

Forma de Jordan de la derivada de Fréchet de funciones matriciales

MIGUEL A. MARMOLEJO*

Universidad del Valle, Departamento de Matemáticas, Cali, Colombia.

Resumen. En este artículo se presenta una fórmula para evaluar funciones matriciales $f : \mathcal{A} \subset \mathbb{C}^{2 \times 2} \rightarrow \mathbb{C}^{2 \times 2}$, en términos de dos funciones escalares que sólo dependen de la traza y el determinante de $X \in \mathbb{C}^{2 \times 2}$. Se explota el conocimiento de las derivadas de Fréchet de las funciones traza y determinante para determinar la derivada de Fréchet de $f(\cdot)$. Como resultado central, se da la forma canónica de Jordan de la derivada de Fréchet $Df(X) : \mathbb{C}^{2 \times 2} \rightarrow \mathbb{C}^{2 \times 2}$.

Palabras clave: Función matricial, forma canónica de Jordan, derivada de Fréchet.

MSC2010: 15A16, 15A21, 15A24, 47A56.

Jordan form of the Fréchet derivative of matrix functions

Abstract. In this paper we present a formula to evaluate matrix functions $f : \mathcal{A} \subset \mathbb{C}^{2 \times 2} \rightarrow \mathbb{C}^{2 \times 2}$, in terms of two scalar functions that only depend on the trace and the determinant of $X \in \mathbb{C}^{2 \times 2}$. The knowledge of the Fréchet derivatives of the trace and determinant functions is used to determine the Fréchet derivative of $f(\cdot)$. As a central result, Jordan's canonical form of the Fréchet derivative $Df(X) : \mathbb{C}^{2 \times 2} \rightarrow \mathbb{C}^{2 \times 2}$ is given.

Keywords: Matrix function, Jordan canonical form, Fréchet derivative.

1. Introducción

La norma de la derivada de Fréchet de una función matricial $f : \mathcal{A} \subset \mathbb{C}^{n \times n} \rightarrow \mathbb{C}^{n \times n}$ aparece explícitamente en una expresión que da el número de condición relativo de f en X . De manera precisa (ver la Sección 3.1 de Higham [4], o la Sección 3.3 de Higham and Lijing [6]):

$$\text{cond}(f, X) = \frac{\|Df(X)\| \|X\|}{\|f(X)\|},$$

* E-mail: miguel.marmolejo@correounivalle.edu.co

Recibido: 14 de julio del 2017, Aceptado: 2 de abril del 2018.

Para citar este artículo: M.A. Marmolejo, Forma de Jordan de la derivada de Fréchet de funciones matriciales, *Rev. Integr. temas mat.* 36 (2018), No. 1, 1–19. doi: 10.18273/revint.v36n1-2018001.

siendo $\|\cdot\|$ una norma matricial y $\|Df(X)\| := \max_{H \neq 0} \frac{\|[Df(X)](H)\|}{\|H\|}$. Este número mide la sensibilidad de $f(X)$ a pequeños cambios en X . Hay trabajos recientes dedicados a la estimación de cotas para este número de condición en situaciones particulares; ver por ejemplo, Cardoso y Sadeghi [1], Deadman y Relton [2] o Kandolf y Relton [8], entre otros.

Es entonces conveniente disponer de expresiones concretas de la derivada de Fréchet de tales funciones. En este contexto, el problema de investigación 3.11 en Higham [4] consiste en determinar la forma de Jordan de la derivada de Fréchet de funciones matriciales $X \rightarrow f(X)$; $X \in \mathcal{A} \subset \mathbb{C}^{n \times n}$, en términos de la de X . Cuando X es diagonalizable, el Corolario 3.12 en Higham [4] implica que la forma canónica de Jordan de dicha derivada es también diagonal (esto se demuestra en la Sección 6 para funciones analíticas). En este artículo se presenta una fórmula para evaluar funciones de matrices 2×2 , la cual se usa para determinar su derivada de Fréchet. La principal contribución de este trabajo consiste en dar la forma canónica de Jordan de la derivada de Fréchet de funciones matriciales en el caso $n = 2$.

El resto del trabajo está organizado como sigue. En la Sección 2 se dan dos definiciones de la matriz $f(X)$; $X \in \mathbb{C}^{n \times n}$. Para matrices 2×2 se introducen dos funciones de valor escalar, que serán fundamentales para el desarrollo del resto del trabajo. En la Sección 3 se deduce una expresión para $f(X)$, $X \in \mathbb{C}^{2 \times 2}$ (Teorema 3.1); como aplicación, se dan condiciones necesarias y suficientes para que $f(X)g(Y) = g(Y)f(X)$, se presentan fórmulas explícitas para calcular las matrices e^X , $\text{sen}(X)$, $\text{cos}(X)$ y $X^{1/2}$ y, finalmente, se considera la evaluación de funciones de algunas matrices particionadas. La Sección 4 se dedica a la derivada de Fréchet de funciones $X \rightarrow f(X)$; $X \in \mathcal{A} \subset \mathbb{C}^{2 \times 2}$ y, la Sección 5, a dar su forma canónica de Jordan (Teorema 5.1). Por último, en la Sección 6, se considera la derivada de Fréchet de funciones matriciales en el caso en que $f(\cdot)$ es analítica y $X \in \mathbb{C}^{n \times n}$; se demuestra que si X es diagonalizable, entonces la forma canónica de Jordan de la derivada de Fréchet es diagonal (Teorema 6.3).

2. Preliminares

Se establecen aquí la terminología, las definiciones y los resultados básicos, que se utilizarán a lo largo del trabajo.

2.1. Definiciones de $f(X)$

Hay varias maneras equivalentes de definir $f(X)$, $X \in \mathbb{C}^{n \times n}$ (ver, por ejemplo, Rinehart [10], el Capítulo 6 de Horn y Johnson [7], el Capítulo 11 de Golub y Van Loan [3] o el Capítulo 1 de Higham [4]). En este artículo presentamos dos definiciones; la primera, en términos de la forma canónica de Jordan y, la segunda, en términos de la integral de Cauchy.

Supóngase que los distintos valores propios de $X \in \mathbb{C}^{n \times n}$ son $\lambda_1, \lambda_2, \dots, \lambda_k$, es decir, $\sigma(X) = \{\lambda_1, \lambda_2, \dots, \lambda_k\}$ es el espectro de X , que $m_X(\lambda) = (\lambda - \lambda_1)^{r_1}(\lambda - \lambda_2)^{r_2} \dots (\lambda - \lambda_k)^{r_k}$ es su polinomio mínimo, y que X tiene forma canónica de Jordan $J = P^{-1}XP$, donde

$$J = \text{diag}(J_1, J_2, \dots, J_p),$$

$$J_i = J_i(\lambda_i) = \begin{bmatrix} \lambda_i & 1 & & & \\ & \lambda_i & \ddots & & \\ & & \ddots & \ddots & \\ & & & \ddots & 1 \\ & & & & \lambda_i \end{bmatrix} \in \mathbb{C}^{m_i \times m_i}; \quad i = 1, 2, \dots, p,$$

y $m_1 + m_2 + \dots + m_p = n$. Aunque la matriz P no es única, la matriz de Jordan J es única, salvo el orden en que aparecen los bloques en su diagonal.

Supóngase además que $f(t)$ es una función con valores escalares, de variable real o compleja t . Con estas notaciones se establecen las siguientes definiciones (ver Higham [4], Definiciones 1.1 y 1.2, p. 3).

Definición 2.1. Se dice que f está definida en el espectro de X , si existen los valores

$$f^{(j)}(\lambda_i); \quad i = 1, 2, \dots, k; \quad j = 1, 2, \dots, r_i,$$

denominados valores de f en el espectro de X . Aquí, $f^{(j)}(\cdot)$ indica la derivada j -ésima de f .

