Analytic Functions of Matrices

雀, E-mail Address

June 7, 2023

Abstract

In this paper, we will expand the conventional power series to a matrix power series, find its condition for convergence, and finally redefine the domain of an analytic function to the set of square matrices whose spectral radius is smaller than the original radius of convergence of the power series.

1 The Cayley-Hamilton Theorem : Revisited

Cayley-Hamilton Theorem은 $M_n(\mathbb{C})$의 임의의 행렬 A와 그 특성다항식 $p_A(\lambda)$에 대하여 $p_A(A) = \mathbf{0}_{n\times n}$이 성립한다는 정리이다. A의 n개의 고윳값들을 $\lambda_1, \lambda_2, \ldots, \lambda_n$이라 하면

$$\det(\lambda I_n - A) = p_A(\lambda) = (\lambda - \lambda_1)(\lambda - \lambda_2)\cdots(\lambda - \lambda_n)$$

에서

$$(-1)^n \det(A) = \det(-A) = p_A(0) = (-1)^n \lambda_1 \lambda_2 \cdots \lambda_n$$

이므로

$$\det(A) = \prod_{i=1}^{n} \lambda_i = \lambda_1 \lambda_2 \cdots \lambda_n$$

이다. 또한 Jordan Decomposition Theorem에 의해 $J = P^{-1}AP$가 되도록 하는 가역행렬 P가 존재하고, 여기서 J는 다음의 block diagonal matrix이며, 각각의 J_i는 주대각선의 A의 고윳값 λ_i, 주대각선 바로 위에는 1, 나머지에는 0이 놓인
정사각행렬이다.

$$J = \begin{bmatrix} J_1 & & & \\ & J_2 & & \\ & & \ddots & \\ & & & J_p \end{bmatrix}, \qquad J_i = \begin{bmatrix} \lambda_i & 1 & & \\ & \lambda_i & \ddots & \\ & & \ddots & 1 \\ & & & \lambda_i \end{bmatrix}$$

따라서 $A = PJP^{-1}$이고 tr의 성질에 의해

$$\operatorname{tr}(A) = \operatorname{tr}(PJP^{-1}) = \operatorname{tr}(JP^{-1}P) = \operatorname{tr}(J) = \sum_i \operatorname{tr}(J_i)$$

가 성립한다. 또한 각각의 J_i에 대하여 $\operatorname{tr}(J_i)$는 고윳값 λ_i의 대수적 중복도 m_i에 대하여 $m_i\lambda_i$이므로 $\operatorname{tr}(A)$는 다음과 같이 구해진다.

$$\operatorname{tr}(A) = \sum_{i=1}^{n} \lambda_i = \lambda_1 + \lambda_2 + \cdots + \lambda_n$$

여기서 λ_i는 중복을 허용한 고윳값들이다.

Jordan Decomposition Theorem

Every square complex matrix A is similar to a block diagonal matrix

$$J = \begin{bmatrix} J_1 & & & \\ & J_2 & & \\ & & \ddots & \\ & & & J_p \end{bmatrix}$$

where each block J_i is a square matrix of the form

$$J_i = \begin{bmatrix} \lambda_i & 1 & & \\ & \lambda_i & \ddots & \\ & & \ddots & 1 \\ & & & \lambda_i \end{bmatrix}$$

and each λ_i is an eigenvalue of A. Hence, there exists an invertible matrix P such that $P^{-1}AP = J$ is such that the only non-zero entries of J are on the diagonal and the super-diagonal. J is called the **Jordan normal form** of A, and each J_i is called a **Jordan block** of A.

한편 A의 임의의 고윳값 λ_i와 이때의 고유벡터 $\mathbf{v}_i$에 대하여

$$A\mathbf{v}_i = \lambda_i \mathbf{v}_i$$

이므로

$$\begin{aligned} A^2\mathbf{v}_i &= A(\lambda_i\mathbf{v}_i) = \lambda_i(A\mathbf{v}_i) = \lambda_i^2\mathbf{v}_i \\ A^3\mathbf{v}_i &= A(\lambda_i^2\mathbf{v}_i) = \lambda_i^2(A\mathbf{v}_i) = \lambda_i^3\mathbf{v}_i \\ &\vdots \\ A^k\mathbf{v}_i &= A(\lambda_i^{k-1}\mathbf{v}_i) = \lambda_i^{k-1}(A\mathbf{v}_i) = \lambda_i^k\mathbf{v}_i \end{aligned}$$

