

Schwerdtfeger formula for matrix functions

OANA-LILIANA CHENDER (BROAINA)
(CLUJ-NAPOCA)

ABSTRACT. The classical Schwerdtfeger formula gives a closed form for $f(A)$, where f is a holomorphic function and $A \in M_n(\mathbb{C})$, in terms of the spectrum of the matrix A and of the Frobenius covariants. In this paper, using the Schwerdtfeger formula, we obtain an extended version for the matrix function $f(tA)$, where t is a positive real number. We present the MATHEMATICA implementation for the Frobenius covariants and for the obtained formula. Also, we discuss the connections to the general Rodrigues problem involving the coefficients of $f(A)$ when it is reduced to a polynomial of A .

1 Introduction

During recent years, there has been a considerably amount of research concerning the functions of matrices as they provide solutions to systems of linear differential equations. For example, the solution of $y' = Ay$ with $y(0) = x$ is given by $y(t) = \exp(tA)x$, where $\exp(tA)$ is the exponential of the matrix tA defined in (5.1). Similarly, we would like to find explicit forms for $\cos(tA)$, $\sin(tA)$, and in general for $f(tA)$. Schwerdtfeger (see [5]) proves that for any holomorphic function f and for any matrix $A \in M_n(\mathbb{C})$ the

[◇]Oana-Liliana Chender (Broaina), Babeş-Bolyai University Cluj-Napoca Romania
email: oanalily@gmail.com

formula (3.1) holds, where μ is the number of distinct eigenvalues λ_j of A , r_j is the multiplicity of λ_j and A_j are the Frobenius covariants of A . Here Γ_j is a smooth closed curve around the complex number λ_j . In this note we use formula (3.1) to derive formula (3.4) and to study some matrix functions and to obtain some applications ([1], [3], [4] and [6]).

In this request we show that by using Schwerdtfeger formula (3.1), one obtains the formula (3.4) which has the same computational cost than (3.1). In Section 4 we present the MATHEMATICA implementation for the Frobenius covariants and for formula (3.4). Section 5 is devoted to the general Rodrigues problem and in Section 6 we present a general method of determining the Rodrigues coefficients by using the Hermite interpolation polynomial.

2 Matrix Functions

The function f defined on the positive real numbers $f : \mathbb{R}^+ \rightarrow \mathbb{C}$, such that f is integrable on $[0, T]$ for all $T > 0$, and there exist constants $\alpha \in \mathbb{R}$ and $M > 0$ such that

$$|f(t)| \leq Me^{\alpha t} \text{ for } t \geq 0$$

is called **exponentially bounded**.

The **Laplace transform** of the exponentially bounded function f is the function $\mathfrak{L}(f) = \tilde{f}$ defined by:

$$\mathfrak{L}(f)(s) = \tilde{f}(s) = \int_0^\infty e^{-st} f(t) dt,$$

for all complex numbers s such that $\operatorname{Re}(s) \geq \alpha$. For $f(t) = e^{iat}$ we have

$$\tilde{f}(s) := \int_0^\infty e^{-st} e^{iat} dt = \frac{1}{s - ia}.$$

The function $f : U \rightarrow \mathbb{C}$ is **holomorphic**, where $U \subseteq \mathbb{C}$ is an open set, if it is complex differentiable, that is the derivative exist at every point $z_0 \in U$. Also, the function $f' : U \rightarrow \mathbb{C}$ is the derivative of f and f is an anti-derivative of f' on U .

A function f is **analytic** in $B(a, R) \subseteq \mathbb{C}$ if there exists $a_k \in \mathbb{C}$ such that for all $z \in B(a, R)$ we have

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

Theorem 2.1 (Cauchy integral formula) *If $f : D \subseteq \mathbb{C} \rightarrow \mathbb{C}$ is a holomorphic function, then*

$$(2.1) \quad f(a) = \frac{1}{2\pi i} \int_{\delta} \frac{f(z)}{z - a} dz$$

where $a \in D$ and δ is a closed curve in D around a .

The vector space of $n \times n$ complex matrices, denoted by $M_n(\mathbb{C})$, is a complete normed space with respect to the norm

$$\|A\| = \max_{x \neq 0} \frac{\|Ax\|}{\|x\|}.$$

Let $A \in M_n(\mathbb{C})$. One says that λ is an **eigenvalue** of A if there is a nonzero $x \in \mathbb{R}^n$ such that $Ax = \lambda x$. The set $\sigma(A)$ is known as the set of eigenvalues of A or the **spectrum** of A .