Definición 2.2. Supóngase que f está definida en el espectro de X . Se define $f(X) := P \text{diag}(f(J_1), f(J_2), \dots, f(J_p)) P^{-1}$, donde el bloque diagonal $f(J_i)$ es la matriz $\mathbb{C}^{m_i \times m_i}$ dada por

$$f(J_i(\lambda_i)) = \begin{bmatrix} f(\lambda_i) & f'(\lambda_i) & \frac{1}{2}f''(\lambda_i) & \cdot & \cdot & \frac{1}{(m_i-1)!}f^{(m_i-1)}(\lambda_i) \\ 0 & f(\lambda_i) & f'(\lambda_i) & \cdot & \cdot & \frac{1}{(m_i-2)!}f^{(m_i-2)}(\lambda_i) \\ \cdot & \cdot & \cdot & \ddots & \ddots & \cdot \\ \cdot & \cdot & \cdot & \ddots & \ddots & \cdot \\ 0 & 0 & 0 & \cdot & \cdot & f(\lambda_i) \end{bmatrix}.$$

La matriz $f(X)$ no depende de la forma de Jordan usada (ver Horn y Johnson [7], Teorema 6.2.9-(b), p. 412). Para funciones multivaluadas como $f(t) = t^{1/2}$ y $f(t) = \log(t)$, se sobreentiende que sólo una rama se escoge en la evaluación de los bloques correspondientes a un mismo λ_i ; $i = 1, 2, \dots, k$ (ver el Ejemplo 4 del Apartado 3.5). En este caso, $f(X)$ se denomina función matricial primaria.

La definición más corta y elegante de función matricial es la que sigue (Definición 1.11, p. 8, en Higham [4]).

Definición 2.3. Sea $f(t)$ una función analítica dentro y sobre un contorno cerrado Γ que encierra a λ_i ; $i = 1, 2, \dots, k$. Se define

$$f(X) := \frac{1}{2\pi i} \int_{\Gamma} f(z)(zI - X)^{-1} dz.$$

Cuando se pueden aplicar, las definiciones anteriores son equivalentes (ver Horn y Johnson [7], Teorema 6.2.28, p. 427, o Rinehart [10], Teorema 1, p. 405).

Propiedades generales sobre $f(X)$ pueden ser consultadas en Horn y Johnson [7] (Teorema 6.2.9, p. 412), o en Higham [4] (Sección 1.3). Se resaltan las siguientes, que serán usadas en la Sección 3.

Proposición 2.4. *Suponga que f está definida en el espectro de X y que Q es una matriz invertible. Entonces,*

1. $f(Q^{-1}XQ) = Q^{-1}f(X)Q$;
2. Si $X = \text{diag}(X_{11}, X_{22}, \dots, X_{mm})$ es diagonal por bloques, entonces

$$f(X) = \text{diag}(f(X_{11}), f(X_{22}), \dots, f(X_{mm})).$$

2.2. Las funciones escalares $\eta, \xi : \mathbb{C}^{2 \times 2} \rightarrow \mathbb{C}$

En lo que sigue, $\mathbb{C}^{2 \times 2}$ denota el espacio de Banach de las matrices complejas 2×2 con la norma de Frobenius. La traza, el determinante y la matriz adjunta de $X \in \mathbb{C}^{2 \times 2}$ se denotarán por $\text{tr}(X)$, $\det(X)$ y $\text{adj}(X)$, respectivamente. Cuando existe, la derivada de Fréchet de una función $k : \mathbb{C}^{2 \times 2} \rightarrow \mathbb{C}$ en el punto $X \in \mathbb{C}^{2 \times 2}$ es la única transformación lineal $[Dk(X)] : \mathbb{C}^{2 \times 2} \rightarrow \mathbb{C}$ tal que para todo $H \in \mathbb{C}^{2 \times 2}$ se cumple

$$k(X + H) = k(X) + [Dk(X)](H) + o(\|H\|).$$

Considérense ahora las funciones $\eta, \xi : \mathbb{C}^{2 \times 2} \rightarrow \mathbb{C}$ definidas por

$$\eta(X) = \frac{\text{tr}(X)}{2}, \quad \xi(X) = \eta^2(X) - \det(X).$$

Se sabe que η es lineal y que para cada matriz invertible Q se cumplen $\eta(X) = \eta(Q^{-1}XQ)$ y $\xi(X) = \xi(Q^{-1}XQ)$. De inmediato se verifican las identidades

$$\xi(X)I = (X - \eta(X)I)^2, \tag{1}$$

$$\xi(tX) = t^2\xi(X), \tag{2}$$

donde I es la matriz idéntica y $t \in \mathbb{C}$. Nótese que los valores propios de X son precisamente $\eta(X) + \sqrt{\xi(X)}$ y $\eta(X) - \sqrt{\xi(X)}$. Aquí, $\sqrt{\xi(X)}$ indica la raíz cuadrada principal de $\xi(X)$.

De otra parte, es bien conocido que las funciones $X \rightarrow \text{tr}(X)$ y $X \rightarrow \det(X)$ son continuamente diferenciables según Fréchet, y que para $H \in \mathbb{C}^{2 \times 2}$ se cumplen las igualdades

$$[D\text{tr}(X)](H) = \text{tr}(H), \quad [D\det(X)](H) = \text{tr}(H\text{adj}(X)).$$

De esto se sigue que las funciones $X \rightarrow \eta(X)$ y $X \rightarrow \xi(X)$ son continuamente diferenciables según Fréchet, y que para $H \in \mathbb{C}^{2 \times 2}$,

$$\begin{aligned} [D\eta(X)](H) &= \eta(H), \\ [D\xi(X)](H) &= 2\eta(X)\eta(H) - \text{tr}(H\text{adj}(X)). \end{aligned} \tag{3}$$

También, de la ecuación (1) se obtiene

$$\begin{aligned} [D\xi(X)I](H) &= [D\xi(X)](H)I \\ &= (X - \eta(X)I)(H - \eta(H)I) \\ &\quad + (H - \eta(H)I)(X - \eta(X)I). \end{aligned} \tag{4}$$

3. Funciones matriciales para matrices 2×2

En esta sección se muestra que si $f(\cdot)$ está definida en el espectro de $X \in \mathbb{C}^{2 \times 2}$, entonces $f(X)$ es un polinomio de la forma

$$f(X) = \alpha_f(X)I + \beta_f(X)(X - \eta(X)I).$$

Esta expresión se usa para establecer algunas propiedades de $f(X)$ y para dar fórmulas explícitas a fin de calcular las matrices e^X , $\text{sen}(X)$, $\text{cos}(X)$ y $X^{1/2}$. Por último, se muestra la utilidad de esta expresión en la evaluación de funciones de algunas matrices particionadas.

De acuerdo con lo establecido en la Sección 2, los valores propios de una matriz $X \in \mathbb{C}^{2 \times 2}$ son $\lambda_1 = \eta(X) + \sqrt{\xi(X)}$ y $\lambda_2 = \eta(X) - \sqrt{\xi(X)}$. Sea $J = P^{-1}XP$ la forma canónica de Jordan de X . Si se describen P por columnas y P^{-1} por filas, entonces

$$P = [u_1 \quad u_2], \quad P^{-1} = \begin{bmatrix} v_1^T \\ v_2^T \end{bmatrix}, \quad I = u_1v_1^T + u_2v_2^T = \begin{bmatrix} v_1^T u_1 & v_1^T u_2 \\ v_2^T u_1 & v_2^T u_2 \end{bmatrix}.$$

Utilizando la Definición 2.2, a continuación se da una expresión para $f(X)$. Se escribe $\eta \equiv \eta(X)$ y $\xi \equiv \xi(X)$.

Caso 1. $J = \text{diag}(\lambda_1, \lambda_2)$; $\lambda_1 \equiv \eta + \sqrt{\xi} \neq \eta - \sqrt{\xi} \equiv \lambda_2$, i.e., $\xi \neq 0$. En este caso, $X = \lambda_1 u_1 v_1^T + \lambda_2 u_2 v_2^T$, $f(X) = P \text{diag}(f(\lambda_1), f(\lambda_2)) P^{-1}$, y se puede escribir

$$\begin{aligned} f(X) &= f(\lambda_1)u_1v_1^T + f(\lambda_2)u_2v_2^T \\ &= \frac{f(\lambda_1) + f(\lambda_2)}{2}I + \frac{f(\lambda_1) - f(\lambda_2)}{2(\lambda_1 - \lambda_2)} \left[X - \frac{(\lambda_1 + \lambda_2)}{2}I \right] \\ &= \frac{f(\eta + \sqrt{\xi}) + f(\eta - \sqrt{\xi})}{2}I + \frac{f(\eta + \sqrt{\xi}) - f(\eta - \sqrt{\xi})}{2\sqrt{\xi}} \left[X - \eta I \right]. \end{aligned}$$

Caso 2. $J = \lambda I \equiv \eta I$; $\xi = 0$. De inmediato se ve que $f(X) = f(\lambda)I \equiv f(\eta)I$.

