1

RAÍCES ANALÍTICAS DE MATRICES INVERTIBLES CUYAS ENTRADAS SON FUNCIONES ANALÍTICAS

ANA ISABEL JULIO TORRES

PROYECTO DE GRADO PRESENTADO COMO REQUISITO PARA OPTAR AL TÍTULO DE MATEMÁTICO

UNIVERSIDAD DE CARTAGENA
FACULTAD DE CIENCIAS EXACTAS Y NATURALES
CARTAGENA
2007

RAÍCES ANALÍTICAS DE MATRICES INVERTIBLES CUYAS ENTRADAS SON FUNCIONES ANALÍTICAS

ANA ISABEL JULIO TORRES

PROYECTO DE GRADO PRESENTADO COMO REQUISITO PARA OPTAR AL TÍTULO DE MATEMÁTICO

ASESOR
PEDRO PABLO ORTEGA PALENCIA

UNIVERSIDAD DE CARTAGENA
FACULTAD DE CIENCIAS EXACTAS Y NATURALES
CARTAGENA
2007

AGRADECIMIENTOS

Dedico este trabajo a mis padres: A mi **madre querida**, *gracias por tu apoyo, confianza y comprensión.*
A mi **padre**, *que siempre ha sido el mejor de mis amigos, le doy gracias por estar siempre a mi lado apoyandome y brindandome lo mejor.*
A los dos muchas gracias por el esfuerzo que dia a dia realizan para complacer todas mis necesidades.

Le doy gracias a mi novio **Roberto Diaz** *por su apoyo incondicional, gracias por ser tan especial.*

Además, le doy gracias a todas aquellas personas que de alguna u otra forma hicieron que todo esto fuera posible.

Finalmente, agradezco mucho a mi asesor **Pedro Ortega** *por contar siempre con su ayuda para la realización de este trabajo.*

Índice

1. INTRODUCCIÓN

La teoría del álgebra lineal, más exactamente el espacio de las matrices, es quizás uno de las teorías más estudiadas en el campo de las matemáticas. Un ejemplo de esto es lo desarrollado en este trabajo, el cual consiste en estudiar una propiedad de las raíces m-ésimas de cualquier matriz dada, más precisamente se buscan condiciones bajo las cuales las raíces m-ésimas resultan ser matrices analíticas.

Inicialmente se definen los términos más significativos que se utilizarán a lo largo del trabajo y se desarrolla un ejemplo que enfoca el resultado del teorema principal dado en el capítulo 6 ; esto se lleva a cabo en el capítulo 2.

En el capítulo 3 se muestra una condición necesaria y suficiente para garantizar la analiticidad de las raíces cuadradas de cualquier matriz triangular superior invertible de orden 2×2 dada. Además, se establece un corolario que surge como consecuencia inmediata del teorema 3.1. El caso general en que la matriz dada sea de orden $n \times n$ requiere de algunos resultados preliminares; estos resultados son presentados en los capítulos 4 y 5.

En el capítulo 6 como ya se mencionó, se establece el teorema principal y se dan algunos resultados importantes de este trabajo.

Finalmente, se presenta un ejemplo que surge como consecuencia de algunos resultados dados en los capítulos 5 y 6.

Este resultado ha sido estudiado recientemente por los doctores **Leiba Rodman** e **Ilya Spitkovsky** en el artículo titulado **Analytic Roots of Invertible Matrix Functions**, Publicado en agosto de 2005 en la revista **Electronic Journal of Linear Algebra**.

2. DEFINICIONES PRINCIPALES Y EJEMPLO

A continuación se darán principalmente las definiciones fundamentales que se utilizarán a lo largo de este trabajo y se realizará un ejemplo que enfoca básicamente el problema a desarrollar.

Definición 2.1. *(Conjunto conexo)*
*Un conjunto se dice **conexo**, si dados dos puntos cualesquiera del conjunto estos pueden unirse mediante un arco que se encuentra totalmente dentro del conjunto.*

Definición 2.2. *(Dominio)*
*Un **dominio** de $\mathbb{C}$ es un conjunto abierto, conexo y no vacío de $\mathbb{C}$.*

Definición 2.3. *(Dominio simplemete conexo)*
*Un domino $\mathbb{G}$ se dice **simplemente conexo**, si cada punto interior de todo contorno cerrado que se encuentra completamente dentro de $\mathbb{G}$ pertenece también a $\mathbb{G}$. Intuitivamente se dice que un dominio que no tiene huecos es un dominio simplemente conexo.*

Definición 2.4. *(Función analítica)*
*Sea $\mathbb{G} \subseteq \mathbb{C}$ un dominio simplemente conexo y $f : \mathbb{G} \longrightarrow \mathbb{C}$ una función compleja. Sea $z_0 \in \mathbb{G}$, se dice que f es **analítica** en z_0, si existe una vecindad $\mathbf{V}$ de z_0 y una serie de potencias $\sum_{k=0}^{\infty} a_k(z-z_0)^k$, la cual converge en $\mathbf{V}$ tal que:*

$$f(z) = \sum_{k=0}^{\infty} a_k(z-z_0)^k, \quad z \in (\mathbf{V} \cap \mathbb{G}).$$

Definición 2.5. *(Singularidad)*
*Un punto z_0 se dice que es una **singularidad** de una función f, si f no es analítica en z_0.*

Definición 2.6. *(Polo de una función)*
*Una singularidad z_0 de una función f, se dice un **polo** de orden n para f si:*

$$\lim_{z \to z_0} (z-z_0)^n f(z) = M,$$

donde M es un número finito diferente de cero y n es un entero positivo.

Definición 2.7. *(Función meromorfa)*
*Sea $\mathbb{G} \subseteq \mathbb{C}$ un dominio simplemente conexo y $f : \mathbb{G} \longrightarrow \mathbb{C}$ una función compleja. Se dice que f es **meromorfa** en z_0, si f es analítica en z_0 ó si z_0 es un polo de f.*

Definición 2.8. *(Matriz de función analítica)*
*Una **matriz de función analítica** no es más que una matriz cuyas entradas son funciones analíticas.*

Definición 2.9. *(Matriz de función meromorfa)*
*Una **matriz de función meromorfa** no es más que una matriz cuyas entradas son funciones meromorfas.*

Definición 2.10. *(Raíces m-ésimas analíticas)*
*Las **raíces m-ésimas analíticas** de una matriz de función analítica $A(z)$ con $z \in \mathbb{G}$, son matrices de funciones $B(z)$, las cuales son analíticas en $\mathbb{G}$ y satisfacen la ecuación $B^m(z) = A(z)$, para todo $z \in \mathbb{G}$, $m \geq 2$, $m \in \mathbb{Z}$.*

Nota 2.1. *Sean $f : \mathbb{G} \longrightarrow \mathbb{C}$ y $g : \mathbb{G} \longrightarrow \mathbb{C}$ dos funciones analíticas. Recordemos que bajo estas hipótesis se tiene que:*

- *La suma, el producto y la multiplicación por escalar $f + g$, $f \cdot g$ y αf, con $\alpha \in \mathbb{C}$, respectivamente, son analíticos en $\mathbb{G}$.*

- *El cociente f/g es analítico en $\mathbb{G}$, excepto en los puntos z de $\mathbb{G}$ en los cuales $g(z)$ se anula.*

Nota 2.2. *En las definiciones anteriores siempre se considera $\mathbb{G}$ como un dominio simplemente conexo; en todo lo que sigue, $\mathbb{G}$ siempre denotará un dominio simplemente conexo ó un intervalo de la recta real.*

Se han considerado ya algunas de las definiciones principales que se usarán a lo largo de este trabajo. Ahora, se desarrollará un ejemplo para el caso de una matriz de función analítica $A(z)$ de orden 2×2.

Ejemplo 2.1. *Sea $\mathbb{G}$ un dominio simplemente conexo tal que $-1, 0 \in \mathbb{G}$ y $-\frac{1}{2} \notin \mathbb{G}$. Considérese la matriz de función analítica*

$$A(z) = \begin{bmatrix} (z + \frac{1}{2})^2 & 1 \\ 0 & \frac{1}{4} \end{bmatrix}, z \in \mathbb{G}. \tag{1}$$

Obsérvese claramente que $A(z)$ es invertible sobre $\mathbb{G}$, puesto que $-\frac{1}{2} \notin \mathbb{G}$, entonces $det(A(z)) \neq 0$ para todo $z \in \mathbb{G}$. Así, $A(z)$ será una matriz de función analítica e invertible.

Con este ejemplo se trata de proyectar a cualquier lector interesado en el tema lo que se desarrollará durante todo este trabajo, lo cual es:
Hallar raíces m-ésimas analíticas de una matriz $A(z)$ dada, la cual es analítica e invertible. En esta oportunidad se mostrará que para la matriz dada en (1) solo existen 4 raices cuadradas, pero ninguna es analítica; más precisamente son meromorficas.

En efecto: Sea $X(z) = \begin{bmatrix} x_{11}(z) & x_{12}(z) \\ x_{21}(z) & x_{22}(z) \end{bmatrix}, z \in \mathbb{G}$ tal que $X^2(z) = A(z)$. Entonces,

$$\begin{bmatrix} x_{11}^2(z)+x_{12}(z)x_{21}(z) & x_{12}(z)(x_{11}(z)+x_{22}(z)) \\ x_{21}(z)(x_{11}(z)+x_{22}(z)) & x_{22}^2(z)+x_{12}(z)x_{21}(z) \end{bmatrix} = \begin{bmatrix} (z+\frac{1}{2})^2 & 1 \\ 0 & \frac{1}{4} \end{bmatrix},$$

de donde se tiene el sistema de cuatro ecuaciones con cuatro incógnitas:

$$\begin{aligned} x_{11}^2(z)+x_{12}(z)x_{21}(z) &= \left(z+\frac{1}{2}\right)^2 \\ x_{12}(z)(x_{11}(z)+x_{22}(z)) &= 1 \\ x_{21}(z)(x_{11}(z)+x_{22}(z)) &= 0 \\ x_{22}^2(z)+x_{12}(z)x_{21}(z) &= \frac{1}{4}\ . \end{aligned}$$

Al resolver este sistema, se encuentra que:

$$\begin{aligned} x_{11}(z) &= \delta_1(z+\frac{1}{2}) \\ x_{12}(z) &= (\delta_1(z+\frac{1}{2})+\delta_2\frac{1}{2})^{-1} \\ x_{21}(z) &= 0 \\ x_{22}(z) &= \delta_2\frac{1}{2}, \end{aligned}$$

donde δ_1 y δ_2 son independientes de los signos ± 1. Por tanto,

$$X(z) = \begin{bmatrix} \delta_1(z+\frac{1}{2}) & (\delta_1(z+\frac{1}{2})+\delta_2\frac{1}{2})^{-1} \\ 0 & \delta_2\frac{1}{2} \end{bmatrix}. \qquad (2)$$

Al combinar los signos para δ_1 y δ_2 se encuentran las cuatro matrices siguientes:

$$X_1(z) = \begin{bmatrix} z+\frac{1}{2} & \frac{1}{z+1} \\ 0 & \frac{1}{2} \end{bmatrix}, \ \delta_1 = 1 = \delta_2$$

$$X_2(z) = \begin{bmatrix} -(z+\frac{1}{2}) & -\frac{1}{z+1} \\ 0 & -\frac{1}{2} \end{bmatrix}, \ \delta_1 = -1 = \delta_2$$

$$X_3(z) = \begin{bmatrix} z+\frac{1}{2} & \frac{1}{z} \\ 0 & -\frac{1}{2} \end{bmatrix}, \ \delta_1 = 1, \ \delta_2 = -1$$

$$X_4(z) = \begin{bmatrix} -(z+\frac{1}{2}) & -\frac{1}{z} \\ 0 & \frac{1}{2} \end{bmatrix}, \ \delta_1 = -1, \ \delta_2 = 1.$$

Teniendo en cuenta la definición 2.8, es fácil ver que ninguna de las cuatro matrices anteriormente dadas son analíticas, pues la entrada (1,2) de cada una de estas matrices no son analíticas; nótese que son meromorficas ya que 0 *y* -1 *son polos para las funciones* $\pm(\frac{1}{z})$ *y* $\pm(\frac{1}{z+1})$*, respectivamente.*

Nota 2.3. *El ejemplo anterior es de suma importancia, pues muestra que para la matriz dada en (1) existen 4 raíces cuadradas meromorficas; lo cual, será analizado en el teorema principal de este trabajo, el cual se establece en el capítulo 6 "* **teorema 6.1** *". Mas exactamente, en el inciso (b) de dicho teorema se justifica lo obtenido en el ejemplo 2.1.*

En la siguiente observación se mostrará que para puntos suficientemente cercanos a cualquier z_0 fijo del dominio $\mathbb{G}$, siempre existe una raíz m-ésima analítica para la matriz $A(z)$ de función analítica e invertible de orden $n \times n$ dada. En efecto:

Observación 1. *Sean* $z_0 \in \mathbb{G}$ *fijo y* $z \in \mathbb{G}$ *suficientemente cercano a* z_0*. Sea* $f : \mathbb{C} \longrightarrow \mathbb{C}$ *analítica en* z_0*. Luego,*

$$f(z) = \sum_{k=0}^{\infty} a_k (z - z_0)^k. \tag{3}$$

Sea $T(z)$ *una matriz que transforma a* $A(z)$ *en descomposición de jordan así:*

$$A(z) = T(z)J(z)T^{-1}(z),$$

con

$$J(z) = \begin{bmatrix} J_{m_1}(\lambda_1(z)) & & & \\ & J_{m_2}(\lambda_2(z)) & & 0 \\ 0 & & \ddots & \\ & & & J_{m_l}(\lambda_l(z)) \end{bmatrix}$$

$$= diag(J_{m_1}(\lambda_1(z)), \cdots, J_{m_l}(\lambda_l(z))),$$

donde $\lambda_1(z), \cdots, \lambda_l(z)$ son los autovalores de $A(z)$(no necesariamente distintos) y m_j es el tamaño del j-ésimo bloque de jordan asociado con $\lambda_j(z)$. Así:

$$J_{m_j}(\lambda_j(z)) = \begin{bmatrix} \lambda_j(z) & 1 & & & \\ & \lambda_j(z) & 1 & & 0 \\ & & \ddots & & \\ 0 & & & \lambda_j(z) & 1 \\ & & & & \lambda_j(z) \end{bmatrix}$$

$$:= \lambda_j(z)I + S_{m_j}\ , \qquad (4)$$

donde

$$S_{m_j} = \begin{bmatrix} 0 & 1 & & & \\ & 0 & 1 & 0 & \\ & & \ddots & & \\ & 0 & & 0 & 1 \\ & & & & 0 \end{bmatrix}$$

y $\sum_{j=1}^{l} m_j = n.$

Luego, por (4) y en virtud de la ***fórmula binomial de newton*** *se tiene:*

$$(J_{m_j}(\lambda_j(z)) - z_0 I)^k = \sum_{v=0}^{k} \binom{k}{v} (\lambda_j(z) - z_0)^{k-v} S_{m_j}^{v}. \qquad (5)$$

Nota 2.4. *Se puede verificar fácilmente que la potencia $S_{m_j}^{v} \cong 0$ siempre que $v \geq m_j$.*

Ahora bien, puesto que

$$f(J_{m_j}(\lambda_j(z))) = \sum_{k=0}^{\infty} a_k (J_{m_j}(\lambda_j(z)) - z_0 I)^k \qquad \text{en virtud de (3)}$$

$$= \sum_{k=0}^{\infty} a_k \sum_{v=0}^{k} \binom{k}{v} (\lambda_j(z) - z_0)^{k-v} S_{m_j}^v. \qquad \text{por (5)} \quad (6)$$

Ahora, para v y j fijos se tiene que:

$$\sum_{k=0}^{\infty} a_k \binom{k}{v} (\lambda_j(z) - z_0)^{k-v} S_{m_j}^v = \sum_{k=0}^{\infty} a_k \frac{k!}{(k-v)!v!} (\lambda_j(z) - z_0)^{k-v} S_{m_j}^v$$

$$= \sum_{k=0}^{\infty} a_k \frac{k(k-1)\cdots(k-v+1)}{v!} (\lambda_j(z) - z_0)^{k-v} S_{m_j}^v$$

$$= \frac{1}{v!} f^{(v)}(\lambda_j(z)).$$

Definición 2.11.
Sea $A(z)$ una matriz de función analítica e invertible de orden $n \times n$ sobre el dominio $\mathbf{G}$, con $\sigma(A(z)) = \{\lambda_1(z), \cdots, \lambda_l(z)\}$ y forma de jordan

$$J(z) = T^{-1}(z) A(z) T(z) = diag(J_{m_1}(\lambda_1(z)), \cdots, J_{m_l}(\lambda_l(z))).$$

Supóngase que la función $f : \mathbb{C} \longrightarrow \mathbb{C}$ es $m_j - 1$ veces diferenciable en $\lambda_j(z)$ para $j = 1, \cdots, l$. Entonces, la matriz de función analítica $f(A(z))$ está definida por

$$f(A(z)) = T(z) f(J(z)) T^{-1}(z),$$

donde $f(J(z)) = diag(f(J_{m_1}(\lambda_1(z))), \cdots, f(J_{m_l}(\lambda_l(z))))$, con

$$f(J_{m_j}(\lambda_j(z))) = \sum_{v=0}^{m_j - 1} \frac{1}{v!} f^{(v)}(\lambda_j(z)) S_{m_j}^v.$$

En virtud del teorema de los residuos, considerando $Ind_\Gamma(\lambda(z)) = 1$, para todo z suficientemente cercano a z_0 se tiene:

$$\frac{f^{(v)}(\lambda_j(z))}{v!} = \frac{1}{2\pi i} \int_\Gamma \frac{f(\lambda(z))}{(\lambda(z) - \lambda_j(z))^{v+1}} \, d\lambda(z), \qquad (7)$$

donde Γ es un contorno simple cerrado rectificable tal que todos los autovalores de $A(z_0)$ están en el interior de Γ y $\lambda(z) \neq \lambda_j(z)$ para todo z suficientemente cercano a z_0.