Theorem 2.2 (Cauchy functional calculus) *Let $f : D \subseteq \mathbb{C} \rightarrow \mathbb{C}$ be a holomorphic function. If $A \in M_n(\mathbb{C})$ such that $\sigma(A) \subset D$, then*

$$f(A) = \frac{1}{2\pi i} \int_{\delta^*} f(\omega) (\omega I - A)^{-1} d\omega$$

where δ^* is a closed curve winding once around $\sigma(A)$.

The following lemmas are straight forward results from linear algebra.

Lemma 2.1 *If $t \in \mathbb{R}$, then $\sigma(tA) = t\sigma(A)$.*

Lemma 2.2 *Let $f : [0, \infty) \rightarrow M_n(\mathbb{C})$ given by $f(t) = e^{tA}$ for $A \in M_n(\mathbb{C})$ fixed. Then*

$$\hat{f}(\lambda) = (\lambda I_n - A)^{-1}.$$

3 The Schwerdtfeger Formula

In 1961, Schwerdtfeger (see [5]) proves that for any holomorphic function f and for any matrix $A \in M_n(\mathbb{C})$ the following formula holds

$$(3.1) \quad f(A) = \sum_{j=1}^{\mu} A_j \sum_{k=0}^{r_j-1} \frac{1}{k!} f^{(k)}(\lambda_j) (A - \lambda_j I_n)^k,$$

where μ is the number of distinct eigenvalues λ_j of A , r_j is the multiplicity of λ_j and

$$(3.2) \quad A_j = \frac{1}{2\pi i} \int_{\Gamma_j} (sI_n - A)^{-1} ds = \frac{1}{2\pi i} \int_{\Gamma_j} \mathfrak{L}[e^{tA}](s) ds$$

are the Frobenius covariants of A . Here Γ_j is a smooth closed curve around the complex number λ_j .

To compute the Frobenius covariants we use the Penrose generalized inverse.

Theorem 3.1 (Penrose generalized inverse matrix) *If $A \in M_n(\mathbb{C})$, then there exist a unique matrix $X \in M_n(\mathbb{C})$ that satisfies the following equations:*

- (i) $AXA = A$;
- (ii) $XAX = X$;
- (iii) $(AX)^* = AX$;
- (iv) $(XA)^* = XA$,

where B^* is the conjugate transpose of B .

The matrix X is said to be the **generalized inverse** of A , and it is denoted by $A^\dagger$.

Theorem 3.2 (Penrose) *If $A \in M_n(\mathbb{C})$ and $\lambda_j \in \sigma(A)$, then*

$$A_j = (F_{\lambda_j} E_{\lambda_j})^\dagger,$$

where $E_{\lambda_j} = ((I_n - (A - \lambda_j I_n)^{r_j})^\dagger (A - \lambda_j I_n)^{r_j})$ and $F_{\lambda_j} = (I_n - (A - \lambda_j I_n)^{r_j})((A - \lambda_j I_n)^{r_j})^\dagger$.

Lemma 3.1 *Let $A \in M_k(\mathbb{C})$. If $t > 0$ then*

$$(3.3) \quad (tA)_j = A_j.$$

Now, we present the extended Schwerdtfeger formula.

Theorem 3.3 *Let $f : D \subseteq \mathbb{C} \rightarrow \mathbb{C}$ be holomorphic. If $A \in M_n(\mathbb{C})$ and $t \geq 0$, then*

$$(3.4) \quad f(tA) = \sum_{j=1}^{\mu} A_j \sum_{k=0}^{r_j-1} \frac{1}{k!} f^{(k)}(t\lambda_j) t^k (A - \lambda_j I_n)^k,$$

where μ denotes the number of different eigenvalues λ_j of A , r_j is the multiplicity of λ_j , and

$$(3.5) \quad A_j := \frac{1}{2\pi i} \int_{\Gamma_j} (sI_n - A)^{-1} ds = \frac{1}{2\pi i} \int_{\Gamma_j} \mathfrak{L}[e^{tA}](s) ds$$

are the Frobenius covariants of A . Here Γ_j is a closed smooth curve enclosing λ_j only.

