详细解答

STEP I 1994，第 2 题

纯数学 · 对数微分法

STEP I 1994，第 2 题

继续阅读之前，请先看题目（原题为英文）。题目载于 STEP Support Programme 的试题资料库：

step.maths.org/questions/94-s1-q2

关键思路

当变量在底数而指数固定时，可以用幂法则（power rule）；当底数固定而变量在指数时，可以用指数函数的求导法则（exponential rule）。一旦 x 同时出现在这两个位置，两条法则都不能直接套用。解决办法是把所有式子写成 e 的幂，或者等价地，先取对数（logarithm），再作隐函数微分（implicit differentiation）：对数能把难缠的幂化为乘积，而乘积正是我们懂得处理的。

全题假设 $x > 0$（否则 x^x 和 $\ln x$ 都没有意义），而在 (ii) 中底数 a 必须为正数。

解答

我们会反复使用的唯一工具：对 $u > 0$，

$$u^v = \mathrm{e}^{v \ln u},$$

所以若 $y = u^v$，则 $\ln y = v \ln u$，再用链式法则（chain rule）求导，得 $\dfrac{1}{y}\dfrac{\mathrm{d}y}{\mathrm{d}x} = \dfrac{\mathrm{d}}{\mathrm{d}x}(v \ln u)$。

第 (i) 部分

这里指数是常数，所以用的就是普通的幂法则：

$$\frac{\mathrm{d}}{\mathrm{d}x}\, x^a = a\, x^{a-1}.$$

如果你想知道为何它对任意实数 a（而不只是整数）都成立，可以写 $x^a = \mathrm{e}^{a \ln x}$；由链式法则得 $\mathrm{e}^{a \ln x} \cdot \frac{a}{x} = a\, x^a / x = a\, x^{a-1}$。

第 (ii) 部分

现在底数是常数。写成 $a^x = \mathrm{e}^{x\ln a}$，其中 $\ln a$ 只是一个常数，

$$\frac{\mathrm{d}}{\mathrm{d}x}\,a^x = \ln a\;\mathrm{e}^{x\ln a} = a^x \ln a.$$

请留意这与 (i) 有多大分别：答案并不是 $x\,a^{x-1}$。

第 (iii) 部分

设 $y = x^x$。取对数，得 $\ln y = x\ln x$。两边求导，右边用积法则（product rule），

$$\frac{1}{y}\frac{\mathrm{d}y}{\mathrm{d}x} = \ln x + x\cdot\frac{1}{x} = 1 + \ln x,$$

所以

$$\frac{\mathrm{d}}{\mathrm{d}x}\,x^x = x^x\,(1+\ln x).$$

想一想 为甚么不能直接用幂法则，得出 $x\cdot x^{x-1} = x^x$?

幂法则假设指数是常数，而指数函数的求导法则假设底数是常数。这里两者都在变。事实上，答案恰好是两个“错误”答案之和：把指数当作固定，得 $x\cdot x^{x-1} = x^x$；把底数当作固定，得 $x^x\ln x$；而真正的导数是 $x^x + x^x\ln x$。这其实是多元链式法则（multivariable chain rule）的变相：当 $y = u^v$ 时，每个变量各自贡献它的变化率。

第 (iv) 部分

设 $y = x^{(x^x)}$，这是一座幂塔，其指数本身就是 x^x。取对数，

$$\ln y = x^x \ln x.$$

用积法则对右边求导，其中 x^x 的导数取自 (iii)：

$$\begin{aligned}\frac{1}{y}\frac{\mathrm{d}y}{\mathrm{d}x} &= x^x(1+\ln x)\cdot\ln x + x^x\cdot\frac{1}{x}\\ &= x^x\left(\frac{1}{x} + \ln x + (\ln x)^2\right).\end{aligned}$$

乘以 $y = x^{(x^x)}$，并把 x 的幂合并，

$$\frac{\mathrm{d}}{\mathrm{d}x}\,x^{(x^x)} = x^{(x^x)}\,x^x\left(\frac{1}{x} + \ln x + (\ln x)^2\right) = x^{x^x+x}\left(\frac{1}{x} + \ln x + (\ln x)^2\right).$$

等价地，把 $\frac{1}{x}$ 收进括号内，得 $x^{x^x+x-1}(1 + x\ln x + x(\ln x)^2)$。

第 (v) 部分

想一想 $(x^x)^x$ 不就是 (iv) 中的式子吗?

