풀이

STEP III 1998, 문제 5

순수수학 • 행렬 지수함수

STEP III 1998, 5번 문제

더 읽어 나가기 전에 문제부터 읽어 보자. 이 문제는 STEP Support Programme의 문제 데이터베이스에 있다:

step.maths.org/questions/98-s3-q5

핵심 아이디어

행렬의 무한급수(infinite series)는 처음 보면 무시무시하지만, 여기 나오는 두 행렬은 모두 거듭제곱(power)이 아주 단순한 규칙으로 반복된다. $\mathbf{M}$의 거듭제곱은 i의 거듭제곱처럼 순환하고, $\mathbf{N}$의 거듭제곱은 제곱부터 모두 영행렬이 된다. 모든 거듭제곱을 알고 나면 급수는 익숙한 스칼라 급수들로 나뉘거나(아니면 아예 유한 개의 항에서 끝나고), 그로부터 각 변환(transformation)이 무엇인지 바로 읽어 낼 수 있다. 마지막 소문항은 직접적인 행렬 계산이고, 그 기하학적 의미는 각 변환이 어떤 직선을 그대로 두는지를 물어봄으로써 드러난다.

풀이

행렬 $\mathbf{M}$

곱을 전개하면

$$\mathbf{M}^2 = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}\begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} = \begin{pmatrix} -1 & 0 \\ 0 & -1 \end{pmatrix} = -\mathbf{I},$$

이므로 요구한 결과가 나온다. 따라서 모든 정수 $k \geq 0$에 대하여

$$\mathbf{M}^{2k} = (\mathbf{M}^2)^k = (-1)^k\mathbf{I}, \qquad \mathbf{M}^{2k+1} = (-1)^k\mathbf{M}.$$

이다. 급수를 $\theta\mathbf{M}$의 짝수 거듭제곱과 홀수 거듭제곱으로 나누면

$$\begin{aligned}\exp(\theta\mathbf{M}) &= \sum_{k=0}^{\infty}\frac{\theta^{2k}}{(2k)!}\mathbf{M}^{2k} + \sum_{k=0}^{\infty}\frac{\theta^{2k+1}}{(2k+1)!}\mathbf{M}^{2k+1}\\ &= \Big(\sum_{k=0}^{\infty}\frac{(-1)^k\theta^{2k}}{(2k)!}\Big)\mathbf{I} + \Big(\sum_{k=0}^{\infty}\frac{(-1)^k\theta^{2k+1}}{(2k+1)!}\Big)\mathbf{M}\\ &= \cos\theta\,\mathbf{I} + \sin\theta\,\mathbf{M} = \begin{pmatrix}\cos\theta & -\sin\theta\\ \sin\theta & \cos\theta\end{pmatrix}.\end{aligned}$$

이다. (이렇게 나누어도 되는 이유는, 행렬 급수의 각 성분이 실수의 급수이고 그 항의 크기가 $|\theta|^r/r!$ 이하로 유계이므로 절대수렴(converge absolutely)하여 항의 순서를 바꾸어 묶을 수 있기 때문이다.)

? **생각해 보기** $\cos\theta\,\mathbf{I} + \sin\theta\,\mathbf{M}$을 전에 어디에서 본 적이 있는가?

$\mathbf{M}^2 = -\mathbf{I}$이므로 행렬 $\mathbf{M}$은 i와 똑같이 행동하고, 위의 계산은 $e^{i\theta} = \cos\theta + i\sin\theta$의 행렬 버전이다. 기하학적으로 $\exp(\theta\mathbf{M})$은 $(1,0)$을 $(\cos\theta, \sin\theta)$로, $(0,1)$을 $(-\sin\theta, \cos\theta)$로 보낸다. 즉 이것은 **원점을 중심으로 평면을 각 θ만큼 시계 반대 방향으로 돌리는 회전(rotation)**이다. (실제로 i를 곱하는 것과 마찬가지로 $\mathbf{M}$ 자체는 $\frac{\pi}{2}$만큼의 회전이다.)

행렬 N

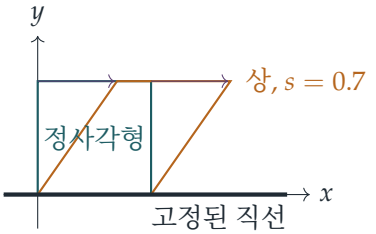
이번에는

$$\mathbf{N}^2 = \begin{pmatrix}0 & 1\\ 0 & 0\end{pmatrix}\begin{pmatrix}0 & 1\\ 0 & 0\end{pmatrix} = \begin{pmatrix}0 & 0\\ 0 & 0\end{pmatrix},$$

이므로 모든 $r \geq 2$에 대하여 $\mathbf{N}^r = \mathbf{0}$이고, 급수는 두 항에서 끝난다.

$$\exp(s\mathbf{N}) = \mathbf{I} + s\mathbf{N} = \begin{pmatrix}1 & s\\ 0 & 1\end{pmatrix}.$$

이 행렬은 (x, y)를 $(x + sy,\ y)$로 보낸다. 모든 점은 x축에 평행하게, 그 높이 y에 비례하는 거리만큼 움직이고, x축 위의 점은 전혀 움직이지 않는다. 따라서 $\exp(s\mathbf{N})$은 x축을 고정점(fixed point)들의 직선으로 갖는 **x축에 평행한 층밀림(shear)**이며, 단위정사각형을 넓이가 같은 평행사변형(parallelogram)으로 보낸다(행렬식(determinant)이 1이다).



언제 교환 가능한가?