Proof. Applying formula (3.1) for matrix tA we have

$$f(tA) = \sum_{j=1}^n (tA)_j \sum_{k=0}^{r_j-1} \frac{1}{k!} f^{(k)}(t\lambda_j) (tA - t\lambda_j I_n)^k = \sum_{j=1}^{\mu} A_j \sum_{k=0}^{r_j-1} \frac{1}{k!} f^{(k)}(t\lambda_j) t^k (A - \lambda_j I_n)^k,$$

where we have used Lemma 2.1 and Lemma 3.1. ■

4 The MATHEMATICA Implementation

In the first code we have the MATHEMATICA implementation for the Frobenius covariants of a given matrix A defined by

$$(4.1) \quad A_j = \frac{1}{2\pi i} \int_{\Gamma_j} (sI - A)^{-1} ds = \frac{1}{2\pi i} \int_{\Gamma_j} \mathfrak{L}[e^{tA}](s) ds,$$

for a given $\lambda_j \in \sigma(A)$.

The second code shows the MATHEMATICA implementation for the formula (3.4)

Notice that the line 21 can be adjusted to use different holomorphic functions.

MATHEMATICA code for calculating A_i provided A is a given matrix

```
LEM=Tally [Eigenvalues[A]];
{p,q}=Dimensions[LEM];
n=Dimensions[A][[1]];
For[i=1,i <= p, i++,
  Ep[i]=
    IdentityMatrix[n]-
      PseudoInverse[
        MatrixPower[A-LEM[[i]][[1]]*IdentityMatrix[n],
          LEM[[i]][[2]]].MatrixPower[
            A-LEM[[i]][[1]]*IdentityMatrix[n], LEM[[i]][[2]]];
  Fp[i]=
    IdentityMatrix[
      n-(MatrixPower[A-LEM[[i]][[1]]*IdentityMatrix[n],
        LEM[[i]][[2]]).PseudoInverse[
          MatrixPower[A-LEM[[i]][[1]]*IdentityMatrix[n],
            LEM[[i]][[2]]];
  K[i]=PseudoInverse[Fp[i].Ep[i]]
];
```

MATHEMATICA code for calculating $f(tM)$ provided M is a given matrix

```
LEM=Tally [Eigenvalues[M]];
{p,q}=Dimensions[LEM];
n=Dimensions[M][[1]];
For[i=1,i <= p, i++,
  Ep[i]=
    IdentityMatrix[n]-
      PseudoInverse[
        MatrixPower[M-LEM[[i]][[1]]*IdentityMatrix[n],
          LEM[[i]][[2]]].MatrixPower[
            M-LEM[[i]][[1]]*IdentityMatrix[n], LEM[[i]][[2]]];
  Print["Ep",i,""]=", Ep[i]];
  Fp[i]=
    IdentityMatrix[
      n-(MatrixPower[M-LEM[[i]][[1]]*IdentityMatrix[n],
        LEM[[i]][[2]]).PseudoInverse[
          MatrixPower[M-LEM[[i]][[1]]*IdentityMatrix[n],
            LEM[[i]][[2]]];
  Print["Fp",i,""]=", Fp[i]];
  K[i]=PseudoInverse[Fp[i].Ep[i]];
  Print["K",i,""]=", K[i]];
];
f[_t]:=Exp[t];
maxdev=Max[Part[LEM, All, 2]];
For[i=0, i <= maxdev, i++, Derivada[i]=Derivative[i][f][y]];
FA=Sum[
  Sum[t^k/Factorial[k]*(Derivada[k] /.
    y->t*LEM[[i]][[1]])*(K[i].(MatrixPower
      [M-LEM[[i]][[1]]*IdentityMatrix[n],k])), {k,
    0, LEM[[i]][[2]]-1}], {i,1,p}];
MatrixForm[FA]
```

5 The Rodrigues Coefficients of Matrix Functions

The exponential map $\exp : gl(n, \mathbb{C}) = M_n(\mathbb{C}) \rightarrow GL(n, \mathbb{C})$, where $GL(n, \mathbb{C})$ denotes the Lie group of real invertible $n \times n$ matrices, is defined by

$$(5.1) \quad \exp(X) = \sum_{k=0}^{\infty} \frac{1}{k!} X^k.$$

According to the well-known Hamilton-Cayley-Frobenius Theorem, it follows that every power X^k , $k \geq n$, is a linear combination of powers $X^0, X^1, \dots, X^{n-1}$, hence we can write

$$(5.2) \quad \exp(X) = \sum_{k=0}^{n-1} a_k(X) X^k,$$

where the real coefficients $a_0(X), \dots, a_{n-1}(X)$ are uniquely defined and depend on the matrix X .