Ahora, sea $J_{m_j}(\lambda_j(z))$ un bloque de jordan asociado con $\lambda_j(z)$ y sea $\lambda(z) \neq \lambda_j(z)$ con

$$|\lambda(z) - \lambda_j(z)| > \max_{j=1,\cdots,l}\{m_j - 1\}\ , \quad \forall z \in \mathrm{G}.$$

Luego,

$$(\lambda(z)I - J_{m_j}(\lambda_j(z)))^{-1} = ((\lambda(z) - \lambda_j(z))I - S_{m_j})^{-1} \qquad \textit{por (4)}$$

$$= (\lambda(z) - \lambda_j(z))^{-1}\left(I - \frac{S_{m_j}}{\lambda(z) - \lambda_j(z)}\right)^{-1} \qquad \bigstar$$

Pero ahora bien,

$$\left\|\frac{S_{m_j}}{\lambda(z) - \lambda_j(z)}\right\| = \frac{\|S_{m_j}\|}{|\lambda(z) - \lambda_j(z)|}$$

$$= \frac{m_j - 1}{|\lambda(z) - \lambda_j(z)|}$$

$$\leq \frac{\max_{j=1,\cdots,l}\{m_j - 1\}}{|\lambda(z) - \lambda_j(z)|}$$

$$< \frac{|\lambda(z) - \lambda_j(z)|}{|\lambda(z) - \lambda_j(z)|} = 1.$$

Luego, en $\bigstar$ se tendrá que:

$$(\lambda(z)I - J_{m_j}(\lambda_j(z)))^{-1} = \frac{1}{\lambda(z) - \lambda_j(z)} \sum_{v=0}^{m_j-1} \left(\frac{1}{\lambda(z) - \lambda_j(z)}\, S_{m_j}\right)^v.$$

Esto último se deduce de la ***serie geométrica*** *y de la nota 2.4. Luego,*

$$\frac{1}{2\pi i}\int_\Gamma f(\lambda(z))(\lambda(z)I - J_{m_j}(\lambda_j(z)))^{-1}d\lambda(z) =$$

$$= \frac{1}{2\pi i}\int_\Gamma f(\lambda(z))\sum_{v=0}^{m_j-1}\frac{S^v_{m_j}}{(\lambda(z)-\lambda_j(z))^{v+1}}\,d\lambda(z)$$

$$= \sum_{v=0}^{m_j-1}\frac{1}{2\pi i}\int_\Gamma \frac{f(\lambda(z))}{(\lambda(z)-\lambda_j(z))^{v+1}}S^v_{m_j}\,d\lambda(z)$$

$$= \sum_{v=0}^{m_j-1}\frac{f^{(v)}(\lambda_j(z))}{v!}S^v_{m_j} \qquad \text{en virtud de (7)}$$

$$= f(J_{m_j}(\lambda_j(z))). \tag{8}$$

Esto último se tiene en virtud de la definición 2.11.
Ahora bien, por la misma definición se tiene:

$$f(A(z)) = T(z)f(J(z))T^{-1}(z)$$

$$= T(z)diag(f(J_{m_1}(\lambda_1(z))),\cdots,f(J_{m_l}(\lambda_l(z))))T^{-1}(z)$$

$$= T(z)\begin{bmatrix} f(J_{m_1}(\lambda_1(z))) & & 0 \\ & \ddots & \\ 0 & & f(J_{m_l}(\lambda_l(z))) \end{bmatrix}T^{-1}(z)$$

$$= \frac{1}{2\pi i}\int_\Gamma f(\lambda(z))\,T(z)\begin{bmatrix} (\lambda(z)I - J_{m_1}(\lambda_1(z)))^{-1} & & 0 \\ & \ddots & \\ 0 & & (\lambda(z)I - J_{m_l}(\lambda_l(z)))^{-1} \end{bmatrix}T^{-1}(z)\,d\lambda(z).$$

Esta última igualdad en virtud de (8).

$$= \frac{1}{2\pi i}\int_\Gamma f(\lambda(z))\, T(z) \begin{bmatrix} (\lambda(z)I - J_{m_1}(\lambda_1(z))) & & 0 \\ 0 & \ddots & \\ & & (\lambda(z)I - J_{m_l}(\lambda_l(z))) \end{bmatrix}^{-1} T^{-1}(z)\, d\lambda(z)$$

$$= \frac{1}{2\pi i}\int_\Gamma f(\lambda(z))\, T(z)\Big[\lambda(z)I - diag(J_{m_1}(\lambda_1(z)), \cdots, J_{m_l}(\lambda_l(z)))\Big]^{-1} T^{-1}(z)\, d\lambda(z)$$

$$= \frac{1}{2\pi i}\int_\Gamma f(\lambda(z))\, \Big(T(z)\Big[\lambda(z)I - diag(J_{m_1}(\lambda_1(z)), \cdots, J_{m_l}(\lambda_l(z)))\Big]T^{-1}(z)\Big)^{-1} d\lambda(z)$$

$$= \frac{1}{2\pi i}\int_\Gamma f(\lambda(z))\, \Big[\lambda(z)I - T(z)diag(J_{m_1}(\lambda_1(z)), \cdots, J_{m_l}(\lambda_l(z)))T^{-1}(z)\Big]^{-1} d\lambda(z)$$

$$= \frac{1}{2\pi i}\int_\Gamma f(\lambda(z))\, (\lambda(z)I - A(z))^{-1}\, d\lambda(z).$$

En resumen, se ha probado que para cualquier función analítica $f : \mathbb{C} \longrightarrow \mathbb{C}$ y una matriz $A(z)$ dada, se puede definir una matriz de función analítica así:

$$f(A(z)) = \frac{1}{2\pi i}\int_\Gamma f(\lambda(z))\, (\lambda(z)I - A(z))^{-1}\, d\lambda(z). \tag{9}$$

Por consiguiente, considerando a $f : \mathbb{C} \longrightarrow \mathbb{C}$ dada por $f(\lambda) = \lambda^{1/m}$ (la cual es analítica) en (9), se tiene que:

$$X(z) = A^{1/m}(z) = \frac{1}{2\pi i}\int_\Gamma \lambda^{1/m}(z)\, (\lambda(z)I - A(z))^{-1}\, d\lambda(z),$$

con lo cual, claramente $X^m(z) = A(z)$. Esto muestra que localmente, es decir en una vecindad de todo punto $z_0 \in \mathbf{G}$ dado, siempre existe una raíz m-ésima analítica para $A(z)$.

Esta observación es la base en la cual se fundamenta este trabajo, puesto que permite afirmar que el espacio solución de las raíces m-ésimas de una matriz $A(z)$ dada, en vecindades pequeñas de puntos fijados sobre el dominio $\mathbf{G}$ nunca es vacío.

3. MATRICES DE FUNCIONES 2×2

En el presente capítulo se establecen algunos resultados sobre matrices de funciones analíticas de orden 2×2 y sus raíces analíticas.

Definición 3.1. ***(Ceros mayorizados)***
Se dice que los ceros de una función analítica $a(z)$, con $z \in \mathbb{G}$, están ***mayorizados*** *por los ceros de una función analítica $b(z)$, con $z \in \mathbb{G}$, si todo z_0 cero de $a(z)$ es también cero de $b(z)$ y la multiplicidad de z_0 como un cero de $a(z)$ no excede a la multiplicidad de z_0 como un cero de $b(z)$; ó en otras palabras, si el cociente $(b(z)/a(z))$ es analítico en $\mathbb{G}$.*

Teorema 3.1.
Una matriz triangular superior de funciones analíticas de orden 2×2

$$A(z) = \begin{bmatrix} a_{11}(z) & a_{12}(z) \\ 0 & a_{22}(z) \end{bmatrix}, \quad z \in \mathbb{G} \tag{10}$$

con entradas diagonales no nulas en todo $\mathbb{G}$, admite una raíz cuadrada analítica si y solo si al menos una de las funciones $\sqrt{a_{11}(z)} \pm \sqrt{a_{22}(z)}$ tienen sus ceros en $\mathbb{G}$ mayorizados por los ceros de $a_{12}(z)$.

Nota 3.1. *Si en el enunciado del teorema se eligen los opuestos aditivos de $\sqrt{a_{11}(z)}$ y $\sqrt{a_{22}(z)}$, el teorema sigue siendo claramente cierto.*

Demostración. **(Teorema 3.1)**
Inicialmente se considera el caso cuando la entrada $a_{12}(z)$ es identicamente nula ($a_{12}(z) \cong 0$, para todo $z \in \mathbb{G}$); esto es cuando

$$A(z) = \begin{bmatrix} a_{11}(z) & 0 \\ 0 & a_{22}(z) \end{bmatrix}, \quad z \in \mathbb{G}.$$

En este caso, las matrices $X(z)$ tales que $X^2(z) = A(z)$ vienen dadas por:

$$X(z) = \begin{bmatrix} \pm\sqrt{a_{11}(z)} & 0 \\ 0 & \pm\sqrt{a_{22}(z)} \end{bmatrix}.$$

Pero ahora bien, La entrada (1,2) de la matriz $X(z)$ es exactamente $a_{12}(z)/(\sqrt{a_{11}(z)} \pm \sqrt{a_{22}(z)}) \cong 0$. Así que $X(z)$ obviamente es analítica puesto que $a_{12}(z)/(\sqrt{a_{11}(z)} \pm \sqrt{a_{22}(z)}) \cong 0$ también lo es y esto último es exactamente equivalente a decir que al menos una de las dos funciones $\sqrt{a_{11}} \pm \sqrt{a_{22}}$ tiene sus ceros mayorizados por los ceros de a_{12}.
Resta mostrar el caso en que $a_{12}(z)$ no es identicamente nula en $\mathbb{G}$.

En efecto: Sea $A(z) = \begin{bmatrix} a_{11}(z) & a_{12}(z) \\ 0 & a_{22}(z) \end{bmatrix}, \ z \in \mathbb{G}$ y supóngase que existe

$X(z) = \begin{bmatrix} x_{11}(z) & x_{12}(z) \\ x_{21}(z) & x_{22}(z) \end{bmatrix}$ tal que $X^2(z) = A(z)$. Entonces,

$$\begin{bmatrix} x_{11}^2(z) + x_{12}(z)x_{21}(z) & x_{12}(z)(x_{11}(z) + x_{22}(z)) \\ x_{21}(z)(x_{11}(z) + x_{22}(z)) & x_{22}^2(z) + x_{12}(z)x_{21}(z) \end{bmatrix} = \begin{bmatrix} a_{11}(z) & a_{12}(z) \\ 0 & a_{22}(z) \end{bmatrix},$$

Luego, se obtiene el sistema:

$$x_{11}^2(z) + x_{12}(z)x_{21}(z) = a_{11}(z) \tag{11}$$
$$x_{12}(z)(x_{11}(z) + x_{22}(z)) = a_{12}(z) \tag{12}$$
$$x_{21}(z)(x_{11}(z) + x_{22}(z)) = 0 \tag{13}$$
$$x_{22}^2(z) + x_{12}(z)x_{21}(z) = a_{22}(z). \tag{14}$$

Al resolver este sistema de cuatro ecuaciones con cuatro incógnitas se obtiene que $X(z)$ debe ser de la forma:

$$X(z) = \begin{bmatrix} \pm\sqrt{a_{11}(z)} & a_{12}(z)/\pm(\sqrt{a_{11}(z)}) + \pm(\sqrt{a_{22}(z)}) \\ 0 & \pm\sqrt{a_{22}(z)} \end{bmatrix}. \tag{15}$$

Ahora bien, puesto que $a_{12}(z)$ no es identicamente nulo en **G**, entonces en (13) se deduce que $x_{21}(z)$ debe ser identicamente nulo, en otras palabras, lo que se ha probado es que la raíz cuadrada analítica $X(z)$ de la matriz triangular $A(z)$ que se está considerando es también triangular superior como se muestra en (15).
Así las cosas, una raíz cuadrada analítica de $A(z)$ existe si y solo si por lo menos una de las ecuaciones:

$$x_{12}(z)(\sqrt{a_{11}(z)} \pm \sqrt{a_{22}(z)}) = a_{12}(z)$$

tiene solución analítica $x_{12}(z)$ en **G**, es decir, que el cociente

$$a_{12}(z)/(\sqrt{a_{11}(z)} \pm \sqrt{a_{22}(z)})$$

sea analítico en **G**. Esto último es equivalente a la condición de mayorización de ceros de la función $a_{12}(z)$ sobre las funciones $\sqrt{a_{11}(z)} \pm \sqrt{a_{22}(z)}$. ■

17

Corolario 3.2.
Sea $A(z)$ una matriz de función analítica de orden 2×2 sobre $\mathbb{G}$. Supóngase que los autovalores $\lambda_1(z)$, $\lambda_2(z)$ de $A(z)$ para todo $z \in \mathbb{G}$, pueden ser enumerados tal que $\lambda_1(z)$ y $\lambda_2(z)$ sean funciones analíticas en $\mathbb{G}$. Si por lo menos se tiene que una de las ramas de la función raíz cuadrada

$$Tr(A(z)) \pm 2\sqrt{det(A(z))} \neq 0,$$

para todo $z \in \mathbb{G}$, entonces existe una raíz cuadrada analítica de $A(z)$.

Nota 3.2. *En el capítulo 5 se verificará que: Dada una matriz de función analítica $A(z)$ de orden $n \times n$, las siguientes afirmaciones son equivalentes:*

(a) Los autovalores $\lambda_1(z), \cdots, \lambda_n(z)$ de $A(z)$ pueden ser enumerados tal que para todo $z \in \mathbb{G}$, $\lambda_1(z), \cdots, \lambda_n(z)$ son funciones analíticas.

(b) Existe una matriz de función analítica invertible $S(z)$ tal que $S(z)^{-1}A(z)S(z)$ es triangular superior para todo $z \in \mathbb{G}$.

Por consiguiente, no se pierde generalidad en suponer que $A(z)$ dada en las hipótesis del corolario a demostrar es triangular superior.

