6.5 從公理出發的微分法

工作紙

題目

STEP II 1992，第 4 題

繼續閱讀之前，請先看題目（原題為英文）。題目載於 STEP Support Programme 的試題資料庫：

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

提示

請先自行嘗試解題。如果遇到困難，請逐一閱讀提示，每讀完一個便回到題目再試。標有停一停，試一試的提示是關鍵步驟。每個提示之後均留有空位供你演算。

想一想　法則 (i) 就是積法則。積法則怎可能告訴你關於常數的任何事情？（試把常數看成一個乘積。）

階段 1 常數（constant）

涵蓋：常數

想一想 法則 (i) 是積法則（product rule）。它怎能告訴你關於常數的事情？試把 1 看成 1×1。

停一停，試一試 取 $f = g = 1$，應用法則 (i)。然後用法則 (iv) 處理其他任何常數。

階段 2　冪（power）

涵蓋：冪 x^2 和 x^3

想一想　寫成 $x^2 = x \cdot x$ 和 $x^3 = x^2 \cdot x$，並運用法則 (i) 和 (iii)。

停一停，試一試　用數學歸納法（induction）證明 $\Diamond x^n = nx^{n-1}$。

階段 3 **所有多項式（polynomial）**

涵蓋：任何多項式

? **想一想** 把一般的多項式寫成 $a_0 + a_1 x + \cdots + a_m x^m$，並運用法則 (ii) 和 (iv)。

背後的數學

導子

在一個代數系統上，若運算 D 具有可加性（additive），並且滿足積法則 $D(fg) = fD(g) + gD(f)$，便稱為導子。這道題說明，在多項式上，一個把常數變成零的導子完全由 $D(x)$ 決定：事實上 $D(h) = h'(x)\,D(x)$，所以唯一的自由度就是多項式 $D(x)$ 的選擇。法則 (iii) 規定了 $D(x) = 1$。導子遍佈數學各個領域：在幾何中，曲面上的向量場（vector field）可以定義為該曲面上的函數的一個導子（它沿着向量場的方向對函數求導）；在代數中，一個結構的導子量度它的「無窮小對稱」（infinitesimal symmetries）；在物理中，量子力學的交換子（commutator）$[A, B] = AB - BA$ 對每個變元而言都是一個導子。

沒有極限的微積分

在極限完全沒有意義的場合，代數式的導數依然行得通。在有限域（finite field）上，例如模質數 p 的算術，並沒有極限可言，但多項式的形式導數（formal derivative）仍由同樣的法則定義，而且仍然能夠偵測重根（repeated root）：多項式 f 有重根，當且僅當 f 與 f' 有公因式。舉例說，我們就是這樣證明當 $p \nmid n$ 時 $x^n - 1$ 沒有重根的；要在有限域上定義 Section 6.1 的分圓多項式（cyclotomic polynomial），就需要這個事實。編碼理論和密碼學也用到同樣的想法。

自動微分

以下是一個現代應用。創造一個數 ε，使得 $\varepsilon^2 = 0$ 但 $\varepsilon \neq 0$，然後用**對偶數**（dual number）$a + b\varepsilon$ 來做算術。這時 $(a + b\varepsilon)(c + d\varepsilon) = ac + (ad + bc)\varepsilon$：$\varepsilon$ 的部分恰好遵從積法則。因此對任何多項式，

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

而在對偶數 $x + \varepsilon$ 處計算 f，就能同時求出函數值和導數，結果完全準確，既不用極限，也沒有因微小差值而產生的捨入誤差（圖 6.5.1）。這稱為**前向模式自動微分**（forward-mode automatic differentiation）。它的近親「反向模式」（reverse mode），就是用來訓練擁有數十億個參數的神經網絡（neural network）的算法（**反向傳播**，backpropagation）：這類系統每學習一次，都是在把積法則和鏈式法則（chain rule）當作導子，一步一步地應用。

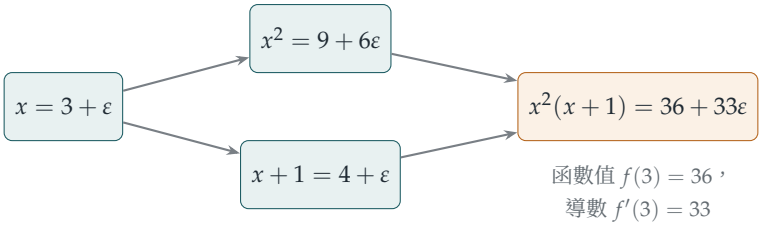

圖 6.5.1　用對偶數（$\varepsilon^2 = 0$）對 $f(x) = x^2(x+1)$ 在 $x = 3$ 處作自動微分。每一步都遵從題目中的法則：$(3+\varepsilon)^2 = 9 + 6\varepsilon$，以及 $(9 + 6\varepsilon)(4 + \varepsilon) = 36 + (9 + 24)\varepsilon$。$\varepsilon$ 的係數就是 $f'(3) = 3 \cdot 9 + 2 \cdot 3 = 33$，計算結果完全準確。

延伸練習

1. **另一個導子。**假設 $\Diamond$ 滿足法則 (i)、(ii) 和 (iv)，但 $\Diamond x = x$ 而不是 1。證明 $\Diamond \mathrm{h}(x) = x\,\mathrm{h}'(x)$。$\Diamond$ 會把 x^n 變成甚麼？

2. **商法則（quotient rule）。**把 $\Diamond$ 推廣到有理函數（rational function）f/g，並假設法則 (i) 仍然成立。把法則 (i) 應用於 $\mathrm{g} \cdot \frac{\mathrm{f}}{\mathrm{g}} = \mathrm{f}$，從而推導出商法則。

3. **重根。**證明若 $(x - a)^2$ 整除多項式 f，則 $x - a$ 整除 f'。利用這個結果證明 $x^n - 1$ 沒有重根。

4. **對偶數。**設 $f(x) = (x^2 + 1)^3$，用對偶數計算 $f(2)$ 和 $f'(2)$，並用鏈式法則驗算。

延伸閱讀

- **Peter Cameron, *Introduction to Algebra***（Oxford University Press）。大學一年級程度的環、多項式和域，並包括形式導數。
- **Andreas Griewank & Andrea Walther, *Evaluating Derivatives***（SIAM）。自動微分的標準參考書。程度較深。
- **Michael Nielsen, *Neural Networks and Deep Learning***（網上免費閱讀）。第 2 章清楚講解反向傳播。

完整的引導式解答見第 6.5 節：

https://beyondstep.keithmcnulty.org/zh-hant/ch06/derivations.html

每星期一道 STEP 題目，完整講解：The STEP Up，以電郵免費發送，訂閱網址：keithmcnulty.substack.com