From this formula, it follows that $\exp(X)$ is a polynomial of X . The problem to find a reasonable formula for $\exp(X)$ is reduced to the problem to determine the coefficients $a_0(X), \dots, a_{n-1}(X)$. We will call this general question, the **Rodrigues problem**, and the numbers $a_0(X), \dots, a_{n-1}(X)$ the **Rodrigues coefficients** of the exponential map with respect to the matrix $X \in M_n(\mathbb{R})$.

The origin of this problem is the classical Rodrigues formula obtained in 1840 for the special orthogonal group $SO(3)$:

$$\exp(X) = I_3 + \frac{\sin \theta}{\theta} X + \frac{1 - \cos \theta}{\theta^2} X^2,$$

where $\sqrt{2}\theta = \|X\|$ is the Frobenius norm of the matrix X . From the many arguments pointing out the importance of this formula we mention here the study of the rigid body rotation in $\mathbb{R}^3$, and the parametrization of the rotations in $\mathbb{R}^3$.

The general idea of construction of such kind of matrix function is to consider an analytic function $f(z) = \alpha_0 + \alpha_1 z + \dots + \alpha_m z^m + \dots$, such that

the induced series $\alpha_0 I_n + \alpha_1 X + \cdots + \alpha_m X^m + \cdots$ is convergent in $M_n(\mathbb{C})$. Then, via Hamilton-Cayley-Frobenius theorem we can write a reduced form for this matrix $f(X)$, that is

$$(5.3) \quad f(X) = \sum_{k=0}^{n-1} a_k^{(f)}(X) X^k.$$

We call the above relation, the **Rodrigues formula** with respect to f . The numbers $a_0^{(f)}(X), \dots, a_{n-1}^{(f)}(X)$ are the **Rodrigues coefficients** of the map f with respect to the matrix $X \in M_n(\mathbb{C})$. Clearly, the real coefficients $a_0^{(f)}(X), \dots, a_{n-1}^{(f)}(X)$ are uniquely defined, they depend on the matrix X , and $f(X)$ is a polynomial of X with coefficients depending on X .

An important property of the Rodrigues coefficients is the invariance under the matrix conjugacy, i.e. the following result holds: For every invertible matrix U the following relations hold

$$(5.4) \quad a_k^{(f)}(UXU^{-1}) = a_k^{(f)}(X), k = 0, \dots, n-1.$$

The Frobenius covariants A_j are polynomials in A , that is we have $A_j = p_j(A)$, $j = 1, \dots, \mu$. Expanding $(A - \lambda_j I_n)^k$ in Schwerdtfeger formula (3.1) we get

$$(5.5) \quad \sum_{k=0}^{n-1} a_k^{(f)}(A) A^k = \sum_{j=1}^{\mu} p_j(A) \sum_{k=0}^{m_j-1} \frac{1}{k!} f^{(k)}(\lambda_j) \sum_{s=0}^k (-1)^s \lambda_j^s A^{k-s}$$

Identifying the coefficient of A^k in (5.5) we obtain the Rodrigues coefficients $a_k^{(f)}(A)$, for $k = 0, \dots, n-1$, but this way involves to many computations.

6 Computing the Rodrigues coefficients via polynomial interpolation

The **minimal polynomial** of $A \in M_n(\mathbb{C})$ is defined to be the unique monic polynomial ψ_A of lowest degree such that $\psi_A(A) = O_n$. The existence of ψ_A is easily proved. A key property is that the minimal polynomial divides any other polynomial p for which $p(A) = O_n$. Indeed, by polynomial long

division any such p can be written as $p = \psi q + r$, where the degree of the remainder r is less than that of ψ_A . But $O_n = p(A) = \psi_A(A)q(A) + r(A) = r(A)$, and this contradicts the minimality of the degree of ψ_A unless $r = 0$. Hence $r = 0$ and ψ_A divides p .

Lemma 6.1 *For complex polynomials p and q and $A \in M_n(\mathbb{C})$ we have $p(A) = q(A)$ if and only if p and q take the same values on the spectrum of A .*

Proof. Suppose that two polynomials p and q satisfy $p(A) = q(A)$. Then $d = p - q$ is zero at A so is divisible by the minimal polynomial ψ_A . In other words, d takes only the value zero on the spectrum of A , that is, p and q takes the same values on the spectrum of A .

Conversely, suppose p and q take the same values on the spectrum of A . Then $d = p - q$ is zero on the spectrum of A and so must be divisible by the minimum polynomial ψ_A . Hence $d = \psi_A h$ for some polynomial h , and since $d(A) = \psi_A(A)r(A) = O_n$, it follows that $p(A) = q(A)$. ■

Thus it is a property of polynomials that the matrix $p(A)$ is completely determined by the values of p on the spectrum of A . It is natural to generalize this property to arbitrary functions and define $f(A)$ in such a way that $f(A)$ is completely determined by the values of f on the spectrum of A .