Demostración. **(Corolario 3.2)**
Teniendo en cuenta la nota 3.2, se puede asumir que:

$$A(z) = \begin{bmatrix} a_{11}(z) & a_{12}(z) \\ 0 & a_{22}(z) \end{bmatrix}.$$

Ahora, si por lo menos se tiene que una de las ramas de la función raíz cuadrada $Tr(A(z)) \pm 2\sqrt{det(A(z))} \neq 0$, para todo $z \in \mathbb{G}$; entonces,

$$a_{11}(z) + a_{22}(z) \pm 2\sqrt{a_{11}(z)a_{22}(z)} \neq 0,$$
$$(\sqrt{a_{11}(z)})^2 \pm 2\sqrt{a_{11}(z)a_{22}(z)} + (\sqrt{a_{22}(z)})^2 \neq 0,$$
$$(\sqrt{a_{11}(z)} \pm \sqrt{a_{22}(z)}\ )^2 \neq 0, \ \forall z \in \mathbb{G}.$$

En consecuencia, al menos una de las funciones $\sqrt{a_{11}(z)} \pm \sqrt{a_{22}(z)}$ será distinta de cero para todo $z \in \mathbb{G}$. Ahora bien, si $\sqrt{a_{11}(z)} + \sqrt{a_{22}(z)} \neq 0$, (el caso cuando $\sqrt{a_{11}(z)} - \sqrt{a_{22}(z)} \neq 0$ es análogo) como $a_{11}(z)$ y $a_{22}(z)$ para todo $z \in \mathbb{G}$ son funciones analíticas, entonces $\sqrt{a_{11}(z)} + \sqrt{a_{22}(z)}$ será analítico y así el cociente $a_{12}(z)/(\sqrt{a_{11}(z)} + \sqrt{a_{22}(z)})$ también será analítico puesto que $a_{12}(z)$ lo es. Pero recuérdese que decir que el cociente $a_{12}(z)/(\sqrt{a_{11}(z)} + \sqrt{a_{22}(z)})$ es analítico es equivalente a decir que los ceros de la función analítica $\sqrt{a_{11}(z)} + \sqrt{a_{22}(z)}$ están mayorizados por los ceros de la función analítica $a_{12}(z)$ y así por el teorema 3.1 se tendrá que $A(z)$ admite una raíz cuadrada analítica lo cual completa la prueba. ∎

Ejemplo 3.1. *Sea $A(z)$ una matriz de función analítica de orden 2×2 sobre un dominio $\mathbb{G}$, dada por:*

$$A(z) = \begin{bmatrix} z-1 & z^2 \\ 0 & z-1 \end{bmatrix}, \; z \in \mathbb{G} \text{ y } 1 \notin \mathbb{G}.$$

Obsérvese que para este caso $Tr(A(z)) - 2\sqrt{det(A(z))} = 0$ para todo $z \in \mathbb{G}$; por tal razón se verifica que $Tr(A(z)) + 2\sqrt{det(A(z))} \neq 0$ para todo $z \in \mathbb{G}$.
En efecto: Es fácil ver que

$$\begin{aligned} Tr(A(z)) + 2\sqrt{det(A(z))} &= 2z - 2 + 2\sqrt{(z-1)^2} \\ &= 4z - 4 \neq 0, \; \forall z \in \mathbb{G}, \end{aligned}$$

puesto que $1 \notin \mathbb{G}$. Así, por el corolario 3.2, existe una raíz cuadrada analítica de $A(z)$. Fácilmente se puede verificar que existen dos raíces cuadradas analíticas para $A(z)$ las cuales están dadas por:

$$X_1(z) = \begin{bmatrix} \sqrt{z-1} & \frac{z^2}{2\sqrt{z-1}} \\ 0 & \sqrt{z-1} \end{bmatrix},$$

$$X_2(z) = \begin{bmatrix} -\sqrt{z-1} & -\frac{z^2}{2\sqrt{z-1}} \\ 0 & -\sqrt{z-1} \end{bmatrix}.$$

4. PRELIMINARES ALGEBRAICOS

A continuación se prueban algunos resultados principales de manera general, considerando un campo $\mathbb{F}$ cualquiera; resultados que serán de mucha utilidad para abordar el caso de matrices de funciones analíticas de orden $n \times n$.

Definición 4.1. ***(Autovalor y Autovector)***
Sea $\mathbf{V}$ *un espacio vectorial sobre un campo* $\mathbf{F}$ *y sea* $\mathbf{S}$ *un subespacio de* $\mathbf{V}$. *Sea* $T : \mathbf{S} \longrightarrow \mathbf{V}$ *una aplicación lineal. Un escalar* $\lambda \in \mathbf{F}$ *se dice* ***autovalor*** *de* T, *si satisface que* $T(x) = \lambda x$. *Al elemento* x *se le dice* ***autovector*** *de* T *asociado al autovalor* λ.

Definición 4.2. ***(Espectro de una Aplicación Lineal)***
Sea $\mathbf{V}$ *un espacio vectorial sobre un campo* $\mathbf{F}$ *y sea* $\mathbf{S}$ *un subespacio de* $\mathbf{V}$. *Sea* $T : \mathbf{S} \longrightarrow \mathbf{V}$ *una aplicación lineal. se define el* ***espectro*** *de* T $(\sigma(T))$ *por:*

$$\sigma(T) = \{\lambda \in \mathbb{F} : T(x) = \lambda x\}.$$

A continuación se enuncia un resultado conocido como el **teorema de la aplicación espectral**, el cual será usado en el teorema 4.2.

Teorema 4.1. ***(Teorema de la aplicación espectral)***
Sea $\mathbf{F}$ *un campo y* A *una matriz de orden* $n \times n$ *sobre* $\mathbb{F}$. *Considérese* f *como una función polinomica sobre la clausura algebraica de* $\mathbb{F}$ $(\overline{\mathbb{F}})$. *Se define*

$$\sigma(f(A)) = \{f(\lambda) : \lambda \in \sigma(A)\},$$

donde $\sigma(A)$ *denota el espectro de* A *como en la definición 4.2. Entonces, se satisface que* $\sigma(f(A)) = f(\sigma(A))$.

Teorema 4.2.
Sea $\mathbf{F}$ *un campo y sea* $A = [a_{i,j}]_{i,j=1}^{n}$ *una matriz triangular superior de orden* $n \times n$ *sobre el campo* $\mathbf{F}$. *Supóngase que:*

$$rango(A - \lambda I) \geq n - 1, \ \forall \lambda \in \mathbf{F}, \tag{16}$$

entonces, toda matriz X *de orden* $n \times n$ *sobre* $\mathbf{F}$ *tal que* $X^m = A$ *para algún entero positivo* m, *debe ser triangular superior.*

Demostración.
Pasando a la clausura algebraica de $\mathbf{F}$, se puede asumir que $\mathbf{F}$ es algebraicamente cerrado. Ahora bien, sea X una matriz de orden $n \times n$ sobre $\mathbf{F}$ tal que $X^m = A$ para algún entero positivo m. Considérese la función f dada en el teorema de la aplicación espectral como $f(x) = x^m$ para todo $x \in \overline{\mathbf{F}}$. Luego, en virtud del teorema de la aplicación espectral se tiene que $f(\sigma(X)) = \sigma(f(X))$; esto es,

$$f(\sigma(X)) = \sigma(X^m) = \sigma(A).$$

2C

Ahora, se verifica que:

$$f(\sigma(X)) = \{\lambda^m : \lambda \in \sigma(X)\}.$$

En efecto:

$$\begin{aligned} \beta \in f(\sigma(X)) &\iff \exists \lambda \in \sigma(X) : \beta = f(\lambda) \\ &\iff \exists \lambda \in \sigma(X) : \beta = \lambda^m \\ &\iff \beta \in \{\lambda^m : \lambda \in \sigma(X)\}. \end{aligned}$$

En resumen.

$$f(\sigma(X)) = \{\lambda^m : \lambda \in \sigma(X)\} = \sigma(A).$$

Ahora, puesto que A es triangular superior, entonces a_{11} es un autovalor de A; por tanto, existe un $\lambda_0 \in \sigma(X)$ tal que $\lambda_0^m = a_{11}$. Ahora bien, si x es un autovector de X correspondiente al autovalor λ_0, entonces

$$\begin{aligned} Ax &= X^m x \\ &= X^{m-1}(Xx) \\ &= X^{m-1}(\lambda_0 x) \\ &= \lambda_0(X^{m-1}x) \\ &\vdots \\ &= \lambda_0^m x \\ &= a_{11}x. \end{aligned} \tag{17}$$

Esto muestra que x es un autovector de A asociado al autovalor a_{11}.
De (17) se tiene que $(A - a_{11}I)x = 0$. Puesto que $a_{11} \in \mathbf{F}$, entonces por hipótesis $rango(A - a_{11}I) \geq n - 1$, más precisamente $rango(A - a_{11}I) = n - 1$.
Ahora bien, $(A - a_{11}I)x = 0$ es exactamente

$$\begin{bmatrix} 0 & a_{12} & \cdots & a_{1n} \\ 0 & a_{22} - a_{11} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & a_{nn} - a_{11} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ \vdots \\ 0 \end{bmatrix},$$

de donde se tendrá que:

$$\sum_{i=2}^{n} a_{1i}x_i = 0,$$

lo cual implica que $x_i = 0$ para todo $i = 2, 3, \cdots, n$. Luego,

$$x = x_1 \begin{bmatrix} 1 & 0 & \cdots & 0 \end{bmatrix}^T.$$

Por otra parte, puesto que se ha dicho que x es un autovector de X asociado al autovalor λ_0, entonces, $Xx = \lambda_0 x$; esto es,

$$\begin{bmatrix} x_{11} & x_{12} & \cdots & x_{1n} \\ x_{21} & x_{22} & \cdots & x_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ x_{n1} & x_{n2} & \cdots & x_{nn} \end{bmatrix} \begin{bmatrix} x_1 \\ 0 \\ \vdots \\ 0 \end{bmatrix} = \begin{bmatrix} \lambda_0 x_1 \\ 0 \\ \vdots \\ 0 \end{bmatrix},$$

entonces,

$$x_1 \begin{bmatrix} x_{11} & x_{21} & \cdots & x_{n1} \end{bmatrix}^T = x_1 \begin{bmatrix} \lambda_0 & 0 & \cdots & 0 \end{bmatrix}^T,$$

de donde $x_{11} = \lambda_0$ y $x_{i1} = 0, \forall i = 2, 3, \cdots, n$.
Esto último muestra que la primera columna de X tiene todas sus entradas cero excepto posiblemente la entrada superior.

De igual forma como se realizó con a_{11}, se puede decir que:
Puesto que a_{22} es un autovalor de A, entonces, existe un $\lambda_1 \in \sigma(X)$ tal que $\lambda_1^m = a_{22}$. Ahora, si x es un autovector de X correspondiente al autovalor λ_1, entonces

$$Ax = X^m x = a_{22}x. \tag{18}$$

Esto muestra que x es un autovector de A asociado al autovalor a_{22}.
De (18) se tiene que $(A - a_{22}I)x = 0$. Puesto que $a_{22} \in \mathbf{F}$, entonces por hipótesis $rango(A - a_{22}I) \geq n - 1$.
Ahora bien, $(A - a_{22}I)x = 0$ es exactamente

$$\begin{bmatrix} a_{11} - a_{22} & a_{12} & a_{13} & \cdots & a_{1n} \\ 0 & 0 & a_{23} & \cdots & a_{2n} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & a_{nn} - a_{22} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ \vdots \\ 0 \end{bmatrix},$$

de donde

$$(a_{11} - a_{22})x_1 + \sum_{i=2}^{n} a_{1i}x_i = 0$$
$$\sum_{i=3}^{n} a_{2i}x_i = 0.$$

De la última igualdad se tiene que $x_i = 0\ \forall i = 3, 4, \cdots, n$. Por tanto,

$x = \begin{bmatrix} x_1 & x_2 & 0 & \cdots & 0 \end{bmatrix}^T$, donde x_1 y x_2 son escalares cualesquiera distintos de cero en el campo $\mathbf{F}$, los cuales satisfacen la ecuación:

$$(a_{11} - a_{22})x_1 + a_{12}x_2 = 0.$$

Ahora bien, puesto que x es un autovector de X asociado al autovalor λ_1, entonces, $Xx = \lambda_1 x$; esto es,

$$\begin{bmatrix} x_{11} & x_{12} & \cdots & x_{1n} \\ 0 & x_{22} & \cdots & x_{2n} \\ 0 & x_{32} & \cdots & x_{3n} \\ \vdots & \vdots & \ddots & \vdots \\ 0 & x_{n2} & \cdots & x_{nn} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ 0 \\ \vdots \\ 0 \end{bmatrix} = \lambda_1 \begin{bmatrix} x_1 \\ x_2 \\ 0 \\ \vdots \\ 0 \end{bmatrix},$$

de donde $x_{i2}x_2 = 0$, $\forall i = 3, 4, \cdots, n$. Así entonces, se deduce que $x_{i2} = 0$ $\forall i = 3, 4, \cdots, n$. Esto último muestra que la segunda columna de X tiene todas sus entradas cero excepto posiblemente la dos primeras entradas superiores.

Hasta este momento se ha mostrado que

$$X = \begin{bmatrix} x_{11} & x_{12} & \cdots & x_{1n-1} & x_{1n} \\ 0 & x_{22} & \cdots & x_{2n-1} & x_{2n} \\ 0 & 0 & \cdots & x_{3n-1} & x_{3n} \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \cdots & x_{n-1n-1} & x_{n-1n} \\ 0 & 0 & \cdots & x_{nn-1} & x_{nn} \end{bmatrix}.$$

Continuando con el proceso, como a_{n-1n-1} es un autovalor de A, existe un $\lambda_s \in \sigma(X)$ tal que $\lambda_s^m = a_{n-1n-1}$. Ahora, si x es un autovector de X correspondiente al autovalor λ_s, entonces

$$Ax = X^m x = \lambda_s^m x = a_{n-1n-1}x. \tag{19}$$

Esto muestra que x es un autovector de A asociado al autovalor a_{n-1n-1}.
De (19) se tiene que $(A - a_{n-1n-1}I)x = 0$.
Ahora bien, $(A - a_{n-1n-1}I)x = 0$ es exactamente

$$\begin{bmatrix} a_{11} - a_{n-1n-1} & a_{12} & \cdots & a_{1n-1} & a_{1n} \\ 0 & a_{22} - a_{n-1n-1} & \cdots & a_{2n-1} & a_{2n} \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \cdots & 0 & a_{n-1n} \\ 0 & 0 & \cdots & 0 & a_{nn} - a_{n-1n-1} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_{n-1} \\ x_n \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ \vdots \\ 0 \\ 0 \end{bmatrix},$$

de donde

$$\begin{cases} (a_{11} - a_{n-1n-1})x_1 + \sum\limits_{i=2}^{n} a_{1i}x_i = 0 \\ (a_{22} - a_{n-1n-1})x_2 + \sum\limits_{i=3}^{n} a_{2i}x_i = 0 \\ a_{n-1n}x_n = 0. \end{cases} \tag{20}$$

De la última igualdad se tiene que $x_n = 0$. Por tanto, todo autovector x asociado al autovalor λ_s es tal que

$$x = \begin{bmatrix} x_1 & x_2 & \cdots & x_{n-1} & 0 \end{bmatrix}^T,$$

donde las variables x_i, con $i = 1, 2, \cdots, n-1$ satisfacen el sistema (20).
Puesto que se ha dicho que x es un autovector de X asociado al autovalor λ_s, entonces $Xx = \lambda_s x$, esto es,

$$\begin{bmatrix} x_{11} & x_{12} & \cdots & x_{1n-1} & x_{1n} \\ 0 & x_{22} & \cdots & x_{2n-1} & x_{2n} \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \cdots & x_{n-1n-1} & x_{n-1n} \\ 0 & 0 & \cdots & x_{nn-1} & x_{nn} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_{n-1} \\ 0 \end{bmatrix} = \lambda_s \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_{n-1} \\ 0 \end{bmatrix},$$

de donde $x_{nn-1}x_{n-1} = 0$, lo cual implica que $x_{nn-1} = 0$. Así se completa la prueba, pues si X es tal que $X^m = A$, entonces se debe tener que X es de la forma

$$X = \begin{bmatrix} x_{11} & x_{12} & \cdots & x_{1n-1} & x_{1n} \\ 0 & x_{22} & \cdots & x_{2n-1} & x_{2n} \\ 0 & 0 & \cdots & x_{3n-1} & x_{3n} \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \cdots & x_{n-1n-1} & x_{n-1n} \\ 0 & 0 & \cdots & 0 & x_{nn} \end{bmatrix}.$$

■

Observación 2. *Sea $X = [x_{ij}]_{i,j=1}^n$ una matriz triangular superior de orden $n \times n$, usando inducción sobre el tamaño de la matriz " n " y luego usando inducción sobre la potencia " m " de la matriz X se puede verificar que las entradas de la matriz X^m están dadas por:*

$$[X^m]_{p,r_p} = \left(\sum_{s=0}^{m-1} x_{pp}^s x_{r_p r_p}^{m-1-s} \right) x_{pr_p} +$$
$$+ P_{m,r_p}(x_{pp}, x_{pp+1}, \cdots, x_{pr_p-1}, x_{p+1p+1}, \cdots, x_{p+1r_p}, \cdots, x_{r_p r_p}),$$

donde $p = 1, 2, \cdots, n-1$ y $r_p = p+1, \cdots, n$. Además,

$$P_{m,r_p}(x_{pp}, x_{pp+1}, \cdots, x_{pr_p-1}, x_{p+1p+1}, \cdots, x_{p+1r_p}, \cdots, x_{r_p r_p})$$

es un cierto polinomio homogéneo de grado m con coeficientes enteros.
Por ejemplo, para la entrada $[X^m]_{1n}$ se tiene que:

$$P_{m,n}(x_{11}, x_{12}, \cdots, x_{1n-1}, x_{22}, \cdots, x_{2n}, \cdots, x_{nn}),$$

es un polinomio homogéneo de grado m, donde intervienen $(\frac{n^2+n-2}{2})$ variables x_{ij}, $1 \leq i \leq j \leq n$, $(i,j) \neq (1,n)$.