Caso 3. $J = \begin{bmatrix} \lambda & 1 \\ 0 & \lambda \end{bmatrix} \equiv \begin{bmatrix} \eta & 1 \\ 0 & \eta \end{bmatrix}$; $\xi = 0$. En este caso, $X = \eta I + u_1 v_2^T$ y $f(X) = P \begin{bmatrix} f(\eta) & f'(\eta) \\ 0 & f(\eta) \end{bmatrix} P^{-1}$. Por tanto,

$$f(X) = f(\eta)I + f'(\eta)u_1v_2^T = f(\eta)I + f'(\eta)(X - \eta I).$$

En resumen, se tiene el siguiente teorema.

Teorema 3.1. Si $f(\cdot)$ está definida en el espectro de $X \in \mathbb{C}^{2 \times 2}$, entonces

$$\begin{aligned} f(X) &= \alpha(X)I + \beta(X)(X - \eta(X)I) \\ &\equiv \alpha_f(X)I + \beta_f(X)(X - \eta(X)I), \end{aligned} \tag{5}$$

donde

$$\alpha(X) \equiv \alpha_f(X) = \frac{f(\eta(X) + \sqrt{\xi(X)}) + f(\eta(X) - \sqrt{\xi(X)})}{2},$$

$$\beta(X) \equiv \beta_f(X) = \begin{cases} \frac{f(\eta(X) + \sqrt{\xi(X)}) - f(\eta(X) - \sqrt{\xi(X)})}{2\sqrt{\xi(X)}}, & \xi(X) \neq 0, \\ f'(\eta(X)), & \xi(X) = 0. \end{cases}$$

Cuando sea claro el contexto, se escribe $\alpha(X)$ y $\beta(X)$, en lugar de $\alpha_f(X)$ y $\beta_f(X)$.

Observación 3.2. De la Fórmula (5) se siguen fácilmente los siguientes hechos.

1. Si $f(X)$ es invertible, entonces

$$[f(X)]^{-1} = \frac{1}{\alpha^2(X) - \xi(X)\beta^2(X)} [\alpha(X)I - \beta(X)(X - \eta(X)I)].$$

2. Para cada $t \in \mathbb{C}$ se verifican las relaciones $\eta(X + tI) = \eta(X) + t$ y $\xi(X + tI) = \xi(X)$. Por tanto, si $f(\cdot)$ y $f'(\cdot)$ son funciones t -periódicas, entonces $f(X + tI) = f(X)$, siempre que estas matrices estén definidas. Por ejemplo: $\text{sen}(X + 2\pi I) = \text{sen}(X)$, $\text{cos}(X + 2\pi I) = \text{cos}(X)$ y $e^{X+2\pi iI} = e^X$.
3. Como $\eta(-X) = -\eta(X)$ y $\xi(-X) = \xi(X)$, si para cada t en el dominio de $f(\cdot)$ se cumple $f(-t) = \pm f(t)$, entonces $f(X) = \pm f(X)$.
4. Sean \mathcal{O} un subconjunto abierto de \mathbb{R} o \mathbb{C} y \mathcal{A} el conjunto abierto de las X tales que $\eta(X) + \sqrt{\xi(X)}$ y $\eta(X) - \sqrt{\xi(X)}$ están en \mathcal{O} . Si $f(\cdot)$ es continuamente diferenciable en \mathcal{O} , entonces las funciones $X \rightarrow \alpha_f(X)$ y $X \rightarrow \beta_f(X)$, $X \in \mathcal{A}$, son continuas. Resulta entonces que la función $X \rightarrow f(X)$ es continua en \mathcal{A} (ver también Horn y Johnson [7], Teorema 6.2.27-(1), p.425).

3.1. Conmutatividad

Sean $X, Y \in \mathbb{C}^{2 \times 2}$.

1. Sean $f(\cdot)$ y $g(\cdot)$ funciones tales que $f(X)$ y $g(Y)$ están definidas. Escribiendo

$$f(X) = \alpha_f(X)I + \beta_f(X)(X - \eta(X)I),$$

$$g(Y) = \alpha_g(Y)I + \beta_g(Y)(Y - \eta(Y)I),$$

se deduce que $f(X)g(Y) - g(Y)f(X) = \beta_f(X)\beta_g(Y)[XY - YX]$. De esto se sigue que $f(X)g(Y) = g(Y)f(X)$ si y sólo si $\beta_f(X) = 0$ o $\beta_g(Y) = 0$ o $XY = YX$. Nótese que cuando $\beta_f(X) = 0$ se obtiene $f(X) = \alpha_f(X)I$, que es un múltiplo de la matriz identidad. Ahora,

$$\beta_f(X) = 0 \Leftrightarrow \begin{cases} f(\eta(X) + \sqrt{\xi(X)}) = f(\eta(X) - \sqrt{\xi(X)}), & \xi(X) \neq 0, \\ f'(\eta(X)) = 0, & \xi(X) = 0. \end{cases}$$

2. Puesto que $\eta(XY) = \eta(YX)$ y $\xi(XY) = \xi(YX)$, si $f(XY)$ está definida, entonces $f(YX)$ está definida, y

$$f(XY) - f(YX) = \beta(XY)[XY - YX].$$

Por tanto $f(XY) = f(YX)$ si y sólo si, $\beta(XY) = 0$ ó $XY = YX$. También (confróntese con el Corolario 1.34, p. 21, de Higham [4]),

$$\begin{aligned} Yf(XY) &= Y[\alpha(XY)I + \beta(XY)(XY - \eta(XY)I)] \\ &= [\alpha(YX)I + \beta(YX)(YX - \eta(YX)I)]Y \\ &= f(YX)Y. \end{aligned}$$

3.2. Ejemplos

En este apartado se usa la Fórmula (5) para dar formas explícitas para evaluar e^X , $\text{sen}(X)$, $\cos(X)$ y $X^{1/2}$; $X \in \mathbb{C}^{2 \times 2}$. Propiedades generales sobre estas matrices aparecen en Higham [4], Golub y Van Loan [3] y Horn y Johnson [7], entre otros. Métodos numéricos para evaluar funciones matriciales pueden consultarse en Higham y Al-Mohy [5], así como las referencias ahí citadas. En estos ejemplos se escribe $\eta \equiv \eta(X)$, $\xi \equiv \xi(X)$.

Ejemplo 3.3. La función $X \rightarrow e^X$,

$$e^X = \begin{cases} e^\eta \left[\cosh(\sqrt{\xi})I + \frac{\text{senh}(\sqrt{\xi})}{\sqrt{\xi}}(X - \eta I) \right], & \xi \neq 0, \\ e^\eta \left[I + (X - \eta I) \right], & \xi = 0. \end{cases}$$

Cuando X es real, i.e., $X \in \mathbb{R}^{2 \times 2}$, entonces e^X es real. En efecto, usando las identidades $\cosh(ix) = \cos(x)$ y $\text{senh}(ix) = i \text{sen}(x)$, se obtiene

$$e^X = \begin{cases} e^\eta \left[\cosh(\sqrt{\xi})I + \frac{\text{senh}(\sqrt{\xi})}{\sqrt{\xi}}(X - \eta I) \right], & \xi > 0, \\ e^\eta \left[I + (X - \eta I) \right], & \xi = 0, \\ e^\eta \left[\cos(\sqrt{-\xi})I + \frac{\text{sen}(\sqrt{-\xi})}{\sqrt{-\xi}}(X - \eta I) \right], & \xi < 0. \end{cases}$$

Ejemplo 3.4. La función $X \rightarrow \text{sen}(X)$,

$$\text{sen}(X) = \begin{cases} \text{sen}(\eta)\cos(\sqrt{\xi})I + \frac{\cos(\eta)\text{sen}(\sqrt{\xi})}{\sqrt{\xi}}(X - \eta I), & \xi \neq 0, \\ \text{sen}(\eta)I + \cos(\eta)(X - \eta I), & \xi = 0. \end{cases}$$

Si X es real, i.e., $X \in \mathbb{R}^{2 \times 2}$, entonces $\text{sen}(X)$ es real. En efecto, las identidades $\cos(ix) = \text{cosh}(x)$ y $\text{sen}(ix) = i \text{senh}(x)$, permiten escribir la correspondiente expresión real.

