

Mathematical Analysis

枫聆

2021 年 2 月 5 日

目录

| | | |
|----------|--------------------------|-----------|
| 1 | 实数 | 2 |
| 1.1 | 实数连续性 | 2 |
| 2 | 极限论 | 3 |
| 2.1 | 数列极限 | 3 |
| 2.2 | 无穷小量 | 4 |
| 2.3 | 区间套 | 6 |
| 2.4 | 收敛原理 | 7 |
| 3 | 一元函数 | 9 |
| 3.1 | 连续函数的性质 | 9 |
| 3.2 | 函数连续性和间断 | 10 |
| 3.3 | 函数连续性在计算极限时的应用 | 11 |
| 4 | 导数及微分 | 12 |
| 4.1 | 常用导数求法和表示 | 12 |
| 4.2 | 函数的增量公式 | 13 |
| 4.3 | 复合函数的导数 | 14 |
| 4.4 | 高阶导数及高阶微分 | 15 |

实数

实数连续性

由于在有理数上划分存在一种边界无法确定的情况，即把数轴上所有有理数划分为 $A|A'$ ，其中要求 A 中所有的有理数都小于 A' 中的有理数，在 A 中无最大有理数，且 A' 中无最小有理数，这种情况下无法确定划分两者的边界，所以引入了无理数的概念，约定上面这种特殊的划分情况定义了某个无理数的 α ，让这个 α 代替缺少的界数，把它插在了 A 里面一切数 a 和 A' 里面一切数 a' 中间。

用上面这种思路来理解有理数也是可以的，对任意一有理数 r 存在两种确定它的划分，还是前面的划分方式即 $a < r$ 在下组 A 中， $a > r$ 在上组 A' 中，而有理数 r 本身可能含于 A 或者 A' ，如果在 A 中，即 A 中有最大有理数，反之在 A' 中，则有最小有理数。为了确定起见，在提及确定有理数 r 的时候，常把其置于固定的一组，即 A 和 A' 任选一个，以后一直用它，在这里取 a 在上组。

实数之间的序关系，用划分它集合对应的包含关系来描述，在有理数里面已经有这样的性质了，再看一下无理数，定义划分 $A|A'$ 确定无理数 α ，划分无理数 $B|B'$ 确定 β ，即对应下述三种关系

1. $\alpha = \beta$, A 和 B 重合, A' 和 B' 重合.
2. $\alpha > \beta$, A 包含 A' .
3. $\alpha < \beta$, A' 包含 A .

还有一个传递关系 $\alpha > \beta, \beta > \gamma$, 则 $\alpha > \gamma$, 这些性质都比较容易证明。

Lemma 1.1. 对于不论怎样地两个实数 α 和 β , 其中 $\alpha > \beta$, 恒有一个位于它们中间的有理数 r : $\alpha > r > \beta$.

证明. 这个性质更强了, 两个实数 ($\alpha > \beta$) 之间不仅有实数, 还有有理数。来证明一下, 定义 α 对应 $A|A'$ 有理数域上的划分, β 对应 $B|B'$, 因为 $\alpha > \beta$, 所以有 A 包含 B , 所以可以在 A 上取一点有理数 r 它不含于 B , 于是它属于 B' , 使得 $\beta \leq r < \alpha$, A 里面没有最大数 (按照前面的统一), 所以把 r 取的大一点就可以把等号去掉。□

开始进攻戴德金基本定理)

Theorem 1.2. 对于实数域内的任一划分 $A|A'$ 必有产生这划分的实数 β 存在, β 或是下组 A 中最大数, 或是上组 A' 中最小数。

证明. 首先还是先把实数域上的划分规定先拿出来, 定义 A 和 A' 是两个非空集合, 每一个实数必落在 A 或者 A' 其中一个里面, 且 A 里面的数都大于 A' 里面的数。

将 A 里面的一切有理数记为 A , A' 里面的一切有理数记为 A' , 容易证明这样 $A|A'$ 是一个有理数域上划分, 划分确定了一个实数 β . 它应该落在 A 或者 A' 中, 假设它落在 A 上, 则它是 A 中的最大数, 假设它不是最大数, 则还存在一个 α_0 使得 $\alpha_0 > \beta$, 根据前面的 lemma 两个实数之间又可以确定一个有理数 $\alpha_0 > r > \beta$, 与前提有理数划分的界数矛盾, 所以 β 是 A 中最大数。□

极限论

数列极限

数列，整序傻傻分不清....