Teorema 4.3.
Sea $A = [a_{ij}]_{i,j=1}^{n}$, $a_{ij} \in \mathbf{F}$ una matriz triangular superior invertible y supóngase que existen las raíces m-ésimas $\sqrt[m]{a_{jj}} \in \mathbf{F}$ para cada $j = 1, \cdots, n$. Supóngase además que se satisface la condición (16) del teorema 4.2. Entonces, existen al menos m raíces m-ésimas distintas de A con entradas en $\mathbf{F}$.

Demostración.
Supóngase que existen las raíces m-ésimas $\sqrt[m]{a_{jj}} \in \mathbf{F}$. Se asume que:

$$a_{ii} = a_{jj} \Longrightarrow \sqrt[m]{a_{ii}} = \sqrt[m]{a_{jj}} \; ; \tag{21}$$

es decir, dados dos elementos iguales en la diagonal principal de $A(z)$ se eligen sus raíces m-ésimas iguales. Ahora bien, Puesto que A satisface las condiciones del teorema 4.2, entonces se tiene que si X es tal que $X^m = A$, entonces X debe ser triangular superior; por tanto, es suficiente construir las entradas de X cuando $i < j$. Por hipótesis

$$x_{ii} = \sqrt[m]{a_{ii}} \neq 0, \quad \forall i = 1, \cdots, n. \tag{22}$$

Por la observación 2 se sabe que:

$$\begin{aligned}[X^m]_{p,r_p} = a_{pr_p} &= \left(\sum_{s=0}^{m-1} x_{pp}^{s} x_{r_pr_p}^{m-1-s}\right) x_{pr_p} + \\ &+ P_{m,r_p}(x_{pp}, x_{pp+1}, \cdots, x_{pr_p-1}, x_{p+1p+1}, \cdots, x_{p+1r_p}, \cdots, x_{r_pr_p}).\end{aligned} \tag{23}$$

Por otra parte, en virtud de (21) y (22) se tiene que:

$$x_{ii} \neq x_{jj} \Longrightarrow x_{ii}^m \neq x_{jj}^m \; ,$$

con lo cual

$$\sum_{s=0}^{m-1} x_{ii}^{s} x_{jj}^{m-1-s} = \frac{x_{ii}^m - x_{jj}^m}{x_{ii} - x_{jj}} \neq 0, \; \forall i,j = 1, 2, \cdots, n.$$

Por consiguiente, $\left(\sum_{s=0}^{m-1} x_{pp}^{s} x_{r_pr_p}^{m-1-s}\right) \neq 0$, para todo $p = 1, 2, \cdots, n-1$ y $r_p = p+1, \cdots, n$. Así entonces, de (23) se puede obtener que:

$$x_{pr_p} = \frac{a_{pr_p} - P_{m,r_p}(x_{pp}, \cdots, x_{pr_p-1}, x_{p+1p+1}, \cdots, x_{p+1r_p}, \cdots, x_{r_pr_p})}{\left(\sum_{s=0}^{m-1} x_{pp}^{s} x_{r_pr_p}^{m-1-s}\right)} \; , \tag{24}$$

donde $p = 1, \cdots, n-1$ y $r_p = p+1, \cdots, n$.

El resultado obtenido en (24) permite calcular todas las entradas de X por encima de la diagonal principal tal que $X^m = A$.

Nótese que en el desarrollo de la demostración no se ha enfocado el hecho de que existen al menos m raíces m-ésimas distintas para A, pero esto se deduce fácilmente con solo observar que el conjunto formado por todas las entradas diagonales de las matrices X tales que $X^m = A$ $(\{x_{jj} : x_{jj} = \sqrt[m]{a_{jj}}\ \}_{j=1}^{n})$, tiene mn elementos distintos; luego, eligiendo n elementos(por permutaciones sin reemplazamiento) de dicho conjunto para formar una diagonal principal asociada a una matriz X tal que $X^m = A$ y así, usando (24) se calculan el resto de elementos por encima de la diagonal principal para la matriz X. De esta forma, se obtienen al menos m raíces m-ésimas distintas para A. ■

Ejemplo 4.1. *Sea* $\mathbb{F} = \mathbb{R}$ *y* $A = \begin{bmatrix} 4 & 2 \\ 0 & 4 \end{bmatrix}$ *una matriz triangular superior invertible. Obviamente existen* $\sqrt{a_{ii}} \in \mathbb{R}$ *para cada* $i = 1, 2$. *Entonces, se verificará que existen al menos dos raíces cuadradas distintas para la matriz* A *con entradas en* $\mathbb{R}$. *En efecto:*
Puesto que A *satisface el teorema 4.2, entonces, toda matriz* X *de orden* 2×2 *sobre* $\mathbb{R}$ *tal que* $X^2 = A$ *es triangular superior; luego, sea* $X = \begin{bmatrix} x_{11} & x_{12} \\ 0 & x_{22} \end{bmatrix}$ *tal que*

$$X^2 = \begin{bmatrix} x_{11}^2 & x_{12}(x_{11} + x_{22}) \\ 0 & x_{22}^2 \end{bmatrix} = \begin{bmatrix} 4 & 2 \\ 0 & 4 \end{bmatrix} = A,$$

entonces, se deduce que:

$$x_{11} = \pm 2 = x_{22} \quad y \quad x_{12} = \frac{2}{\pm 2 + \pm 2}.$$

Por consiguiente, observando que la entrada (1,2) de la matriz X *no existe cuando se eligen las raíces cuadradas opuestas, entonces se puede ver que solo existen dos raíces cuadradas distintas para* A *dadas por:*

$$X_1 = \begin{bmatrix} 2 & \frac{1}{2} \\ 0 & 2 \end{bmatrix} \quad y \quad X_2 = \begin{bmatrix} -2 & -\frac{1}{2} \\ 0 & -2 \end{bmatrix}.$$

Corolario 4.4.
Sea $A = [a_{ij}]_{i,j=1}^{n}$, $a_{ij} \in \mathbb{F}$ *una matriz triangular superior invertible y supóngase que existen las raíces* m*-ésimas* $\sqrt[m]{a_{jj}} \in \mathbb{F}$ *para cada* $j = 1, \cdots, n$. *Supóngase además que se satisface la condición (16) del teorema 4.2. Entonces, existen a lo sumo* m^n *raíces* m*-ésimas distintas de* A *con entradas en* $\mathbb{F}$.

Demostración.
Puesto que A satisface las hipótesis del teorema 4.2, entonces se tiene que toda matriz X de orden $n \times n$ sobre el campo $\mathbb{F}$ tal que $X^m = A$ debe ser triangular superior. Considérese el conjunto

$$Z_i := \{x_{ii} : x_{ii} = \sqrt[m]{a_{ii}}\}_{i=1}^{n}.$$

Nótese que para cada $i = 1, \cdots, n$, el conjunto Z_i posee m elementos distintos.
Ahora, sea

$$M := \{X : X^m = A\}$$

$$= \{X : X = \begin{bmatrix} x_{11} & & & & \\ 0 & x_{22} & & * & \\ \vdots & \vdots & \ddots & & \\ 0 & 0 & \cdots & x_{n-1n-1} & \\ 0 & 0 & \cdots & 0 & x_{nn} \end{bmatrix},$$

donde $x_{ii} \in Z_i$ para todo $i = 1, \cdots, n$ y los elementos sobre la diagonal principal están dados por (24) $\}$.

Nótese que los elementos sobre la diagonal principal de cada matriz M dependen de la elección hecha para los elementos de la diagonal principal. Además, puesto que para cada $i = 1, \cdots, n$ Z_i tiene m elementos distintos, entonces al aplicar permutaciones con reemplazamiento sobre la diagonal principal, M tendrá a lo sumo m^n matrices distintas.
Ahora bien,debido a la forma que tiene los elementos sobre la diagonal principal a saber

$$x_{pr_p} = \frac{a_{pr_p} - P_{m,r_p}(x_{pp}, \cdots, x_{pr_p-1}, x_{p+1p+1}, \cdots, x_{p+1r_p}, \cdots, x_{r_pr_p})}{\left(\sum_{s=0}^{m-1} x_{pp}^s x_{r_pr_p}^{m-1-s}\right)},$$

entonces, se deduce que si el denominador $\left(\sum_{s=0}^{m-1} x_{pp}^s x_{r_pr_p}^{m-1-s}\right) \neq 0$ para todo los x_{ii} que se eligen sobre la diagonal principal, entonces M tiene exactamente m^n raíces m-ésimas distintas; si por lo menos una de las sumas $\left(\sum_{s=0}^{m-1} x_{pp}^s x_{r_pr_p}^{m-1-s}\right) = 0$ se disminuyen el número de elementos de M.

En resumen, M no tiene más de m^n raíces m-ésimas distintas para A. ■

Ejemplo 4.2. *Sea* $\mathbb{F} = \mathbb{R}$ *y* $A = \begin{bmatrix} 1 & 3 & 4 \\ 0 & 5 & \sqrt{2} \\ 0 & 0 & 8 \end{bmatrix}$ *una matriz triangular superior invertible. Obviamente existen las raíces* $\sqrt{a_{ii}} \in \mathbb{R}$ *para cada* $i = 1, 2, 3$.
En virtud del corolario (4.4), se verificará que existen a lo sumo ocho raíces cuadradas distintas para la matriz A *con entradas en* $\mathbb{R}$*. En efecto:*
En virtud de que A *satisface el teorema 4.2, entonces, toda matriz* X *de orden* 3×3 *sobre* $\mathbb{R}$ *tal que* $X^2 = A$ *es triangular superior. Además, en virtud de (24),*

$$x_{pr_p} = \frac{a_{pr_p} - P_{m,r_p}(x_{pp}, \cdots, x_{pr_p-1}, x_{p+1p+1}, \cdots, x_{p+1r_p}, \cdots, x_{r_pr_p})}{\left(\sum_{s=0}^{m-1} x_{pp}^s x_{r_pr_p}^{m-1-s}\right)},$$

donde en este caso $p = 1,\ 2$ *y* $r_p = p + 1,\ 3$.

Así,

$$x_{12} = \frac{a_{12} - P_{2,2}(x_{11}, x_{22})}{x_{11} + x_{22}} \ ,$$

$$x_{13} = \frac{a_{13} - P_{2,3}(x_{11}, x_{12}, x_{22}, x_{23}, x_{33})}{x_{11} + x_{33}} \ ,$$

$$x_{23} = \frac{a_{23} - P_{2,3}(x_{22}, x_{33})}{x_{22} + x_{33}} \ ,$$

donde se puede verificar que los polinomios

$$P_{2,2}(x_{11}, x_{22})$$

y

$$P_{2,3}(x_{22}, x_{33})$$

son idénticamente nulos. Luego, las matrices X *tales que* $X^2 = A$ *están dadas por:*

$$X = \begin{bmatrix} \pm 1 & \frac{1}{\pm 1 + \pm\sqrt{5}} & \frac{4 - P_{2,3}}{\pm 1 + \pm\sqrt{8}} \\ 0 & \pm\sqrt{5} & \frac{\sqrt{2}}{\pm\sqrt{5} + \pm\sqrt{8}} \\ 0 & 0 & \pm\sqrt{8} \end{bmatrix}.$$

Al combinar los signos, se verifica que existen 2^3 *raíces cuadradas distintas, lo cual enfoca lo afirmado en el corolario (4.4).*

5. PRELIMINARES ANALÍTICOS

En el presente capítulo se enuncian algunos resultados importantes acerca de la analiticidad de los autovalores.

Acontinuación se darán unas definiciones y un lema que serán de mucha utilidad para la prueba del teorema 5.1

Definición 5.1. *(Dominio Smith)*
Un ***dominio smith*** *es un anillo $\mathcal{R}$ conmutativo con unidad sin divisores de cero tal que toda matriz X con entradas en $\mathcal{R}$ puede ser transformada vía $X \longrightarrow EXF$ (donde E y F son matrices invertibles sobre $\mathcal{R}$) a una forma diagonal $d(x_1, \cdots, x_p)$, donde $x_1, \cdots, x_p \in \mathcal{R}/\{0\}$ tal que x_j es divisible por x_{j+1} en $\mathcal{R}$ para $j = 1, \cdots, p-1$.*

Definición 5.2. *(Anillo de funciones analíticas)*
El conjunto $\mathcal{A}(\mathbb{G})$ de las funciones analíticas definidas sobre un dominio $\mathbb{G}$, junto con las operaciones de adición y multiplicación usual de las funciones forma un anillo y se llama ***anillo de funciones analíticas*** *sobre $\mathbb{G}$.*

Lema 5.1.
El anillo $\mathcal{A}(\mathbb{G})$ de funciones analíticas sobre $\mathbb{G}$ es un dominio smith.

Demostración.
Sea $A(z)$ una matriz de función analítica en $\mathcal{A}(\mathbb{G})$. Supóngase que $A(z)$ es de orden $n \times n$ y rango r. El lema es claramente cierto cuando $n = 1$. Supóngase entonces que el resultado es cierto para matrices de orden menor que n.
Supóngase que el rango de $A(z)$ es menor que n, entonces existe al menos un vector x, analítico en $\mathbb{G}$, para el cual $Ax \equiv 0$ en $\mathbb{G}$; y se puede asegurar que x no es nulo, combinando algunos factores comunes de los coeficientes, cuando son expresados en términos de la base canonica $e_1, \cdots, e_n$. Ahora, existe una matriz de función Q_0 analítica e invertible en $\mathbb{G}$, para la cual $x = Q_0 e_n$, luego, $AQ_0 e_n \equiv 0$; así, los elementos en la última columna de AQ_0 son todos ceros. Ahora, considerando la matriz conjugada $\overline{Q}_0 \overline{A}$ se puede ver de la misma forma, que existe una matriz de función anlítica e invertible $\overline{P}_0$ en $\mathbb{G}$, la cual es tal que los coeficientes en la última columna de $\overline{Q}_0 \overline{A} \overline{P}_0$ son también todos ceros. Se sigue entonces que $P_0 A Q_0$ tiene ceros tanto su última columna como su última fila. Por consiguiente, $P_0 A Q_0$ puede ser considerada como una matriz de orden $n-1$, relativa a la base $e_1, \cdots, e_{n-1}$. Existen por tanto, en virtud de la hipótesis dos matrices de funciones P_1 y Q_1 de orden $n-1$, las cuales son analíticas e invertibles en $\mathbb{G}$ y son tales que $P_1 P_0 A Q_0 Q_1$ satisface la definición (5.1) en la base $e_1, \cdots, e_{n-1}$. Extendiendo a P_1 y Q_1 a la base original $e_1, \cdots, e_n$, por medio de las condiciones $P_1 e_n = e_n = Q_1 e_n$, entonces se tendrá la definición (5.1) para $A(z)$ en la base original $e_1, \cdots, e_n$. ■

Teorema 5.1.
Sea $\mathbb{G}$ un dominio en $\mathbb{C}$, ó un intervalo en $\mathbb{R}$. Sea $A(z)$ una matriz de función analítica de orden $n \times n$ sobre $\mathbb{G}$. Las siguientes afirmaciones son equivalentes:

(a) Los autovalores $\lambda_1(z), \lambda_2(z), \cdots, \lambda_n(z)$ de $A(z)$ pueden ser enumerados tal que, para todo $z \in \mathbb{G}$, $\lambda_1(z), \lambda_2(z), \cdots, \lambda_n(z)$ resulten ser funciones analíticas.

(b) Existe una matriz de función analítica invertible $\mathbb{S}(z)$ tal que $\mathbb{S}^{-1}(z)A(z)\mathbb{S}(z)$ es una matriz triangular superior, para todo $z \in \mathbb{G}$.