Ejemplo 3.5. La función $X \rightarrow \cos(X)$,

$$\cos(X) = \begin{cases} \cos(\eta)\cos(\sqrt{\xi})I - \frac{\sin(\eta)\sin(\sqrt{\xi})}{\sqrt{\xi}}(X - \eta I), & \xi \neq 0, \\ \cos(\eta)I - \sin(\eta)(X - \eta I), & \xi = 0. \end{cases}$$

Cuando X es real, $\cos(X)$ es real. Para dar la fórmula correspondiente, basta usar las identidades $\cos(ix) = \cosh(x)$ y $\sin(ix) = i \sinh(x)$.

Utilizando las fórmulas anteriores, cálculos directos permiten comprobar las identidades $\sin^2(X) + \cos^2(X) = I$ y $\cos(X) + i \sin(X) = e^{iX}$. Si además se tiene en cuenta que $\eta(2X) = 2\eta(X)$ y $\xi(2X) = 4\xi(X)$, entonces es fácil comprobar las identidades $\cos^2(X) - \sin^2(X) = \cos(2X)$ y $2\sin(X)\cos(X) = \sin(2X)$.

Ejemplo 3.6. La función $X \rightarrow X^{1/2}$.

Si $X = 0$, entonces $X^{1/2} = 0$. Para $X \neq 0$ se tiene la siguiente expresión general:

$$X^{1/2} = \begin{cases} \frac{(\eta + \sqrt{\xi})^{1/2} + (\eta - \sqrt{\xi})^{1/2}}{2}I + \frac{(\eta + \sqrt{\xi})^{1/2} - (\eta - \sqrt{\xi})^{1/2}}{2\sqrt{\xi}}(X - \eta I), & \xi \neq 0, \\ \eta^{1/2}I + \frac{1}{2\eta^{1/2}}(X - \eta I), & \xi = 0; \eta \neq 0. \end{cases}$$

Para $X = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}$, $\eta = \xi = 0$, $X^{1/2}$ no está definida. Cuando $X \in \mathbb{R}^{2 \times 2}$ no tiene valores propios reales negativos, entonces $X^{1/2}$ toma valores reales. En efecto, con esta condición

$$X^{1/2} = \begin{cases} \pm \left[\sqrt{\eta}I + \frac{1}{2\sqrt{\eta}}(X - \eta I) \right], & \xi = 0; \eta > 0, \\ \pm \left[\frac{\sqrt{\eta + \sqrt{\xi}} \pm \sqrt{\eta - \sqrt{\xi}}}{2}I + \frac{\sqrt{\eta + \sqrt{\xi}} \mp \sqrt{\eta - \sqrt{\xi}}}{2\sqrt{\xi}}(X - \eta I) \right], & \xi > 0; \eta - \sqrt{\xi} > 0, \\ \pm \left[aI + \frac{1}{2a}(X - (a^2 - b^2)I) \right], & \xi < 0, \end{cases}$$

donde, en el caso $\xi < 0$ (X tiene dos valores propios complejos conjugados), se ha escrito: $(\eta + \sqrt{\xi})^{1/2} = a + ib$; $a, b \in \mathbb{R}$, $(\eta - \sqrt{\xi})^{1/2} = a - ib$; $\eta = a^2 - b^2$ y $\sqrt{\xi} = 2abi$. Además, en este caso, $X^{1/2}$ toma los valores complejos $\pm i \left[bI - \frac{1}{2b}(X - (a^2 - b^2)I) \right]$ (Confrontar con el Lema 6.4 de Higham [4]).

3.3. Matrices particionadas

En este apartado se muestra que la Fórmula (5) sirve para el cálculo de funciones de matrices particionadas con estructura especial. En particular, si $X = \text{diag}(X_{11}, X_{22}, \dots, X_{mm})$ y las matrices X_{ii} , $i = 1, 2, \dots, m$ tienen tamaño menor que 3, entonces el cálculo de $f(X)$ es inmediato, pues, por la Proposición 2.4,

$$f(X) = \text{diag}(f(X_{11}), f(X_{22}), \dots, f(X_{mm})).$$

Supóngase ahora que $X, Y \in \mathbb{C}^{2 \times 2}$.

1. Si $M = \begin{bmatrix} X & Y \\ Y & X \end{bmatrix}$ y $P = P^{-1} = \frac{1}{\sqrt{2}} \begin{bmatrix} I & I \\ I & -I \end{bmatrix}$, entonces

$$P^{-1}MP = \begin{bmatrix} X+Y & 0 \\ 0 & X-Y \end{bmatrix}.$$

La Proposición 2.4 conduce a

$$\begin{aligned} f(M) &= P \begin{bmatrix} f(X+Y) & 0 \\ 0 & f(X-Y) \end{bmatrix} P^{-1} \\ &= \frac{1}{2} \begin{bmatrix} f(X+Y) + f(X-Y) & f(X+Y) - f(X-Y) \\ f(X+Y) - f(X-Y) & f(X+Y) + f(X-Y) \end{bmatrix}. \end{aligned}$$

2. Si $M = \begin{bmatrix} X & -Y \\ Y & X \end{bmatrix}$ y $P = \begin{bmatrix} iI & I \\ I & iI \end{bmatrix}$, entonces $P^{-1} = \frac{1}{2} \begin{bmatrix} -iI & I \\ I & -iI \end{bmatrix}$ y

$$P^{-1}MP = \begin{bmatrix} X+iY & 0 \\ 0 & X-iY \end{bmatrix}.$$

Ahora, por la Proposición 2.4,

$$\begin{aligned} f(M) &= P \begin{bmatrix} f(X+iY) & 0 \\ 0 & f(X-iY) \end{bmatrix} P^{-1} \\ &= \frac{1}{2} \begin{bmatrix} f(X+iY) + f(X-iY) & i[f(X+iY) - f(X-iY)] \\ i[f(X-iY) - f(X+iY)] & f(X+iY) + f(X-iY) \end{bmatrix}. \end{aligned}$$

4. La derivada de Fréchet de funciones matriciales 2×2

En adelante, \mathcal{O} es un conjunto abierto de \mathbb{R} ó \mathbb{C} , $f(\cdot)$ es una función escalar tres veces continuamente diferenciable en \mathcal{O} y \mathcal{A} es el conjunto abierto de las $X \in \mathbb{C}^{2 \times 2}$ tales que $\eta(X) + \sqrt{\xi(X)}$ y $\eta(X) - \sqrt{\xi(X)}$ están en \mathcal{O} . La derivada de Fréchet de $f : \mathcal{A} \rightarrow \mathbb{C}^{2 \times 2}$ en el punto $X \in \mathcal{A}$ es la única transformación lineal $Df(X) : \mathbb{C}^{2 \times 2} \rightarrow \mathbb{C}^{2 \times 2}$ tal que para $H \in \mathbb{C}^{2 \times 2}$,

$$f(X+H) = f(X) + [Df(X)](H) + o(\|H\|).$$

Bajo las condiciones sobre $f(\cdot)$ estipuladas arriba, la derivada de Fréchet $[Df(X)](H)$ existe y es continua en las variables X y H ; ver Higham [4] (Teorema 3.8 pg. 60). Ahora se da una fórmula para esta derivada.

