习题讨论课05答案：导数与高阶导数

★号（越）多表示题目（越）难

一、导数与微分

【定义】

- 微分 $\mathrm{d}f(x_0)$ 是线性函数（正比例函数），满足

$$f(x_0+h)-f(x_0)=\mathrm{d}f(x_0)(h)+o(h),\quad h\to 0.$$

- 导数 $f'(x_0)$ 是微分 $\mathrm{d}f(x_0)$ 的比例系数：

$$\mathrm{d}f(x_0)(h)=f'(x_0)h,\quad h\in\mathbb{R},$$

  传统上，人们习惯用变量表示函数，$y=f(x)$，所以 $\mathrm{d}y|_{x_0}$ 就是 $\mathrm{d}f(x_0)$，因此上式也可以写成

$$\mathrm{d}y|_{x_0}=f'(x_0)\mathrm{d}x|_{x_0}.$$

  这里 x 被看成是函数 $\mathbb{R}\ni x\mapsto x$，而 $\mathrm{d}x|_{x_0}$ 是线性函数 $\mathbb{R}\ni h\mapsto h\in\mathbb{R}$. 对可微函数 f 定义域中的任何 x，

$$\mathrm{d}y=f'(x)\mathrm{d}x.$$

- 导数 $f'(x_0)$ 是差商极限

$$f'(x_0)=\lim_{x\to x_0}\frac{f(x)-f(x_0)}{x-x_0}=\lim_{h\to 0}\frac{f(x_0+h)-f(x_0)}{h}.$$

  物理意义：平均变化率的极限=瞬时变化率，瞬时速度；几何意义：割线斜率的极限=切线斜率。

  在 Leibniz 发明的符号体系中，导数被写成 $\left.\frac{\mathrm{d}y}{\mathrm{d}x}\right|_{x_0}$，这是受微分 $\mathrm{d}y=f'(x_0)\mathrm{d}x$ 的启发。

- 函数图像 $y=f(x)$ 在点 $(x_0,f(x_0))$ 处的**切线**是

$$y-f(x_0)=f'(x_0)(x-x_0),\quad x\in\mathbb{R}.$$

  平面曲线 $(x(t),y(t))$ 在点 $(x(t_0),y(t_0))$ 处的**切线**是

$$\begin{cases}x-x(t_0)=x'(t_0)(t-t_0),\\ y-y(t_0)=y'(t_0)(t-t_0),\end{cases},\qquad t\in\mathbb{R}.$$

  （这里要求**切向量** $(x'(t_0),y'(t_0))$ 不为零向量，即两个导数不全为零），也可以写成

$$\frac{x-x(t_0)}{x'(t_0)}=\frac{y-y(t_0)}{y'(t_0)},$$

  这是一个比例式，比例式中规定：当分母为零时，分子也为零。

  过曲线上的点 P 且与 P 处的切线垂直的直线叫曲线在点 P 处的**法线**。

【求导法则——转化，把可微性和导数归结为已知的可微性和导数】

根据函数的运算结构拆解。以下结论中等式左端的可微性是由右端被求导的函数的可微性保证的，另外要注意验证求导的其他先决条件（商的导数、反函数求导）

1. 线性：$(\lambda u+\mu v)'=\lambda u'+\mu v'$.

2. Leibniz公式：$(uv)'=u'v+uv'$

   微分
   $$\mathrm{d}(uv)=v\mathrm{d}u+u\mathrm{d}u.$$
   常见错误：甚至连 Leibniz 本人最初都误以为
   $$\mathrm{d}(uv)=\mathrm{d}u\mathrm{d}v.$$
   这个等式不成立：左边是线性的；右边是两个线性函数乘积，不是线性的。

3. 除法：$v\neq 0$ 时，$(\frac{u}{v})'=\frac{u'v-uv'}{v^2}$

4. 链索法则：$(u\circ v)'=(u'\circ v)\cdot v'$，右边的 $\circ v$ 是为了换元：把中间求导变量换为最终求导自变量，
   $$(u(v(x)))'=u'(v(x))v'(x).$$

5. 反函数求导：当 $u'\neq 0$ 且反函数 u^{-1} 连续时，$\left(u^{-1}\right)'=\frac{1}{u'\circ u^{-1}}$，右边的 $\circ u^{-1}$ 是为了换元，
   $$(u^{-1})'(y)=\frac{1}{u'(x)}=\frac{1}{u'(x(y))},\quad y=u(x),x=u^{-1}(y).$$