이고 A^k의 고윳값은 λ_i^k가 되어 다음이 성립한다. $(k \in \mathbb{N})$

$$\mathrm{tr}(A^k) = \sum_{i=1}^{n} \lambda_i^k = \lambda_1^k + \lambda_2^k + \cdots + \lambda_n^k$$

따라서 이러한 관계를 이용하여 A의 특성다항식 $p_A(\lambda)$의 모든 항의 계수를 $\det(A)$와 $\mathrm{tr}(A)$를 이용하여 구할 수 있으며, λ^{n-k}의 계수 c_{n-k}는 complete exponential Bell polynomial을 이용하여 다음과 같이 쓸 수 있다.

$$c_{n-k} = \frac{(-1)^k}{k!} B_k(s_1, -1!s_2, 2!s_3, \ldots, (-1)^{k-1}(k-1)!s_k)$$

Exterior Algebra의 개념을 이용하여 이를 더 직접적인 형태로 바꾸면 다음과 같이 쓸 수 있다.

$$p_A(\lambda) = \sum_{k=0}^{n} \lambda^{n-k} \frac{(-1)^k}{k!} \begin{vmatrix} \mathrm{tr}(A) & k-1 & 0 & \cdots & 0 \\ \mathrm{tr}(A^2) & \mathrm{tr}(A) & k-2 & \cdots & 0 \\ \vdots & \vdots & & \ddots & \vdots \\ \mathrm{tr}(A^{k-1}) & \mathrm{tr}(A^{k-2}) & & \cdots & 1 \\ \mathrm{tr}(A^k) & \mathrm{tr}(A^{k-1}) & & \cdots & \mathrm{tr}(A) \end{vmatrix}$$

계수들은 뉴턴 항등식이나 Faddeev-LeVerrier 알고리즘을 이용하여 직접 도출할 수 있다. 2×2 행렬의 특성다항식이

$$p_A(\lambda) = \lambda^2 - \mathrm{tr}(A)\lambda + \det(A)$$

임은 잘 알려져 있지만, 위 알고리즘을 이용하면 임의의 $n \times n$ 행렬의 특성다항식을 오직 $\mathrm{tr}(A^k)$만을 이용하여 구할 수 있다. 예를 들어, $A \in M_3(\mathbb{C})$일 때 A의 특성다항식은

$$\boxed{\begin{aligned} p_A(\lambda) = \lambda^3 - \mathrm{tr}(A)\lambda^2 &+ \frac{1}{2}\left[(\mathrm{tr}(A))^2 - \mathrm{tr}(A^2)\right]\lambda \\ &- \frac{1}{6}\left[(\mathrm{tr}(A))^3 - 3\,\mathrm{tr}(A)\,\mathrm{tr}(A^2) + 2\,\mathrm{tr}(A^3)\right] \end{aligned}}$$

이며, 앞서 $p_A(0) = (-1)^n \det(A) = -\det(A)$임을 보였으므로 이를 이용하여 $\det(A)$를 다음과 같이 쓸 수 있다.

$$\det(A) = \frac{1}{6}\left[(\mathrm{tr}(A))^3 - 3\,\mathrm{tr}(A)\,\mathrm{tr}(A^2) + 2\,\mathrm{tr}(A^3)\right]$$

4×4 행렬 A에 대해서 특성다항식 $p_A(\lambda)$는

$$\boxed{\begin{aligned} p_A(\lambda) = \lambda^4 - \mathrm{tr}(A)\lambda^3 &+ \frac{1}{2}\left[(\mathrm{tr}(A))^2 - \mathrm{tr}(A^2)\right]\lambda^2 \\ &- \frac{1}{6}\left[(\mathrm{tr}(A))^3 - 3\,\mathrm{tr}(A)\,\mathrm{tr}(A^2) + 2\,\mathrm{tr}(A^3)\right]\lambda + \det(A) \end{aligned}}$$

와 같이 구해지며, 여기서 $\det(A)$는

$$\det(A) = \frac{1}{24}\left[(\mathrm{tr}(A))^4 - 6\,\mathrm{tr}(A^2)(\mathrm{tr}(A))^2 + 3(\mathrm{tr}(A^2))^2 + 8\,\mathrm{tr}(A^3)\,\mathrm{tr}(A) - 6\,\mathrm{tr}(A^4)\right]$$

이다. 따라서 4×4 행렬 A에 대한 Cayley-Hamilton Theorem은 다음과 같다.