Demostración.
La prueba de este teorema consiste en verificar que $(a) \Longrightarrow (b)$ y $(b) \Longrightarrow (a)$. En efecto: Supóngase que (a) se satisface. Sea $\lambda_1(z)$ un autovalor analítico de $A(z)$ y considérese la matriz de función $B(z) := A(z) - \lambda_1(z)I$. Obviamente $B(z)$ así definida es analítica. Por tanto, $B(z) \in \mathcal{A}(\mathbb{G})$. Luego, en virtud de la definición 5.1 existen matrices $E(z)$ y $F_1(z)$ en $\mathcal{A}(\mathbb{G})$ de orden $n \times n$ las cuales son invertibles y satisfacen:

$$B(z) = E(z)diag(x_1(z), x_2(z), \cdots, x_n(z))F_1(z), \quad \forall z \in \mathbb{G}, \tag{25}$$

donde $x_1(z), x_2(z), \cdots x_n(z)$ resultan ser funciones escalares analíticas. Ahora bien, puesto que $\lambda_1(z)$ es un autovalor asociado a $A(z)$, entonces $A(z)x = \lambda_1(z)x$, para un cierto vector $x \neq 0$. Luego, $(A(z) - \lambda_1(z)I)x = 0$; con lo cual $B(z) = A(z) - \lambda_1(z)I$ será no invertible para todo $z \in \mathbb{G}$; en otras palabras, $detB(z) = 0$ para todo $z \in \mathbb{G}$. Entonces, por (25) puesto que $E(z)$ y $F_1(z)$ son invertibles se debe tener que al menos una de las funciones $x_1, \cdots, x_n$ es idénticamente nula. Por ejemplo, sea $x_1 \equiv 0$. Entonces,

$$\begin{aligned}
F_1(z)A(z)F_1^{-1}(z) &= F_1(z)(B(z) + \lambda_1(z)I)F_1^{-1}(z) \\
&= F_1(z)B(z)F_1^{-1}(z) + \lambda_1(z)F_1(z)F_1^{-1}(z) \\
&= F_1(z)B(z)F_1^{-1}(z) + \lambda_1(z)I \\
&= F_1(z)[E(z)diag(0, x_2, x_3, \cdots, x_n)F_1(z)]F_1^{-1}(z) + \lambda_1(z)I \\
&= F_1(z)E(z)diag(0, x_2, x_3, \cdots, x_n) + \lambda_1(z)I \\
&= F_1(z)E(z)\begin{bmatrix} 0 & 0 & \cdots & 0 \\ 0 & x_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & x_n \end{bmatrix} + \lambda_1(z)\begin{bmatrix} 1 & 0 & \cdots & 0 \\ 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 1 \end{bmatrix} \\
&= \begin{bmatrix} \lambda_1(z) & * \\ 0 & A_1(z) \end{bmatrix}.
\end{aligned}$$

3C

En resumen,

$$F_1(z)A(z)F_1^{-1}(z) = \begin{bmatrix} \lambda_1(z) & * \\ 0 & A_1(z) \end{bmatrix}, \tag{26}$$

donde $A_1(z)$ es una matriz de función analítica de orden $(n-1) \times (n-1)$. Esto muestra que $A(z)$ es semejante a la matriz dada en (26). Por tanto, $A(z)$ y la matriz dada en (26) tienen los mismos autovalores; así, eligiendo el autovalor $\lambda_2(z)$ de $A(z)$ se debe tener que

$$\begin{bmatrix} \lambda_1(z) & * \\ 0 & A_1(z) \end{bmatrix} x = \lambda_2(z)x, \tag{27}$$

para algún vector $x \neq 0$; donde además $A_1(z) = \begin{bmatrix} A_{11}(z) \\ A_{21}(z) \\ \vdots \\ A_{n-11}(z) \end{bmatrix}$ en la cual,

$A_{i1}(z)$ denota la i-ésima fila de $A_1(z)$ para cada $i = 1, \cdots, n-1$.

Luego, de (27) se obtiene:

$$\begin{bmatrix} \lambda_1(z) & * \end{bmatrix} x = \lambda_2(z)x_1,$$

siendo $x = [x_1 \cdots x_n]^T$ y el sistema:

$$\begin{cases} A_{11}(z)y &= \lambda_2(z)x_2 \\ A_{21}(z)y &= \lambda_2(z)x_3 \\ & \vdots \\ A_{n-1}(z)y &= \lambda_2(z)x_n, \end{cases} \tag{28}$$

donde " y " es un vector de orden $(n-1) \times 1$, dado por $y = [x_2 \cdots x_n]^T$.
Ahora bien, el sistema (28) puede ser escrito en la forma $A_1(z)y = \lambda_2(z)y$, es decir, $\lambda_2(z)$ resulta ser un autovalor para $A_1(z)$ (más generalmente $\lambda_2(z), \cdots, \lambda_n(z)$ resultan ser autovalores de $A_1(z)$)(Recuerde que una matriz de orden $n \times n$ tiene a lo sumo n autovalores distintos).

En forma similar al proceso anterior, considérese $C(z) = A_1(z) - \lambda_2(z)I$, donde I denota la matriz identica de orden $(n-1) \times (n-1)$. $C(z)$ así definida es analítica. Por tanto, $C(z) \in \mathcal{A}(\mathbb{G})$; luego, existen matrices invertibles $H(z)$ y $M(z)$ de orden $(n-1) \times (n-1)$ en $\mathcal{A}(\mathbb{G})$ tal que

$$C(z) = H(z)diag(r_1(z), r_2(z), \cdots, r_{n-1}(z))M(z), \quad \forall z \in \mathbb{G}.$$

Ahora bien, puesto que $detC(z) = 0$ para todo $z \in \mathbb{G}$, entonces se debe tener que al menos una de las funciones $r_1(z), r_2(z), \cdots, r_{n-1}(z)$ es identicamente nula;

supóngase que $r_1(z) \equiv 0$. Entonces

$$\begin{aligned}
M(z)A_1(z)M^{-1}(z) &= M(z)(C(z)+\lambda_2(z)I)M^{-1}(z) \\
&= M(z)C(z)M^{-1}(z)+\lambda_2(z)I \\
&= M(z)(H(z)diag(0,r_2(z),\cdots,r_{n-1}(z))M(z))M^{-1}(z)+\lambda_2(z)I \\
&= M(z)H(z)diag(0,r_2(z),\cdots,r_{n-1}(z))+\lambda_2(z)I \\
&= M(z)H(z)\begin{bmatrix} 0 & 0 & \cdots & 0 \\ 0 & r_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & r_{n-1} \end{bmatrix} + \lambda_2(z)\begin{bmatrix} 1 & 0 & \cdots & 0 \\ 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 1 \end{bmatrix} \\
&= \begin{bmatrix} \lambda_2(z) & * \\ 0 & A_2(z) \end{bmatrix}.
\end{aligned}$$

En resumen,

$$M(z)A_1(z)M^{-1}(z) = \begin{bmatrix} \lambda_2(z) & * \\ 0 & A_2(z) \end{bmatrix} \tag{29}$$

donde $A_2(z)$ es una matriz de función analítica de orden $(n-2)\times(n-2)$. Esto muestra que $A_1(z)$ es semejante a la matriz dada en (29).

Por otro lado, considérese $1 \in \mathcal{A}(\mathrm{G})$ como la función escalar unidad. Entonces, definiendo

$$F_2(z) = \begin{bmatrix} 1 & * \\ 0 & M(z) \end{bmatrix} \quad y \quad F_2^{-1}(z) = \begin{bmatrix} 1 & * \\ 0 & M^{-1}(z) \end{bmatrix},$$

se tiene al multiplicar en (26) por $F_2(z)$ a izquierda y por $F_2^{-1}(z)$ a derecha lo siguiente:

$$F_2(z)F_1(z)A(z)F_1^{-1}(z)F_2^{-1}(z) = \begin{bmatrix} 1 & * \\ 0 & M(z) \end{bmatrix} F(z)A(z)F^{-1}(z) \begin{bmatrix} 1 & * \\ 0 & M^{-1}(z) \end{bmatrix}$$

$$= \begin{bmatrix} 1 & * \\ 0 & M(z) \end{bmatrix} \begin{bmatrix} \lambda_1(z) & * \\ 0 & A_1(z) \end{bmatrix} \begin{bmatrix} 1 & * \\ 0 & M^{-1}(z) \end{bmatrix}$$

$$= \begin{bmatrix} \lambda_1(z) & * \\ 0 & M(z)A_1(z) \end{bmatrix} \begin{bmatrix} 1 & * \\ 0 & M^{-1}(z) \end{bmatrix}$$

$$= \begin{bmatrix} \lambda_1(z) & * \\ 0 & M(z)A_1(z)M^{-1}(z) \end{bmatrix}$$

$$= \begin{bmatrix} \lambda_1(z) & & * \\ & \lambda_2(z) & \\ 0 & & A_2(z) \end{bmatrix}, \text{ [\textit{en virtud de} (29)]}$$

Por consiguiente, se ha probado que

$$F_2(z)F_1(z)A(z)F_1^{-1}(z)F_2^{-1}(z) = \begin{bmatrix} \lambda_1(z) & & * \\ & \lambda_2(z) & \\ 0 & & A_2(z) \end{bmatrix},$$

donde $A_2(z)$ es una matriz de función analítica de orden $(n-2) \times (n-2)$.

Continuando con el mismo proceso $(n-2)$ veces se obtiene el siguiente resultado:

$$\Big(\prod_{i=n-2}^{1} F_i(z)\Big) A(z) \Big(\prod_{i=1}^{n-2} F_i^{-1}(z)\Big) = \begin{bmatrix} \lambda_1(z) & & & & \\ & \lambda_2(z) & & * & \\ & & \ddots & & \\ & 0 & & \lambda_{n-2}(z) & \\ & & & & R(z) \end{bmatrix}, \qquad (30)$$

donde $R(z)$ es una matriz de función analítica de orden 2×2.

Finalmente considérese $D(z) = R(z) - \lambda_{n-1}(z)I$, donde I denota la matriz identica de orden 2×2. $D(z)$ así definida es analítica. Por tanto, $D(z) \in \mathcal{A}(\mathbb{G})$; luego, existen matrices invertibles $J(z)$ y $K(z)$ de orden 2×2 en $\mathcal{A}(\mathbb{G})$ tal que

$$D(z) = J(z)diag(t_1, t_2)K(z), \quad \forall z \in \mathbb{G}.$$

Ahora bien, puesto que $detD(z) = 0$ para todo $z \in \mathbb{G}$, entonces se debe tener que al menos una de las funciones t_1, t_2 es identicamente nula; supóngase que

$t_1 \equiv 0$. Entonces

$$\begin{aligned}
K(z)R(z)K^{-1}(z) &= K(z)(D(z) + \lambda_{n-1}(z)I)K^{-1}(z) \\
&= K(z)D(z)K^{-1}(z) + \lambda_{n-1}(z)I \\
&= K(z)(J(z)diag(0, t_2)K(z))K^{-1}(z) + \lambda_{n-1}(z)I \\
&= K(z)J(z)diag(0, t_2) + \lambda_{n-1}(z)I \\
&= K(z)J(z)\begin{bmatrix} 0 & 0 \\ 0 & t_2 \end{bmatrix} + \begin{bmatrix} \lambda_{n-1}(z) & 0 \\ 0 & \lambda_{n-1}(z) \end{bmatrix} \\
&= \begin{bmatrix} \lambda_{n-1}(z) & * \\ 0 & r_{22}(z) \end{bmatrix}.
\end{aligned}$$

Por consiguiente,

$$K(z)R(z)K^{-1}(z) = \begin{bmatrix} \lambda_{n-1}(z) & * \\ 0 & r_{22}(z) \end{bmatrix}. \tag{31}$$

Esto muestra que $R(z)$ es semejante a la matriz dada en (31).

Ahora bien, sea

$$F_{n-1}(z) = \begin{bmatrix} 1 & & & & \\ & 1 & & * & \\ & & \ddots & & \\ & 0 & & 1 & \\ & & & & K(z) \end{bmatrix}$$

y

$$F_{n-1}^{-1}(z) = \begin{bmatrix} 1 & & & & \\ & 1 & & * & \\ & & \ddots & & \\ & 0 & & 1 & \\ & & & & K^{-1}(z) \end{bmatrix}.$$

Luego, se tiene al multiplicar en (30) a la izquierda por $F_{n-1}(z)$ y a la derecha por $F_{n-1}^{-1}(z)$ lo siguiente:

$$\left(\prod_{i=n-1}^{1} F_i(z)\right) A(z) \left(\prod_{i=1}^{n-1} F_i^{-1}(z)\right) = \begin{bmatrix} \lambda_1(z) & & & * \\ & \lambda_2(z) & & \\ & & \ddots & \\ 0 & & \lambda_{n-2}(z) & \\ & & & K(z)R(z)K^{-1}(z) \end{bmatrix}$$

$$= \begin{bmatrix} \lambda_1(z) & & & & * \\ & \lambda_2(z) & & & \\ & & \ddots & & \\ & & & \lambda_{n-2}(z) & & \\ 0 & & & & \lambda_{n-1}(z) & \\ & & & & & r_{22}(z) \end{bmatrix}. \qquad (32)$$

Esto muestra que $A(z)$ es semejante a la matriz dada en (32); por tanto, tienen igual determinante con lo cual se concluye que $r_{22}(z) = \lambda_n(z)$. Por último, considerando

$$S(z) = \prod_{i=n-1}^{1} F_i(z)$$

se concluye que $S(z)$ es invertible (puesto que cada $F_i(z)$ lo es, para todo $i = 1, \cdots, n-1$) y además satisface

$$S(z)A(z)S^{-1}(z) = \begin{bmatrix} \lambda_1(z) & & * \\ & \ddots & \\ 0 & & \lambda_n(z) \end{bmatrix}.$$

Así, se tiene la implicación $(a) \Longrightarrow (b)$.

Resta probar la implicación $(b) \Longrightarrow (a)$. En efecto:

Supóngase que (b) se satisface. Entonces, para toda matriz de función analítica $A(z)$ de orden $n \times n$, existe una matriz de función analítica invertible $S(z)$ tal que $S^{-1}(z)A(z)S(z)$ es triangular superior. Puesto que $S(z)$ es analítica, entonces $S^{-1}(z)$ también lo es y puesto que $A(z)$ por hipótesis es analítica, entonces la matriz triangular superior dada por $S^{-1}(z)A(z)S(z)$ resulta ser analítica; con lo cual, los elementos en la diagonal principal de la matriz triangular superior $S^{-1}(z)A(z)S(z)$ son analíticos, los cuales son precisamente los autovalores de $A(z)$. Así se concluye que los autovalores de $A(z)$ pueden ser enumerados tal que estos resulten ser funciones analíticas. Esto completa la demostración. ■

Antes de enunciar una condición suficiente para la analiticidad de los autovalores, se introduce una definición conocida de curva diferenciable ; dada así:

Definición 5.3. ***(Curva diferenciable)***
Sea I un intervalo en $\mathbb{R}$. Una curva $\gamma : I \longrightarrow \mathbb{R}^n$, se dice ***diferenciable*** *si su derivada $\dot{\gamma}(t)$ existe y es continua para todo $t \in I$.*

Proposición 5.2.
Sea $A(z)$ una matriz de función analítica de orden $n \times n$ sobre un intervalo real $\mathbb{G}$. Si para todo $z \in \mathbb{G}$, los autovalores $\lambda_1(z), \cdots, \lambda_n(z)$ de $A(z)$ pertenecen a una curva diferencible $\Gamma \subset \mathbb{C}$ fija, entonces $\lambda_1(z), \cdots, \lambda_n(z)$ pueden ser enumerados tal que sean funciones analíticas de $z \in \mathbb{G}$.

Demostración.
Sea $z_0 \in \mathbb{G}$. Los autovalores $\lambda_j(z)$ de $A(z)$ pueden ser escritos como series de potencias fraccionales en una vecindad de z_0 así:

$$\lambda_j(z) = \sum_{k=0}^{\infty} a_{jk}(z-z_0)^{k/p} \ ; \quad a_{jk} \in \mathbb{C}.$$

Supóngase que $p \geq 2$ y que el menor de los coeficientes a_{jk}, con k no múltiplo de p, es no nulo. Sea k_0 el más pequeño de los enteros, no múltiplo de p, tal que $a_{jk_0} \neq 0$. Entonces, haciendo que $z \to z_0$ por la derecha y por la izquierda se tiene que:

$$\lim_{z \to z_0^+} [\lambda_j(z) - \lambda_j(z_0)](z-z_0)^{-k_0/p} = a_{jk_0} \ .$$

$$\lim_{z \to z_0^-} [\lambda_j(z) - \lambda_j(z_0)](z-z_0)^{-k_0/p} = (-1)^{k_0/p} a_{jk_0}.$$

Claramente, los números a_{jk_0} y $(-1)^{k_0/p} a_{jk_0}$ cada uno coincide con la dirección tangencial de Γ en el punto $\lambda_j(z_0)$, o son opuestos en esta dirección. Por tanto $(-1)^{k_0/p}$ debe ser real, lo cual es absurdo. Así, se ha probado que los autovalores de $A(z)$ son analíticos en una vecindad de z_0. Puesto que $z_0 \in \mathbb{G}$ fue elegido arbitrario entonces, se puede concluir que para todo $z \in \mathbb{G}$, los autovalores $\lambda_1(z), \cdots, \lambda_n(z)$ de $A(z)$ son analíticos. ∎

6. RESULTADO PRINCIPAL

En este capítulo se presenta el resultado principal de este trabajo y algunos resultados que garantizan la existencia de raíces m-ésimas analíticas de una matriz dada.