Definition 2.1. 若对于每一整数 ε , 不论它怎样小, 恒有序号 N , 使在 $n > N$ 时, 一切 x_n 的指满足不等式

$$|x_n - a| < \varepsilon$$

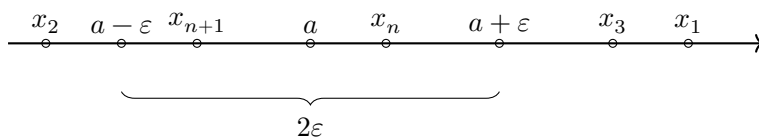
, 则称常数 a 为整序变量 $x = x_n$ 的极限.

a 是整序变量的极限这一事实, 记成:

$$\lim x_n = a \text{ 或者 } \lim x = a$$

, 也可以说这个序列收敛于 a

有一个很有趣的几何解释在这里,



以 a 点为中心的线段不论取的多小 (其长度为 2ε), 一切 x_n 点从某点起, 必全部落在这线段之内, 这样在线段之外一定只有有限长度个点了, 表示极限的点 a 表示整序变量的数值的点的凝聚中心.

无穷小量

Definition 2.2. 极限为零的整序变量 x_n 称为无穷小量, 或简称无穷小.

这里有一个有趣的命题.

Proposition 2.3. 无限个无穷小之积不一定是无穷小.

它是一个自然语言的命题, 所以里面有一些争议. 先看一个一般性构造证明.

证明. 取一系列数列:

$$\begin{aligned}\{a_1\}: & 1, \frac{1}{2}, \frac{1}{3}, \frac{1}{4}, \frac{1}{5}, \frac{1}{6}, \dots \\ \{a_2\}: & 1, 2, \frac{1}{3}, \frac{1}{4}, \frac{1}{5}, \frac{1}{6}, \dots \\ \{a_3\}: & 1, 1, 3^2, \frac{1}{4}, \frac{1}{5}, \frac{1}{6}, \dots \\ \{a_4\}: & 1, 1, 1, 4^3, \frac{1}{5}, \frac{1}{6}, \dots \\ & \vdots\end{aligned}$$

数列 $\{a_k\}$ 的第 n 项记为 $a_k(n)$, 其通项公式为

$$a_k(n) = \begin{cases} 1, & n < k \\ k^{k-1}, & n = k \\ \frac{1}{n}, & n > k \end{cases}$$

显然对于任意的 $k \in \mathbb{Z}^+$ 都满足 $\lim_{n \rightarrow \infty} a_k(n) = 0$. 所以这一系列数列都是无穷小. 那么这一系列数列乘积的第 n 项为

$$\begin{aligned}\prod_{k=1}^{\infty} a_k(n) &= \left(\prod_{k=1}^{n-1} a_k(n) \right) a_n(n) \left(\prod_{k=n+1}^{\infty} a_k(n) \right) \\ &= \left(\frac{1}{n} \right)^{n-1} \cdot n^{n-1} \cdot 1^{\infty} = 1.\end{aligned}$$

因此 $\lim_{n \rightarrow \infty} \prod_{k=1}^{\infty} a_k(n) = 1$. □

但是有一个奇怪现象是什么呢? $\prod_{k=1}^n a_k(n)$ 中有一项 $a_n(n) = n^{n-1}$ 是一个无穷大, 并不是无穷小. 奇怪的东西乱入了.

