Miscelánea Matemática **51** (2010) 29–40 SMM

Seno y coseno de una matriz

Rafael Prieto Curiel
Instituto Tecnológico Autónomo de México ITAM
rafaelprietocuriel@yahoo.com

1. Introducción

En muchas áreas de las matemáticas resulta de gran utilidad definir funciones matriciales y las funciones más sencillas en las que podemos pensar son polinomios de matrices cuadradas. Evaluar un polinomio en matrices cuadradas es sólo cuestión de sustituir las potencias de la matriz, considerando la constante del polinomio como un múltiplo de la matriz identidad.

Sin duda alguna, una de las funciones matriciales más conocidas es la exponencial. Esta función se define como la única solución al sistema de ecuaciones

$$\dot{\Phi}(t) = A\Phi(t)$$

junto con la condición inicial $\Phi(0) = I$. Expresamos $f(t) = e^{tA}$ como la solución. Gracias a esa propiedad de la matriz exponencial, sabemos que la solución al sistema de ecuaciones diferenciales

$$\dot{x}(t) = Ax,$$

con la condición inicial $x(0) = x_0$ es

$$x(t) = e^{tA}x_0,$$

pues vemos que cumple la ecuación diferencial y las condiciones iniciales.

Esta clase de sistemas diferenciales aparecen con frecuencia al resolver una ecuación diferencial homogénea de orden n con coeficientes constantes de la forma

$$y^{(n)}(t) = \alpha_0 y(t) + \alpha_1 y'(t) + \ldots + \alpha_{n-1} y^{(n-1)}(t),$$

que se puede expresar como un sistema de n ecuaciones diferenciales de la forma

$$\frac{d}{dt}\begin{pmatrix} y \\ y' \\ \vdots \\ y^{(n-1)} \end{pmatrix} = \begin{pmatrix} 0 & 1 & 0 & \dots \\ 0 & 0 & 1 & \dots \\ & & \ddots & \\ \alpha_0 & \alpha_1 & \dots & \alpha_{n-1} \end{pmatrix}\begin{pmatrix} y \\ y' \\ \vdots \\ y^{(n-1)} \end{pmatrix}.$$

La existencia de la función exponencial nos lleva a preguntarnos por la existencia de otras funciones matriciales, y si estas tienen alguna aplicación en el estudio de ecuaciones diferenciales.

Al igual que las funciones escalares, existen diversas formas para definir alguna función matricial, y en particular para llegar al seno y coseno de una matriz lo haremos a partir de tres distintos métodos: a partir de una serie de potencias, como la solución de una ecuación diferencial y mediante otras funciones.

2. Funciones trigonométricas como una serie

Decimos que una función es *analítica en* t_0 si existe $R > 0$ tal que

$$f(t) = \sum_{k=0}^{\infty} \alpha_k (t - t_0)^k$$

es convergente en $|t - t_0| < R$. Al intervalo $|t - t_0| < R$, donde se garantiza la existencia de $f(t)$, se le llama *intervalo de convergencia.*

Extendiendo esta función a matrices, definimos

$$f(A) = \sum_{k=0}^{\infty} \alpha_k (A - t_0 I)^k.$$

Esta definición tiene el inconveniente de que se requiere garantizar la existencia de $f(A)$, es decir, la convergencia de la serie; para ello analizaremos dos casos dependiendo si la matriz A es o no diagonalizable.

2.1. Matrices diagonalizables

Decimos que una matriz es diagonalizable si existen matrices P invertible y D diagonal tales que se puede escribir

$$A = PDP^{-1},$$

donde

$$D = \begin{pmatrix} d_1 & 0 & \dots & 0 \\ 0 & d_2 & \dots & 0 \\ \vdots & & \ddots & 0 \\ 0 & 0 & \dots & d_n \end{pmatrix},$$

una matriz con los eigenvalores $\{d_1, d_2, \dots, d_n\}$ en la diagonal.