Teorema 6.1. *(TEOREMA PRINCIPAL)*
Sea $\mathbb{G}$ un dominio en $\mathbb{C}$ y sea $A(z)$, $z \in \mathbb{G}$, una matriz de función analítica invertible. Supóngase que los autovalores de $A(z)$ pueden ser enumerados tal que ellos sean funciones analíticas en $\mathbb{G}$. Se fija un entero $m \geq 2$. Entonces:

(a) Existen al menos m matrices de funciones meromorficas $X_j(z)$ distintas de orden $n \times n$, $j = 1, \cdots, m$, tal que $X_j^m(z) = A(z)$.

(b) Si se adiciona, que para toda función analítica $\lambda(z)$ al menos uno de los subdeterminantes de $A(z) - \lambda(z)I$ de orden $(n-1) \times (n-1)$ no es identicamente nulo, entonces, existen a lo sumo m^n matrices de funciones meromorficas $X_j(z)$ distintas de orden $n \times n$, con $j = 1, \cdots, m$, tal que $X_j^m(z) = A(z)$.

Demostración.
Sea $\mathbb{G}$ un dominio en $\mathbb{C}$ y sea $A(z)$, $z \in \mathbb{G}$, una matriz de función analítica invertible. Supóngase además que los autovalores pueden ser enumerados tal que ellos sean funciones analíticas en $\mathbb{G}$. Entonces por el teorema 5.1 (b) se puede reducir $A(z)$ del caso general al caso en que sea triangular superior; es decir, se puede asumir sin perdida de generalidad que $A(z)$ es triangular superior.

Considérese [1]$\mathbb{F}$ como el campo de las funciones meromorficas de $\mathbb{G}$ en $\mathbb{C}$. Puesto que $A(z)$ es triangular superior, entonces, la diagonal principal está conformada por sus autovalores; esto es $a_{jj}(z) = \lambda_j(z)$ para todo $j = 1, \cdots, n$. Luego, cada $\lambda_j(z)$ es analítico para todo $z \in \mathbb{G}$ y todo $j = 1, \cdots, n$. Además, $\sqrt[m]{a_{jj}(z)} = \sqrt[m]{\lambda_j(z)} \in \mathbb{F}$, para todo $j = 1, \cdots, n$ por ser composición de funciones meromorficas. Así las cosas, en virtud del teotema 4.3 se tendrá que existen al menos m matrices de funciones $X_j(z)$ distintas de orden $n \times n$, tal que $X_j^m(z) = A(z)$; las cuales resultarán siendo meromorficas puesto que sus entradas pertenecen al campo $\mathbb{F}$.

Ahora, se verifica la afirmación (b).
Supóngase que para toda función analítica $\lambda(z)$ al menos uno de los subdeterminantes de $A(z) - \lambda(z)I$ de orden $(n-1) \times (n-1)$ no es identicamente nulo. Se verifica ahora, que esta hipótesis implica que:

$$rango(A(z) - \lambda(z)I) \geq n - 1 \qquad \forall \lambda(z) \in \mathbb{F}.$$

[1]Si $\mathbb{F}$ denota el conjunto de las funciones meromorfas en $\mathbb{G}$, $\mathbb{F}$ junto con las operaciones de adición y multiplicación usuales entre funciones es un cuerpo conmutativo (o, un campo, véase [3], [9])

CO153

En efecto: Supóngase que existe $\lambda(z) \in \mathbf{F}$ para el cual se tiene que:

$$rango(A(z) - \lambda(z)I) < n - 1 \qquad \forall z \in \mathbb{G}.$$

Esto implica que al menos dos columnas de la matriz $A(z) - \lambda(z)I$ son linealmente dependiente. Por consiguiente, cualquier submatriz de $A(z) - \lambda(z)I$ de orden $(n-1) \times (n-1)$ tiene determinante igual a cero, puesto que toda submatriz contiene al menos una columna linealmente dependiente.
Pero ahora bien, si el $\lambda(z)$ que se ha supuesto, es autovalor de $A(z)$, entonces $\lambda(z) = a_{ii}(z)$ para algún $i = 1, \cdots, n$, así que $\lambda(z)$ es analítico sobre $\mathbb{G}$, lo cual contradice las hipótesis dadas en b).
Por otro lado, si $\lambda(z) \in \mathbf{F}$ no es autovalor de $A(z)$, supóngase $S(z)$ cualquier submatriz de $A(z) - \lambda(z)I$ de orden $(n-1) \times (n-1)$. Obviamente $S(z)$ contiene en algunas de sus entradas elementos de la forma $a_{ii} - \lambda(z)$ para algunos $i = 1, \cdots, n$. Ahora bien, puesto que $detS(z) = 0$ y siempre $detS(z)$ depende de todas las entradas de $S(z)$, entonces se tendrá que $\lambda(z)$ debe ser una función acotada para todo $z \in \mathbb{G}$, con lo cual $\lambda(z)$ será analítico y nuevamente se contradicen las hipótesis dadas en (b).
Por consiguiente, lo anteriormente supuesto es falso. Por tanto se tiene que

$$rango(A(z) - \lambda(z)I) \geq n - 1 \qquad \forall \lambda(z) \in \mathbb{F}.$$

Así se tienen todas las hipótesis del corolario 4.4 y en consecuencia existen a lo sumo m^n matrices de funciones meromorficas $X_j(z)$ distintas de orden $n \times n$, con $j = 1, \cdots, m$ tal que $X_j^m(z) = A(z)$. ■

Los resultados que se enuncian a continuación establecen el problema de existencia de raíces m-ésimas **analíticas** para una matriz analítica $A(z)$ dada, para la cual se supone que sus autovalores son analíticos. Para esto se debilitan un poco las hipótesis del teorema principal en el sentido de que se adiciona la condición de mayorización de ceros estudiada en el capítulo 3. Más concretamente se enuncia el primer resultado así:

Teorema 6.2.
Sea $A(z) = [a_{ij}(z)]_{i=1}^n$ una matriz de función analítica triangular superior e invertible en $\mathbb{G}$. Sea $X(z) := [x_{ij}(z)]_{i,j=1}^n$ una matriz triangular superior tal que $X^m(z) = A(z)$ y cuyos elementos por encima de la diagonal principal están dados por (24). Supóngase además, que los ceros de

$$\sum_{s=0}^{m-1} x_{pp}^s(z) x_{r_p r_p}^{m-1-s}(z)$$

están mayorizados por los ceros de

$$a_{pr_p}(z) - P_{m,r_p}(x_{pp}, x_{pp+1}, \cdots, x_{pr_p-1}, x_{p+1p+1}, \cdots, x_{p+1r_p}, \cdots, x_{r_p r_p}),$$

entonces, la matriz triangular superior $X(z) := [x_{ij}(z)]_{i,j=1}^n$ es una raíz m-ésima analítica de $A(z)$.

Demostración.
Sea $A(z) = [a_{ij}(z)]_{i,j=1}^{n}$ una matriz de función analítica triangular superior e invertible en $\mathbb{G}$. Puesto que los elementos por encima de la diagonal principal de $X(z)$ están dados por (24), esto es:

$$x_{pr_p} = \frac{a_{pr_p} - P_{m,r_p}(x_{pp}, \cdots, x_{pr_p-1}, x_{p+1p+1}, \cdots, x_{p+1r_p}, \cdots, x_{r_pr_p})}{\left(\sum_{s=0}^{m-1} x_{pp}^{s} x_{r_pr_p}^{m-1-s}\right)},$$

con $p = 1, \cdots, n-1$ y $r_p = p+1, \cdots, n$ y por hipótesis se tiene que los ceros de

$$\sum_{s=0}^{m-1} x_{pp}^{s} x_{r_pr_p}^{m-1-s}$$

están mayorizados por los ceros de

$$a_{pr_p} - P_{m,r_p}(x_{pp}, \cdots, x_{pr_p-1}, x_{p+1p+1}, \cdots, x_{p+1r_p}, \cdots, x_{r_pr_p}),$$

entonces, se tendrá que las funciones $x_{pr_p}(z)$, con $p = 1, \cdots, n-1$ y $r_p = p+1, \cdots, n$ son analíticas para todo z en $\mathbb{G}$; es decir, todos los elementos que constituyen a la matriz $X(z)$ son analíticos y así $X(z)$ tal que $X^m(z) = A(z)$ será una raiz m-ésima analítica para $A(z)$. ∎

Corolario 6.3.
Sea $A(z) = [a_{ij}(z)]_{i,j=1}^{n}$ una matriz de función analítica triangular superior e invertible en $\mathbb{G}$. Si las raíces m-ésimas analíticas $x_{jj}(z) = \sqrt[m]{a_{jj}(z)}$, con $j = 1, \cdots, n$ pueden ser elegidas tal que

$$a_{ii} = a_{jj} \Longrightarrow \sqrt[m]{a_{ii}} = \sqrt[m]{a_{jj}},$$

con lo cual,

$$\sum_{s=0}^{m-1} x_{ii}^{s}(z) x_{jj}^{m-1-s}(z) = \frac{x_{ii}^{m}(z) - x_{jj}^{m}(z)}{x_{ii}(z) - x_{jj}(z)} \neq 0 \quad \forall z \in \mathbb{G}.$$

Entonces, $A(z)$ tiene una raíz m-ésima analítica.

Demostración.
Por hipótesis la suma $\sum_{s=0}^{m-1} x_{ii}^{s}(z) x_{jj}^{m-1-s}(z)$ es analítica y no nula para todo $z \in \mathbb{G}$ y todo $i, j = 1, \cdots, n$. Luego, el cociente

$$x_{pr_p} = \frac{a_{pr_p} - P_{m,r_p}(x_{pp}, \cdots, x_{pr_p-1}, x_{p+1p+1}, \cdots, x_{p+1r_p}, \cdots, x_{r_pr_p})}{\left(\sum_{s=0}^{m-1} x_{pp}^{s} x_{r_pr_p}^{m-1-s}\right)},$$

con $p = 1, \cdots, n-1$ y $r_p = p+1, \cdots, n$ será analítico; es decir, los ceros de

$$\sum_{s=0}^{m-1} x_{pp}^{s} x_{r_pr_p}^{m-1-s}$$

están mayorizados por los ceros de

$$a_{pr_p} - P_{m,r_p}(x_{pp}, \cdots, x_{pr_p-1}, x_{p+1p+1}, \cdots, x_{p+1r_p}, \cdots, x_{r_pr_p}).$$

Así, en virtud del teorema 6.2 se tiene que $A(z)$ admite una raíz m-ésima analítica. ■

A continuación se dará un resultado que generaliza el teorema 3.1, el cual se enuncia así:

Teorema 6.4.
Sea $A(z) = \begin{bmatrix} a_{11}(z) & a_{12}(z) \\ 0 & a_{22}(z) \end{bmatrix}$, con $z \in \mathbb{G}$ una matriz de función analítica triangular superior de orden 2×2, con entradas diagonales no nulas para todo $z \in \mathbb{G}$. Entonces, $A(z)$ tiene una raíz m-ésima analítica si y solo si al menos una de las m^2 funciones

$$\sum_{s=0}^{m-1} \left(\sqrt[m]{a_{11}(z)} \right)^s \left(\sqrt[m]{a_{22}(z)} \right)^{m-1-s},$$

(donde cualquiera de los m valores de las raíces m-ésimas pueden ser escogidos independientemente para $\sqrt[m]{a_{11}(z)}$ y $\sqrt[m]{a_{22}(z)}$) tienen sus ceros en $\mathbb{G}$ mayorizados por los ceros de $a_{12}(z)$.

Demostración.
Sea $X(z) = \begin{bmatrix} x_{11}(z) & x_{12}(z) \\ 0 & x_{22}(z) \end{bmatrix}$, para todo $z \in \mathbb{G}$ tal que $X^m(z) = A(z)$. Por inducción sobre la potencia m se verifica que:

$$X^m(z) = \begin{bmatrix} x_{11}^m(z) & x_{12}(z)\left(\sum_{s=0}^{m-1} x_{11}^s(z) x_{22}^{m-1-s}(z)\right) \\ 0 & x_{22}^m(z) \end{bmatrix}, \quad \forall z \in \mathbb{G}. \qquad (33)$$

En efecto:
Los casos $m = 2$ y $m = 3$ se deducen fácilmente.
Supóngase ahora que se satisface el caso $m-1$, esto es:

$$X^{m-1}(z) = \begin{bmatrix} x_{11}^{m-1}(z) & x_{12}(z)\left(\sum_{s=0}^{m-2} x_{11}^s(z) x_{22}^{m-2-s}(z)\right) \\ 0 & x_{22}^{m-1}(z) \end{bmatrix}, \quad \forall z \in \mathbb{G}.$$

Ahora bien, puesto que $X^m(z) = X^{m-1}(z)X(z)$, entonces

$$X^m(z) = \begin{bmatrix} x_{11}^{m-1}(z) & x_{12}(z)\left(\sum_{s=0}^{m-2} x_{11}^s(z) x_{22}^{m-2-s}(z)\right) \\ 0 & x_{22}^{m-1}(z) \end{bmatrix} \begin{bmatrix} x_{11}(z) & x_{12}(z) \\ 0 & x_{22}(z) \end{bmatrix}$$

$$= \begin{bmatrix} x_{11}^m(z) & x_{12}(z)\left(\sum_{s=0}^{m-1} x_{11}^s(z) x_{22}^{m-1-s}(z)\right) \\ 0 & x_{22}^m(z) \end{bmatrix}.$$

En virtud de la expresión (33) ya verificada y puesto que $X^m(z) = A(z)$, entonces, se deduce que:

$$x_{11}(z) = \sqrt[m]{a_{11}(z)},$$
$$x_{22}(z) = \sqrt[m]{a_{22}(z)} \qquad y$$
$$x_{12}(z)\Big(\sum_{s=0}^{m-1} x_{11}^{s}(z)x_{22}^{m-1-s}(z)\Big) = a_{12}(z).$$

De la última igualdad se puede decir que $A(z)$ tiene una raíz m-ésima analítica si y solo si al menos una de las m^2 funciones

$$\sum_{s=0}^{m-1} x_{11}^{s}(z)x_{22}^{m-1-s}(z) = \sum_{s=0}^{m-1} \big(\sqrt[m]{a_{11}(z)}\,\big)^{s}\big(\sqrt[m]{a_{22}(z)}\,\big)^{m-1-s},$$

admite una solución analítica $x_{12}(z)$ en $\mathbb{G}$; esto es que el cociente

$$x_{12}(z) = \frac{a_{12}(z)}{\sum_{s=0}^{m-1}\big(\sqrt[m]{a_{11}(z)}\,\big)^{s}\big(\sqrt[m]{a_{22}(z)}\,\big)^{m-1-s}}$$

sea analítico para todo $z \in \mathbb{G}$; o en otras palabras, que la suma

$$\sum_{s=0}^{m-1}\big(\sqrt[m]{a_{11}(z)}\,\big)^{s}\big(\sqrt[m]{a_{22}(z)}\,\big)^{m-1-s}$$

tenga sus ceros en $\mathbb{G}$ mayorizados por los ceros de $a_{12}(z)$. ∎

7. EJEMPLO

En esta sección se da un ejemplo donde el dominio $\mathbb{G}$ que se considera es un intervalo real y se supone que los autovalores de la matriz $A(z)$ dada, son números reales positivos. Además, se puede observar que este ejemplo es una consecuencia inmediata de algunos resultados dados en capítulos anteriores.

Lema 7.1.
$\mathbb{R}^+$ es una curva diferenciable.

Demostración.
Considérese una parametrización para $\mathbb{R}^+$ dada por:

$$\begin{aligned} \alpha : (-\infty, 0) &\longrightarrow \mathbb{R}^2 \\ t &\longrightarrow (t^2, 0). \end{aligned}$$

Claramente $\alpha(t)$ es una curva diferenciable para todo $t \in (-\infty, 0)$, puesto que su derivada

$$\dot{\alpha}(t) = (2t, 0),$$

existe y es continua para todo $t \in (-\infty, 0)$. Esto muestra claramente que se satisface sobre $\alpha(t)$ la definición 5.3 y así, $\mathbb{R}^+$ será una curva diferenciable. ■

Teorema 7.1.
Sea $A(z)$ una matriz de función analítica invertible de orden $n \times n$ sobre un intervalo real $\mathbb{G}$. Si los autovalores de $A(z)$ son positivos para todo $z \in \mathbb{G}$, entonces $A(z)$ tiene raíces m-ésimas analíticas para todo entero positivo m.

