풀이

STEP II 1987, 문제 9

순수수학 • 직교행렬

> **STEP II 1987, 9번 문제**
>
> 더 읽어 나가기 전에 문제부터 읽어 보자. 이 문제는 STEP Support Programme의 문제 데이터베이스에 있다:
>
> step.maths.org/questions/87-s2-q9

핵심 아이디어

$\mathbf{B}$를 전치(transpose)하면 조건 $\mathbf{B}^{\mathrm{T}}\mathbf{B} = \mathbf{I}$의 양쪽 바깥 끝에 역행렬(inverse)이 놓이게 된다. 그 역행렬들은 가역(invertible)이므로, 정보를 잃지 않고 왼쪽에 $\mathbf{I} - \mathbf{A}^{\mathrm{T}}$를, 오른쪽에 $\mathbf{I} - \mathbf{A}$를 곱할 수 있다. 그러면 $\mathbf{A}$와 $\mathbf{A}^{\mathrm{T}}$에 대한 다항식 항등식(polynomial identity)이 남는데, 이것은 거의 곧바로 간단해진다. 주의할 점은 "필요충분조건(if and only if)"을 보여야 하므로 모든 단계가 거꾸로도 성립해야 한다는 것이다.

풀이

사용하는 표준적인 결과는 다음과 같다.

- $(\mathbf{XY})^{\mathrm{T}} = \mathbf{Y}^{\mathrm{T}}\mathbf{X}^{\mathrm{T}}$이고 전치는 선형(linear)이므로 $(\mathbf{I} \pm \mathbf{A})^{\mathrm{T}} = \mathbf{I} \pm \mathbf{A}^{\mathrm{T}}$이다($\mathbf{I}^{\mathrm{T}} = \mathbf{I}$이므로).
- $\mathbf{X}$가 비특이(non-singular)이면 $\mathbf{X}^{\mathrm{T}}$도 비특이이고, $(\mathbf{X}^{-1})^{\mathrm{T}} = (\mathbf{X}^{\mathrm{T}})^{-1}$이다.

1단계: $\mathbf{B}$의 전치. $\mathbf{I} - \mathbf{A}$가 비특이이므로 그 전치 $\mathbf{I} - \mathbf{A}^{\mathrm{T}}$도 비특이이다. 곱의 순서를 뒤집으면

$$\mathbf{B}^{\mathrm{T}} = \left((\mathbf{I} - \mathbf{A})^{-1}\right)^{\mathrm{T}}(\mathbf{I} + \mathbf{A})^{\mathrm{T}} = (\mathbf{I} - \mathbf{A}^{\mathrm{T}})^{-1}(\mathbf{I} + \mathbf{A}^{\mathrm{T}}).$$

따라서

$$\mathbf{B}^{\mathrm{T}}\mathbf{B} = (\mathbf{I} - \mathbf{A}^{\mathrm{T}})^{-1}\,(\mathbf{I} + \mathbf{A}^{\mathrm{T}})(\mathbf{I} + \mathbf{A})\,(\mathbf{I} - \mathbf{A})^{-1}.$$

2단계: 되돌릴 수 있게 역행렬 없애기. $\mathbf{P} = \mathbf{I} - \mathbf{A}^{\mathrm{T}}$, $\mathbf{Q} = \mathbf{I} - \mathbf{A}$, $\mathbf{M} = (\mathbf{I} + \mathbf{A}^{\mathrm{T}})(\mathbf{I} + \mathbf{A})$로 두면 $\mathbf{B}^{\mathrm{T}}\mathbf{B} = \mathbf{P}^{-1}\mathbf{M}\mathbf{Q}^{-1}$이다. 다음이 성립함을 보이자.

$$\mathbf{P}^{-1}\mathbf{M}\mathbf{Q}^{-1} = \mathbf{I} \quad \Longleftrightarrow \quad \mathbf{M} = \mathbf{P}\mathbf{Q}.$$