The Cayley-Hamilton Theorem for 4×4 Matrices

$$\begin{aligned} \mathbf{0}_{4\times4} = p_A(A) = A^4 - \mathrm{tr}(A)A^3 &+ \frac{1}{2}\left[(\mathrm{tr}(A))^2 - \mathrm{tr}(A^2)\right]A^2 \\ &- \frac{1}{6}\left[(\mathrm{tr}(A))^3 - 3\,\mathrm{tr}(A)\,\mathrm{tr}(A^2) + 2\,\mathrm{tr}(A^3)\right]A + \det(A)I_4 \end{aligned}$$

2 A Direct Proof of the Cayley-Hamilton Theorem

넘어가기에 앞서 Cayley-Hamilton Theorem의 직접적인 증명을 제시하도록 하겠다. (아래의 방법 외에도 여러 가지 증명 방법이 알려져 있다.)

$n \times n$ 행렬 A에 대하여 $B = \operatorname{adj}(\lambda I_n - A)$라 정의하자. 여기서 adj는 고전적 수반 행렬 연산자 (Classical Adjoint / Adjugate)이다. 또한 임의의 정사각행렬 $C \in M_n(\mathbb{C})$에 대하여 $C \operatorname{adj}(C) = \det(C) I_n$임이 알려져 있으므로 다음이 성립한다.

$$(\lambda I_n - A)B = \det(\lambda I_n - A)I_n = p_A(\lambda)I_n$$

여기서 $p_A(\lambda)$는 A의 특성다항식이다. B의 각 원소는 λ에 관한 다항식이므로 각 λ^i을 골라 내어 다음과 같이 쓸 수 있다.

$$B = \sum_{i=0}^{n-1} \lambda^i B_i$$

여기서 B_i는 온전히 숫자(스칼라)로만 이루어진 행렬이다. 따라서

$$\begin{aligned} p_A(\lambda)I_n = (\lambda I_n - A)B &= (\lambda I_n - A)\sum_{i=0}^{n-1} \lambda^i B_i \\ &= \sum_{i=0}^{n-1} \lambda^{i+1} B_i - \sum_{i=0}]^{n-1} \lambda^i AB_i \\ &= \lambda^n B_{n-1} + \sum_{i=1}^{n-1} \lambda^i (B_{i-1} - AB_i) - AB_0 \end{aligned}$$

한편

$$p_A(\lambda)I_n = \lambda^n I_n + \lambda^{n-1} c_{n-1} I_n + \cdots + \lambda c_1 I_n + c_0 I_n$$

이고, 위 식 전개의 등호가 성립하기 위해서는 각 λ^i에 곱해진 각각의 행렬이 모두 같아야 한다. 즉, 다음이 성립해야 한다.

$$B_{n-1} = I_n, \quad B_{i-1} - AB_i = c_i I_n \ (1 \le i \le n-1), \quad -AB_0 = c_0 I_n$$

위 식의 양변에 각각 A^i을 곱하여 합하면 다음을 얻는다.

$$A^n B_{n-1} + \sum_{i=1}^{n-1} (A^i B_{i-1} - A^{i+1} B_i) - AB_0 = A^n + c_{n-1} A^{n-1} + \cdots + c_1 A + c_0 I_n$$

이때 좌변은 telescoping sum의 형태가 되어 $\mathbf{0}_{n\times n}$이 되고, 우변은 $p_A(A)$가 되어 최종적으로 다음을 얻는다.

$$p_A(A) = \mathbf{0}_{n\times n}$$

이로써 일반적인 $n \times n$ 행렬에 대한 Cayley-Hamilton Theorem이 증명되었다. ■

3 Convergence of a Matrix Power Series

멱급수 표현이 다음과 같은 해석함수 f가 주어져 있다고 하자.

$$f(z) = \sum_{k=0}^{\infty} a_k z^k \tag{1}$$

$M_n(\mathbb{C})$의 한 행렬 A에 대하여 A의 특성방정식 $p_A(\lambda)$는 n차 복소계수 다항식이다. 따라서 $f(z)$를 $p_A(z)$로 나눈 몫과 나머지를 각각 $q(z)$, $r(z)$라 하면

$$f(z) = p_A(z)q(z) + r(z)$$

이고 Cayley-Hamilton Theorem에 의해 $p_A(A) = \mathbf{0}_{n\times n}$이므로 다음을 얻는다.

