6.5 공리에서 출발하는 미분

학습지

문제

STEP II 1992, 4번 문제

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

step.maths.org/questions/92-s2-q4

힌트

먼저 혼자 힘으로 문제를 풀어 보자. 막히면 힌트를 **하나씩** 읽고, 힌트를 하나 읽을 때마다 다시 문제로 돌아가자. **잠깐, 직접 해 보기**라고 표시된 힌트가 핵심 단계이다. 각 힌트 뒤에는 풀이를 적을 공간이 있다.

> **생각해 보기** 규칙 (i)은 곱의 법칙이다. 곱의 법칙이 어떻게 상수에 대해 무언가를 알려 줄 수 있을까? (상수를 곱으로 생각해 보자.)

1단계 상수

다루는 내용: 상수

생각해 보기 규칙 (i)은 곱의 법칙(product rule)이다. 이 규칙이 상수(constant)에 대해 무엇을 알려 줄 수 있을까? 1을 1×1로 생각해 보아라.

잠깐, 직접 해 보기 $f = g = 1$로 놓고 규칙 (i)을 적용한다. 그다음 다른 상수에 대해서는 규칙 (iv)를 사용한다.

2단계 거듭제곱

다루는 내용: 거듭제곱 x^2과 x^3

생각해 보기 $x^2 = x \cdot x$, $x^3 = x^2 \cdot x$로 쓰고 규칙 (i)과 (iii)을 사용한다.

잠깐, 직접 해 보기 수학적 귀납법(induction)으로 $\Diamond x^n = nx^{n-1}$임을 증명한다.

3단계 모든 다항식

다루는 내용: 임의의 다항식

생각해 보기 일반적인 다항식(polynomial)을 $a_0 + a_1 x + \cdots + a_m x^m$으로 쓰고 규칙 (ii)와 (iv)를 사용한다.

숨은 수학

도분

어떤 대수적 체계 위의 연산 D가 덧셈을 보존하고 곱의 법칙 $D(fg) = fD(g) + gD(f)$를 만족하면, 이를 **도분**이라고 한다. 이 문제는 다항식 위에서 상수를 0으로 보내는 도분이 $D(x)$에 의해 완전히 결정된다는 것을 보여 준다. 실제로 $D(h) = h'(x)\,D(x)$이므로, 남은 자유는 다항식 $D(x)$를 고르는 것뿐이다. 규칙 (iii)은 $D(x) = 1$로 정한다. 도분은 수학 전반에 나타난다. 기하학에서는 곡면 위의 **벡터장**(vector field)을 그 곡면 위의 함수들에 작용하는 도분으로 정의할 수 있다(벡터장은 함수를 그 장의 방향으로 미분한다). 대수학에서는 어떤 구조의 도분이 그 구조의 "무한소 대칭"을 측정한다. 물리학에서는 양자역학의 교환자(commutator) $[A, B] = AB - BA$가 각 변수에 대해 도분이다.

극한 없는 미적분

대수적인 도함수는 극한이 전혀 의미가 없는 상황에서도 작동한다. 소수 p를 법으로 하는 산술과 같은 유한체(finite field) 위에서는 극한이 없지만, 다항식의 형식적 도함수(formal derivative)는 같은 규칙으로 정의되며 여전히 중근(repeated root)을 찾아낸다. 다항식 f가 중근을 가질 필요충분조건은 f와 f'이 공통인수를 가지는 것이다. 예를 들어 $p \nmid n$일 때 $x^n - 1$이 중근을 갖지 않음을 이런 방법으로 보이며, 이 사실은 Section 6.1의 원분다항식(cyclotomic polynomial)을 유한체 위에서 정의하는 데 필요하다. 같은 아이디어가 부호이론(coding theory)과 암호학(cryptography)에서도 쓰인다.