Using Lemma 6.1, we obtain the following result: Let f be defined on the spectrum of $A \in M_n(\mathbb{C})$ and let ψ_A be the minimal polynomial of A . Then $f(A) = p(A)$, where p is the polynomial of degree less than $\sum_{i=1}^s n_i = \deg \psi_A$ that satisfies the interpolation conditions

$$(6.1) \quad p^{(j)}(\lambda_i) = f^{(j)}(\lambda_i), j = 0, \dots, n_i - 1, i = 1, \dots, s.$$

There is a unique such p and it is known as the **Hermite interpolating polynomial**.

The Hermite interpolating polynomial p is given by the Lagrange-Hermite formula

$$(6.2) \quad p(t) = \sum_{i=1}^s \left[\left(\sum_{j=0}^{n_i-1} \frac{1}{j!} \phi_i^{(j)}(\lambda_i) (t - \lambda_i)^j \right) \prod_{j \neq i} (t - \lambda_j)^{n_j} \right],$$

where $\phi_i(t) = f(t) / \prod_{j \neq i} (t - \lambda_j)^{n_j}$. For a matrix with distinct eigenvalues we have $n_i \equiv 1$ and $s = n$, hence this formula reduces to the familiar Lagrange form

$$(6.3) \quad p(t) = \sum_{i=1}^n f(\lambda_i) l_i(t), \quad l_i(t) = \prod_{j=1, j \neq i}^n \frac{t - \lambda_j}{\lambda_i - \lambda_j}.$$

Using this technique we obtain a different way to recover formulas derived in the paper [2] for the general linear group $GL(n, \mathbb{R})$ and the special orthogonal group $SO(n, \mathbb{R})$, with the exponential map and the Cayley transform and $n = 2, 3, 4$.

An elegant alternative to (6.2) is the Newton divided difference form

$$(6.4) \quad p(t) = f[x_1] + f[x_1, x_2](t - x_1) + f[x_1, x_2, x_3](t - x_1)(t - x_2) + \dots \\ + f[x_1, x_2, \dots, x_m](t - x_1)(t - x_2) \dots (t - x_{m-1}),$$

where $m = \deg \psi_A$ and the set $x_{i=1}^m$ comprises the distinct eigenvalues $\lambda_1, \dots, \lambda_s$ with λ_i having multiplicity n_i . Here the $f[\dots]$ denote divided differences on the mentioned points. Another polynomial q for which $f(A) = q(A)$ is given by (6.4) with $m = n$ and $x_{i=1}^m$ the set of all n eigenvalues of A :

$$(6.5) \quad q(t) = f[\lambda_1] + f[\lambda_1, \lambda_2](t - \lambda_1) + f[\lambda_1, \lambda_2, \lambda_3](t - \lambda_1)(t - \lambda_2) + \dots \\ + f[\lambda_1, \lambda_2, \dots, \lambda_n](t - \lambda_1)(t - \lambda_2) \dots (t - \lambda_{n-1}).$$

This polynomial is independent of the Jordan structure of A and is in general of higher degree than p . However, the properties of divided differences ensure that q and p take the same values on the spectrum of

A , so $q(A) = p(A) = f(A)$.

References

- [1] D. Andrica, I.N. Casu, *Lie Groups, Exponential Map and Geometric Mechanics* (Romanian), Cluj University Press, 2008.
- [2] D. Andrica, O.L. Chender (Broaina), *A New Way to Compute the Rodrigues Coefficients of Functions of the Lie Groups of Matrices* in "Essays in Mathematics and its Applications" in the Honor of Vladimir Arnold, (Th.H. Rassias, P.M. Pardalos, Eds.), Springer, 2016, pp. 1-24.
- [3] J.H. Gallier, *Geometric Methods and Application, for Computer Science and Engineering*, TAM, Vol. 38, Springer, 2011.
- [4] J. Gallier, D. Xu, *Computing exponentials of skew-symmetric matrices and logarithms of orthogonal matrices*, International Journal of Robotics and Automation, Vol. 17, No. 4, 2002, 2-11.
- [5] R.A. Horn, Ch.R. Johnson, *Matrix Analysis*, Cambridge University Press, 1990.
- [6] S. Mondal, *Surjectivity of Maps Induced on Matrices by Polynomials and Entire Functions*, The Amer. Math. Monthly, 124, March 2017, 260-264.