A_1 & A_2 & \cdots & A_m \\ A_m & A_1 & \cdots & A_{m-1} \\ \vdots & \vdots & \ddots & \vdots \\ A_2 & A_3 & \cdots & A_1 \end{pmatrix}.$$

Es decir, son matrices circulantes donde cada elemento es una matriz.

Ejemplo 4.3. La matriz

$$\begin{pmatrix} a & b & e & f \\ c & d & g & h \\ e & f & a & b \\ g & h & c & d \end{pmatrix}$$

es un bloque circulante con $m = n = 2$. Si

$$A_1 = \begin{pmatrix} a & b \\ c & d \end{pmatrix}, \; A_2 = \begin{pmatrix} e & f \\ g & h \end{pmatrix}$$

entonces la matriz de arriba es $\mathrm{circ}(A_1, A_2)$.

Por supuesto ella no tiene que ser una matriz circulante en el sentido ordinario, claramente si $a \neq d$ esta matriz nunca será circulante.

Denotaremos el conjunto de matrices circulantes por bloques de tipo (m, n) por $\mathcal{BC}_{(m,n)}$.

Teorema 4.4. *$A \in \mathcal{BC}_{(m,n)}$ si y sólo si*

$$A(\Pi_m \otimes I_n) = (\Pi_m \otimes I_n)A.$$

Prueba. $\Pi_m \otimes I_n \in \mathcal{BC}_{(m,n)}$ y está dada por

$$\Pi_m \otimes I_n = \begin{pmatrix} O_n & I_n & O_n & \cdots & O_n \\ O_n & O_n & I_n & \cdots & O_n \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ O_n & O_n & O_n & \cdots & I_n \\ I_n & O_n & O_n & \cdots & O_n \end{pmatrix}_{mn \times mn} \qquad ;$$

por otro lado $A \in \mathcal{BC}_{(m,n)}$ si y sólo si

$$A = \begin{pmatrix} A_1 & A_2 & \cdots & A_m \\ A_m & A_1 & \cdots & A_{m-1} \\ \vdots & \vdots & \ddots & \vdots \\ A_2 & A_3 & \cdots & A_1 \end{pmatrix}$$

y se usa ahora la multiplicación formal (por bloques) de matrices para obtener la igualdad. □

Como

$$\begin{aligned} I_m \otimes A_1 &= \mathrm{diag}(A_1, A_1, \ldots, A_1), \\ \Pi_m \otimes A_2 &= \mathrm{circ}(0, A_2, \ldots, 0), \\ \Pi_m^2 \otimes A_3 &= \mathrm{circ}(0, 0, A_3, 0, \ldots, 0), \ldots \end{aligned}$$

podemos escribir $\mathrm{circ}(A_1, A_2, \ldots, A_m)$ como

$$\mathrm{circ}(A_1, A_2, \ldots, A_m) = \sum_{j=0}^{m-1} (\Pi_m^j \otimes A_{j+1}).$$

Ahora con esta expresión para $\mathrm{circ}(A_1, A_2, \ldots, A_m)$ es muy fácil probar, de otra forma, el Teorema 4.4.

Observación 4.5. Bloques circulantes del mismo tipo no necesariamente conmutan, por ejemplo

$$\begin{pmatrix} A & 0 \\ 0 & A \end{pmatrix} \begin{pmatrix} B & 0 \\ 0 & B \end{pmatrix} = \begin{pmatrix} AB & 0 \\ 0 & AB \end{pmatrix}$$

y

$$\begin{pmatrix} B & 0 \\ 0 & B \end{pmatrix} \begin{pmatrix} A & 0 \\ 0 & A \end{pmatrix} = \begin{pmatrix} BA & 0 \\ 0 & BA \end{pmatrix}$$

conmutan sólo si $AB = BA$.

En realidad esto es consecuencia de un hecho más general que probaremos ahora.