($\Rightarrow$) 왼쪽에 $\mathbf{P}$를, 오른쪽에 $\mathbf{Q}$를 곱하면 $\mathbf{M} = \mathbf{P}\mathbf{I}\mathbf{Q} = \mathbf{P}\mathbf{Q}$이다.
($\Leftarrow$) 왼쪽에 $\mathbf{P}^{-1}$을, 오른쪽에 $\mathbf{Q}^{-1}$을 곱하면 $\mathbf{P}^{-1}\mathbf{M}\mathbf{Q}^{-1} = \mathbf{P}^{-1}\mathbf{P}\mathbf{Q}\mathbf{Q}^{-1} = \mathbf{I}$이다.

> ? **생각해 보기** 왜 그냥 "양변에 곱하고" 끝낼 수는 없을까?

등식에 행렬을 곱하면 언제나 참인 등식이 나오지만, 일반적으로 그 과정을 되돌릴 수는 없다. $\mathbf{X}\mathbf{Y} = \mathbf{X}\mathbf{Z}$에서 $\mathbf{X}$가 가역이 아니면 $\mathbf{Y} = \mathbf{Z}$라고 결론지을 수 없다. 여기서 역방향이 성립하는 것은 오로지 $\mathbf{P}$와 $\mathbf{Q}$가 역행렬을 갖기 때문이며, 바로 이 지점에서 $\mathbf{I} - \mathbf{A}$가 비특이라는(따라서 $\mathbf{I} - \mathbf{A}^{\mathrm{T}}$도 비특이라는) 가정이 쓰인다. 또한 행렬의 곱셈은 교환법칙(commutative law)이 성립하지 않으므로 $\mathbf{P}$는 왼쪽에, $\mathbf{Q}$는 오른쪽에 두어야 한다는 점에도 주의하자.

3단계: 전개하기. 인수의 순서를 주의 깊게 지키면서 $\mathbf{M} = \mathbf{P}\mathbf{Q}$의 양변을 전개하면

$$(\mathbf{I} + \mathbf{A}^{\mathrm{T}})(\mathbf{I} + \mathbf{A}) = \mathbf{I} + \mathbf{A} + \mathbf{A}^{\mathrm{T}} + \mathbf{A}^{\mathrm{T}}\mathbf{A},$$
$$(\mathbf{I} - \mathbf{A}^{\mathrm{T}})(\mathbf{I} - \mathbf{A}) = \mathbf{I} - \mathbf{A} - \mathbf{A}^{\mathrm{T}} + \mathbf{A}^{\mathrm{T}}\mathbf{A}.$$

$\mathbf{I}$항과 $\mathbf{A}^{\mathrm{T}}\mathbf{A}$항이 양변에 모두 나타나므로, 빼면

$$\mathbf{M} = \mathbf{P}\mathbf{Q} \quad \Longleftrightarrow \quad 2(\mathbf{A} + \mathbf{A}^{\mathrm{T}}) = \mathbf{O} \quad \Longleftrightarrow \quad \mathbf{A} + \mathbf{A}^{\mathrm{T}} = \mathbf{O}.$$

이 각각은 모두 되돌릴 수 있다. 양변에 같은 행렬을 더하거나 빼는 것, 그리고 0이 아닌 스칼라(scalar) 2 또는 $\frac{1}{2}$를 곱하는 것이기 때문이다.

결론. 1–3단계를 이어 붙이면, 모든 고리가 동치이므로

$$\mathbf{B}^{\mathrm{T}}\mathbf{B} = \mathbf{I} \iff (\mathbf{I} + \mathbf{A}^{\mathrm{T}})(\mathbf{I} + \mathbf{A}) = (\mathbf{I} - \mathbf{A}^{\mathrm{T}})(\mathbf{I} - \mathbf{A}) \iff \mathbf{A} + \mathbf{A}^{\mathrm{T}} = \mathbf{O},$$

이고, 이것이 증명해야 할 내용이다. 말로 하면, $\mathbf{B}$가 직교(orthogonal)행렬일 필요충분조건은 $\mathbf{A}$가 반대칭(skew-symmetric)행렬인 것이다.