6. 对数求导法，把乘除、乘方开方、指数运算转化为加减、乘除四则运算。
   $$\frac{uv}{w}=\mathrm{e}^{\ln u+\ln v-\ln w},\quad u^v=\mathrm{e}^{v\ln u}.$$
   （本质：把复杂函数尽可能改写为基本初等函数的复合与四则运算形式。）
   为什么可以无视函数值的正负？

【Leibniz导数符号的优越性及不足之处】

在链索法则和反函数求导中使用Leiniz的符号，$z=u(y),y=v(x)$ 可微，则
$$\frac{\mathrm{d}z}{\mathrm{d}x}=\frac{\mathrm{d}z}{\mathrm{d}y}\frac{\mathrm{d}y}{\mathrm{d}x}.$$
$$\frac{\mathrm{d}x}{\mathrm{d}y}=\frac{1}{\frac{\mathrm{d}y}{\mathrm{d}x}},\qquad x=v^{-1}(y).$$
这样的符号即使对小学生来说也是极其友好的。但要注意其中求导的位置。另外，这样美好的事情到二阶及以上的导数就不对了。因此不要被这些看似美好的形式所迷惑，只有这样你才有可能成为一个 sophomore（更有智慧的人）.

例 1. 设 f 在闭区间 $[\alpha,\beta]$ 上连续，在开区间 $I=(\alpha,\beta)$ 内的可微。

1. 证明：对任意 $x_0\in I$，若数列 $a_n,b_n\in I$ 满足 $a_n\le x_0\le b_n$，$a_n<b_n$，$\lim\limits_{n\to+\infty}a_n=x_0=\lim\limits_{n\to+\infty}b_n$，则
$$\lim_{n\to+\infty}\frac{f(b_n)-f(a_n)}{b_n-a_n}=f'(x_0);$$
2. 对任意 $a_1,b_1\in I,a_1<b_1$，构造区间套 $[a_n,b_n]$ 使得
$$\frac{f(b_{n+1})-f(a_{n+1})}{b_{n+1}-a_{n+1}}\le\frac{f(b_n)-f(a_n)}{b_n-a_n},\quad\forall n\ge 1,$$
并且
$$\lim_{n\to+\infty}a_n=\lim_{n\to+\infty}b_n.$$
3. 利用1和2的结论证明，对任意 $a,b\in I$，$a<b$，都存在 $\xi\in[a,b]$ 使得
$$f'(\xi)\le\frac{f(b)-f(a)}{b-a}.$$
4. 利用3证明：若对任意 $x\in I$，$f'(x)\ge 0$，则 f 在 $[\alpha,\beta]$ 上单调不减；若对任意 $x\in I$，$f'(x)>0$，则 f 在 $[\alpha,\beta]$ 上严格增。
5. 利用4证明：若对任意 $x\in I$，$f'(x)=0$，则 f 在 $[\alpha,\beta]$ 上为常数。

证明. (1) 对任意$\varepsilon>0$，存在 $\delta>0$ 使得对任意 $|x-x_0|<\delta$，都有 $x\in I$，且
$$-\varepsilon|x-x_0|\le f(x)-f(x_0)-f'(x_0)(x-x_0)\ge\varepsilon|x-x_0|.$$
存在 N 使得对任意 $n\ge N$，都有
$$x_0-\delta<a_n\le x_0\le b_n<x_0+\delta,$$
从而
$$-\varepsilon(x_0-a_n)\le f(a_n)-f(x_0)-f'(x_0)(a_n-x_0)\le\varepsilon(x_0-a_n),$$
$$-\varepsilon(b_n-x_0)\le f(b_n)-f(x_0)-f'(x_0)(b_n-x_0)\le\varepsilon(b_n-x_0),$$
以上两行不等式相减，得到
$$-\varepsilon(b_n-a_n)\le f(b_n)-f(a_n)-f'(x_0)(b_n-a_n)\le\varepsilon(b_n-a_n).$$
即
$$-\varepsilon\le\frac{f(b_n)-f(a_n)}{b_n-a_n}-f'(x_0)\le\varepsilon.$$
因此
$$\lim_{n\to+\infty}\frac{f(b_n)-f(a_n)}{b_n-a_n}=f'(x_0).$$

(2) 当区间 $[a_n, b_n]$ 已造好，我们来造下一个区间。

若
$$\frac{f(\frac{a_n+b_n}{2})-f(a_n)}{\frac{a_n+b_n}{2}-a_n} \le \frac{f(b_n)-f(a_n)}{b_n-a_n},$$
则取 $a_{n+1}=a_n, b_{n+1}=\frac{a_n+b_n}{2}$.