Vemos que en este caso

$$\begin{aligned} f(A) &= \sum_{k=0}^{\infty} \alpha_k (A - t_0 I)^k \\ &= \sum_{k=0}^{\infty} \alpha_k (PDP^{-1} - t_0 I)^k \\ &= \sum_{k=0}^{\infty} \alpha_k P (D - t_0 I)^k P^{-1} \\ &= P \left(\sum_{k=0}^{\infty} \alpha_k (D - t_0 I)^k \right) P^{-1}. \end{aligned}$$

Podemos ver que los únicos términos en la suma serán los de la diagonal; además, en cada una de las entradas de la diagonal tenemos

$$\sum_{k=0}^{\infty} \alpha_k (d_j - t_0)^k,$$

donde d_j es el eigenvalor j-ésimo. Podemos garantizar la convergencia de la serie si $|d_j - t_0| < R$, por lo que sabemos que $f(A)$ existe si todos los eigenvalores están en el intervalo de convergencia de f y además, una expresión para $f(A)$ es

$$f(A) = P \begin{pmatrix} f(d_1) & 0 & \dots & 0 \\ 0 & f(d_2) & \dots & 0 \\ \vdots & & \ddots & 0 \\ 0 & 0 & \dots & f(d_n) \end{pmatrix} P^{-1}.$$

2.2. Matrices no diagonalizables y los bloques de Jordan

Desafortunadamente, no todas las matrices son diagonalizables, por lo que, para garantizar la existencia de $f(A)$ necesitamos otro argumento. Sabemos que todas las matrices se pueden expresar de la forma

$$A = PJP^{-1},$$

donde J es una matriz de la forma

$$J = \begin{pmatrix} J_1 & & & \\ & J_2 & & \\ & & \ddots & \\ & & & J_m \end{pmatrix},$$

con m bloques de Jordan en la diagonal, donde cada bloque de Jordan es una matriz de la forma

$$J_i = \begin{pmatrix} \lambda_i & 1 & \dots & 0 \\ 0 & \lambda_i & \ddots & 0 \\ \vdots & & \ddots & 1 \\ 0 & 0 & \dots & \lambda_i \end{pmatrix}$$

con el eigenvalor λ_i en la diagonal, y unos en la diagonal superior. Vemos que en este caso,

$$A^k = \left(PJP^{-1}\right)^k = PJ^kP^{-1}.$$

Al calcular J^k obtenemos

$$J^k = \begin{pmatrix} J_1 & & & \\ & J_2 & & \\ & & \ddots & \\ & & & J_l \end{pmatrix}^k = \begin{pmatrix} J_1^k & & & \\ & J_2^k & & \\ & & \ddots & \\ & & & J_l^k \end{pmatrix}.$$

Tomaremos entonces cada uno de los bloques por separado. Es sencillo probar por inducción (tomando en cuenta que si $k < j$ entonces $\binom{k}{j} = 0$) que

$$\begin{aligned}
(J_i)^k &= \begin{pmatrix} \lambda_i & 1 & \dots & 0 \\ 0 & \lambda_i & \dots & 0 \\ \vdots & & \ddots & 1 \\ 0 & 0 & \dots & \lambda_i \end{pmatrix}^k \\
&= \begin{pmatrix} \lambda_i^k & k(\lambda_i)^{k-1} & \frac{k(k-1)}{2}(\lambda_i)^{k-2} & \dots & \binom{k}{j}(\lambda_i)^{k-j} \\ 0 & \lambda_i^k & k(\lambda_i)^{k-1} & \dots & \\ \vdots & & \ddots & \ddots & \vdots \\ 0 & 0 & & \dots & \lambda_i^k \end{pmatrix} \\
&= \begin{pmatrix} \binom{k}{0}\lambda_i^k & \binom{k}{1}(\lambda_i)^{k-1} & \binom{k}{2}(\lambda_i)^{k-2} & \dots & \binom{k}{j}(\lambda_i)^{k-j} \\ 0 & \binom{k}{0}\lambda_i^k & \binom{k}{1}(\lambda_i)^{k-1} & \dots & \\ \vdots & & \ddots & \ddots & \vdots \\ 0 & 0 & & \dots & \binom{k}{0}\lambda_i^k \end{pmatrix},
\end{aligned}$$

donde j es el tamaño del bloque de Jordan.