A partir de la expresión $f(X) = \alpha_f(X)I + \beta_f(X)(X - \eta(X)I)$, usando las reglas de derivación, se llega a que, para cada $H \in \mathbb{C}^{2 \times 2}$,

$$\begin{aligned} [Df(X)](H) &= [D\alpha_f(X)](H)I + [D\beta_f(X)](H)(X - \eta(X)I) \\ &\quad + \beta_f(X)(H - \eta(H)I). \end{aligned} \tag{6}$$

Esta fórmula se completa cuando se expliciten las derivadas de Fréchet de las funciones $X \rightarrow \alpha_f(X)$ y $X \rightarrow \beta_f(X)$. Los cálculos de estas derivadas, que se presentan en el apéndice de este artículo, conducen a los siguientes resultados:

$$[D\alpha_f(X)](H) = \begin{cases} \alpha_{f'}(X)[D\eta(X)](H) + \frac{1}{2}\beta_{f'}(X)[D\xi(X)](H), & \xi(X) \neq 0, \\ f'(\eta(X))[D\eta(X)](H) + \frac{f''(\eta(X))}{2}[D\xi(X)](H), & \xi(X) = 0. \end{cases}$$

$$[D\beta_f(X)](H) = \begin{cases} \beta_{f'}(X)[D\eta(X)](H) + \left[\frac{\alpha_{f'}(X) - \beta_f(X)}{2\xi(X)}\right][D\xi(X)], & \xi(X) \neq 0, \\ f''(\eta(X))[D\eta(X)](H) + \frac{f'''(\eta(X))}{6}[D\xi(X)](H), & \xi(X) = 0. \end{cases}$$

Por ejemplo, para $X = \begin{bmatrix} \eta + \sqrt{\xi} & 0 \\ 0 & \eta - \sqrt{\xi} \end{bmatrix}$, $\eta, \xi \in \mathbb{C}$, $\xi \neq 0$, y $H = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}$ se obtiene:

$$[D\eta(X)](H) = 1/2; \quad [D\xi(X)](H) = \sqrt{\xi}; \quad [D\alpha_f(X)](H) = \frac{f'(\eta + \sqrt{\xi})}{2};$$

$$[D\beta_f(X)](H) = \frac{f'(\eta + \sqrt{\xi})}{2\sqrt{\xi}} - \frac{f(\eta + \sqrt{\xi}) - f(\eta - \sqrt{\xi})}{4\xi}.$$

Reemplazando en la fórmula (6) se llega a

$$\begin{aligned} [Df(X)](H) &= \frac{f'(\eta + \sqrt{\xi})}{2} I \\ &+ \left[\frac{f'(\eta + \sqrt{\xi})}{2\sqrt{\xi}} - \frac{f(\eta + \sqrt{\xi}) - f(\eta - \sqrt{\xi})}{4\xi} \right] \begin{bmatrix} \sqrt{\xi} & 0 \\ 0 & -\sqrt{\xi} \end{bmatrix} \\ &+ \left[\frac{f(\eta + \sqrt{\xi}) - f(\eta - \sqrt{\xi})}{2\sqrt{\xi}} \right] \begin{bmatrix} 1/2 & 0 \\ 0 & -1/2 \end{bmatrix} \\ &= \begin{bmatrix} f'(\eta + \sqrt{\xi}) & 0 \\ 0 & 0 \end{bmatrix} = f'(\eta + \sqrt{\xi})H = f'(X)H = Hf'(X). \end{aligned}$$

Observación 4.1. De la Fórmula (4) se sigue que

$$[D\xi(X)](H) = 2(H - \eta(H)I)(X - \eta(X)I) + [X, H],$$

donde $[X, H] \equiv XH - HX$. Usando esta representación se obtiene

$$[Df(X)](H) = Hf'(X) + [X, H]R(X), \quad (7)$$

siendo

$$R(X) = \begin{cases} \left(\frac{f'(X) - \beta_f(X)I}{2\xi(X)}\right)(X - \eta(X)I), & \xi(X) \neq 0, \\ \frac{f''(\eta(X))}{2} + \frac{f'''(\eta(X))}{6}(X - \eta(X)I), & \xi(X) = 0. \end{cases}$$

Así las cosas, si $XH = HX$ ó $f(x) = ax + b$, $a, b \in \mathbb{C}$; entonces $[Df(X)](H) = f'(X)H = Hf'(X)$.

5. Forma canónica de Jordan de la derivada de Fréchet

Sea $J = P^{-1}XP$ la forma canónica de Jordan de $X \in \mathbb{C}^{2 \times 2}$, donde $P = [u_1 \ u_2]$ y $P^{-1} = \begin{bmatrix} v_1^T \\ v_2^T \end{bmatrix}$. Es fácil ver que $\mathcal{B} := \{H_{ij} = u_i v_j^T; i, j = 1, 2\}$ es una base de $\mathbb{C}^{2 \times 2}$. La matriz de la transformación lineal $[Df(X)]$ en esta base se denotará por $[Df(X)]_{\mathcal{B}}$.

Se deduce fácilmente de la fórmula (6) que para cada $H \in \mathbb{C}^{2 \times 2}$

$$P^{-1}[Df(X)](H)P = [Df(P^{-1}XP)](P^{-1}HP) = [Df(J)](P^{-1}HP). \quad (8)$$

Ahora, si $H = H_{ij} = u_i v_j^T$, entonces $P^{-1}HP = e_i e_j^T = E_{ij}$, donde $I = [e_1 \ e_2]$, es decir, $\{E_{ij}; i, j = 1, 2\}$ es la base estándar de $\mathbb{C}^{2 \times 2}$. Para calcular $[Df(J)](E_{ij})$, podemos usar la Fórmula (6) o determinar directamente las derivadas direccionales:

$$\lim_{t \rightarrow 0} \frac{f(J + tE_{ij}) - f(J)}{t}.$$

De cualquier forma se obtienen los siguientes resultados.

Caso 1. $J = \text{diag}(\eta + \sqrt{\xi}, \eta - \sqrt{\xi})$; $\xi \neq 0$. En este caso (confronte con el Teorema 3.11; p. 62 de Higham [4]),

$$\begin{aligned} [Df(J)](E_{11}) &= f'(\eta + \sqrt{\xi})E_{11}; & [Df(J)](E_{12}) &= \frac{f(\eta + \sqrt{\xi}) - f(\eta - \sqrt{\xi})}{2\sqrt{\xi}} E_{12}; \\ [Df(J)](E_{21}) &= \frac{f(\eta + \sqrt{\xi}) - f(\eta - \sqrt{\xi})}{2\sqrt{\xi}} E_{21}; & [Df(J)](E_{22}) &= f'(\eta - \sqrt{\xi})E_{22}. \end{aligned}$$

Esto significa que la matriz de $[Df(X)]$ en la base \mathcal{B} es la matriz diagonal:

$$[Df(X)]_{\mathcal{B}} = \text{diag}(d_{11}, d_{22}, d_{33}, d_{44}),$$

donde

$$d_{11} = f'(\eta + \sqrt{\xi}); \quad d_{22} = \frac{f(\eta + \sqrt{\xi}) - f(\eta - \sqrt{\xi})}{2\sqrt{\xi}} = d_{33}; \quad d_{44} = f'(\eta - \sqrt{\xi}).$$

Caso 2. $J = \eta I$; $\xi = 0$. En este caso, $[Df(J)](E_{ij}) = f'(\eta)(E_{ij})$ y la matriz que representa a $[Df(X)]$ en la base \mathcal{B} (y en cualquier otra base de $\mathbb{C}^{2 \times 2}$) es la matriz diagonal

$$[Df(X)]_{\mathcal{B}} = f'(\eta)I.$$

Caso 3. $J = \begin{bmatrix} \eta & 1 \\ 0 & \eta \end{bmatrix}; \xi = 0$. En este caso

$$\begin{aligned} [Df(J)](E_{11}) &= f'(\eta)E_{11} + \frac{f''(\eta)}{2}E_{12}; & [Df(J)](E_{12}) &= f'(\eta)E_{12}; \\ [Df(J)](E_{21}) &= \frac{f''(\eta)}{2}E_{11} + \frac{f'''(\eta)}{6}E_{12} + f'(\eta)E_{21} + \frac{f''(\eta)}{2}E_{22}; \\ [Df(J)](E_{22}) &= \frac{f''(\eta)}{2}E_{12} + f'(\eta)E_{22}. \end{aligned}$$

Esto quiere decir que la matriz de $[Df(X)]$ en la base \mathcal{B} es:

$$[Df(X)]_{\mathcal{B}} = \begin{bmatrix} f'(\eta) & 0 & f''(\eta)/2 & 0 \\ f''(\eta)/2 & f'(\eta) & f'''(\eta)/6 & f''(\eta)/2 \\ 0 & 0 & f'(\eta) & 0 \\ 0 & 0 & f''(\eta)/2 & f'(\eta) \end{bmatrix}.$$

La forma canónica de Jordan de la matriz $[Df(X)]$ depende de $f'(\eta)$ y $f''(\eta)$.