再看一个经典的无穷个无穷小之和:

$$\lim_{n \rightarrow \infty} \left(\frac{1}{n^2} + \frac{2}{n^2} + \frac{3}{n^2} + \cdots + \frac{n}{n^2} \right) = \lim_{n \rightarrow \infty} \sum_{k=1}^n \frac{k}{n^2} = \lim_{n \rightarrow \infty} \frac{1+2+3+\cdots+n}{n^2} = \lim_{n \rightarrow \infty} \frac{\frac{1}{2}n(n+1)}{n^2} = \frac{1}{2}.$$

定义 $b_k(n) = \frac{k}{n^2}$, 则 $b_n(n) = \frac{n}{n^2}$ 还是无穷小. 两个东西对比一下, 你可能需要重新定义命题.

区间套

Lemma 2.4. 给定单调增数列 x_n 和单调减数列 y_n , 且恒有

$$x_n < y_n,$$

若其差 $y_n - x_n$ 趋向于 0, 则它们有公共有限极限:

$$c = \lim x_n = \lim y_n.$$

常用形式是取一个闭区间 $[a, b]$, 然后在取一个区间 $[a', b']$ 使得

$$a \leq a' < b' \leq b$$

, 则 $[a', b']$ 是套在 $[a, b]$ 里面的.

设有一区间套的无穷序列

$$[a_1, b_1], [a_2, b_2], \dots, [a_n, b_n], \dots$$

后一个总是套在前面一个内, 并且在 n 增大时这些区间的长度趋向于 0:

$$\lim b_n - a_n = 0.$$

则区间的两端点 a_n 和 b_n 趋于共同极限

$$c = \lim a_n = \lim b_n.$$

收敛原理

Theorem 2.5. 数列变量 x_n 有有限极限的充分必要条件是: 对于每一个数 $\varepsilon > 0$, 存在序号 N , 使得 $n > N$ 及 $n' > N$, 不等式

$$|x_n - x_{n'}| < \varepsilon$$

成立.

记录一下用分割法证明其必要性.

证明. 设前提条件都已经满足, 在全体实数域下构造一个划分. 对于任何实数 α , 若 x_n 从某项其满足不等式

$$x_n > \alpha,$$

则取这种实数 α 归入下组 A , 其余的 (即不落在 A 里面的) 一起实数归入上组 A' .

首先我们来说明这样确实产生了一个实数上的划分. 由前提条件, 对于任意数 $\varepsilon > 0$ 及其对应的 N . 若 $n > N$ 及 $n' > N$, 则下面不等式成立

$$x_{n'} - \varepsilon < x_n < x_{n'} + \varepsilon$$

. 现在我们可以看到每一个数 $x_{n'} - \varepsilon$ 都是小于 x_n 的, 所以它归入下组 A . 另一方面 $x_{n'} + \varepsilon$ 都大于 x_n , 所以 $x_{n'} + \varepsilon$ 放不进去 A , 那它只能归入 A' 了, 所以 A 和 A' 都是非空的. 我们的划分方式对于每一个数 α 和确定序列 x_n , 要么它属于 A 或者属于 A' . 同时 A 中实数都小于 A' 的实数. 如果 $\alpha > \alpha', \alpha \in A, \alpha' \in A'$, 则 x_n 从某一项其也都大于 α' , 这样就产生矛盾了. 所以确实产生了一个实数上的划分.

根据戴德金基本定理, 有实数 a 存在它是这两组数之间的界数, 即

$$\alpha \leq a \leq \alpha'.$$

我们注意到当 $n > N$ 时, $x_{n'} - \varepsilon$ 是一个 α , 而 $x_{n'} + \varepsilon$ 是一个 α' . 所以我们有

$$x_{n'} - \varepsilon \leq a \leq x_{n'} + \varepsilon$$

. 即 $|x_{n'} - a| \leq \varepsilon$, 所以 $\lim x_n = a$. □

Theorem 2.6. *B.Bolzano-C.Weierstrass.* 任何有界数列内恒能选出收敛于有限极限的部分极限.(有界序列一定有收敛子列)