Vemos entonces que al aplicar $f(J_i)$ llegamos a

$$
\begin{aligned}
f(J_i) &= \sum_{k=0}^{\infty} \alpha_k (J_i - t_0 I)^k \\
&= \sum_{k=0}^{\infty} \alpha_k \begin{pmatrix} \lambda_i - t_0 & 1 & \dots & 0 \\ 0 & \lambda_i - t_0 & \dots & 0 \\ \vdots & & \ddots & 1 \\ 0 & 0 & \dots & \lambda_i - t_0 \end{pmatrix}^k \\
&= \sum_{k=0}^{\infty} \alpha_k \begin{pmatrix} \binom{k}{0}(\lambda_i - t_0)^k & \binom{k}{1}(\lambda_i - t_0)^{k-1} & \dots & \binom{k}{j}(\lambda_i - t_0)^{k-j} \\ 0 & \binom{k}{0}(\lambda_i - t_0)^k & \dots & \\ \vdots & & \ddots & \vdots \\ 0 & 0 & \dots & \binom{k}{0}(\lambda_i - t_0)^k \end{pmatrix},
\end{aligned}
$$

y al ver la entrada p, q, con $p \leq q$, y tomando $r = q - p$

$$
\begin{aligned}
[f(J_i)]_{p,q} &= \sum_{k=0}^{\infty} \alpha_k \binom{k}{q-p} (\lambda_i - t_0)^{k-p+q} \\
&= \frac{1}{r!} \sum_{k=0}^{\infty} \alpha_k k! \left[\frac{1}{(k-r)!} (\lambda_i - t_0)^{k-r} \right] \\
&= \frac{1}{r!} \sum_{k=0}^{\infty} \alpha_k k! \frac{d}{d\lambda_i} \left[\frac{1}{(k-r+1)!} (\lambda_i - t_0)^{k-r+1} \right] \\
&= \frac{1}{r!} \sum_{k=0}^{\infty} \alpha_k k! \frac{d^2}{d\lambda_i^2} \left[\frac{1}{(k-r+2)!} (\lambda_i - t_0)^{k-r+2} \right] \\
&\vdots \\
&= \frac{1}{r!} \sum_{k=0}^{\infty} \alpha_k k! \frac{d^r}{d\lambda_i^r} \left[\frac{1}{(k)!} (\lambda_i - t_0)^{k} \right] \\
&= \frac{1}{r!} \sum_{k=0}^{\infty} \alpha_k \frac{d^r}{d\lambda_i^r} (\lambda_i - t_0)^k .
\end{aligned}
$$

Suponiendo que la función es diferenciable y que $f^{(n)}(\lambda_i)$ existe para una n suficientemente grande, vemos que

$$\begin{aligned}[f(J_i)]_{p,q} &= \frac{1}{r!}\frac{d^r}{d\lambda_i^r}\left(\sum_{k=0}^{\infty}\alpha_k(\lambda_i - t_0)^k\right) \\ &= \frac{1}{r!}\frac{d^r}{d\lambda_i^r}f(\lambda_i) \\ &= \frac{1}{(q-p)!}f^{(q-p)}(\lambda_i),\end{aligned}$$

de donde concluimos dos cosas: podemos asegurar que la serie es convergente y por lo tanto $f(A)$ existe si $f^{(r_i)}(\lambda_i)$ existe para todos los eigenvalores, donde r_i es el tamaño del i-ésimo bloque de Jordan, y además

$$f(J_i) = \begin{pmatrix} f(\lambda_i) & \frac{f'(\lambda_i)}{1!} & \frac{f^{(2)}(\lambda_i)}{2!} & \dots & \frac{f^{(j)}(\lambda_i)}{j!} \\ 0 & f(\lambda_i) & & \dots & \frac{f^{(j-1)}(\lambda_i)}{(j-1)!} \\ \vdots & & & \ddots & \vdots \\ 0 & 0 & & \dots & f(\lambda_i) \end{pmatrix}.$$

Es importante mencionar que $f(A)$ puede existir aunque los eigenvalores estén fuera del intervalo de convergencia de f, por ejemplo, si definimos la función:

$$f(A) = (I - A)^{-1}$$

tiene como serie de potencias asociada

$$f(A) = \sum_{k=0}^{\infty} A^k,$$

la cual es convergente sólo si los eigenvalores de la matriz A son en valor absoluto menores que uno, sin embargo, la matriz $(I - A)$ tiene inversa si los eigenvalores de A son distintos de uno.