若
$$\frac{f(\frac{a_n+b_n}{2})-f(a_n)}{\frac{a_n+b_n}{2}-a_n} > \frac{f(b_n)-f(a_n)}{b_n-a_n},$$
则必然有
$$\frac{f(b_n)-f(\frac{a_n+b_n}{2})}{b_n-\frac{a_n+b_n}{2}} \le \frac{f(b_n)-f(a_n)}{b_n-a_n},$$
此时取 $a_{n+1}=\frac{a_n+b_n}{2}, b_{n+1}=b_n$.

这样构造的区间套即满足要求。

(3) 取$a_1=a, b_1=b$，并按(2)构造区间套。由有界闭区间套定理得到 $\xi \in \bigcap\limits_{n\ge 1}[a_n, b_n]$，于是 $\xi = \lim\limits_{n\to+\infty} a_n = \lim\limits_{n\to+\infty} b_n$. 于是由(1)知
$$f'(\xi) = \lim_{n\to+\infty}\frac{f(b_n)-f(a_n)}{b_n-a_n} \le \frac{f(b)-f(a)}{b-a}.$$

(4) 由(3)知，任取 $a, b \in I$，$a<b$，都存在 $\xi\in[a,b]$ 使得
$$\frac{f(b)-f(a)}{b-a} \ge f'(\xi) \ge 0.$$
所以 $f(b)\ge f(a)$. 即 f 在开区间 $I=(\alpha,\beta)$ 内单调不减。

任取 $x\in(\alpha,\beta)$ 以及任意 $u\in(\alpha,x), v\in(x,\beta)$。则
$$f(u)\le f(x)\le f(v).$$
让 $u\to\alpha^+, v\to\beta^-$，由于 f 在 α 和 β 处连续，所以
$$f(\alpha)=\lim_{u\to\alpha^+}f(u)\le f(x)\le \lim_{v\to\beta^-}f(v)=f(\beta).$$
因此 f 在闭区间 $[\alpha,\beta]$ 上单调不减。

如果 $f'>0$，则 $f(b)>f(a)$. 即 f 在 (α,β) 内严格增。

任取 $x,y\in(\alpha,\beta)$ 以及 任意 $u\in(\alpha,x), v\in(y,\beta)$，则
$$f(u)<f(x)<f(y)<f(v).$$
于是
$$f(\alpha)=\lim_{u\to\alpha^+}f(u)\le f(x)<f(y)\le \lim_{v\to\beta^-}f(v)=f(\beta).$$
所以 f 在闭区间 $[\alpha,\beta]$ 上严格增。

(5) 因为 $f'=0$，所以 $f'\le 0$，由(4)知 $-f$ 在 $[\alpha,\beta]$ 上单调不减，从而 f 在 $[\alpha,\beta]$ 上单调不增；另一方面，$f'\ge 0$，由(4)知 f 在 $[\alpha,\beta]$ 上单调不减。所以 f 在 $[\alpha,\beta]$ 上为常数。 □

例 2. 设 $A > 1$.

$$f(x) = \begin{cases} x + Ax^2 \sin \frac{1}{x}, & x \neq 0; \\ 0, & x = 0. \end{cases}$$

1. 证明：f 可微，但 f' 在 $x = 0$ 处不连续，并讨论 f' 在 $x = 0$ 处间断类型。

2. 证明：$f'(0) > 0$，但在 $x = 0$ 的任意邻域内，f 都不是单调函数。

3. 例1中如果不假定 $a_n \leq x_0 \leq b_n$，则例1结论不成立。

证明. (1) 对 $x \neq 0$，$x + Ax^2 \sin \frac{1}{x}$ 是初等函数，有任意阶导数，且

$$f'(x) = 1 + 2Ax \sin \frac{1}{x} + Ax^2 \cos \frac{1}{x} \left(-\frac{1}{x^2} \right) = 1 - A \cos \frac{1}{x} + 2Ax \sin \frac{1}{x}.$$

它在 $x \to 0^{\pm}$ 时无极限，$x = 0$ 是 f' 的第二类间断点。

$$f'(0) = \lim_{x \to 0} \frac{x + Ax^2 \sin \frac{1}{x}}{x} = 1.$$

因此

$$f'(x) = \begin{cases} 1 - A \cos \frac{1}{x} + 2Ax \sin \frac{1}{x}, & x \neq 0; \\ 1, & x = 0. \end{cases}$$

(2) $f'(0) = 1 > 0$. 取 $x_n = \frac{1}{2n\pi}, y_n = \frac{1}{2n\pi + \pi}$，则

$$f'(x_n) = 1 - A < 0, \quad f'(y_n) = 1 + A > 0,$$

于是在包含 x_n 的任何开区间 I_n 内，f 都不是单调不减的；在包含 y_n 的任何开区间 J_n 内，f 都不是单调不增的。

对任意 $\delta > 0$，取 n 足够大使 $x_n, y_n \in (0, \delta)$，于是 f 在区间 $[0, \delta)$ 中 f 都不是单调的。同理可证 f 在区间 $(-\delta, 0]$ 中也不是单调的。

(3) 由 $f'(x_n) < 0$ 知，存在 $0 < a_n < x_n < b_n < 2x_n$ 使得 $\frac{f(b_n) - f(a_n)}{b_n - a_n} < 0$.