证明. 假设一切 x_n 都位于界限 a 与 b 之间. 现在将 $[a, b]$ 分为两半, 则必有一半含有无限多个数列 x_n 里面的元素. 因为不是这样则数列 x_n 就有有限多个了. 用 $[a_1, b_1]$ 表示其中含有无限多个数列元素的那半个区间 (若两个区间都含有无限多个数列元素, 任取一个即可). 类似的取在区间 $[a_1, a_2]$ 分出它的一半 $[a_2, b_2]$, 使得它也含有

无限多个 x_n . 继续这种取法至无穷, 在第 k 次分出的区间 $[a_k, b_k]$ 内同样也有无穷多个 x_n . 此外, 第 k 个区间的长度为

$$\frac{b-a}{2^k}.$$

可以看到这个长度趋于 0 的, 然后把区间套用在这里, 就可以知道 a_k 和 b_k 是趋于某个公共极限 c .

也就是说从上面我们分出来的区间里面, 都挑一个元素出来, 构成的子列是取趋于某个极限的. □

利用 BC 定理, 我们可以尝试把前面对收敛原理的证明可以写的更简单.

证明. 假设前提条件满足, 即对任意的 $\varepsilon > 0$, 当 $n > N, n' > N$ 时 $|x_n - x_{n'}| < \varepsilon$. 也就是

$$x_{n'} - \varepsilon < x_n < x_{n'} + \varepsilon.$$

可以看到 x_n 是有界的, 可以想象把这个界限拉长把前 N 个 x_n 也包含进来.

假设从 $\{x_n\}$ 里面挑出来某个子列 $\{x_{n_k}\}$ 收敛于 c , 即

$$|x_{n_k} - c| < \varepsilon.$$

当 $n_k > N, n > N$ 时, 我们有

$$|x_n - x_{n_k}| < \varepsilon.$$

两式联立有

$$|x_n - c| < 2\varepsilon.$$

所以 x_n 也是收敛于 c 的. □

一元函数

连续函数的性质

Lemma 3.1. E.Borel. 若闭区间 $[a, b]$ 被一个开区间的无穷系 $\Sigma = \{\sigma\}$ 所覆盖, 则恒能从 Σ 里面选出有限子系

$$\Sigma^* = \sigma_1, \sigma_2, \dots, \sigma_n,$$

它同样可以覆盖全区间 $[a, b]$.

证明. H.lebesgue's. 定义 x^* 为区间 $[a, b]$ 中使得区间 $[a, x^*]$ 能用有限个区间 σ 来覆盖的点. x^* 肯定是存在, 因为 a 就是, 只要找一个包含 a 的开区间 σ 就行, 这样想的话, 又可以找到一群, σ 中接近 a 的都是这样的点.

所以我们的任务是证明 b 也是这样的一个 x^* . 因为一切 $x^* \leq b$, 故亦有

$$\sup\{x^*\} = c \leq b.$$

, 因为 c 也是 $[a, b]$ 中一点, 同样可以找到包含它的开区间 $\sigma_0 = (\alpha, \beta)$. 但依据上确界的性质, 我们还可以找到 x^* 使得 $\alpha < x^* \leq c$. 所以现在把 σ_0 来在加到有限个区间 σ 里面去, 现在就可以覆盖区间 $[a, c]$, 也就是说上确界 c 也是 x^* .

而且 c 是不能小于 b 的, 如果 c 小于 b , 如果是这样 c 和 β 也可以找到一点 x^* , 这与 c 是上确界矛盾的. 这样, 必须有 $c = b$, 即 b 也是属于 x^* . 所以 $[a, b]$ 可以被有限覆盖. \square

为什么 (a, b) 不是紧致的呢? 考虑 $(a + \frac{b-a}{n}, b)$, 任何一个 (a, b) 的真子集都可以被它覆盖, 但是它不能有限覆盖 (a, b) , 因为如果有限就意味着我们能找到一个最大的开区间属于 (a, b) , 实际上这样的开区间并不存在. 但是如果加上 a 和 b , 这种方式已经无法覆盖 a 和 b 两点了.