不是。根据指数定律（law of indices）$(u^m)^n = u^{mn}$，我们有 $(x^x)^x = x^{x\cdot x} = x^{x^2}$，而 (iv) 中的指数是 x^x，不是 x^2。两者碰巧在 $x = 1$ 和 $x = 2$ 时相等（在 $x = 2$ 时都等于 16），但在 $x = 3$ 时，一个是 3^{27}，另一个只是 3^9。

因此设 $y = x^{x^2}$。于是 $\ln y = x^2 \ln x$，由积法则得

$$\frac{1}{y}\frac{\mathrm{d}y}{\mathrm{d}x} = 2x\ln x + x^2\cdot\frac{1}{x} = x\,(1+2\ln x).$$

因此

$$\frac{\mathrm{d}}{\mathrm{d}x}\,(x^x)^x = x^{x^2}\cdot x\,(1+2\ln x) = x^{x^2+1}\,(1+2\ln x).$$

（你也可以直接对 $(x^x)^x$ 求导：$\ln y = x\ln(x^x) = x^2\ln x$，结果一样。）

验算

- 在 $x = 1$ 时，(iii)、(iv) 和 (v) 的导数全都等于 1：例如在 (iv) 中，$1^{1+1}(1+0+0) = 1$。用中心差分（central difference）$\big(f(1.000001) - f(0.999999)\big)/0.000002$ 计算，结果吻合。
- x^x 在 $x = \mathrm{e}^{-1}$ 处有极小值，该处 $1+\ln x = 0$；而 $(x^x)^x = x^{x^2}$ 在 $x = \mathrm{e}^{-1/2}$ 处有极小值，该处 $1+2\ln x = 0$。两者都可以在图像计算机上看到。
- 我们在 $0.2 \le x \le 2.5$（以及 $0.2 \le a \le 4$）的范围内随机取了 2000 个点，把每条公式与数值导数比较；相对误差始终小于 10^{-9}。

常见错误

- 写成 $\frac{\mathrm{d}}{\mathrm{d}x}x^x = x\cdot x^{x-1}$ 或 $\frac{\mathrm{d}}{\mathrm{d}x}a^x = x\,a^{x-1}$：幂法则要求指数是常数。
- 把 (iv) 和 (v) 当作同一个函数。幂运算不满足结合律（associative）：一般而言 $x^{(x^x)} \neq (x^x)^x$。
- 对 $\ln y$ 求导之后，忘记乘回 y。

STEP 之外

当 a 是无理数时，例如 $x^{\sqrt{2}}$，x^a 究竟是甚么意思？大学里的答案是：先用幂级数（power series）定义 exp，再把 ln 定义为它的反函数，然后对 $u > 0$ 定义 $u^v = \mathrm{e}^{v \ln u}$；这样一来，所有指数定律和上面的导数公式都成为定理。Abbott 的 *Understanding Analysis*，或 Bartle 与 Sherbert 的 *Introduction to Real Analysis*，都对此有严谨的处理。(iii) 中“两个错误答案之和”，正是对偏导数（partial derivative）和多元链式法则的初步一瞥。

Beyond STEP 详细解答：https://beyondstep.keithmcnulty.org/zh-hans/。解答文字采用 CC0 1.0；题目本身版权属 Cambridge University Press & Assessment 所有（©），本文件不予转载。由机器从英文翻译而成；尚未经审阅。