자동 미분

현대적인 응용을 하나 살펴보자. $\varepsilon^2 = 0$이지만 $\varepsilon \neq 0$인 수 ε을 새로 만들어 내고, **이원수**(dual numbers) $a + b\varepsilon$으로 산술을 한다. 그러면 $(a + b\varepsilon)(c + d\varepsilon) = ac + (ad + bc)\varepsilon$이다. 즉 ε 부분은 정확히 곱의 법칙을 따른다. 따라서 임의의 다항식에 대해

$$f(x + \varepsilon) = f(x) + f'(x)\,\varepsilon,$$

이고, f에 이원수 $x + \varepsilon$을 대입하여 계산하면 함숫값과 미분계수가 동시에, 정확하게 계산된다. 극한도 필요 없고, 작은 차이에서 생기는 반올림 오차도 없다(그림 6.5.1). 이를 **전진 모드 자동 미분**(forward-mode automatic differentiation)이라고 한다. 그 사촌인 후진 모드는 수십억 개의 매개변수를 가진 신경망(neural network)을 훈련시키는 데 쓰이는 알고리즘(**역전파**(backpropagation))이다. 이런 시스템이 학습할 때마다, 시스템은 곱의 법칙과 연쇄법칙(chain rule)을 도분으로서 한 단계씩 적용하고 있는 것이다.

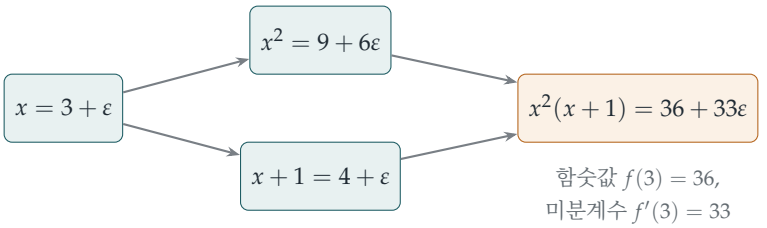

그림 6.5.1 $f(x) = x^2(x+1)$의 $x = 3$에서의 이원수($\varepsilon^2 = 0$)를 이용한 자동 미분. 각 단계는 이 문제의 규칙을 따른다. $(3+\varepsilon)^2 = 9 + 6\varepsilon$이고, $(9+6\varepsilon)(4+\varepsilon) = 36 + (9+24)\varepsilon$이다. ε의 계수는 $f'(3) = 3 \cdot 9 + 2 \cdot 3 = 33$이며, 정확하게 계산된다.

더 도전하기

1 **다른 도분.** $\Diamond$가 규칙 (i), (ii), (iv)를 만족하되, 1 대신 $\Diamond x = x$라고 하자. $\Diamond \mathrm{h}(x) = x\,\mathrm{h}'(x)$임을 보여라. $\Diamond$는 x^n을 무엇으로 보내는가?

2 **몫의 미분법.** 규칙 (i)이 여전히 성립한다고 가정하고 $\Diamond$를 유리함수(rational function) f/g로 확장하라. $\mathrm{g} \cdot \frac{\mathrm{f}}{\mathrm{g}} = \mathrm{f}$에 규칙 (i)을 적용하여 몫의 미분법(quotient rule)을 유도하라.

3 **중근.** $(x-a)^2$이 다항식 f를 나누면 $x-a$가 f′을 나눔을 보여라. 이를 이용하여 $x^n - 1$이 중근을 갖지 않음을 보여라.

4 **이원수.** 이원수를 이용하여 $f(x) = (x^2+1)^3$에 대해 $f(2)$와 $f'(2)$를 계산하고, 연쇄법칙으로 검산하라.

더 읽을거리

- **Peter Cameron, *Introduction to Algebra*** (Oxford University Press). 대학 1학년 수준에서 환, 다항식, 체를 다루며 형식적 도함수도 소개한다.
- **Andreas Griewank & Andrea Walther, *Evaluating Derivatives*** (SIAM). 자동 미분에 관한 표준 참고서이다. 고급 수준.
- **Michael Nielsen, *Neural Networks and Deep Learning*** (온라인 무료 공개). 2장에서 역전파를 명쾌하게 설명한다.

안내에 따른 전체 풀이는 제 6.5절에 있다:

https://beyondstep.keithmcnulty.org/ko/ch06/derivations.html

매주 STEP 문제 하나를 끝까지 풀어 본다. The STEP Up은 이메일로 무료로 받아 볼 수 있다: keithmcnulty.substack.com