Demostración.
En virtud de que los autovalores de $A(z)$ son positivos para todo $z \in \mathbb{G}$; esto es, $\lambda_j(\mathbb{G}) \subset \mathbb{R}^+$, con $j = 1, \cdots, n$; entonces, por lo mostrado en el lema 7.1, dichos autovalores pertenecen a una curva diferenciable fija; luego, por la proposición 5.2 se tiene que los autovalores de $A(z)$ pueden ser enumerados tal que ellos sean funciones analíticas. Así, en virtud del teorema 5.1 existe una matriz de función analítica invertible $S(z)$ tal que $S^{-1}(z)A(z)S(z)$ resulta ser una matriz triangular superior para todo $z \in \mathbb{G}$; con lo cual, se puede suponer sin perdida de generalidad que la matriz $A(z)$, dada como hipótesis en el teorema, es una matriz triangular superior. Con lo anteriormente dicho, supóngase que $A(z)$ es de la forma:

$$A(z) = \begin{bmatrix} \lambda_1(z) & & & \\ & \lambda_2(z) & & * \\ & & & \\ & 0 & \ddots & \\ & & & \lambda_n(z) \end{bmatrix},$$

donde $\lambda_j(z) > 0$, para todo $z \in \mathbb{G}$ y toda $j = 1, \cdots, n$. Ahora bien, eligiendo las raíces m-ésimas positivas de las entradas diagonales de $A(z)$ se tiene que

$$\sum_{s=0}^{m-1} x_{ii}^{s}(z) x_{jj}^{m-1-s}(z) \neq 0, \quad \forall z \in \mathbb{G}.$$

Por tanto, en virtud del corolario 6.3 $A(z)$ tendrá raíces m-ésimas analíticas para todo entero positivo m. ■

4

Referencias

[1] Apostol. T. M, **Calculus vol** I **y** II. Editorial Reverté, S.A.

[2] Charris, J; De Castro, R Y Varela, J: Fundamentos del Analisis Complejo de una Variable Volumen 8. Editorial Guadalupe Ltda.

[3] E. A. Grove & G. Ladas, Introducction to Complex Variable, Houghton Miffin, 1974.

[4] Fraleigh Percy, A: Algebra Abstracta. Tercera edición. Editorial Addison Wessley Iberoamericana, S. A.

[5] J. H. M. Wedderburn. On matrices whose coefficients are functions of a single variable. Trans. of Amer. Math. Soc. 16:328-3332, 1915.

[6] L. Rodman. Matrix functions, in: Handbook of Algebra. Vol. 1 (1996), 117-154; Correction and Addition, Handbook of Algebra. Vol. 3,63. North-Holland, Amsterdam, 2003.

[7] Rodman, L y Spitkovsky, I. M: Analytic Roots of Invertible Matrix Funtions. Electronic Journal of Linear Algebra. Volume 13, pp. 187-196, August 2005.

[8] Serge, L: Algebra lineal. Segunda edición. Editorial Addison-Wessley Publishing Company.

[9] W. R. Derrick, Introductory Complex Analysis, and Applications, Academic Press, 1974.

Electronic Journal of Linear Algebra ISSN 1081-3810
A publication of the International Linear Algebra Society
Volume 13, pp. 187-196, August 2005

ELA

ANALYTIC ROOTS OF INVERTIBLE MATRIX FUNCTIONS*

LEIBA RODMAN† AND ILYA M. SPITKOVSKY†

Abstract. Various conditions are developed that guarantee existence of analytic roots of a given analytic matrix function with invertible values defined on a simply connected domain.

Key words. Analytic matrix functions, Roots.

AMS subject classifications. 47A56.

1. Introduction. Let G be a simply connected domain in the complex plain $\mathbb{C}$, or an open interval of the real line. By an analytic, resp., a meromorphic, matrix function in G we mean a matrix whose entries are (scalar) analytic, resp., meromorphic functions in G. Unless specified otherwise, G will denote a fixed simply connected domain, or a fixed real interval.

Let $A(z)$, $z \in G$ be an $n \times n$ matrix function analytic and invertible in G. In this paper we study analytic mth roots of A, that is, matrix functions B which are analytic in G and satisfy the equation $B(z)^m = A(z)$ for all $z \in G$. Here $m \geq 2$ is positive integer. Of course, it is a well known fact from complex analysis that for $n = 1$ there are exactly m analytic mth root functions.

However, in the matrix case not much is known. See, for example, [2]. This is somewhat surprising, especially because the problem is a natural one.

To start with, consider an example.

EXAMPLE 1.1. Let G be a simply connected domain such that $-1, 0 \in G$ and $-\frac{1}{2} \notin G$. Consider the analytic matrix function

$$A(z) = \begin{bmatrix} (z+\frac{1}{2})^2 & 1 \\ 0 & \frac{1}{4} \end{bmatrix}, \quad z \in G.$$

Clearly $A(z)$ is invertible on G. An easy algebraic computation shows that there are only four meromorphic (in G) functions $B(z)$ such that $B(z)^2 = A(z)$, given by the formula

$$B(z) = \begin{bmatrix} \delta_1(z+\frac{1}{2}) & (\delta_1(z+\frac{1}{2}) + \frac{1}{2}\delta_2)^{-1} \\ 0 & \frac{1}{2}\delta_2 \end{bmatrix}, \quad z \in G, \tag{1.1}$$

where δ_1 and δ_2 are independent signs ± 1. None of the four functions $B(z)$ is analytic in G. □

One can use functional calculus to compute square roots of an analytic invertible $n \times n$ matrix function $A(z)$, as follows. Fix $z_0 \in G$. For $z \in G$ sufficiently close to z_0,

*Received by the editors 21 February 2005. Accepted for publication 15 August 2005. Handling Editor: Harm Bart.

†Department of Mathematics, College of William and Mary, Williamsburg, Virginia 23187-8795, USA (lxrodm@math.wm.edu, ilya@math.wm.edu).

Electronic Journal of Linear Algebra ISSN 1081-3810
A publication of the International Linear Algebra Society
Volume 13, pp. 187-196, August 2005

ELA

use the formula

$$B(z) = \frac{1}{2\pi i}\int_\Gamma (\lambda I - A(z))^{-1}(\lambda)^{1/m} d\lambda, \tag{1.2}$$

where Γ is a simple closed rectifiable contour such that all eigenvalues of $A(z_0)$ are inside Γ, the origin is outside of Γ, and $(\lambda)^{1/m}$ is an analytic branch of the mth root function. Clearly, $B(z)^m = A(z)$. Formula (1.2) shows that locally, i.e., in a neighborhood of every given $z_0 \in G$, an analytic mth root of $A(z)$ always exists.

In the next section we treat the case of 2×2 analytic matrix functions and their analytic roots, using a direct approach. The general case of $n \times n$ matrix functions requires some preliminary results which are presented in Sections 3 and 4. Our main results are stated and proved in Section 5. Finally, in the last Section 6 we collect several corollaries concerning analytic roots of analytic matrix functions.

2. 2×2 matrix functions. We start with triangular 2×2 matrix functions. We say that the zeroes of a (scalar) analytic function $a(z)$, $z \in G$, are *majorized* by the zeroes of an analytic function $b(z)$, $z \in G$, if every zero z_0 of $a(z)$ is also a zero of $b(z)$, and the multiplicity of z_0 as a zero of $a(z)$ does not exceed the multiplicity of z_0 as a zero of $b(z)$, or in other words, if the quotient $b(z)/a(z)$ is analytic in G.

THEOREM 2.1. *A 2×2 analytic triangular matrix function*

$$A(z) = \begin{bmatrix} a_{11}(z) & a_{12}(z) \\ 0 & a_{22}(z) \end{bmatrix}, \quad z \in G \tag{2.1}$$

with diagonal entries non-zero everywhere in G admits an analytic square root if and only if at least one of the functions $\sqrt{a_{11}} + \sqrt{a_{22}}$ and $\sqrt{a_{11}} - \sqrt{a_{22}}$ has its zeroes in G majorized by the zeroes of a_{12}.

Here by $\sqrt{a_{11}}$ and $\sqrt{a_{22}}$ we understand either of the two branches of the analytic square root scalar function; the statement of the theorem does not depend on this choice.

Proof. The case when a_{12} is identically equal to zero is trivial: the condition on the majorization of zeroes holds automatically, and the square roots of A are delivered by the formula

$$B = \begin{bmatrix} \pm\sqrt{a_{11}} & 0 \\ 0 & \pm\sqrt{a_{22}} \end{bmatrix}$$

(this exhausts all the possibilities if the diagonal entries of A are not identical, and many more analytic square roots exist otherwise). So, it remains to consider the case of a_{12} not equal zero identically.

If $B(z) = [b_{ij}]_{i,j=1}^2$ is the analytic square root of $A(z)$ in G, then, in particular,

$$(b_{11} + b_{22})b_{12} = a_{12} \text{ and } (b_{11} + b_{22})b_{21} = 0. \tag{2.2}$$

Since a_{12} is not identically zero in G, neither is $b_{11} + b_{22}$. From the analyticity of functions involved it follows from the second equation in (2.2) that b_{21} is identically

Electronic Journal of Linear Algebra ISSN 1081-3810
A publication of the International Linear Algebra Society
Volume 13, pp. 187-196, August 2005

ELA

4t

zero. In other words, the analytic quadratic roots of the triangular matrix under consideration are triangular as well.

But an upper triangular B is a square root of A if and only if

$$b_{11}^2 = a_{11}, \; b_{22}^2 = a_{22},$$

and the first equation of (2.2) holds. The analytic square root of A therefore exists if and only if at least one of the equations

$$b_{12}(\sqrt{a_{11}} + \sqrt{a_{22}}) = a_{12}$$

and

$$b_{12}(\sqrt{a_{11}} - \sqrt{a_{22}}) = a_{12}$$

has an analytic in G solution b_{12}. This, in turn, is equivalent to the condition on the zeroes of the functions a_{12} and $\sqrt{a_{11}} \pm \sqrt{a_{22}}$ mentioned in the statement of the theorem. □

COROLLARY 2.2. *Let $A(z)$ be a 2×2 analytic matrix function on G. Assume that the eigenvalues $\lambda_1(z), \lambda_2(z)$ of $A(z)$, for every $z \in G$, can be enumerated so that $\lambda_1(z)$ and $\lambda_2(z)$ become analytic functions on G. If for at least one branch of the square root function we have*

$$\mathrm{tr}(A(z)) + 2\sqrt{\det(A(z))} \neq 0$$

for every $z \in G$, then there exists an analytic square root of $A(z)$.

For the proof use Theorem 2.1 and 4.1 (to be proved in a later section); the latter theorem allows us to assume that $A(z)$ is upper triangular.

3. Algebraic preliminaries.

THEOREM 3.1. *Let F be a field, and let $A = [a_{i,j}]_{i,j=1}^n$ be an $n \times n$ upper triangular matrix over the field. Assume that*

$$\mathrm{rank}(A - \lambda I) \geq n - 1 \qquad \forall \quad \lambda \in F. \tag{3.1}$$

Then every $n \times n$ matrix X over F such that $X^m = A$ for some positive integer m must be upper triangular.

Proof. Passing to the algebraic closure of F, we may assume that F is algebraically closed. Let X be such that $X^m = A$. By the spectral mapping theorem (which is easily seen using the Jordan form of X)

$$\{\lambda^m \; : \; \lambda \in \sigma(X)\} = \sigma(A).$$

In particular, there exists $\lambda_0 \in \sigma(X)$ such that $\lambda_0^m = a_{1,1}$. If x is a corresponding eigenvector of X, then

$$Ax = X^m x = \lambda_0^m x = a_{1,1} x.$$

Thus, x is an eigenvector of A corresponding to the eigenvalue $a_{1,1}$. But the condition (3.1) implies that A is nonderogatory: Only one eigenvector (up to a nonzero scalar

47

Electronic Journal of Linear Algebra ISSN 1081-3810
A publication of the International Linear Algebra Society
Volume 13, pp. 187-196, August 2005

ELA

multiple) for every eigenvalue. So x is a scalar multiple of $[1\ 0\ \dots\ 0]^T$, and therefore the first column of X (except possibly for the top entry) consists of zeroes. Now use induction on n to complete the proof. □

The same result holds for any polynomial equation

$$\sum_{j=1}^{m} c_j X^j = A,$$

rather than $X^m = A$.

Theorem 3.1 says nothing about existence of mth roots X of A. A necessary condition (under the hypotheses of Theorem 3.1) is obviously that mth roots of the diagonal elements of A exist in F. From now on we assume that F has characteristic zero (or more precisely that the characteristic of F does not divide m). It turns out that under this assumption, the necessary condition for existence of the mth roots of invertible nonderogatory matrices is also sufficient.

We proceed by induction on the size n of matrices. Let $X = [x_{i,j}]_{i,j=1}^n$ be an $n \times n$ upper triangular matrix: $x_{i,j} = 0$ if $i > j$. Then (using induction on m for example) one verifies that the $(1, n)$ entry of X^m has the form

$$[X^m]_{1,n} = \left(\sum_{j=0}^{m-1} x_{1,1}^j x_{n,n}^{m-1-j}\right) x_{1,n} + P_{m,n}(x_{1,1}, x_{1,2}, \dots, x_{1,n-1}, x_{2,2}, \dots, x_{2,n}, \dots x_{n,n}),$$

where

$$P_{m,n} = P_{m,n}(x_{1,1}, x_{1,2}, \dots, x_{1,n-1}, x_{2,2}, \dots, x_{2,n}, \dots x_{n,n})$$

is a certain homogeneous polynomial of degree m with integer coefficients of the $(n^2 + n - 2)/2$ variables $x_{i,j}$, $1 \le i \le j \le n$, $(i, j) \ne (1, n)$. For example,

$$P_{3,3} = (x_{1,1}x_{1,2} + x_{1,2}x_{2,2})x_{2,3} + x_{1,2}x_{2,3}x_{3,3}.$$

THEOREM 3.2. *Let $A = [a_{i,j}]_{i,j=1}^n$, $a_{i,j} \in F$, be an upper triangular invertible matrix, and assume that there exist mth roots $\sqrt[m]{a_{j,j}} \in F$, for $j = 1, 2, \dots, n$. Then there exist at least m distinct mth roots of A with entries in F.*

Proof. Select $\sqrt[m]{a_{j,j}}$ so that

$$a_{i,i} = a_{j,j} \implies \sqrt[m]{a_{i,i}} = \sqrt[m]{a_{j,j}}. \tag{3.2}$$

Construct the elements of the upper triangular $n \times n$ matrix $X = [b_{i,j}]_{i,j=1}^n$ such that $X^m = A$ by induction on $j - i$. For the base of induction, let

$$b_{i,i} = \sqrt[m]{a_{i,i}} \ne 0, \quad i = 1, 2, \dots, n. \tag{3.3}$$

48

Electronic Journal of Linear Algebra ISSN 1081-3810
A publication of the International Linear Algebra Society
Volume 13, pp. 187-196, August 2005

ELA

If $b_{i,j}$ with $j - i < k$ are already constructed, we let $b_{1,k+1}, b_{2,k+2}, \ldots, b_{n-k,n}$ be defined by the equalities

$$a_{1,k+1} = \left(\sum_{j=0}^{m-1} b_{1,1}^{j} b_{k+1,k+1}^{m-1-j} \right) b_{1,k+1} + P_{m,k+1}(b_{1,1}, b_{1,2}, \ldots, b_{1,k}, b_{2,2}, \ldots, b_{2,k+1}, \ldots b_{k+1,k+1}), \tag{3.4}$$

and so on, the last equality being

$$a_{n-k,n} = \left(\sum_{j=0}^{m-1} b_{n-k,n-k}^{j} b_{n,n}^{m-1-j} \right) b_{n-k,n} + P_{m,k+1}(b_{n-k,n-k}, b_{n-k,n-k+1}, \ldots, b_{n-k,n-1}, b_{n-k+1,n-k+1}, \ldots, b_{n-k+1,n}, \ldots b_{n,n}). \tag{3.5}$$

Condition (3.2) guarantees that

$$\sum_{j=0}^{m-1} b_{i,i}^{j} b_{j,j}^{m-1-j} \neq 0, \quad i,j = 1, \ldots, n,$$

and therefore equalities (3.4)–(3.5) can be uniquely solved for $b_{1,k+1}, \ldots, b_{n-k,n}$. The proof is completed. □

COROLLARY 3.3. *Let $A = [a_{i,j}]_{i,j=1}^{n}$, $a_{i,j} \in F$, be an upper triangular invertible matrix. Assume that there exist mth roots $\sqrt[m]{a_{j,j}} \in F$, for $j = 1, 2, \ldots, n$. Assume furthermore that the condition* (3.1) *is satisfied. Then there exist not more than m^n distinct mth roots of A with entries in F.*

For the proof combine Theorem 3.1 and (the proof of) Theorem 3.2.