Podemos ver que en la forma en que hemos definido las funciones aplicadas a una matriz, se cumple que $f(A) + g(A) = (f + g)(A)$, lo cual no es de sorprender, pero otra propiedad que resulta interesante en esta definición es que

$$f(A)g(A) = g(A)f(A) = (fg)(A),$$

si ambas funciones existen. Verificar esta propiedad en matrices diagonalizables es sólo cuestión de sustituir ambas funciones. Para probarlo en el caso de bloques de Jordan, vemos que si $f(J_i)$ y $g(J_i)$ son matrices de la forma

$$f(J_i) = \begin{pmatrix} f_1 & f_2 & \dots & f_n \\ 0 & f_1 & & \vdots \\ \vdots & & \ddots & f_2 \\ 0 & 0 & \dots & f_1 \end{pmatrix}$$

y

$$g(J_i) = \begin{pmatrix} g_1 & g_2 & \dots & g_n \\ 0 & g_1 & & \vdots \\ \vdots & & \ddots & g_2 \\ 0 & 0 & \dots & g_1 \end{pmatrix},$$

entonces la entrada (p,q) de $f(A)g(A)$, con $p \leq q$ es

$$[f(A)g(A)]_{(p,q)} = \sum_{k=0}^{q-p} f_{k+1}g_{1+q-p-k}$$

y por otro lado, la entrada

$$[g(A)f(A)]_{(p,q)} = \sum_{m=0}^{q-p} g_{m+1}f_{1+q-p-m}.$$

Haciendo el cambio de variable $k = q-p-m$ llegamos a

$$\begin{aligned}[g(A)f(A)]_{(p,q)} &= \sum_{m=0}^{q-p} g_{m+1}f_{1+q-p-m} \\ &= \sum_{k=0}^{q-p} g_{1+q-p-k}f_{k+1} = [f(A)g(A)]_{(p,q)},\end{aligned}$$

de donde concluimos que en general las matrices con esa estructura conmutan. Además, si primero hacemos la multiplicación de las funciones y luego la aplicamos a alguna matriz, vemos que

$$[(fg)(A)]_{(p,q)} = \frac{(fg)^{(q-p)}}{(q-p)!}.$$

Aplicando la regla de Leibniz llegamos a

$$
\begin{aligned}
&= \frac{1}{(q-p)!}\sum_{k=0}^{q-p}\binom{q-p}{k} f^{(k)} g^{(q-p-k)} \\
&= \frac{1}{(q-p)!}\sum_{k=0}^{q-p}\frac{(q-p)!}{(q-p-k)!k!} f^{(k)} g^{(q-p-k)} \\
&= \sum_{k=0}^{q-p}\left(\frac{f^{(k)}}{k!}\right)\left(\frac{g^{(q-p-k)}}{(q-p-k)!}\right) \\
&= \sum_{k=0}^{q-p} f_k g_{q-p-k} \\
&= \left[f(A)g(A)\right]_{(p,q)},
\end{aligned}
$$

por lo que (sin importar si la matriz es o no diagonalizable) si $f(A)$ y $g(A)$ existen entonces

$$f(A)g(A) = (fg)(A) = g(A)f(A).$$

2.3. Funciones trigonométricas

Definimos las dos funciones trigonométricas básicas (seno y coseno) como

$$\operatorname{sen}(tA) = \sum_{k=0}^{\infty}\frac{(-1)^k (tA)^{2k+1}}{(2k+1)!},$$

y el coseno como

$$\cos(tA) = \sum_{k=0}^{\infty}\frac{(-1)^k (tA)^{2k}}{(2k)!}.$$

Sabemos que el intervalo de convergencia de ambas series es infinito, por lo que el seno y coseno de una matriz con entradas reales siempre existen. Además, si $A = PDP^{-1}$, una matriz diagonalizable, entonces