$$f(A) = r(A)$$

즉, A에 관한 임의의 해석함수가 수렴할 때, 그 함숫값은 A에 관한 $n-1$차 이하의 다항식으로 나타낼 수 있다. 이제 해석함수가 수렴하기 위한 조건, 즉 행렬 멱급수의 수렴반경을 구해보도록 하겠다. 식 (1)의 멱급수의 수렴반경을 R이라 하고, A의 n개의 고윳값을 중복을 허락하여 $\lambda_1, \lambda_2, \ldots, \lambda_n$이라 하자. 이때 A의 spectral radius $\rho(A) = \max\{|\lambda_1|, |\lambda_2|, \ldots, |\lambda_n|\}$에 대하여 다음 정리가 성립한다.

Convergence of a Matrix Power Series

수렴반경이 R인 멱급수

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

과 고윳값이 $\lambda_1, \lambda_2, \ldots, \lambda_n$이고 spectral radius가 $\rho(A)$인 $M_n(\mathbb{C})$의 행렬 A를 고려하자. 이때 행렬 멱급수

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

이 수렴하기 위한 필요조건은

$$\rho(A) < R$$

인 것이다. 또한, $\rho(A) = R$인 경우 스칼라 멱급수와 마찬가지로 판단이 어려우며 직접 대입을 통해 수렴성을 확인해야 한다. 즉, 복소평면 상의 중심이 원점이고 반지름의 길이가 R인 원에 대하여 A의 복소 고윳값들이 모두 이 원 안에 들어올 때 위의 행렬 멱급수가 수렴한다. 고윳값들 중 단 하나라도 원 밖으로 벗어난다면 위 행렬 멱급수는 발산하며, 원의 테두리에 놓인 경우 직접 대입을 통해 수렴성을 확인해야 한다.

Proof

먼저 A의 한 고윳값 λ_i와 이때의 고유벡터 $\mathbf{v}_i$에 대하여 $|\lambda_i| > R$일 때 $f(A)$가 발산함을 증명하자. $A\mathbf{v}_i = \lambda_i \mathbf{v}_i$이므로 $A^k \mathbf{v}_i = \lambda_i^k \mathbf{v}_i$이고

$$\left[\sum_{k=0}^{N} a_k A^K\right] \mathbf{v}_i = \sum_{k=0}^{N} a_k A^k \mathbf{v}_i = \sum_{k=0}^{N} a_k \lambda_i^k \mathbf{v}_i = \left[\sum_{k=0}^{N} a_k \lambda_i^k\right] \mathbf{v}_i$$

이며, $|\lambda_i| > R$이므로 $N \to \infty$일 때 우변의 급수는 발산한다. 따라서 좌변의 급수도 발산하게 되어 $f(A)$는 발산한다. ($\mathbf{v}_i$는 $\mathbb{C}^n$의 벡터에 불과하므로 유한한 값을 가지고, 급수의 수렴성에 영향을 미치지 않는다.)

이제 A가 $A = PDP^{-1}$로 대각화 가능한 경우를 살펴보자. $D = \mathrm{diag}(\lambda_1, \lambda_2, \ldots, \lambda_n)$이므로

$$\begin{aligned}
\sum_{k=0}^{N} a_k A^k &= \sum_{k=0}^{N} a_k (PDP^{-1})^k \\
&= \sum_{k=0}^{N} a_k P D^k P^{-1} \\
&= P \left[\sum_{k=0}^{N} a_k \,\mathrm{diag}(\lambda_1^k, \lambda_2^k, \ldots, \lambda_n^k) \right] P^{-1} \\
&= P \,\mathrm{diag} \left[\sum_{k=0}^{N} a_k \lambda_1^k, \ \ \sum_{k=0}^{N} a_k \lambda_2^k, \ \ \ldots, \ \ \sum_{k=0}^{N} a_k \lambda_n^k \right] P^{-1} \\
&\xrightarrow[N \to +\infty]{} P \,\mathrm{diag}(f(\lambda_1), f(\lambda_2), \ldots, f(\lambda_n)) P^{-1}
\end{aligned}$$

이며, 모든 λ_i에 대하여 $|\lambda_i| < R$이므로 각각의 $f(\lambda_i)$가 수렴하여 마지막 식의 대각행렬 또한 잘 정의된다는 것을 알 수 있다. 즉, $f(A)$는 수렴하고 잘 정의된다.