Teorema 4.6. *Sean $A = \mathrm{circ}(A_1, \ldots, A_m)$, $B = \mathrm{circ}(B_1, \ldots, B_m)$ en $\mathcal{BC}_{(m,n)}$. Si $A_j B_k = B_k A_j$ para $k, j = 1, \ldots, n$, entonces $AB = BA$.*

Prueba. Usamos la representación dada anteriormente

$$A = \sum_{j=0}^{m-1} \Pi_m^j \otimes A_{j+1}, \quad B = \sum_{k=0}^{m-1} \Pi_m^k \otimes B_{k+1}$$

luego

$$AB = \sum_{j,k=0}^{m-1} (\Pi_m^j \otimes A_{j+1})(\Pi_m^k \otimes B_{k+1}) = BA.$$

$\square$

Teorema 4.7. *$A \in \mathcal{BC}_{(m,n)}$ si y sólo si es de la forma*

$$A = (V_m \otimes V_n)\,\mathrm{diag}(M_1, M_2, \ldots, M_m)(V_m \otimes V_n)^{-1},$$

donde

$$V_n \equiv V_n(w_1, \ldots, w_n), \quad V_m \equiv V_m(\varepsilon_1, \ldots, \varepsilon_m),$$

y w_i, ε_j son las raíces n-simas y m-simas de la unidad, respectivamente, y $M_1, \ldots, M_m$ son matrices de orden n.

Prueba. Como $\Pi_m = \text{circ}(0, 1, 0, \ldots, 0)$ entonces usando resultados de la Proposición 3.7 y el Corolario 3.8 tenemos que

$$\Pi_m = V_m \,\text{diag}(\varepsilon_k)_{k=1}^m V_m^{-1} \text{ y } V_m \equiv V_m(\varepsilon_1, \ldots, \varepsilon_m).$$

Además $A \in \mathcal{BC}_{(m,n)}$ si y sólo si existen $A_1, \ldots, A_m$ de orden n tales que

$$A = (V_m \otimes V_n)(\sum_{j=0}^{m-1} \text{diag}(\varepsilon_k^j)_{k=1}^n \otimes B_j)(V_m \otimes V_n)^{-1},$$

con $B_j = V_n^{-1} A_{j+1} V_n$. Además

$$\sum_{j=0}^{m-1} \text{diag}(\varepsilon_n^j)_{k=1}^n \otimes B_j = \text{diag}(M_1, M_2, \ldots, B_n),$$

donde

$$\begin{aligned} M_1 &= B_0 + \varepsilon_1 B_1 + \cdots + \varepsilon_1^{m-1} B_m, \\ M_2 &= B_0 + \varepsilon_2 B_1 + \cdots + \varepsilon_2^{m-1} B_m, \ldots \\ M_m &= B_0 + \varepsilon_m B_1 + \cdots + \varepsilon_m^{m-1} B_m. \end{aligned}$$

Se completa así la prueba. □

Teorema 4.8. *Sean $A, B \in \mathcal{BC}_{(m,n)}$ y α_k escalares. Entonces A^T, A^*, $\alpha_1 A + \alpha_2 B$, AB, $p(A) = \sum_{k=0}^n \alpha_k A^k$ y A^{-1}, si existen, pertenecen a $\mathcal{BC}_{(m,n)}$.*

Definición 4.9. Sea A una matriz por bloques del tipo (m, n)

$$A = \begin{pmatrix} A_{11} & A_{12} & \cdots & A_{1m} \\ \vdots & \vdots & \ddots & \vdots \\ A_{m1} & A_{m2} & \cdots & A_{mm} \end{pmatrix},$$

cin $m \times m$ bloques de orden n cada uno. Si cada bloque A_{i_j} es a su vez una matriz circulante, diremos que A es una matriz con bloques circulantes y la clase de estas matrices la denotamos por $\mathcal{CB}_{(m,n)}$.

Vamos ahora a dar algunos resultados como los que fueron dados para $\mathcal{BC}_{(m,n)}$.

Teorema 4.10. *$A \in \mathcal{CB}_{(m,n)}$ si y sólo si $A(I_m \otimes \Pi_n) = (I_m \otimes \Pi_n)A$.*

Prueba. $I_m \otimes \Pi_n \in \mathcal{CB}_{(m,n)}$ y $I_m \otimes \Pi_n = \operatorname{diag}(\Pi_n, \Pi_n, \dots, \Pi_n)$. Entonces $A(I_m \otimes \Pi_n) = (A_{jk}\Pi_n)_{1\le d,k\le m}$ y $(I_m \otimes \Pi_n)A = (\Pi_n A_{jk})_{1\le j,k\le m}$. Así $A(I_n \otimes \Pi_n) = (\Pi_m \otimes \Pi_n)$ si y sólo si $A_{jk}\Pi_n = \Pi_n A_{jk}$, para $j,k = 1,\dots,n$ y A_{jk} es un bloque de orden n. Pero esta igualdad vale, de acuerdo con la Proposición 3.2, si y sólo si A_{j_k} es circulante, y como $1 \le j,k \le n$ son arbitrarios cada bloque es circulante. □