검산

▶ 반대칭행렬 $\mathbf{A} = \begin{pmatrix} 0 & t \\ -t & 0 \end{pmatrix}$를 생각하자. 그러면 $\det(\mathbf{I} - \mathbf{A}) = 1 + t^2 \neq 0$이고

$$\mathbf{B} = \frac{1}{1+t^2}\begin{pmatrix} 1-t^2 & 2t \\ -2t & 1-t^2 \end{pmatrix},$$

이는 회전행렬(rotation matrix)이므로($t = \tan\frac{\theta}{2}$로 두면 된다) 예상대로 $\mathbf{B}^{\mathrm{T}}\mathbf{B} = \mathbf{I}$이다.

- 1×1인 경우: $a \neq 1$인 $\mathbf{A} = (a)$에 대하여 $\mathbf{B} = \left(\frac{1+a}{1-a}\right)$이고, $\mathbf{B}^2 = 1$이면 $1 + a = \pm(1 - a)$이어야 하므로 $a = 0$이다(다른 부호는 불가능하다). 이는 $2a = 0$과 일치한다.
- 크기가 1부터 5까지인 무작위 행렬 약 2000개를 정확한 유리수 연산으로 시험해 보았다. 반대칭인 $\mathbf{A}$는 모두 정확히 $\mathbf{B}^{\mathrm{T}}\mathbf{B} = \mathbf{I}$를 주었고, 나머지는 하나도 그렇지 않았다.

흔한 실수

- 순서를 뒤집지 않고 $\mathbf{B}^{\mathrm{T}} = (\mathbf{I} + \mathbf{A}^{\mathrm{T}})(\mathbf{I} - \mathbf{A}^{\mathrm{T}})^{-1}$로 쓰는 것. ($\mathbf{I} + \mathbf{A}^{\mathrm{T}}$와 $(\mathbf{I} - \mathbf{A}^{\mathrm{T}})^{-1}$이 교환 가능하므로 이 식은 우연히 참이지만, 그 사실은 따로 정당화가 필요하다.)
- $(\mathbf{I} + \mathbf{A}^{\mathrm{T}})(\mathbf{I} + \mathbf{A})$를 전개할 때 $\mathbf{A}^{\mathrm{T}}\mathbf{A}$ 대신 $\mathbf{A}\mathbf{A}^{\mathrm{T}}$로 쓰는 것. 이 문제에서는 어느 쪽이든 소거되지만, 양변을 일관되게 전개했을 때에만 그렇다.
- 한 방향만 증명하는 것. 문제는 "필요충분조건"을 묻고 있으므로, 각 단계를 왜 되돌릴 수 있는지 설명해야 한다.

STEP 너머

사상 $\mathbf{A} \mapsto (\mathbf{I} + \mathbf{A})(\mathbf{I} - \mathbf{A})^{-1}$을 **케일리 변환**(*Cayley transform*)이라고 한다. 이 변환은 벡터공간(vector space)을 이루는 반대칭행렬들을 군(group)을 이루는 직교행렬들로 바꾸어, 회전의 유리 매개화(rational parametrisation)를 제공한다. ($\mathbf{A}$가 실수 반대칭행렬이면 $\mathbf{I} - \mathbf{A}$는 자동으로 가역이며, 이렇게 얻어지는 직교행렬은 정확히 고윳값(eigenvalue) -1을 갖지 않는 것들임을 보일 수 있다.) 이것은 리 군(Lie group)이 그 리 대수(Lie algebra)와 어떻게 연결되는지를 처음으로 엿보게 해 준다. 직교행렬과 반대칭행렬에 대해 더 알고 싶다면 스트랭(Strang)의 *Introduction to Linear Algebra*(선형대수학 입문)가 좋은 출발점이다.

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.