이제 행렬 A가 대각화 불가능한 일반적인 경우를 다루기 위해 Jordan-Chevalley Decomposition을 이용하여 $A = X_S + X_N$으로 분해하자. 여기서 X_S는 A와 동일한 고윳값을 가지는 대각화 가능한 행렬이며, X_N은 $X_S X_N = X_N X_S$를 만족하는 nilpotent matrix이다. (이러한 두 행렬 X_S와 X_N이 항상 유일하게 존재한다는 것이 Jordan-Chevalley Decomposition의 내용이며, 증명은 생략하겠다.)

N은 nilpotent이므로 $N^n = \mathbf{0}_{n \times n}$이며, X_S와 X_N 사이에 곱셈의 교환법칙이 성립하므로 이항정리를 적용하여 다음과 같이 쓸 수 있다.

$$A^k = (X_N + X_S)^k = \sum_{r=0}^{k} \binom{k}{r} X_N^r X_S^{k-r} = \sum_{r=0}^{n} \binom{k}{r} X_N^r X_S^{k-r}$$

(단, $r > k$일 때 $\binom{k}{r} = 0$으로 정의한다.) 따라서 n 이상의 임의의 자연수 N에 대하여

$$\begin{aligned}
\sum_{k=0}^{N} a_k A^k &= \sum_{k=0}^{N} a_k \sum_{r=0}^{n} \binom{k}{r} X_N^r X_S^{k-r} \\
&= \sum_{r=0}^{n} X_N^r \sum_{k=r}^{N} a_k \binom{k}{r} X_S^{k-r} \\
&= \sum_{r=0}^{n} X_N^r \sum_{k=0}^{N-r} a_{k+r} \binom{k+r}{r} X_S^k
\end{aligned}$$

여기서 $f_r(z) := \sum_{k \geq 0} a_{k+r} \binom{k+r}{r} z^k$라 정의하면 $\binom{k+r}{r}$은 k에 관한 r차 다항식이므로 급수의 수렴성에 영향을 미치지 않으며, 수렴반경도 R로 유지된다. 또한 앞선 논의에서 대각화 가능한 행렬에 대하여 f가 수렴함을 보였고, X_S는 대각화 가능하므로 $f_r(X_S)$는 수렴한다. 따라서 N을 $+\infty$로 보내면 $f(A)$는 다음 값으로 수렴한다.

$$f(A) = \sum_{r=0}^{n} X_N^r f_r(X_S) = \sum_{r=0}^{n} X_N^r \sum_{k \geq 0} a_{k+r} \binom{k+r}{r} X_S^k$$

따라서 집합 $\{A \in M_n(\mathbb{C}) \mid \rho(A) < R\}$의 임의의 원소 A에 대하여 $f(A)$가 수렴하고, 잘 정의됨이 증명되었다. ■

4 Examples and Applications

이로써 우리는 행렬의 연산을 멱급수와 일반적인 해석함수로까지 확장하여, 행렬을 우리에게 익숙한 스칼라처럼 연산할 수 있는 근거를 얻었다. 다음은 행렬 멱급수의 몇 가지 예시이다.

4.1 Exponentials and Trigonometric Functions

e^z, $\sin z$, $\cos z$, $\tan z$의 멱급수 전개(맥클로린 급수)는 다음과 같다.

$$\begin{aligned}
e^z &= \sum_{n=0}^{\infty} \frac{z^n}{n!} = 1 + \frac{z}{1!} + \frac{z}{2!} + \frac{z}{3!} + \cdots \\
\sin x &= \sum_{n=0}^{\infty} (-1)^n \frac{z^{2n+1}}{(2n+1)!} = z - \frac{z^3}{3!} + \frac{z^5}{5!} - \frac{z^7}{7!} + \cdots \\
\cos z &= \sum_{n=0}^{\infty} (-1)^n \frac{z^{2n}}{(2n)!} = 1 - \frac{z^2}{2!} + \frac{z^4}{4!} - \frac{z^6}{6!} + \cdots \\
\tan z &= \sum_{n=1}^{\infty} \frac{(-1)^{n-1} 2^{2n} (2^{2n}-1) B_{2n}}{(2n)!} z^{2n-1} = z + \frac{1}{3}z^3 + \frac{2}{15}z^5 + \frac{17}{315}z^7 + \cdots
\end{aligned}$$