Teorema 4.11. *$A \in \mathcal{CB}_{(m,n)}$ si y sólo si A es de la forma $A = \sum_{k=0}^{n-1}(A_{k+1} \otimes \Pi_n k)$, donde A_{j+1} son matrices cuadradas de orden m.*

Prueba. $A = (A_{jk}) \in \mathcal{CB}_{m,n}$ si y sólo si para $1 \le j,k \le m$ se tiene

$$A_{jk} = \operatorname{circ}(a_1^{(jk)}, a_2^{(jk)}, \dots, a_n^{(jk)}) = \sum_{l=1}^{n} a_l^{(jk)} \Pi_n^{l-1}.$$

Entonces

$$A = \begin{pmatrix} a_1^{(11)} & I_n & \cdots & a_1^{(1m)} & I_n \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ a_1^{(m1)} & I_n & \cdots & a_1^{(mm)} & I_n \end{pmatrix} + \begin{pmatrix} a_2^{(11)} & \pi_n & \cdots & a_2^{(1m)} & \pi_n \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ a_2^{(m1)} & \pi_n & \cdots & a_2^{(mm)} & \pi_n \end{pmatrix} + \cdots$$

$$\cdots + \begin{pmatrix} a_m^{(11)} & \pi_n^{n-1} & \cdots & a_m^{(1m)} & \pi_n^{n-1} \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ a_m^{(m1)} & \pi_n^{n-1} & \cdots & a_m^{(mm)} & \pi_n^{n-1} \end{pmatrix}.$$

Si $A_1 = (a_1^{(jk)})_{1\le j,k\le m}$, $A_2 = (a_2^{(jk)})_{1\le j,k\le m}, \dots,$ $A_n = (a_n^{(jk)})_{1\le j,k\le m}$, entonces $A = A_1 \otimes I_n + A_2 \otimes \pi_n + \cdots + A_n \otimes \pi_n^{n-1}$. □

Note que cada matriz $A_1, \dots, A_n$ es de tamaño $m \times m$. Veamos ahora cómo podemos diagonalizar por bloques una matriz en $\mathcal{CB}_{(m,n)}$. Si $A \in \mathcal{CB}_{(m,n)}$, $A = (A_{jk})$, $1 \le j,k \le m$ y cada bloque A_{jk} es circulante, entonces cada bloque puede ser diagonalizado.

$$A_{jk} = V_n \operatorname{diag}(h_{jk}(w_k))_{\ell=1}^{n} V_n^{-1}$$

donde $V_n = V_n(w_1, \dots, w_n)$ y $h_{jk}(x) = \sum_{s=1}^{n} a_s^{(jk)} x^{s-1}$. Si hacemos

$$\lambda_\ell^{(jk)} = h_{jk}(w_\ell), \quad 1 \le j,k \le m. \quad 1 \le \ell \le n,$$

entonces

$$A_{jk} = V_n \operatorname{diag}(\lambda_\ell^{(jk)})_{\ell=1}^{n} V_n^{-1}.$$

En otras palabras,

$$\begin{aligned}
A &= \begin{pmatrix} V_n \quad \mathrm{diag}(\lambda_\ell^{(11)})_{\ell=1}^n \quad V_n^{-1} & \cdots & V_n \quad \mathrm{diag}(\lambda_\ell^{(1m)})_{\ell=1}^n \quad V_n^{-1} \\ \vdots & \ddots & \vdots \\ V_n \quad \mathrm{diag}(\lambda_\ell^{(m1)})_{\ell=1}^n \quad V_n^{-1} & \cdots & V_n \quad \mathrm{diag}(\lambda_\ell^{(mm)})_{\ell=1}^n \quad V_n^{-1} \end{pmatrix} \\
&\equiv \mathrm{diag}(V_n)_{\ell=1}^m \begin{pmatrix} \mathrm{diag}(\lambda_\ell^{(11)})_{\ell=1}^n & \cdots & \mathrm{diag}(\lambda_\ell^{(1m)})_{\ell=1}^n \\ \vdots & \ddots & \vdots \\ \mathrm{diag}(\lambda_\ell^{(m1)})_{\ell=1}^n & \cdots & \mathrm{diag}(\lambda_\ell^{(mm)})_{\ell=1}^n \end{pmatrix} \mathrm{diag}(V_n^{-1})_{\ell=1}^n \\
&= (I_m \otimes V_n)(H_{jk})_{1\le j,k\le m}(I_m \otimes V_n)^{-1},
\end{aligned}$$

donde

$$(H_{jk}) = \begin{pmatrix} \mathrm{diag}(\lambda_\ell^{(11)})_{\ell=1}^n & \cdots & \mathrm{diag}(\lambda_\ell^{(1m)})_{\ell=1}^n \\ \vdots & \ddots & \vdots \\ \mathrm{diag}(\lambda_\ell^{(m1)})_{\ell=1}^m & \cdots & \mathrm{diag}(\lambda_\ell^{(mm)})_{\ell=1}^n \end{pmatrix}.$$