4. Analytic preliminaries. It is well-known that eigenvalues of an analytic matrix function need not be analytic. More precisely, there need not exist an enumeration of the eigenvalues at each point that yields analytic functions on the whole domain, or even locally.

Under some additional hypotheses, the analyticity of eigenvalues can be guaranteed. For example, the well-known Rellich's theorem [6], [7], also [4, Chapter S6]), asserts that if G is a real interval and $A(z)$, $z \in G$, is a Hermitian valued analytic matrix function, then the eigenvalues of $A(z)$ are analytic. Results on triangularization of analytic matrix functions under certain additional conditions, see for example [3], also yield as a by-product analyticity of eigenvalues.

We formulate a general result on analyticity of eigenvalues. The result may be known, but we did not find a comparable statement in the literature. The simple connectedness of G is not needed here.

THEOREM 4.1. *Let G be a domain in* **C**, *or an interval in* **R**. *Let $A(z)$ be an analytic $n \times n$ matrix function on G. The following statements are equivalent:*

Electronic Journal of Linear Algebra ISSN 1081-3810
A publication of the International Linear Algebra Society
Volume 13, pp. 187-196, August 2005

(a) *The eigenvalues $\lambda_1(z), \ldots, \lambda_n(z)$ of $A(z)$ can be enumerated so that, for every $z \in G$, $\lambda_1(z), \ldots, \lambda_n(z)$ are analytic functions of $z \in G$.*
(b) *There exists an invertible analytic matrix function $S(z)$ such that $S(z)^{-1}A(z)S(z)$ is upper triangular for every $z \in G$.*

Proof. The implication (b) $\Longrightarrow$ (a) is obvious. Assume (a) holds. Let $\lambda_1(z)$ be an analytic eigenvalue of $A(z)$. Consider the analytic matrix function $B(z) := A(z) - \lambda_1(z)I$.

At this point we use the property (proved in [9]) that the ring $\mathcal{A}(G)$ of analytic functions on G is a *Smith domain.* Recall that a Smith domain is a commutative unital ring $\mathcal{R}$ without divisors of zero such that every matrix X with entries in $\mathcal{R}$ can be transformed via $X \longrightarrow EXF$, where E and F are invertible matrices over $\mathcal{R}$, to a diagonal form $\mathrm{diag}(x_1, \ldots, x_p)$ (possibly bordered by zero rows and/or columns), with $x_1, \ldots, x_p \in \mathcal{R}\setminus\{0\}$ such that x_j is divisible by x_{j+1} in $\mathcal{R}$, for $j = 1, 2, \ldots, p-1$. Therefore we have a representation of $B(z)$ in the form

$$B(z) = E(z)(\mathrm{diag}(x_1, \ldots, x_n))F(z),$$

where $E(z)$ and $F(z)$ are invertible analytic (in G) $n \times n$ matrix functions, and $x_1, \ldots, x_n$ are analytic scalar functions. Since $\det B(z) = 0$ for all $z \in G$, we must have that at least one of the functions $x_1, \ldots, x_n$ is identically zero. Say, $x_1 \equiv 0$. Then

$$F(z)A(z)F(z)^{-1} = F(z)B(z)F(z)^{-1} + \lambda_1(z)I = \begin{bmatrix} \lambda_1(z) & * \\ 0 & A_1(z) \end{bmatrix},$$

where $A_1(z)$ is an $(n-1) \times (n-1)$ analytic matrix function. It is easy to see that the statement (a) holds for $A_1(z)$ (since it holds for $A(z)$). Now we complete the proof by using induction on n, and by applying the induction hypothesis to $A_1(z)$. $\square$

A sufficient condition for analyticity of eigenvalues is that the eigenvalues are contained in a simple differentiable curve, such as the real line (if for example the matrix function is Hermitian valued) or the unit circle (if for example the matrix function is unitary valued). We quote a statement from [8] (Theorem 3.3 there).

PROPOSITION 4.2. *Let $A(z)$ be an $n \times n$ matrix function, analytic on a real interval G. If for every $z \in G$, the eigenvalues $\lambda_1(z), \ldots, \lambda_n(z)$ of $A(z)$ belong to a fixed differentiable curve, then $\lambda_1(z), \ldots, \lambda_n(z)$ can be enumerated so that they become analytic functions of $z \in G$.*

5. Main results. We state the main results of the paper.

THEOREM 5.1. *Let G be a domain in $\mathbf{C}$, and let $A(z)$, $z \in G$, be an invertible analytic matrix function. Assume that the eigenvalues of $A(z)$ can be enumerated so that they form analytic functions in G. Fix an integer $m \geq 2$. Then:*
(a) *There exist at least m distinct meromorphic $n \times n$ matrix functions $B_j(z)$, $j = 1, 2, \ldots, m$, such that $B_j(z)^m = A(z)$.*
(b) *If in addition, for every analytic function $\lambda(z)$ at least one of the $(n-1) \times (n-1)$ subdeterminants of $A(z) - \lambda(z)I$ is not identically zero, then there exist at most m^n distinct matrix functions $B_j(z)$ as in (a).*

50

Electronic Journal of Linear Algebra ISSN 1081-3810
A publication of the International Linear Algebra Society
Volume 13, pp. 187-196, August 2005

Proof. Applying Theorem 4.1 if necessary, we may reduce the general case to that of a triangular matrix function A. Letting the field of scalar meromorphic functions play the role of F, we then derive part (a) from Theorem 3.2. Now observe that for matrix functions A as in (b) condition (3.1) holds. Indeed, suppose that λ is a meromorphic on G function for which $A(z) - \lambda(z)I$ has rank smaller than $n-1$ everywhere on G. Then all the $(n-1) \times (n-1)$ subdeterminants of $A(z) - \lambda(z)I$ vanish on G, and $\lambda(z)$ (as an eigenvalue of $A(z)$) must be bounded together with A (and therefore analytic) on all domains lying strictly inside G. Thus, λ is analytic on G, which is not allowed by (b). It remains to invoke Corollary 3.3. □

The hypothesis in (b) is satisfied, for example, if $A(z)$ is a lower Hessenberg matrix with no identically zero elements on the superdiagonal.

We now turn our attention to analytic mth roots. First of all note that for every invertible $n \times n$ analytic matrix function the existence of an mth analytic root in neighborhood of every point $z_0 \in G$ is guaranteed (cf. Example 1.1). For the existence of analytic (in G) mth roots, it is clear from Theorem 5.1 that, assuming analyticity of eigenvalues, a sufficient condition would involve majorization relations between zeroes of certain analytic functions. In general, these relations are not very transparent; they are implicitly given in the proof of Theorem 3.2. We provide a full description of these relations for 2×2 matrix functions, and an inductive construction for the general $n \times n$ case.

THEOREM 5.2. *Let $A(z) = [a_{i,j}(z)]_{i,j=1}^n$, $a_{i,j}(z) = 0$ if $i > j$, be an upper triangular invertible matrix function, analytic in G. Define inductively the analytic functions $b_{i,j}(z)$, $1 \le i \le j \le n$, by the properties* (3.2) – (3.4), *under the assumptions that the zeroes of*

$$\sum_{j=0}^{m-1} b_{q,q}^j b_{q+k,q+k}^{m-1-j}$$

are majorized by the zeroes of

$$a_{q,q+k} - P_{m,k+1}(b_{q,q}, b_{q,q+1}, \dots, b_{q,q+k-1}, b_{q+1,q+1}, \dots, b_{q+1,q+k}, \dots b_{q+k,q+k}),$$

for $q = 1, 2, \dots, n-k$, and for $k = 1, 2, \dots, n-1$. Then the matrix $B(z) := [b_{i,j}(z)]_{i,j=1}^n$, where $b_{i,j}(z) = 0$ if $i > j$, is an analytic mth root of $A(z)$.

In particular:

COROLLARY 5.3. *Let $A(z) = [a_{i,j}(z)]_{i,j=1}^n$, $a_{i,j}(z) = 0$ if $i > j$, be an upper triangular invertible matrix function, analytic in G. If the analytic mth roots $b_{j,j}(z) = \sqrt[m]{a}_{j,j}(z)$, $(j = 1, 2, \dots, n)$ can be chosen so that*

$$a_{i,i} = a_{j,j} \implies \sqrt[m]{a_{i,i}} = \sqrt[m]{a_{j,j}} \tag{5.1}$$

and

$$\sum_{q=0}^{m-1} b_{i,i}^q(z) b_{j,j}^{m-1-q}(z) \neq 0 \quad \text{for all } z \in G, \quad 1 \le i < j \le n, \tag{5.2}$$

51

Electronic Journal of Linear Algebra ISSN 1081-3810
A publication of the International Linear Algebra Society
Volume 13, pp. 187-196, August 2005

ELA
www.math.technion.ac.il/iic/ela

then $A(z)$ *has an analytic mth root.*

For the proof, simply observe that (5.2) guarantees that the majorization of zeroes as required in Theorem 5.2 is satisfied trivially.

For 2×2 matrices, the result of Theorem 5.2 simplifies considerably, and yields the following generalization of Theorem 2.1:

THEOREM 5.4. *Let*

$$A(z) = \begin{bmatrix} a_{1,1}(z) & a_{1,2}(z) \\ 0 & a_{2,2}(z) \end{bmatrix}, \quad z \in G \tag{5.3}$$

be an analytic 2×2 *upper triangular matrix function with diagonal entries non-zero everywhere in* G. *Then* $A(z)$ *has an analytic mth root if and only if at least one of the* m^2 *functions*

$$\sum_{j=0}^{m-1} \left(\sqrt[m]{a_{1,1}}\right)^j \left(\sqrt[m]{a_{2,2}}\right)^{m-1-j},$$

where any of the m values of the mth root may be chosen independently for $\sqrt[m]{a_{1,1}}$ *and for* $\sqrt[m]{a_{2,2}}$, *has its zeroes in* G *majorized by the zeroes of* $a_{1,2}$.

6. Deductions from the main theorems. We collect in this section several results that can be proved using the main theorems of Section 5.

THEOREM 6.1. *Let* $A(z)$, $z \in G$, *be an invertible analytic* $n \times n$ *matrix, and assume that the number* r *of distinct eigenvalues of the matrix* $A(z)$ *is independent of* $z \in G$. *Then* $A(z)$ *has at least* m^r *distinct analytic mth roots.*

Proof. Since the number of distinct eigenvalues is constant, the eigenvalues may be enumerated so that they become analytic functions (indeed, at a branch point several generally distinct eigenvalues must coalesce). By Theorem 4.1 we may assume that $A(z)$ is upper triangular: $A(z) = [a_{i,j}(z)]_{i,j=1}^n$, where $a_{i,j}(z) = 0$ if $i > j$. Moreover, the proof of Theorem 4.1 shows that the diagonal elements $a_{j,j}$ (actually, the eigenvalues of $A(z)$) may be arranged in clusters:

$$\begin{gathered} a_{1,1}(z) \equiv a_{2,2}(z) \equiv \cdots \equiv a_{k_1,k_1}(z), \ldots, \\ \vdots \\ a_{k_{r-1}+1,k_{r-1}+1}(z) \equiv a_{k_{r-1}+2,k_{r-1}+2}(z) \equiv \cdots \equiv a_{k_r,k_r}(z). \end{gathered}$$

Here $k_1 < k_2 < \cdots < k_r = n$, and $a_{k_j,k_j}(z) \neq a_{k_p,k_p}(z)$ for $j \neq p$ and for all $z \in G$.

Next, we use the well-known property that a linear transformation of a rectangular matrix X defined by $X \mapsto QX - XR$, is invertible provided the square size matrices Q and R have no eigenvalues in common. We apply a suitable transformation of the form

$$A(z) \mapsto S(z)^{-1}A(z)S(z), \tag{6.1}$$

Electronic Journal of Linear Algebra ISSN 1081-3810
A publication of the International Linear Algebra Society
Volume 13, pp. 187-196, August 2005

where

$$S(z) = \begin{bmatrix} I_{k_1 \times k_1} & S_{1,2}(z) & S_{1,3}(z) & \cdots & S_{1,r}(z) \\ 0 & I_{(k_2-k_1)\times(k_2-k_1)} & S_{2,3}(z) & \cdots & S_{2,r}(z) \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & I_{(k_r-k_{r-1})\times(k_r-k_{r-1})} \end{bmatrix},$$

with a $(k_j - k_{j-1}) \times (k_q - k_{q-1})$ analytic matrix function $S_{j,q}(z)$ $(1 \le j < q \le r)$. (We formally put $k_0 = 0$.) A transformation (6.1) can be chosen in such a way that the resulting analytic matrix function

$$B(z) := S(z)^{-1} A(z) S(z)$$

is block diagonal:

$$B(z) = \mathrm{diag}(B_1(z), B_2(z), \ldots, B_r(z)),$$

where $B_j(z)$ is a $(k_j - k_{j-1}) \times (k_j - k_{j-1})$ analytic upper triangular matrix function with $a_{k_{j-1}+1}(z), \ldots, a_{k_j}(z)$ on the main diagonal. Since the functions on the main diagonal of $B_j(z)$ are identical, Corollary 5.3 is applicable, and the result follows. □

It is well-known that if $A(z)$ is an analytic invertible matrix function on a real interval G, then $A(z)$ has an analytic polar decomposition and an analytic singular value decomposition (see, for example, [1], [5]). Both properties follow easily from the fact (which can be deduced from Rellich's theorem, [6], [7]) that the positive definite analytic matrix function $A(z)^* A(z)$ has a positive definite analytic square root. Using Theorem 5.2, a more general statement can be obtained:

THEOREM 6.2. *Let $A(z)$ be $n \times n$ analytic invertible matrix function on a real interval G. If the eigenvalues of $A(z)$ are positive for every $z \in G$, then $A(z)$ has analytic mth roots for every positive integer m.*

Proof. By Proposition 4.2, the eigenvalues of $A(z)$ can be chosen analytic. By Theorem 4.1, we may assume that $A(z)$ is upper triangular. Taking positive mth roots of the diagonal entries of $A(z)$, we see that the hypotheses of Corollary 5.3 are satisfied. An application of Corollary 5.3 completes the proof. □

Acknowledgment. The research was partially supported by NSF grant DMS-0456625. The research of the first author was also partially supported by the Faculty Research Assignment grant from the College of William and Mary.

REFERENCES

[1] A. Bunse-Gerstner, R. Byers, V. Mehrmann, and N. K. Nichols. Numerical computation of an analytic singular value decomposition of a matrix valued function. *Numer. Math.*, 60:1–39, 1991.

[2] H. Flanders. Analytic solutions of matrix equations. *Linear Multilinear Algebra*, 2:241–243, 1974/1975.

[3] H. Gingold and P.-F. Hsieh. Globally analytic triangularization of a matrix function. *Linear Algebra Appl.*, 169:75–101, 1992.

53

Electronic Journal of Linear Algebra ISSN 1081-3810
A publication of the International Linear Algebra Society
Volume 13, pp. 187-196, August 2005

ELA

[4] I. Gohberg, P. Lancaster, and L. Rodman. *Matrix polynomials*. Academic Press, New York, 1982.

[5] V. Mehrmann and W. Rath. Numerical methods for the computation of analytic singular value decompositions. *Electron. Trans. Numer. Anal.*, 1:72–88, 1993.

[6] F. Rellich. Störungstheorie der Spektralzerlenung, I. *Math. Ann.* 113:600–619, 1936.

[7] T. Kato. *Perturbation theory for linear operators*, 2nd edition. Springer-Verlag, Berlin-New York, 1976.

[8] L. Rodman. Matrix functions, in: *Handbook of Algebra*. Vol. 1 (1996), 117–154; Correction and Addition, *Handbook of Algebra*. Vol. 3, 63, North-Holland, Amsterdam, 2003

[9] J.H.M. Wedderburn. On matrices whose coefficients are functions of a single variable. *Trans. of Amer. Math. Soc.* 16:328-332, 1915.

6245