Caso $f''(\eta) = f'''(\eta) = 0$. Es claro que $[Df(X)]_{\mathcal{B}} = f'(\eta)I$.

Caso $f''(\eta) = 0; f'''(\eta) \neq 0$. En la base $\mathcal{B}^* := \{H_{11}, H_{12}, \frac{6}{f'''(\eta)}H_{21}, H_{22}\}$ se tiene

$$[Df(X)]_{\mathcal{B}^*} = \begin{bmatrix} f'(\eta) & 0 & 0 & 0 \\ 0 & f'(\eta) & 1 & 0 \\ 0 & 0 & f'(\eta) & 0 \\ 0 & 0 & 0 & f'(\eta) \end{bmatrix}.$$

Caso $f''(\eta) \neq 0$. En la base $\mathcal{B}^* = \{R_{11}, R_{12}, R_{21}, R_{22}\}$, donde

$R_{11} = \frac{(f''(\eta))^2}{2}H_{12}$, $R_{12} = \frac{f''(\eta)}{2}(H_{11} + H_{22}) + \frac{f'''(\eta)}{6}H_{12}$, $R_{21} = H_{21}$ y $R_{22} = H_{11} - H_{22}$, se tiene

$$[Df(X)]_{\mathcal{B}^*} = \begin{bmatrix} f'(\eta) & 1 & 0 & 0 \\ 0 & f'(\eta) & 1 & 0 \\ 0 & 0 & f'(\eta) & 0 \\ 0 & 0 & 0 & f'(\eta) \end{bmatrix}.$$

El análisis anterior se resume en el siguiente teorema.

Teorema 5.1. *La forma canónica de Jordan de $[Df(X)]$, $X \in \mathbb{C}^{2 \times 2}$ posee las siguientes propiedades:*

- *Es diagonal cuando X es diagonalizable, o cuando X no es diagonalizable pero $f''(\eta) = f'''(\eta) = 0$.*
- *Tiene un bloque de tamaño 2 y dos bloques de tamaño 1 cuando X no es diagonalizable, $f''(\eta) = 0$ y $f'''(\eta) \neq 0$.*

- Tiene un bloque de tamaño 3 y uno de tamaño 1 cuando X no es diagonalizable y $f''(\eta) \neq 0$.

Ejemplo 5.2. Si $f(x) = e^x$, la forma de Jordan de $[Df(X)]$ es diagonal si y sólo si X es diagonalizable. Cuando X tiene forma de Jordan $\begin{bmatrix} \eta & 1 \\ 0 & \eta \end{bmatrix}$, entonces $[Df(X)]$ tiene forma de Jordan

$$\begin{bmatrix} e^\eta & 1 & 0 & 0 \\ 0 & e^\eta & 1 & 0 \\ 0 & 0 & e^\eta & 0 \\ 0 & 0 & 0 & e^\eta \end{bmatrix}.$$

Observación 5.3. En relación con la norma de la derivada de Fréchet, se cumple la siguiente desigualdad:

$$\|[Df(X)]\| \leq k^2(P)\|[Df(J)]\|,$$

donde $k(P) = \|P\|\|P^{-1}\|$. En efecto, de la expresión (8) se obtiene

$$\begin{aligned} \|[Df(X)](H)\| &= \|P[Df(J)](P^{-1}HP)P^{-1}\| \\ &\leq \|P\|\|[Df(J)]\|\|(P^{-1}HP)\| \|P^{-1}\| \\ &\leq \|P\|\|[Df(J)]\|\|P^{-1}\|\|H\|\|P\|\|P^{-1}\| \\ &= k^2(P)\|[Df(J)]\|\|H\|. \end{aligned}$$

Ahora bien, de las expresiones anteriores se puede determinar $\|[Df(J)]\|$. En particular, si $J = \text{diag}(\eta + \sqrt{\xi}, \eta - \sqrt{\xi})$, $\xi \neq 0$, entonces

$$\|[Df(J)]\| = \max\{|f'(\eta + \sqrt{\xi})|, \left| \frac{f(\eta + \sqrt{\xi}) - f(\eta - \sqrt{\xi})}{2\sqrt{\xi}} \right|, |f'(\eta - \sqrt{\xi})|\}.$$

Si $J = \eta I$ o si $J = \begin{bmatrix} \eta & 1 \\ 0 & \eta \end{bmatrix}$ y $0 = f''(\eta) = f'''(\eta)$, entonces $\|[Df(J)]\| = |f'(\eta)|$.

6. Derivada de Fréchet de funciones matriciales de matrices $n \times n$

En esta sección se considera el caso en que $f(\cdot)$ es analítica y X es una matriz $n \times n$. Cuando $f(\cdot)$ es analítica dentro y sobre un contorno cerrado Γ que encierra a λ_i , $i = 1, 2, \dots, k$, la fórmula más compacta de la derivada de Fréchet de $X \rightarrow f(X)$; $X \in \mathbb{C}^{n \times n}$ es la de Stickel (ver Stickel [11], Teorema 1, p. 84):

$$[Df(X)](H) = \frac{1}{2\pi i} \int_{\Gamma} f(z)(zI - X)^{-1}H(zI - X)^{-1}dz, \quad (9)$$

la cual, cuando $XH = HX$, conduce a

$$\begin{aligned} [Df(X)](H) &= H \left[\frac{1}{2\pi i} \int_{\Gamma} f(z)(zI - X)^{-2}dz \right] = \left[\frac{1}{2\pi i} \int_{\Gamma} f(z)(zI - X)^{-2}dz \right] H \\ &= Hf'(X) = f'(X)H. \end{aligned}$$

Más aún, para la j -ésima derivada de Fréchet, $j = 2, 3, \dots$, se obtiene (ver el Teorema 4.1 de Deadman y Relton [2]):

$$[D^{(j)}f(X)](H_1, H_2, \dots, H_j) = \frac{1}{2\pi i} \int_{\Gamma} f(z)R(z)dz, \quad (10)$$

siendo

$$R(z) = \sum_{\sigma} (zI - X)^{-1} H_{\sigma(1)} (zI - X)^{-1} H_{\sigma(2)} \dots (zI - X)^{-1} H_{\sigma(j)} (zI - X)^{-1},$$

donde la suma se hace sobre las $j!$ permutaciones σ de $\{1, 2, \dots, j\}$.

Los siguientes ejemplos ilustran la utilidad de estas fórmulas.

6.1. Ejemplos

En esta subsección $X, H \in \mathbb{C}^{n \times n}$.

Ejemplo 6.1. Supóngase que $M = \begin{bmatrix} X & H \\ 0 & X \end{bmatrix}$. Entonces la Definición 2.3 y la Fórmula (9) conducen a la siguiente identidad (confronte con la fórmula (3.16), p. 60 de Higham [4]):

$$\begin{aligned} f(M) &= \frac{1}{2\pi i} \int_{\Gamma} f(z)[(zI - M)^{-1} dz \\ &= \frac{1}{2\pi i} \int_{\Gamma} f(z) \begin{bmatrix} (zI - X)^{-1} & (zI - X)^{-1} H (zI - X)^{-1} \\ 0 & (zI - X)^{-1} \end{bmatrix} dz \\ &= \begin{bmatrix} f(X) & [Df(X)](H) \\ 0 & f(X) \end{bmatrix}, \end{aligned}$$

la cual se puede usar para obtener $[Df(X)](H)$ a partir de $f(M)$, o para calcular $f(M)$ a partir de $f(X)$ y $[Df(X)](H)$.