Ahora combinemos las dos ideas que recientemente hemos expuesto. Sea A una matriz del tipo (m,n). Si A es circulante por bloques y cada bloque a su vez es circulante diremos que A es de clase $\mathcal{BCCB}_{(m,n)}$.

Ejemplo 4.12.

$$A = \begin{bmatrix} a\ b & c\ d & e\ f \\ b\ a & d\ c & f\ e \\ e\ f & a\ b & c\ d \\ f\ e & b\ a & d\ c \\ c\ d & e\ f & a\ b \\ d\ c & f\ e & b\ a \end{bmatrix} \in \mathcal{BCCB}_{(3,2)}.$$

Una matriz en $\mathcal{BCCB}_{(m,n)}$ no necesariamente es ella misma una matriz circulante.

Sea entonces $A = \mathrm{circ}(A_1,\dots,A_m) \in \mathcal{BCCB}_{(m,n)}$ y cada A_j, para $j = 1,\dots,m$, una matriz circulante de orden n; entonces, como ya se hizo antes,

$$A = \sum_{j=0}^{m-1} \Pi_m^j \otimes A_{j+1} \text{ y } A_{j+1} = V_n\,\mathrm{diag}(\lambda_k^{(j+1)})_{k=1}^n V_n^{-1},$$

donde $\lambda_k^{(j+1)} = h_{j+1}(w_k)$ y h_{j+1} es el polinomio asociado a los coeficientes de A_{j+1}; V_n y w_k $(k = 1,\dots,n)$ son definidos como antes. Por otro lado

$\Pi_m = \operatorname{circ}(0, 1, \ldots, 0)$, por tanto $\Pi_m = V_m \operatorname{diag}(\varepsilon_k)_{k=1}^m V_m^{-1}$ como en la prueba de la Proposición 4.7. Así pues

$$\begin{aligned} A &= \sum_{j=0}^{m-1} (V_m \operatorname{diag}(\varepsilon_k)_{k=1}^m V_m^{-1}) \otimes V_n \operatorname{diag}(\lambda_k^{(j+1)})_{k=1}^n \; x V_n^{-1} \\ &= \sum_{j=0}^{m-1} (V_m \otimes V_n) \operatorname{diag}(\varepsilon_k)_{k=1}^m \otimes \operatorname{diag}(\lambda_k^{(j+1)})_{k=1}^n (V_m \otimes V_n)^{-1} \\ &= (V_m \otimes V_n) \sum_{j=0}^{m} \operatorname{diag}(\varepsilon_k)_{k=1}^m \otimes \operatorname{diag}(\lambda_k^{(j+1)})_{k=1}^n (V_m \otimes V_n)^{-1}. \end{aligned}$$

Teorema 4.13. *Las matrices en $\mathcal{BCCB}_{(m,n)}$ son diagonalizables por la matriz $V_m \otimes V_n$, por tanto todas ellas conmutan. Si los autovalores de los bloques circulantes A_{j+1} $(j = 0, \ldots, m-1)$ están dados por $\{\lambda_k^{(j+1)} = h_{j+1}(w_k),\ k = 1, \ldots, n\}$, la matriz diagonal de autovalores de la matriz en $\mathcal{BCCB}_{(m,n)}$ está dada por*

$$\sum_{j=0}^{m-1} \operatorname{diag}(\varepsilon_k)_{k=1}^m \otimes \operatorname{diag}(\lambda_k^{(j+1)})_{k=1}^m.$$

Recíprocamente, cualquier matriz de la forma $A = (V_m \otimes V_n)\Lambda(V_m \otimes V_n)^{-1}$, donde Λ es una matriz diagonal, está en $\mathcal{BCCB}_{(m,n)}$.