이때 e^z, $\sin z$, $\cos z$의 수렴반경은 ∞이며, $\tan z$의 수렴반경은 $\frac{\pi}{2}$이다. 즉, 임의의 정사각행렬 A에 대하여 e^A, $\sin A$, $\cos A$는 위의 멱급수 표현으로 잘 정의되며, $\tan A$의 경우 A의 고윳값들의 절댓값이 모두 $\frac{\pi}{2}$보다 작은 경우에 수렴한다. 이러한 정의를 이용하여 $(e^A)^{-1} = e^{-A}$, $e^{iA} = \cos A + i \sin A$, $\sin^2 A + \cos^2 A = I_n$ 등의 정리를 증명할 수 있으며, 더 나아가 A와 B 사이에 곱셈의 교환법칙이 성립할 때(즉, $AB = BA$일 때) $e^A e^B = e^{A+B}$, $\sin(A+B) = \sin A \cos B + \cos A \sin B$ 등이 성립함을 보일 수 있다. (이 글에서는 간략한 소개만 하도록 하겠다.)

또한, Jacobi formula를 이용하여 임의의 $n \times n$ 복소 정사각행렬에 대한 다음의 아름다운 정리를 증명할 수 있으며, 역시 마찬가지로 증명은 생략하도록 하겠다.

$$\boxed{\det(e^A) = e^{\mathrm{tr}(A)}}$$

또다른 예시로 2×2 정사각행렬

$$A = \begin{bmatrix} -1 & 4 \\ 2 & 1 \end{bmatrix}$$

에 대하여

$$3 \tan A = A \tan 3$$

이 성립함을 보여 보자. A의 고윳값은 ± 3이므로 $\tan z$에 A를 직접 대입하여 계산하기에는 무리가 있다. 한편 $\sin A$와 $\cos A$를 구하기 위해 A의 거듭제곱을 계산해보면

$$A^2 = \begin{bmatrix} 9 & 0 \\ 0 & 9 \end{bmatrix} = 9I_2$$

이므로 $\sin A$와 $\cos A$를 각각 다음과 같이 계산할 수 있다.

$$\sin A = \sum_{n=0}^{\infty} (-1)^n \frac{A^{2n+1}}{(2n+1)!} = \sum_{n=0}^{\infty} (-1)^n \frac{(9I_2)^n A}{(2n+1)!} = \frac{1}{3} \sum_{n=0}^{\infty} (-1)^n \frac{3^{2n+1}}{(2n+1)!} A = \frac{1}{3} A \sin 3$$

$$\cos A = \sum_{n=0}^{\infty} (-1)^n \frac{A^{2n}}{(2n)!} = \sum_{n=0}^{\infty} (-1)^n \frac{(9I_2)^n}{(2n)!} = \sum_{n=0}^{\infty} (-1)^n \frac{3^{2n}}{(2n)!} A = I_2 \cos 3$$

따라서

$$\tan A = (\sin A)(\cos A)^{-1} = \left[\frac{1}{3} A \sin 3\right] (I_2 \cos 3)^{-1} = \frac{1}{3} A \tan 3$$

이고

$$3 \tan A = A \tan 3$$

을 얻어 증명이 완료되었다. ■

4.2 Applications of the Jordan Canonical Form

앞선 칼럼에서 해석함수 f의 멱급수 전개식이

$$f(z) = \sum_{n=0}^{\infty} c_n z^n$$

과 같이 주어질 때,

$$J = \begin{bmatrix} \lambda & 1 & & \\ & \lambda & \ddots & \\ & & \ddots & 1 \\ & & & \lambda \end{bmatrix}$$

에 대하여 $f(J)$를 다음과 같이 계산할 수 있음을 보였다.