Si, por ejemplo, $X = \begin{bmatrix} \eta & 1 \\ 0 & \eta \end{bmatrix}$ y $H = \begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix}$, entonces

$$M = \begin{bmatrix} \eta & 1 & 0 & 0 \\ 0 & \eta & 1 & 0 \\ 0 & 0 & \eta & 1 \\ 0 & 0 & 0 & \eta \end{bmatrix} \quad \text{y} \quad f(M) = \begin{bmatrix} f(\eta) & f'(\eta) & f''(\eta)/2 & f'''(\eta)/6 \\ 0 & f(\eta) & f'(\eta) & f''(\eta)/2 \\ 0 & 0 & f(\eta) & f'(\eta) \\ 0 & 0 & 0 & f(\eta) \end{bmatrix}.$$

De aquí que $[Df(X)](H) = \begin{bmatrix} f''(\eta)/2! & f'''(\eta)/3! \\ f'(\eta) & f''(\eta)/2! \end{bmatrix}$. De otra parte; cuando $f(x) = e^x$ y X conmuta con H , entonces

$$e^M = \begin{bmatrix} e^X & e^X H \\ 0 & e^X \end{bmatrix} = \begin{bmatrix} e^X & H e^X \\ 0 & e^X \end{bmatrix}.$$

Ejemplo 6.2. Supóngase ahora que

$$M = \begin{bmatrix} X & H & 0 \\ 0 & X & H \\ 0 & 0 & X \end{bmatrix}.$$

Entonces, cuando existe, la matriz $(zI - M)^{-1}$ es

$$\begin{bmatrix} (zI - X)^{-1} & (zI - X)^{-1}H(zI - X)^{-1} & (zI - X)^{-1}H(zI - X)^{-1}H(zI - X)^{-1} \\ 0 & (zI - X)^{-1} & (zI - X)^{-1}H(zI - X)^{-1} \\ 0 & 0 & (zI - X)^{-1} \end{bmatrix}.$$

La Definición 2.3 y la Fórmula (10) llevan a

$$\begin{aligned} f(M) &= \frac{1}{2\pi i} \int_{\Gamma} f(z)[(zI - M)^{-1} dz \\ &= \begin{bmatrix} f(X) & [Df(X)](H) & \frac{1}{2}[D^2f(X)](H, H) \\ 0 & f(X) & [Df(X)](H) \\ 0 & 0 & f(X) \end{bmatrix}, \end{aligned}$$

expresión que permite calcular $[D^2f(X)](H, H)$ a partir de $f(M)$. Como en el ejemplo anterior, cuando $X = \begin{bmatrix} \eta & 1 \\ 0 & \eta \end{bmatrix}$ y $H = \begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix}$, se obtiene

$$\frac{1}{2}[D^2f(X)](H, H) = \begin{bmatrix} f^{(4)}(\eta)/4! & f^{(5)}(\eta)/5! \\ f^{(3)}(\eta)/3! & f^{(4)}(\eta)/4! \end{bmatrix}.$$

Estos ejemplos constituyen una generalización de la matrix $f(J_i)$ que aparece en la Definición 2.2 (ver Najfeld and Havel [9], Teorema 4.13, p. 350).

Se termina esta sección con el siguiente resultado anunciado en la introducción.

Teorema 6.3. *Si X es diagonalizable y $f(\cdot)$ es analítica dentro y sobre un contorno cerrado Γ que encierra el espectro de X , entonces la forma canónica de Jordan de la derivada de Fréchet es diagonal.*

Demostración. Supóngase que $X \in \mathbb{C}^{n \times n}$ es diagonalizable y que $P^{-1}XP = J = \text{diag}(\lambda_1, \lambda_2, \dots, \lambda_n)$, donde

$$P = [u_1 \ u_2 \ \dots \ u_n], \quad [P^{-1}]^T = [v_1 \ v_2 \ \dots \ v_n].$$

Entonces $\mathcal{B} := \{H_{ij} = u_i v_j^T; i, j = 1, 2, \dots, n\}$ es una base de $\mathbb{C}^{n \times n}$ y $P^{-1}H_{ij}P = e_i e_j^T = E_{ij}$, donde $I = [e_1 \ e_2 \ \dots \ e_n]$, es decir, $\{E_{ij}; i, j = 1, 2, \dots, n\}$ es la base estándar de $\mathbb{C}^{n \times n}$. Con estas notaciones se puede escribir

$$\begin{aligned} P^{-1}[Df(X)](H_{ij})P &= \frac{1}{2\pi i} \int_{\Gamma} f(z)[(zI - J)^{-1}E_{ij}(zI - J)^{-1}]dz \\ &= \left[\frac{1}{2\pi i} \int_{\Gamma} f(z)(z - \lambda_i)^{-1}(z - \lambda_j)^{-1} dz \right] E_{ij} \\ &= f[\lambda_i, \lambda_j] E_{ij}, \end{aligned}$$

donde

$$f[\lambda_i, \lambda_j] = \begin{cases} \frac{f(\lambda_j) - f(\lambda_i)}{\lambda_j - \lambda_i}, & \lambda_i \neq \lambda_j, \\ f'(\lambda_i), & \lambda_i = \lambda_j. \end{cases}$$

De esto se sigue que $[Df(X)](H_{ij}) = f[\lambda_i, \lambda_j]H_{ij}$, por lo que la matriz que representa la transformación lineal $[Df(X)](\cdot)$ en la base \mathcal{B} es diagonal. \square

Apéndice: Derivadas de Fréchet de las funciones $\alpha_f(\cdot)$ y $\beta_f(\cdot)$

Aquí se mantiene la notación de la Sección 4; en particular, \mathcal{O} es un conjunto abierto de \mathbb{R} ó \mathbb{C} , $f(\cdot)$ es una función escalar tres veces continuamente diferenciable en \mathcal{O} y \mathcal{A} es el conjunto abierto de las $X \in \mathbb{C}^{2 \times 2}$ tales que $\eta(X) + \sqrt{\xi(X)}$ y $\eta(X) - \sqrt{\xi(X)}$ están en \mathcal{O} . Además, se usa el siguiente hecho: el conjunto abierto \mathcal{D} de las $X \in \mathbb{C}^{2 \times 2}$ tales que $\xi(X) \neq 0$ es denso en $\mathbb{C}^{2 \times 2}$. En efecto, si $X \notin \mathcal{D}$ y $P^{-1}XP = J = \begin{bmatrix} \eta(X) & * \\ 0 & \eta(X) \end{bmatrix}$, $* \in \{0, 1\}$, es la forma canónica de Jordan de X , la matriz

$$X_\epsilon := P \begin{bmatrix} \eta(X) + \epsilon & * \\ 0 & \eta(X) - \epsilon \end{bmatrix} P^{-1}, \quad \epsilon > 0, \quad (11)$$

satisface $\|X - X_\epsilon\| \leq \|P\| \|P^{-1}\| 2\epsilon^2$, es decir, $\lim_{\epsilon \rightarrow 0} X_\epsilon = X$. Nótese que $\eta(X_\epsilon) = \eta(X)$ y $\xi(X_\epsilon) = \epsilon^2$.