Compact means small. It is a peculiar kind of small, but at its heart, compactness is a precise way of being small in the mathematical world. The smallness is peculiar because, as in the example of the open and closed intervals $(0, 1)$ and $[0, 1]$, a set can be made “smaller” (that is, compact) by adding points to it, and it can be made “larger” (non-compact) by taking points away.

函数连续性和间断

函数 $f(x)$ 在点 x_0 处右连续或者左连续

$$f(x_0 + 0) = \lim_{x \rightarrow x_0 + 0} f(x) = f(x_0) \quad (1)$$

$$f(x_0 - 0) = \lim_{x \rightarrow x_0 - 0} f(x) = f(x_0). \quad (2)$$

函数 $f(x)$ 在点 x_0 有右间断或者左间断是指对应的上式不成立. 例如第一个式子不成立则是右间断.

函数连续的充要条件可以变成函数在点 x_0 处连续就是等于说它在这一点是同时左连续和右连续的.

如果 $f(x)$ 在 x_0 处有有限极限 $f(x_0 + 0)$ 或者 $f(x_0 - 0)$ 存在, 但是它们均不等于 $f(x_0)$, 则称 x_0 是这里是一个普通间断点或者第一类间断点 (跃度). 若极限 $f(x_0 + 0)$ 或者 $f(x_0 - 0)$ 是无穷或者根本不存在, 则称 x_0 这里是第二类间断点.

Example 3.2. $f(x)$ 定义在区间 $[0, 1]$ 上: 若 x 是无理数则 $f(x) = 0$; 若 x 是有理数表示为不可约通分数 $\frac{p}{q}$ 则 $f(\frac{p}{q}) = \frac{1}{p}$. 可以得到一个有趣的结论: $f(x)$ 在任一有理数有普通间断点, 任一无理数上连续.

事实上无论 x 取任意数 x_0 , 对于任意的 $\varepsilon > 0$, 要使得 $f(x) < \varepsilon$, 只需要取 $p > \frac{1}{\varepsilon}$. 不满足这样的正整数 p 只有有限多个. 我们找一个 x_0 的邻域 $(x_0 - \delta, x_0 + \delta)$ 把这些点排除在外, 那么所有在这个邻域里面的点 (排除 x_0) 都可以满足 $|f(x)| < \varepsilon$. 即意味着

$$f(x_0 + 0) = f(x_0 - 0) = 0.$$

若 x_0 是一个有理数, 则 x_0 是一个普通间断点. 反之若 x_0 是一个无理数则在 x_0 处连续.

函数连续性在计算极限时的应用

有三个比较重要的极限.

$$1. \lim_{\alpha \rightarrow 0} \frac{\log_a(1+\alpha)}{\alpha} = \log_a e$$

直接用对数函数的性质, 把 α 放到对数函数里面就行, 对数函数里面的极限是 e . 当 $a = e$ 时极限的结果就是很漂亮的 1.

$$2. \lim_{\alpha \rightarrow 0} \frac{a^\alpha - 1}{\alpha} = \ln a$$

遇到这样略微有些复杂的表达式, 直接考虑换元. 让 $\beta = a^\alpha - 1$, 则 $\beta \rightarrow 0$. 原式就变成了 $\lim_{\beta \rightarrow 0} \frac{\beta}{\log_a(\beta+1)}$. 变成了上面我们熟悉样子.

$$3. \lim_{\alpha \rightarrow 0} \frac{(1+\alpha)^\mu - 1}{\alpha} = \mu$$

还是考虑换元, 但是不要换的太彻底, 适可而止即可. $\beta = (1+\alpha)^\mu - 1$, 其中 $\beta \rightarrow 0$. 我们可以得到一个有趣的等式 $\mu \ln(1+\alpha) = \ln(1+\beta)$. 到这里就够了, 不用把 β 把 α 表示出来. 我们把原式现在整理如下

$$\frac{(1+\alpha)^\mu - 1}{\alpha} = \frac{\beta}{\alpha} = \frac{\beta}{\ln(1+\beta)} \cdot \frac{\ln(1+\alpha)}{\alpha}.$$

又变成了我们熟悉的样子, 两边的极限都是 1, 所以最后的整体的极限为 μ .