$c = \cos\theta, \sigma = \sin\theta$라 하자. 그러면

$$\exp(s\mathbf{N})\exp(\theta\mathbf{M}) = \begin{pmatrix} 1 & s \\ 0 & 1 \end{pmatrix}\begin{pmatrix} c & -\sigma \\ \sigma & c \end{pmatrix} = \begin{pmatrix} c+s\sigma & -\sigma+sc \\ \sigma & c \end{pmatrix},$$
$$\exp(\theta\mathbf{M})\exp(s\mathbf{N}) = \begin{pmatrix} c & -\sigma \\ \sigma & c \end{pmatrix}\begin{pmatrix} 1 & s \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} c & sc-\sigma \\ \sigma & s\sigma+c \end{pmatrix}.$$

이다. 둘을 빼면

$$\exp(s\mathbf{N})\exp(\theta\mathbf{M}) - \exp(\theta\mathbf{M})\exp(s\mathbf{N}) = s\sin\theta\begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}.$$

이다. 이것이 모든 s에 대하여 영행렬일 필요충분조건은 $\sin\theta = 0$이다($\sin\theta \neq 0$이면 $s = 1$로 잡으면 된다). 따라서 두 행렬이 모든 s에 대하여 교환 가능(commute)할 필요충분조건은

$$\theta = n\pi, \qquad n \in \mathbb{Z}.$$

이다. 이 값들에 대하여 $\exp(\theta\mathbf{M}) = (-1)^n\mathbf{I}$이고, 이는 항등행렬(identity)이거나 $-\mathbf{I}$이다.

> **생각해 보기** 이 층밀림과 교환 가능한 회전은 왜 반바퀴 회전이거나 아무것도 하지 않는 회전이어야 하는가?

R를 회전, S를 층밀림($s \neq 0$)이라 하고 $SR = RS$라고 가정하자. x축 위의 임의의 점 P를 잡으면 $S(P) = P$이다. 그러면

$$S(R(P)) = R(S(P)) = R(P),$$

이므로 $R(P)$ 역시 층밀림의 고정점이다. 그런데 층밀림이 고정하는 점은 x축 위의 점뿐이다. 따라서 회전은 x축을 자기 자신으로 보내야 하고, 이런 성질을 가진 원점 중심의 회전은 0만큼의 회전(항등변환)이나 π만큼의 회전(반바퀴 회전(half-turn) $(x, y) \mapsto (-x, -y)$)뿐이다. 역으로 이 두 변환은 $\pm\mathbf{I}$이므로 **모든** 선형사상(linear map)과 교환 가능하다. 평면을 반바퀴 돌리는 것과 층밀림을 하는 것은 어느 순서로 하든 같은 결과를 준다. 그 밖의 회전은 층밀림의 고정된 직선을 기울어지게 하므로, 회전하기 전에 층밀림을 하느냐 회전한 뒤에 층밀림을 하느냐에 따라 결과가 달라진다.

검산

- 급수에서 요구하는 대로 $\theta = 0$일 때 회전행렬은 $\mathbf{I}$이고, $s = 0$일 때 층밀림은 $\mathbf{I}$이다.
- 두 행렬식은 모두 1이다($\cos^2\theta + \sin^2\theta$와 $1 \cdot 1$). 넓이를 보존하는 회전과 층밀림이라면 당연히 그래야 한다.

- 수치적으로, $\theta = 1$일 때 급수의 80개 항을 더하면 성분이 $\cos 1 \approx 0.5403$, $\sin 1 \approx 0.8415$와 일치한다. $\theta = 1, s = 1$일 때 교환자(commutator)의 대각성분은 $\pm \sin 1 \approx \pm 0.8415$이며, 이는 $s \sin\theta \,\mathrm{diag}(1, -1)$과 일치한다.

흔한 실수

- 행렬에 대해서도 $\exp(\mathbf{A}) \exp(\mathbf{B}) = \exp(\mathbf{A} + \mathbf{B})$가 성립한다고 가정하는 것. 이 법칙에는 $\mathbf{AB} = \mathbf{BA}$가 필요한데, 여기서는 $\sin\theta = 0$이 아니면 바로 이 조건이 성립하지 않는다.
- π의 모든 정수배가 아니라 $\theta = 0$만(또는 $\theta = 0, \pi$만) 답하는 것.
- $\exp(s\mathbf{N})$을 늘림(stretch)이나 확대(enlargement)라고 부르는 것. 이 변환은 넓이를 보존하고 x축의 각 점을 고정하며, 바로 그 점 때문에 층밀림이다.

STEP 너머

행렬지수함수(matrix exponential)는 **리 대수**(*Lie algebra*)와 **리 군**(*Lie group*)을 잇는 다리이다. $\mathbf{M}$은 회전들의 군을, $\mathbf{N}$은 층밀림들의 군을 생성하며, $\exp(t\mathbf{A})$는 선형 미분방정식(linear differential equation) $\frac{\mathrm{d}}{\mathrm{d}t}\mathbf{x} = \mathbf{Ax}$의 해가 된다. 교환 가능하지 않은 행렬에 대해 $\exp(\mathbf{A}) \exp(\mathbf{B}) = \exp(\mathbf{A} + \mathbf{B})$가 성립하지 않는 정도는 교환자 $\mathbf{AB} - \mathbf{BA}$로 측정되는데, 이는 마지막 소문항에서 만난 아이디어이다. 스트랭(Strang)의 *Introduction to Linear Algebra*(선형대수학 입문)는 행렬지수함수와 그것을 미분방정식에 활용하는 방법을 다룬다.

Beyond STEP 풀이: https://beyondstep.keithmcnulty.org/ko/. 풀이 글은 CC0 1.0이다. 문제 자체의 저작권은 © Cambridge University Press & Assessment에 있으며 여기에 싣지 않는다. 영어 원문을 기계 번역하였으며 아직 검토되지 않았다.