$$
\begin{aligned}
\operatorname{sen}(tA) &= P\operatorname{sen}(tD)P^{-1} \\
&= P\begin{pmatrix} \operatorname{sen}(td_1) & & & \\ & \operatorname{sen}(td_2) & & \\ & & \ddots & \\ & & & \operatorname{sen}(td_n)\end{pmatrix}P^{-1},
\end{aligned}
$$

y de manera similar el coseno. Por otro lado para matrices no diagonalizables, tenemos que $\operatorname{sen}(tA) = P\operatorname{sen}(tJ)P^{-1}$ y el seno, visto en cada uno de los bloques, queda como

$$\operatorname{sen}(tJ_i) = \begin{pmatrix} \operatorname{sen}(t\lambda_i) & t\cos(t\lambda_i) & \dots & \frac{t^{r_i}}{r_i!}\operatorname{sen}^{(r_i)}(t\lambda_i) \\ & \operatorname{sen}(t\lambda_i) & & \vdots \\ & & \ddots & t\cos(t\lambda_i) \\ & & & \operatorname{sen}(t\lambda_i) \end{pmatrix}.$$

Si A es una matriz de números reales, se puede dar el caso en el que los valores propios sean complejos, en cuyo caso tenemos que pensar en las funciones seno y coseno complejos, sin embargo vemos que la serie de potencias de la matriz A es real por lo que al conjugarse las funciones trigonométricas complejas con las matrices P y P^{-1}, también complejas, se obtiene una función real.

3. Funciones trigonométricas como la solución de una ecuación diferencial

Al igual que podemos definir la función exponencial como la solución a un sistema de ecuaciones diferenciales, podemos partir de la ecuación diferencial

$$\ddot{\Phi}(t) = -A^2\Phi(t)$$

con dos diferentes condiciones iniciales $\Phi_1(0) = 0$ y $\Phi_1'(0) = A$ y por otro lado $\Phi_2(0) = I$ y $\Phi_2'(0) = 0$. Gracias al teorema de existencia y unicidad, sabemos que la ecuación tiene solución, por lo que podemos aplicar la transformada de Laplace en ambos lados, y llegar a que

$$s^2\hat{\Phi}(s) - s\Phi(0) - \Phi'(0) = -A^2\hat{\Phi}(s),$$

donde $\hat{\Phi}(s)$ es la transformada de Laplace de $\Phi(t)$ y $\Phi(0)$ y $\Phi'(0)$ son las condiciones iniciales correspondientes. Al tomar s suficientemente grande, se puede garantizar que $I + (A/s)^2$ es invertible, por lo que al despejar llegamos a que

$$\hat{\Phi}(s) = \frac{1}{s}\left[I + \left(\frac{1}{s}A\right)^2\right]^{-1}\Phi(0) + \frac{1}{s^2}\left[I + \left(\frac{1}{s}A\right)^2\right]^{-1}\Phi'(0).$$

Además, se puede ver que

$$\left[I + \left(\frac{1}{s}A\right)^2\right]^{-1} = \sum_{j=0}^{\infty}(-1)^j\left(\frac{1}{s}A\right)^{2j},$$

por lo que llegamos a que

$$\hat{\Phi}(s) = \sum_{j=0}^{\infty} (-1)^j \left(\frac{1}{s}\right)^{2j+1} A^{2j}\Phi(0) + (-1)^j \left(\frac{1}{s}\right)^{2j+2} A^{2j}\Phi'(0).$$

Si tomamos la transformada de Laplace inversa y separamos las condiciones iniciales, llegamos a que

$$\Phi_1(t) = \sum_{j=0}^{\infty} \frac{(-1)^j t^{2j+1} A^{2j+1}}{(2j+1)!} = \operatorname{sen}(tA),$$

y por otro lado

$$\Phi_2(t) = \sum_{j=0}^{\infty} \frac{(-1)^j t^{2j} A^{2j}}{(2j)!} = \cos(tA).$$

Concluímos entonces que en la forma en la que definimos el seno y coseno de una matriz, también se cumple que son la única solución a la ecuación diferencial de la que partimos.

