微分形式的不变性

姜鱼

2025 年 2 月 16 日

我们接触微分，往往都是从导数出发，我们先定义了极限，再将比值的极限定义为导数$\frac{dy}{dx}$。最初，对于求导符号$\frac{dy}{dx}$来说，其中的$\frac{d}{dx}$可视作一个整体，但是随即，我们还是妥协于这种分式的形式，把分子分母独立拿开，变成单独的dx和dy。

$$dy = f'(x)\,dx \tag{1}$$

这叫做微分，可以视为其表达着具体的含义，也可以将其仅作为符号记法。但是记号永远不是最值得关注的部分，微分的意义在于我们可以望文生义地推导，例如在一个简单的解微分方程中：

$$\frac{dy}{dx} = 2txy \tag{2}$$

由于分式可以拆开，我们可以将dx和dy试做整体，很方便地进行移动，若使左右两侧变量统一，则有：

$$\frac{dy}{y} = 2txdx \tag{3}$$

将其左右两侧分别积分，有：

$$\ln y = tx^2 + C \tag{4}$$

最后得到

$$y = e^{tx^2+C} = Ae^{tx^2} \tag{5}$$

如果y和x满足如上的微分方程，我们就可以通过这样的过程，求得y和x的函数关系。初次见面，得益于我们用分式表示导数，并用d和$\int$作为互逆的记号处理无穷小量，或许这种望文生义的写法比较容易被理解和接受，那么我们思考一个更有意思的问题。

如果引入一个参数t作为中间变量，且y与x都有关于t的直接的约束$y = f(t)$和$x = g(t)$，那么此时y和x的函数关系不能用f来表示，而是用新的函数F来表示其复合函数的关系

$$y = F(x) = F[g(t)] \tag{6}$$

则y的微分将变得有趣了起来。先从函数第一定义出发，用变量的思维审视其中的变化关系，考虑到变量y对不同变量x和t的求导，有$\frac{dy}{dt} = f'(t)$和$\frac{dy}{dx} = F'(x)$，则对于不同的微分形式，理应有$dy = f'(t)dt$和$dy = F'(x)dx$，我们的问题是，其是否能形成关联，即是否有

$$dy = f'(t)dt = F'(x)dx \tag{7}$$

这样式子成立呢？答案是成立的。

想要证明这点，从计算着手的过程略有繁杂，我们从$dy=F'(x)dx$出发尝试导出$dy=f'(t)dt$。利用一下对复合函数求导的链式法则$\frac{dy}{dx}=\frac{dy}{dt}\cdot\frac{dt}{dx}$，或者用大家更为熟知而更简洁的$f'(x)=\frac{f'(t)}{g'(x)}$，我们引入变量$t$的构造方式同样作用于此处：

$$dy=F'(x)dx=F'(x)\frac{dx}{dt}\cdot dt \tag{8}$$

对于x的导数同样有$\frac{dx}{dt}=g'(t)$，所以继续改写原式为$dy=F'(x)*g'(t)dt$，最后由链式法则不难导出$F'(x)\cdot g'(t)=f'(t)$，于是得到了$dy=f'(t)dt$，得证$dy=f'(t)dt=F'(x)dx$。这个过程的运算和符号比较繁杂，但是原理实际上是非常清晰的，从$t\to y$的变化率由$t\to x$和$x\to y$的共同作用所构成是很显然的事情。

例题（多个函数之间的求导和微分关系）：我们有三个函数$a=a(x)$，$b=b(x)$和$r=r(x)$，我们利用微分求a随r的变化率

$$\frac{da}{dr}=\frac{a'(x)\,dx}{r'(x)\,dx}=\frac{a'(x)}{r'(x)}$$

同理，对其他函数也有

$$\frac{db}{dr}=\frac{b'(x)\,dx}{r'(x)\,dx}=\frac{b'(x)}{r'(x)}$$

$$\frac{da}{db}=\frac{a'(x)\,dx}{b'(x)\,dx}=\frac{a'(x)}{b'(x)}$$

例题（隐函数的求导）：暂略。

对于一阶导而言，我们可以写出$dy=f'(t)dt=F'(x)dx$，我们将这种性质叫做形式不变性。它的名字很有趣，不论是形式这个词，还是不变性这个词，都是非常抽象的，但是放在这里，却又有一种天然的美感，让我们觉得如此命名是非常合理。微分的记号由莱布尼兹发明，它看起来很优美，当然，如果不出意外的话，会破坏这种美感的意外就要来了。