Prueba. La primera parte de este Teorema es exactamente la discusión que se ha hecho ántes. De tal manera que lo único a verificar es el recíproco. Es decir, si $A = (V_m \otimes V_n)\Lambda(V_m \otimes V_n)^{-1}$ con Λ diagonal, entonces está en $\mathcal{BCCB}_{(m,n)}$. Esto es equivalente a demostrar que $A(\Pi_m \otimes I_n) = (\Pi_m \otimes I_n)A$ y $A(I_m \otimes \Pi_n) = (I_m \otimes \Pi_n)A$. Nosotros probaremos sólo la segunda de estas igualdades. Para demostrar la igualdad en cuestión primero verificaremos las siguientes.

(a) $(V_m^{-1} \otimes V_n^{-1})(I_m \otimes \Pi_n) = (I_m \otimes \Pi_n V_n)(V_m^{-1} \otimes V_m^{-1})$.

(b) $(I_m \otimes \Pi_n)(V_m \otimes V_n) = (V_m \otimes V_n)(I_m \otimes \Pi_n V_n)$.

Para hacer esto usamos el hecho de que $\Pi_n = V_n \operatorname{diag}(w_k)_{k=1}^n V_n^{-1}$, así

$$\begin{aligned} (V_m^{-1} \otimes V_n^{-1})(I_m \otimes \Pi_n) &= (V_m^{-1} \otimes V_n^{-1})(I_m \otimes V_n \operatorname{diag}(w_k)_{k=1}^n V_n^{-1}) \\ &= (I_m \otimes \Pi_n V_n)(V_m^{-1} \otimes V_n^{-1}) \end{aligned}$$

lo cual muestra (a); (b) se obtiene de manera parecida. Pero entonces

$$\begin{aligned} A(I_m \otimes \Pi_n) &= ((V_m \otimes V_n)\Lambda(V_m^{-1} \otimes V_n^{-1}))(I_m \otimes V_n \operatorname{diag}(w_k)_{k=1}^n V_n^{-1}) \\ &= (V_m \otimes V_n)\Lambda(I_m \otimes \Pi_n V_n)(V_m^{-1} \otimes V_n^{-1}) \\ &= (I_m \otimes \Pi_n)A. \end{aligned}$$

□

Teorema 4.14. *Si* $A \equiv A(t) \in \mathcal{BCCB}_{(m,n)}$*, para* t *en un intervalo* I*, y* $A(t) = \operatorname{circ}(A_1(t), \ldots, A_m(t))$*, entonces*

$$A(t) = (V_m \otimes V_n) \sum_{j=0}^{m-1} \operatorname{diag}(\varepsilon_k)_{k=1}^n \otimes \operatorname{diag}(\lambda_k^{(j+1)}(t))_{k=1}^n (V_m \otimes V_n)^{-1},$$

donde

$$\{\lambda_k^{(j+1)}(t) = h_{j+1}(w_k)(t),\ k = 1, \ldots, n\}$$

son los autovalores de los bloques circulantes $A_{j+1}(t)$*,* $j = 0, \ldots, m-1$*;* $h_{j+1}(.)(t)$ *es el polinomio asociado a los coeficientes de* $A_{j+1}(t)$*.*

La prueba de este resultado va similar a la del Teorema 4.13. Un caso interesante de estudiar es aquel donde los coeficientes de la matriz $A(t)$ en $\mathcal{BCCB}_{(m,n)}$ son múltiplos de t; es decir si $A(t) = (a_{ij}(t))$ en $\mathcal{BCCB}_{(m,n)}$; $a_{ij}(t) = \eta_{ij}t$, η_{ij} constantes, $1 \leq i, j \leq mn$.

Las matrices circulantes por bloques del tipo aquí estudiado aparecen en ciertos sistemas de ecuaciones diferenciales con control y la forma de controlar dichos sistemas es usando la particular manera de escribirlas como se ha hecho en el presente trabajo ([5]).

Referencias

[1] Bellman, R. *Introduction to Matrix Analysis*, 2nd ed., McGraw-Hill, New York, 1970.

[2] Davis, P. J. *Circulant Matrices*, John Wiley & sons, New York, 1979.

[3] Gilmore, R. *Lie Groups, Lie Algebras, and some of their Applications*, Wiley, New York, 1974.

[4] Lancaster, P., Tismenetsky, M. *The Theory of Matrices* 2nd ed,, Academic Press, San Diego, 1985.

[5] Lara, T. *Controllability and Applications of C. N. N.*, Ph. D. thesis, Georgia Institute of Technology, Atlanta - Georgia, December 1997.