Derivada de Fréchet de la función $X \rightarrow \alpha_f(X)$.

$$\alpha_f(X) = \frac{f(\eta(X) + \sqrt{\xi(X)}) + f(\eta(X) - \sqrt{\xi(X)})}{2}.$$

Si $X \in \mathcal{A} \cap \mathcal{D}$, i.e., $\xi(X) \neq 0$, y $H \in \mathbb{C}^{2 \times 2}$, entonces las reglas de derivación implican que

$$\begin{aligned} [D\alpha_f(X)](H) &= \frac{f'(\eta(X) + \sqrt{\xi(X)}) + f'(\eta(X) - \sqrt{\xi(X)})}{2} [D\eta(X)](H) \\ &\quad + \frac{f'(\eta(X) + \sqrt{\xi(X)}) - f'(\eta(X) - \sqrt{\xi(X)})}{4\sqrt{\xi(X)}} [D\xi(X)](H) \\ &\equiv \alpha_{f'}(X) [D\eta(X)](H) + \frac{1}{2} \beta_{f'}(X) [D\xi(X)](H). \end{aligned}$$

En especial, para la matriz X_ϵ que aparece en (11)

$$\begin{aligned} [D\alpha_f(X_\epsilon)](H) &= \frac{f'(\eta(X) + \epsilon) + f'(\eta(X) - \epsilon)}{2} [D\eta(X_\epsilon)](H) \\ &\quad + \frac{f'(\eta(X) + \epsilon) - f'(\eta(X) - \epsilon)}{4\epsilon} [D\xi(X_\epsilon)](H). \end{aligned}$$

Supóngase ahora que $X \in \mathcal{A}$ es tal que $\xi(X) = 0$. Sea X_ϵ la matriz en (11) que verifica $\lim_{\epsilon \rightarrow 0} X_\epsilon = X$. Por la continuidad de las funciones involucradas

$$\begin{aligned}\alpha_f(X + H) &= \alpha_f\left(\lim_{\epsilon \rightarrow 0}(X_\epsilon + H)\right) = \lim_{\epsilon \rightarrow 0} \alpha_f(X_\epsilon + H) \\ &= \lim_{\epsilon \rightarrow 0} \left(\alpha_f(X_\epsilon) + [D\alpha_f(X_\epsilon)](H) + o(\|H\|)\right) \\ &= \alpha_f(X) + f'(\eta(X))[D\eta(X)](H) + \frac{f''(\eta(X))}{2}[D\xi(X)](H) + o(\|H\|).\end{aligned}$$

Esto significa que, para $X \in \mathcal{A}$ tal que $\xi(X) = 0$, se verifica

$$[D\alpha_f(X)](H) = f'(\eta(X))[D\eta(X)](H) + \frac{f''(\eta(X))}{2}[D\xi(X)](H).$$

Derivada de Fréchet de la función $X \rightarrow \beta_f(X)$.

$$\beta_f(X) = \begin{cases} \frac{f(\eta(X) + \sqrt{\xi(X)}) - f(\eta(X) - \sqrt{\xi(X)})}{2\sqrt{\xi(X)}}, & \xi(X) \neq 0, \\ f'(\eta(X)), & \xi(X) = 0. \end{cases}$$

Si $X \in \mathcal{A} \cap \mathcal{D}$, entonces para $H \in \mathbb{C}^{2 \times 2}$ se tiene que

$$\begin{aligned}[D\beta_f(X)](H) &= \frac{f'(\eta(X) + \sqrt{\xi(X)}) - f'(\eta(X) - \sqrt{\xi(X)})}{2\sqrt{\xi(X)}}[D\eta(X)](H) \\ &+ \left[\frac{f'(\eta(X) + \sqrt{\xi(X)}) + f'(\eta(X) - \sqrt{\xi(X)})}{4\xi(X)^{3/2}} \sqrt{\xi(X)} \right. \\ &\quad \left. - \frac{f(\eta(X) + \sqrt{\xi(X)}) - f(\eta(X) - \sqrt{\xi(X)})}{4\xi(X)^{3/2}} \right] [D\xi(X)](H) \\ &\equiv \beta_{f'}(X)[D\eta(X)](H) + \left[\frac{\alpha_{f'}(X) - \beta_f(X)}{2\xi(X)} \right] [D\xi(X)](H).\end{aligned}$$

En consecuencia, para la matriz X_ϵ mencionada arriba, se obtiene que

$$\begin{aligned}[D\beta_f(X_\epsilon)](H) &= \frac{f'(\eta(X) + \epsilon) - f'(\eta(X) - \epsilon)}{2\epsilon}[D\eta(X_\epsilon)](H) \\ &+ \left[\frac{\epsilon[f'(\eta(X) + \epsilon) + f'(\eta(X) - \epsilon)]}{4\epsilon^3} \right. \\ &\quad \left. - \frac{f(\eta(X) + \epsilon) - f(\eta(X) - \epsilon)}{4\epsilon^3} \right] [D\xi(X_\epsilon)](H).\end{aligned}$$

Sean ahora $X \in \mathcal{A}$ con $\xi(X) = 0$ y X_ϵ la matriz en (11) que verifica $\lim_{\epsilon \rightarrow 0} X_\epsilon = X$. Por la continuidad de las funciones involucradas y la regla de L'Hôpital, se sigue que

$$\begin{aligned}
\beta_f(X + H) &= \beta_f\left(\lim_{\epsilon \rightarrow 0}(X_\epsilon + H)\right) = \lim_{\epsilon \rightarrow 0} \beta_f(X_\epsilon + H) \\
&= \lim_{\epsilon \rightarrow 0} \left(\beta_f(X_\epsilon) + [D\beta_f(X_\epsilon)](H) + o(\|H\|)\right) \\
&= \beta_f(X) + f''(\eta(X))[D\eta(X)](H) + \frac{f'''(\eta(X))}{6}[D\xi(X)](H) + o(\|H\|).
\end{aligned}$$

Esto quiere decir que para $X \in \mathcal{A}$ tal que $\xi(X) = 0$

$$[D\beta_f(X)](H) = f''(\eta(X))[D\eta(X)](H) + \frac{f'''(\eta(X))}{6}[D\xi(X)](H).$$

7. Conclusiones

En este artículo se ha mostrado que si $f(X)$ está definida, $X \in \mathbb{C}^{2 \times 2}$, entonces

$$f(X) = \alpha_f(X)I + \beta_f(X)(X - \eta(X)I),$$

donde las funciones $\alpha_f(\cdot)$, $\beta_f(\cdot)$ y $\eta(\cdot)$ dependen de la traza y el determinante de X . Se ha usado esta expresión para dar fórmulas explícitas para evaluar las matrices e^X , $\text{sen}(X)$, $\text{cos}(X)$ y $X^{1/2}$, para evaluar funciones de ciertas matrices particionadas y para determinar la derivada de Fréchet de la función $X \rightarrow f(X)$; $X \in \mathcal{A} \subset \mathbb{C}^{2 \times 2}$. Como resultado central, se ha calculado la forma canónica de Jordan de su derivada de Fréchet. La determinación de la forma de Jordan de la derivada de Fréchet en el caso $n \geq 3$ queda abierta y, para funciones analíticas se puede abordar con la fórmula de Stickel (9).

Agradecimientos

El autor agradece las valiosas sugerencias y observaciones hechas por el árbitro del artículo.

Referencias

- [1] Cardoso J.R. and Sadeghi A., “On the conditioning of the matrix-matrix exponentiation”, *Numer. Algorithms* (2017), 1–21
- [2] Deadman E. and Relton S.D., “Taylor’s theorem for matrix functions with applications to condition number estimation”, *Linear Algebra Appl.* 504 (2016), 354–371.
- [3] Golub G. and Van Loan C., *Matrix Computations*, The Johns Hopkins University Press, Baltimore, 1996.
- [4] Higham N.J., *Functions of Matrices: Theory and Computation*, Siam, Philadelphia, 2008.
- [5] Higham and Al-Mohy A.H., “Computing matrix functions”, *Acta Numer.* 19 (2010), 159–208.

- [6] Higham N.J. and Lijing L., “Matrix Functions: A short Course”, en *Ser. Contemp. Appl. Math.* 19, Higher Ed. Press, Beijing, (2015), 1–27.
- [7] Horn R.A. and Johnson C.R., *Topics in Matrix Analysis*, Cambridge University Press, New York, 1991.
- [8] Kandolf P. and Relton S.D., “A block Krylov method to compute the action of the Fréchet derivative of a matrix function on a vector with applications to condition number estimation”, *Siam J. Sci. Comp.* 39 (2017), No. 4, A1416–A1434.
- [9] Najfeld I. and Havel T.F., “Derivatives of the matrix exponential and their computation”, *Adv. in Appl. Math.* 16 (1995), No. 3, 321–375.
- [10] Rinehart R.F., “The Equivalence of Definitions of a Matric Function”, *Amer. Math. Monthly* 62 (1955), No. 6, 395–414.
- [11] Stickel E., “On the Fréchet Derivative of Matrix Functions”, *Linear Algebra Appl.* 91 (1987), 83–88.