例题（二阶微分是否具有形式不变性？）：我们依然承袭之前的记号，设$y=f(t)$和$x=g(t)$。对于一阶导，我们有$\frac{dy}{dx}=\frac{dy}{dt}\cdot\frac{dt}{dx}=\frac{f'(t)}{g'(t)}$，继续令变量$y$对变量$x$求导，并进行一些复杂的运算：

$$\begin{aligned}\frac{d}{dx}\left(\frac{dy}{dx}\right)&=\frac{d}{dt}\left(\frac{f'(t)}{g'(t)}\right)\cdot\frac{dt}{dx}\\&=\frac{f''(t)\,g'(t)-f'(t)\,g''(t)}{\left[g'(t)\right]^2}\cdot\frac{1}{g'(t)}\\&=\frac{f''(t)}{\left[g'(t)\right]^2}-\frac{f'(t)\,g''(t)}{\left[g'(t)\right]^3}\end{aligned}$$

这是利用导数定义求得结果的常规方法，虽然计算过程枯燥复杂，但是必然不会有错。因为与“常规方法”相对的，有一种虽然看似简单，但实际上是错误的想法：二阶微分。既然我们有$d^2f=f''(x)\,dx^2$，好像我们可以直接将其代入求导的分式，然后写出

$$\frac{d^2y}{dx^2}=\frac{f''(t)\,dt^2}{\left[g'(t)\,dt\right]^2}=\frac{f''(t)}{\left[g'(t)\right]^2}$$

有了正确的作为对比，我们当然知道这是错误的。在正确的推导过程中，最后一行等号后面我特意进行了化简，方便与下面的错误结果进行比较。他们都有$\frac{f''(t)}{[g'(t)]^2}$这一项，但是正确的答案后面还有一项减号：$\frac{f'(t)g''(t)}{[g''(t)]^3}$如果读者尝试用导数的定义推导过求导的商法则，则可以注意到缺少的那一部分的来源，如同积法则一般，是为加上分子对于分母额外的影响（简单来说，这里看

作加法，只不过结果是负的，所以是减去）。分母的g'^2可以看作是为满足齐次的一种构造。显然用微分来计算的话，结果上看是缺少了计算这一部分的过程，所以错误。

对这道题的解读确实可以到此为止，因为我们的确解释明白了正确答案和错误答案的正确和错误。但是我想对于各位来说，这种浅尝辄止的回答并不能让人尽兴。错误的角度固然可以从结果的形式分析，但是更深层次的来源究竟是什么呢？

追问的理由，是因为如果我们认可从前把微分当做独立的无穷小量，等式试做无穷小量的关联来看的话，理想中的话，这种计算理应不会出现差错，微分和求导的两种计算方式结果理应殊途同归。即便用微分直接运算，直觉来讲，在推导过程中也许也可能存在一种"修正的步骤"，可以引入或者构造一些东西，使微分趋于正确答案。这种修正的步骤其实在一阶微分的推导过程中出现过，它也并非在最开始对x和t分别微分的时候，结构和形式就是一模一样的，而是引入了一个利用链式法则换元的过程，最终构造了$dy = F'(x) * g'(t)dt = f'(t)$。

但是对于二阶微分，这仿佛并不仅仅是形式上的暂时不完美，而是一种错误，而且错误又是非常彻底的，毫无修正的办法和可能。这令我们十分好奇错误的原因。

首先我们会去思考记号问题，也就是d^2y是否合理？为了想清楚这点，我们又难免追问下去：dy的记法是否合理？思考二阶微分的错误或许有点难，但是先从本就正确的一阶微分开始思考，或许能帮我们找到突破口。$dy = f'(t)dt = F'(x)dx$的形式不变性作为结果的体现，导出了一阶求导和微分之间的关系。那么对于二阶微分，我们可以先尝试写出$d^2y = f''(t)dt^2 = F''(x)dx^2$，然后发现它是错的（验算留做习题）。

为什么对于一阶微分，它是正确的呢？如果我们细品我们推导它的过程中采取的手段，首当其冲的是链式法则，但是链式法则理应同样可以应用在二阶导数，所以它并不是导致二阶微分和一阶微分出现差异的地方。事实上，在我们关注$F(x) = F[g(t)]$的时候，一定能注意到还有一个dy作为它们统一的结果。无论通过什么渠道，所计算的dy都是一致的，这源于$y = f(t) = F(x)$，而能如此写下这个连等式，又离不开两个变量$x = g(t)$的自然约束。一阶导直接反应了同一个作为f和F两种映射法则的结果的y的变化，虽然两种映射f和g对应的两种导数f'和g'依然不同，但是如果仔细观察f'和g'的求导过程，其来源都是对y变化的反馈。

而二阶导，就没有一个y作为统一结果了。我们要各自分析f'和F'的变化率，且理应注意到在自然约束xÆ$g(t)$恒成立的情况下，$f(x)$和$f(t)$的结果是完全不一样的，$F(t)$也肯定不能等同于是$F(x)$。一定要注意到书写为函数的$F(x)$是作为一个映射过程，而不是结果（结果记作了y）。f和F是不同的映射路径，t和x是不同的映射的出发点。因为其共同导向了y，在原函数中，可以通过代入g变换，使两个映射互相转换，f变成F，或者F变成f。但在一阶导处，已经不能有类似的转换了，让f'→g'的操作并不是直接带入$x = h(t)$，这个过程已经可以说明导函数并非可以通过自然约束相关联，我们考虑的二阶导，是作为一阶导的一阶导，作为映射的变化率。