易见 $\lim\limits_{n \to +\infty} a_n = \lim\limits_{n \to +\infty} b_n = 0$，$\lim\limits_{n \to +\infty} \frac{f(b_n) - f(a_n)}{b_n - a_n} = f'(0)$ 不成立（**注意：这个表达与** $\lim\limits_{n \to +\infty} \frac{f(b_n) - f(a_n)}{b_n - a_n} \neq f'(0)$ **是否一样？**）。 □

例 3. 设 $\alpha > 0$, 记

$$f_\alpha(x) = x^{(x^\alpha)}.$$

1. 求 $f'_\alpha(x)$.

2. 证明 $x = 0$ 为 f_α 的可去间断点。

3. 记

$$g(x) = \begin{cases} f_\alpha(x), & x > 0; \\ \lim\limits_{t \to 0^+} f_\alpha(t), & x \leq 0. \end{cases}$$

讨论 g 是否在 $x = 0$ 处可微。在它可微时，讨论 g' 的连续性.

解. (1) $f_\alpha(x)=\mathrm{e}^{x^\alpha\ln x}$ 是初等函数，所以

$$f'_\alpha(x)=\mathrm{e}^{x^\alpha\ln x}\left(x^\alpha\ln x\right)'=\mathrm{e}^{x^\alpha\ln x}\left(\alpha x^{\alpha-1}\ln x+x^{\alpha-1}\right)=x^{x^\alpha+\alpha-1}(\alpha\ln x+1).$$

(2) 因为

$$\lim_{x\to0^+}x^\alpha\ln x=\lim_{t\to+\infty}\frac{-t}{\mathrm{e}^{\alpha t}}=0,\quad x=\mathrm{e}^{-t},$$

所以

$$\lim_{x\to0^+}f_\alpha(x)=\lim_{x\to0^+}\mathrm{e}^{x^\alpha\ln x}=\mathrm{e}^0=1.$$

因此 $x=0$ 为 f_α 的可去间断点。

(3) 由(2)知

$$g(x)=\begin{cases}x^{(x^\alpha)}, & x>0;\\ 1, & x\le0.\end{cases}$$

$$\lim_{x\to0^+}\frac{x^{(x^\alpha)}-1}{x}=\lim_{x\to0^+}\frac{\mathrm{e}^{x^\alpha\ln x}-1}{x^\alpha\ln x}\frac{x^\alpha\ln x}{x}=\begin{cases}0, & \alpha>1;\\ -\infty, & \alpha\le1.\end{cases}$$

所以 $g'_+(0)$ 存在当且仅当 $\alpha>1$. 因此，当 $\alpha\le1$ 时，g 在 $x=0$ 处不可微。

当 $\alpha>1$ 时，$g'_+(0)=g'_-(0)=0$，g 可微，

$$g'(x)=\begin{cases}x^{x^\alpha+\alpha-1}(\alpha\ln x+1), & x>0;\\ 0, & x\le0.\end{cases}$$

由(2)中的中间结论知，$\lim\limits_{x\to0^+}x^{\alpha-1}(\alpha\ln x+1)=0$。所以 $\lim\limits_{x\to0^+}g'(x)=0$，因此 g' 是连续函数。

□

例 4. $f(x)$ 在 $x=a$ 可导, $f(a)\neq0$, 求 $\lim\limits_{x\to\infty}\left(\frac{f\left(a+\frac1x\right)}{f(a)}\right)^x$

解法1. 换底，换元，$y=\frac1x\to0$，

$$\left(\frac{f\left(a+\frac1x\right)}{f(a)}\right)^x=\mathrm{e}^{\frac{\ln f(a+y)-\ln f(a)}{y}}\to\mathrm{e}^{(\ln f(y))'|_{y=a}}=\mathrm{e}^{\frac{f'(a)}{f(a)}}.$$

□

解法2.