导数及微分

导数常用的表示法.

1. $\frac{dy}{dx}$ 或 $\frac{df(x_0)}{dx}$ 莱布尼茨 (G.W.Leibniz);
2. y' 或者 $f'(x_0)$ 拉格朗日 (J.L.Lagrange);
3. Dy 或者 $Df(x_0)$ 柯西 (A.L.Cauchy).

常用导数求法和表示

1. 常函数的导数等于 0. 这个就非常 trivial 了.

2. $y = x^n$, 其中 n 是自然数. $y' = nx^{n-1}$.

$y + \Delta y = (x + \Delta x)^n$ 这个式子二项式展开即可. 即 $x^n + nx^{n-1}\Delta x + \frac{n(n-1)}{1 \cdot 2}x^{n-2}(\Delta x)^2 + \dots$.

3. $y = \frac{1}{x}$. $y' = -\frac{1}{x^2}$. 直接用导数的基本定义就行.

4. $y = \sqrt{x}$. $y' = \frac{1}{2}x^{-\frac{1}{2}}$. 直接用导数的基本定义就行.

5. $y = x^\mu$, 其中 μ 是任意实数. $y' = \mu x^{\mu-1}$.

$\frac{\Delta y}{\Delta x} = \frac{(x+\Delta x)^\mu - x^\mu}{\Delta x} = x^{\mu-1} \cdot \frac{(x+\frac{\Delta x}{x})^\mu - 1}{\frac{\Delta x}{x}}$. 其中左边极限是我们前面写过的一个重要极限值为 μ .

6. $y = a^x$, 其中 $a > 0$. $y' = a^x \cdot \ln a$.

$\frac{\Delta y}{\Delta x} = \frac{a^{x+\Delta x} - a^x}{\Delta x} = a^x \cdot \frac{a^{\Delta x} - 1}{\Delta x}$. 最后等式做边又是我们熟悉的极限.

函数的增量公式

设 $y = f(x)$. 在 x 的定义域上固定一个 x_0 , 用 $\Delta x \leq 0$ 表示 x 的任意增量. 于是对应的函数的增量为

$$\Delta y = \Delta f(x_0) = f(x_0 + \Delta x) - f(x_0).$$

若 $y = f(x)$ 在 x_0 处有有限的导数 $y'_x = f'(x_0)$. 则函数的增量可以表示如下的形式.

$$\Delta f(x_0) = f'(x_0) \cdot \Delta x + \alpha \cdot \Delta x.$$

或者更简短地

$$\Delta y = y'_x \cdot \Delta x + \alpha \cdot \Delta x.$$

式中的 α 是依赖 Δx 的变量, 且随着 Δx 一同趋于零.

这个 α 是怎么来的呢? 在导数的定义中, $\Delta x \rightarrow 0$ 时, 有

$$\frac{\Delta y}{\Delta x} \rightarrow y'_x.$$

故令

$$\alpha = \frac{\Delta y}{\Delta x} - y'_x.$$

这里可以看出来 $\alpha \rightarrow 0$. 通过这个等式把 Δy 表示出来就是前面的等式.

因为 $\alpha \cdot \Delta x$ 是比 Δ 更高阶的无穷小. 故上面的等式还可以改成写

$$\Delta y = y'_x \cdot \Delta x + o(\Delta x).$$

这个式子相对来说就非常简洁了.