所以我们要关注从关注变量的性质，转换到关注求导映射的性质上来。

由于基于不同的变量x和t的函数，求导的映射过程是截然不同的，就如同现实中某一个物理量有与另外两个物理量的直接管理约束，但从变化的角度，基于另外两个物理量的变化关系是截然不同的。把不同的变量看成一个有方向性的东西，这种方向是一种变量到变量之间函数映射的方向，这样一个函数对应一个映射的路径，我们考虑对于同样的出发点x和t，两种不同的路径所体现的两种不同的有关变化率的性质，会为我们导向不同的结果。这个时候就无关于y层面的约束了，因为对于$? = f(t)$，f'只取决于函数f而不会有所改变，而对于$? = F(x)$，F'只取决于函数F而不会有所改变。

我们将其称作求导对函数路径的依赖，我们必须时时刻刻关注我们对什么路径进行求导。

那么到二阶微分，$\frac{dy}{dt}$和$\frac{dy}{dx}$的性质上的不同，注定了$d\left(\frac{dy}{dx}\right)$和$d\left(\frac{dy}{dt}\right)$的不同，所以对于$d^2y$没有所谓的形式不变性，只能作为“求导两次”的记号，对单一路径使用。对于不同的路径，d^y的意义是完全不一样的。

在初学导数时，我想我们所接触的例题应该都不约而同的是求$f(x)=x^2$的导数。在最初，我们不把Δx当成0而放在分母，并在最后导出了一个$2x+\Delta x$的极限，并在此终于把Δx当成了0而忽略。试问这里的忽略，是认为它存在，只是太小而不去计算呢，还是认为它明明就不存在，所以忽略的很正常呢？注意到，求导中真正使我们导出2x的原因，只在于完全平方公式的结构，这又正是二次函数$f(x)=x^2$的性质的体现，和我们所设的Δx的确无关。就算不让Δx→0，我们依然可以顺利导出2x的部分。从这里我们可以完全肯定，Δx的的确确是一种虚设的量，其意义单纯是为了导出函数决定变化量的性质的部分，而不在于增量本身。

借用物理学中的一些概念，我更习惯把趋于0的dx称为虚位移，随着$dx\to 0$而趋于0的dy称为虚功。这种叫法或许更能有效地反应出其真正的内涵。我们引入无穷小的变化量，并非真正用于计算函数，而是用于导出函数的变化性。之所以用“性”这个词，意在更注重其作为一个函数内在蕴含的性质，而不受我们外在所虚设的dx所影响。

有很多人在解释这些事情的时候会说，求导用的记号是分式，但实际上他只是一个记号，只是在一阶微分的时候恰好符合了分式的运算法则。我是很不喜欢这种说辞的，因为从来源上看，一阶的微分恰恰是在我们考虑函数的性质，并由分式表达速率的性质出发，于是自然而然得到的记号法则，所以用“恰好”这个词来解释不变性的理念是一种曲解。我们要是对其追问为什么要使用分式作为记号，亦或者追问这里的分式是否是除法，他们的回答必然会含糊不清，悄悄携带一种偏否定而又不全否定的态度敷衍了事。

我们其实完全可以肯定地说，这里用分式和d的记号并不荒谬，分式就是除法，微分就是无穷小量。反而求导这一理念，$\frac{d}{dx}$的的确确是抽象的记号，是用来表示函数到函数的映射的，由微分衍生出来的高度抽象化的事情。

当然，对于导数的定义而言，除法的确就是除法，是人类研究如何度量速率所找到的方法。但是对于不考虑人类用以描述自然科学的手段，只关注事物本身，它的速率的确应是它内在具有的一个性质。无关于差分还是微商，都是我们用以导出速率这一性质的手段。客观存在的性质都是的确客观存在的。这个时候再去考虑用映射的角度处理函数，进行求导，抽象出性质读取的本质，抛去那些容易产生误解的取极限的手段和过程，不失为一种更遵循客观自然的理解角度。

求导是一个函数向另一个函数的映射，例如由$f(x)=x^2$到$f'(x)=2x$的映射。对于求导符号$\frac{df}{dx}$来说，其中的$\frac{d}{dx}$可视作一个整体，这个过程可以将其看作是$\frac{d}{dx}f\left(x\right)\to f'\left(x\right)$。求导是原函数向导函数的映射，一种把利用形式逻辑将极限运算简单化的操作。而微分往往更重视导数意义的本身，也就是无穷小量变化的比值。我们会把$\frac{df}{dx}$视作df和dx的比值，并且可以拆开，写作

$$df=f'\left(x\right)dx$$

对于二阶导数而言，我们的求导记号是$\frac{d^2f}{dx^2}=f''\left(x\right)$，所以微分的写法可以写作$d^2f=f''\left(x\right)dx^2$，对于高阶导数，我们依此类推。