$$f(J) = \sum_{n=0}^{\infty} c_n J^n = \sum_{n=0}^{\infty} c_n \begin{bmatrix} \lambda^n & \binom{n}{1}\lambda^{n-1} & \binom{n}{2}\lambda^{n-2} & \cdots & \binom{n}{k-1}\lambda^{n-k+1} \\ 0 & \lambda^n & \binom{n}{1}\lambda^{n-1} & \cdots & \binom{n}{k-2}\lambda^{n-k+2} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & \binom{n}{1}\lambda^{n-1} \\ 0 & 0 & 0 & \cdots & \lambda^n \end{bmatrix}$$

$$= \begin{bmatrix} \sum_n c_n \lambda^n & \sum_n c_n \binom{n}{1}\lambda^{n-1} & \sum_n c_n \binom{n}{2}\lambda^{n-2} & \cdots & \sum_n c_n \binom{n}{k-1}\lambda^{n-k+1} \\ 0 & \sum_n c_n \lambda^n & \sum_n c_n \binom{n}{1}\lambda^{n-1} & \cdots & \sum_n c_n \binom{n}{k-2}\lambda^{n-k+2} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & \sum_n c_n \binom{n}{1}\lambda^{n-1} \\ 0 & 0 & 0 & \cdots & \sum_n c_n \lambda^n \end{bmatrix}$$

$$= \boxed{\begin{bmatrix} \frac{f(\lambda)}{0!} & \frac{f'(\lambda)}{1!} & \frac{f''(\lambda)}{2!} & \cdots & \frac{f^{(k-1)}(\lambda)}{(k-1)!} \\ 0 & \frac{f(\lambda)}{0!} & \frac{f'(\lambda)}{1!} & \cdots & \frac{f^{(k-2)}(\lambda)}{(k-2)!} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & \frac{f'(\lambda)}{1!} \\ 0 & 0 & 0 & \cdots & \frac{f(\lambda)}{0!} \end{bmatrix}}$$

Jordan Decomposition Theorem에 의해 임의의 $n \times n$ 복소 정사각행렬에 대하여 가역행렬 P가 존재하여 $J = P^{-1}AP$이고, J는 다음과 같은 block diagonal matrix이다.

$$J = \begin{bmatrix} J_1 & & & \\ & J_2 & & \\ & & \ddots & \\ & & & J_p \end{bmatrix}$$

이때 각각의 J_i는 주대각선에 A의 고윳값 λ_i, 주대각선 바로 위에는 1, 나머지 자리에는 0이 놓인 상삼각행렬이다. $A = PJP^{-1}$이므로 $f(A)$는 일반적으로 다음과 같이 계산할 수 있다.

$$\begin{aligned} f(A) = f(PJP^{-1}) &= \sum_{n=0}^{\infty} c_n (PJP^{-1})^n = \sum_{n=0}^{\infty} c_n P J^n P^{-1} \\ &= P \sum_{n=0}^{\infty} c_n \begin{bmatrix} J_1^n & & & \\ & J_2^n & & \\ & & \ddots & \\ & & & J_p^n \end{bmatrix} P^{-1} \\ &= P \begin{bmatrix} \sum_n c_n J_1^n & & & \\ & \sum_n c_n J_2^n & & \\ & & \ddots & \\ & & & \sum_n c_n J_p^n \end{bmatrix} P^{-1} \\ &= P \begin{bmatrix} f(J_1) & & & \\ & f(J_2) & & \\ & & \ddots & \\ & & & f(J_p) \end{bmatrix} P^{-1} \end{aligned}$$

여기서 각각의 $f(J_i)$는 다음과 같다. (m_i는 고윳값 λ_i의 대수적 중복도이다.)

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

References

[1] Frobenius, G. (1878). "Ueber lineare Substutionen und bilineare Formen". J. Reine Angew. Math. 1878 (84): 1–63. doi:10.1515/crll.1878.84.1.

[2] Abdeljaoued, Jounaidi and Lombardi, Henri (2004). Méthodes matricielles - Introduction à la complexité algébrique, (Mathématiques et Applications, 42) Springer, ISBN 3540202471 .

[3] Humphreys, James E. (1972), Introduction to Lie Algebras and Representation Theory, Springer, ISBN 978-0-387-90053-7

[4] Bhatia, R. (1997). Matrix Analysis. Graduate Texts in Mathematics. Vol. 169. Springer.

[5] Hall, Brian C. (2015), Lie groups, Lie algebras, and representations: An elementary introduction, Graduate Texts in Mathematics, vol. 222 (2nd ed.), Springer, ISBN 978-3-319-13466-6