4. A partir de la función exponencial

Como último método para llegar a las funciones trigonométricas, podemos partir de la función exponencial y, al igual que con funciones escalares, definimos la función

$$f(t) = (e^{itA} + e^{-itA})/2,$$

y la función

$$g(t) = (e^{itA} - e^{-itA})/2i,$$

donde e^{itA} será la función exponencial de A evaluada en un número imaginario obtenido a partir de la serie de la función exponencial. Es fácil comprobar que las potencias pares de la serie para f se agruparán, mientras que las impares se restarán, por lo que se llegará a la serie del coseno, y de una manera similar se llega con la función g a la serie del seno de una matriz.

5. Algunas propiedades de las funciones trigonométricas

Una propiedad que es muy útil en el estudio de ecuaciones diferenciales es

$$\begin{aligned}
\frac{d}{dt}\left[\operatorname{sen}(tA)\right] &= \frac{d}{dt}\left[\sum_{k=0}^{\infty}\frac{(-1)^k(tA)^{2k+1}}{(2k+1)!}\right] \\
&= \sum_{k=0}^{\infty}\frac{d}{dt}\left[\frac{(-1)^k(tA)^{2k+1}}{(2k+1)!}\right] \\
&= \sum_{k=0}^{\infty}A\left[\frac{(-1)^k(tA)^{2k}}{(2k)!}\right] \\
&= A\cos(tA).
\end{aligned}$$

Por otro lado

$$\begin{aligned}
\frac{d}{dt}\left[\cos(tA)\right] &= \frac{d}{dt}\left[\sum_{k=0}^{\infty}\frac{(-1)^k(tA)^{2k}}{(2k)!}\right] \\
&= \sum_{k=0}^{\infty}\frac{d}{dt}\left[\frac{(-1)^k(tA)^{2k}}{(2k)!}\right] \\
&= \sum_{k=1}^{\infty}A\left[\frac{(-1)^k(tA)^{2k-1}}{(2k-1)!}\right] \\
&= -A\operatorname{sen}(tA)
\end{aligned}$$

por lo que ambas funciones cumplen la ecuación diferencial

$$\ddot{x}(t) = -A^2x(t).$$

Además al evaluar las series en cero, vemos que $\cos(0) = I$ y que $\operatorname{sen}(0) = 0$, por lo que concluimos que una solución a la ecuación diferencial $\ddot{x}(t) = -A^2x(t)$ con las condiciones iniciales $x(0) = x_0$ y $x'(0) = Ax_1 = x_2$ es

$$x(t) = \cos(tA)x_0 + \operatorname{sen}(tA)x_1.$$

Además, cumplen algunas identidades que resultan curiosas como que

$$\operatorname{sen}^2(tA) + \cos^2(tA) = I.$$

Una forma sencilla de probarla es definiendo la función $h(t) = \operatorname{sen}^2(tA) + \cos^2(tA)$. Vemos que

$$h(0) = \operatorname{sen}^2(0) + \cos^2(0) = I,$$

por otro lado, si calculamos $h'(t)$ vemos que

$$h'(t) = 2\operatorname{sen}(tA)\cos(tA) - 2\cos(tA)\operatorname{sen}(tA)$$

y gracias a que la multiplicación de funciones conmuta, llegamos a que

$$h'(t) = 0,$$

por lo que es una función constante.

Otra identidad que cumplen las funciones trigonométricas es

$$2\operatorname{sen}(tA)\cos(tA) = \operatorname{sen}(2tA).$$

Demostrar esta propiedad se puede lograr fácilmente utilizando las expresiones función del seno y coseno en términos de la función exponencial, y usando el hecho de que $(e^{tA})(e^{tA}) = e^{2tA}$.

Referencias

[1] W. Boyce y R. DiPrima, *Ecuaciones diferenciales y problemas con valores en la frontera*, Limusa Wiley, 2005.

[2] C.D. Meyer, *Matrix Analysis and Applied Linear Algebra*, S.I.A.M., 2000.

[3] S. H. Friedberg, A. J. Insel, L. E. Spence, *Linear Algebra*, Prentice Hall, 2002.

[4] N. J. Higham, *Functions of matrices: theory and computation*, S.I.A.M., 2008.