复合函数的导数

设函数 $\mu = \varphi(x)$ 在某一点 x_0 处有导数 $u'_x = \varphi'(x_0)$. 函数 $y = f(\mu)$ 在对应的 $\mu_0 = \varphi(x_0)$ 也有导数 $y'_u = f'(u_0)$. 于是复合函数 $y = f(\varphi(x))$ 在 x_0 处亦有导数

$$[f(\varphi(x_0))]' = f'_u(\varphi(x_0)) \cdot \varphi'(x_0).$$

或者

$$\frac{dy}{dx} = \frac{dy}{du} \cdot \frac{du}{dx}.$$

证明. 先根据导数的定义来求 Δy . 给 x 以任意增量 Δx , Δu 是函数 $u = \varphi(x)$ 对应增量, 最后 Δy 是由增量 Δu 所引起的函数 $y = f(u)$ 的增量. 根据函数增量公式我们有

$$\Delta y = y'_u \cdot \Delta u + \alpha \cdot \Delta u.$$

然后两边都除以 Δx

$$\frac{\Delta y}{\Delta x} = y'_u \cdot \frac{\Delta u}{\Delta x} + \alpha \cdot \frac{\Delta u}{\Delta x}.$$

当 $\Delta x \rightarrow 0$ 时, $\frac{\Delta u}{\Delta x}$ 就是 u'_x , 而 $\alpha \cdot \frac{\Delta u}{\Delta x}$ 是一个趋于 0 高阶无穷小. 所以最后

$$\lim_{\Delta \rightarrow 0} \frac{\Delta y}{\Delta x} = y'_u \cdot \lim_{\Delta \rightarrow 0} \frac{\Delta u}{\Delta x} = y'_u \cdot u'_x.$$

□

这个式子也是我们常说的链式法则 (记得在以前在看 mit 的微积分的时候, 那个代课老师在讲链式法则的时候, 突然拿了一条真的链子出来. 说这个法则的强大在于让我们挣脱了链子的束缚...), 可以推广到任意有限个函数复合的情形. 我们可以尝试来推一下二阶链式法则.

$$\begin{aligned} \left(\frac{dy}{dx}\right)' &= \left(\frac{dy}{du}\right)' \cdot \frac{du}{dx} + \frac{dy}{du} \cdot \left(\frac{du}{dx}\right)' \\ &= \frac{d^2y}{du^2} \cdot \frac{du}{dx} \cdot \frac{du}{dx} + \frac{dy}{du} \cdot \frac{d^2u}{dx^2} \\ &= \frac{d^2y}{du^2} \cdot \left(\frac{du}{dx}\right)^2 + \frac{dy}{du} \cdot \frac{d^2u}{dx^2} \end{aligned} \tag{3}$$

高阶导数及高阶微分

二阶微分记为:

$$d^2y = d(dy).$$

二阶微分的微分记为:

$$d^3y = d(d^2y).$$

一般地说, 函数 $y = f(x)$ 的 $(n-1)$ 阶微分的微分称为函数 $y = f(x)$ 的 n 阶微分

$$d^n y = d(d^{n-1}y).$$

在求高阶微分时很重要的一件事, 是要记住 dx 是不依赖于 x 的任意的数, 关于 x 而微分时必须把它看成常数因子. 在这种情形, 将有

$$\begin{aligned} d^2y &= d(dy) = d(y' \cdot dx) = dy' \cdot dx = (y'' \cdot dx) \cdot dx = y'' \cdot dx^2, \\ d^3y &= d(d^2y) = d(y'' \cdot dx^2) = dy'' \cdot dx^2 = (y''' \cdot dx) \cdot dx^2 = y''' \cdot dx^3. \end{aligned}$$

很容易可以猜出普遍规律是

$$d^n y = y^{(n)} \cdot dx^n.$$

由它可以进一步推得

$$y^{(n)} = \frac{d^n y}{dx^n}.$$

但是高阶微分没有形式不变性, 即若 $x = \varphi(t)$, 于是 y 可以看成 t 的复合函数 $y = f(\varphi(t))$. 它关于 t 的一阶微分可以写成

$$dy = y'_x \cdot dx,$$

其中 $dx = x'_t \cdot dt$. 再求它关于 t 的二阶微分

$$d^2y = d(y'_x \cdot dx) = dy'_x \cdot dx + y'_x \cdot d(dx) = y'' \cdot dx + y'_x \cdot d^2x.$$

这才是二阶微分的一般形式. 之前的高阶微分形式 x 是自变量, 所以 $d^2x = 0$.