$$\begin{aligned}\lim_{x\to\infty}\left(\frac{f\left(a+\frac1x\right)}{f(a)}\right)^x&=\lim_{x\to\infty}\left(1+\frac{f(a+\frac1x)-f(a)}{f(a)}\right)^{\frac{f(a)}{f(a+\frac1x)-f(a)}\frac{f(a+\frac1x)-f(a)}{f(a)\frac1x}}\\&=\left(\lim_{x\to\infty}\left(1+\frac{f(a+\frac1x)-f(a)}{f(a)}\right)^{\frac{f(a)}{f(a+\frac1x)-f(a)}}\right)^{\lim\limits_{x\to\infty}\frac{f(a+\frac1x)-f(a)}{f(a)\frac1x}}\\&=\mathrm{e}^{\frac{f'(a)}{f(a)}}\end{aligned}$$

□

解法3.

$$\begin{aligned}\left(\frac{f\left(a+\frac{1}{x}\right)}{f(a)}\right)^x &= \mathrm{e}^{x\ln\left(1+\frac{f(a+\frac{1}{x})-f(a)}{f(a)}\right)}\\ &= \mathrm{e}^{x\ln\left(1+\frac{f'(a)\frac{1}{x}+o\left(\frac{1}{x}\right)}{f(a)}\right)}\\ &= \mathrm{e}^{x\left(\frac{f'(a)}{f(a)}\frac{1}{x}+o\left(\frac{1}{x}\right)\right)} = \mathrm{e}^{\frac{f'(a)}{f(a)}+o(1)} \to \mathrm{e}^{\frac{f'(a)}{f(a)}}, \quad x\to\infty.\end{aligned}$$

□

问题：上面哪个解法有错误？请做出相应的修改。然后思考一下：你认为哪个解法更好一些。

例 5. 设 $a>b>0$. 考虑具有相同形状的一族椭圆

$$E_t:\frac{x^2}{a^2t}+\frac{y^2}{b^2t}=1,\quad t>0.$$

设 $P(x_1,y_1)$ 是椭圆 E_t 上一点，它既不在椭圆 E_t 的长轴上，也不在椭圆 E_t 的短轴上。证明：存在唯一的 $s>0$ 使得椭圆E_t 在点 (x_1,y_1) 处的法线与椭圆 E_s 相切。

证明. 在方程

$$\frac{x^2}{a^2t}+\frac{y^2}{b^2t}=1$$

中，视 y 为关于 x 的函数，两边对 x 在 $x=x_1$ 处求导数，得到

$$\frac{2x_1}{a^2t}+\frac{2y_1y'(x_1)}{b^2t}=0.$$

解得

$$y'(x_1)=-\frac{b^2x_1}{a^2y_1},$$

从而椭圆 E_t 在点 P 处的切线为

$$y-y_1=-\frac{b^2x_1}{a^2y_1}(x-x_1),$$

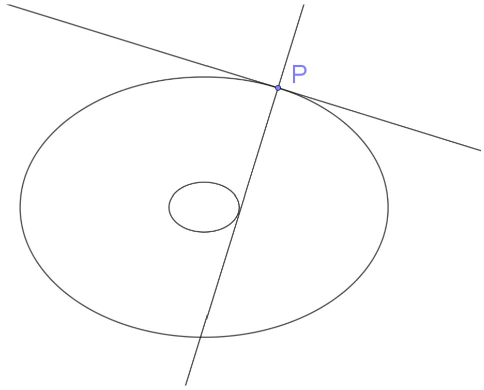
法线为

$$y-y_1=\frac{a^2y_1}{b^2x_1}(x-x_1).$$

该法线是椭圆 E_s 在 (x_2,y_2) 处的切线，当且仅当

$$y_2-y_1=\frac{a^2y_1}{b^2x_1}(x_2-x_1)=-\frac{b^2x_2}{a^2y_2}(x_2-x_1).$$

因此 $a^4y_1y_2+b^4x_1x_2=0$.



于是得到方程组

$$\begin{cases} a^2y_1(x_2-x_1)-b^2x_1(y_2-y_1)=0, \\ b^4x_1x_2+a^4y_1y_2=0, \\ s=\dfrac{x_2^2}{a^2}+\dfrac{y_2^2}{b^2}. \end{cases}$$

前两个方程是关于 x_2, y_2 的线性方程组，其系数矩阵

$$\begin{pmatrix} a^2y^1 & -b^2x_1 \\ b^4x_1 & a^4y_1 \end{pmatrix}$$

可逆，所以 x_2, y_2 有唯一解，由第三个方程知满足条件的 s 存在且唯一。 □

注：这是二元函数求极值时的最速下降法的几何背景。

例 6. 1. 证明旋轮线 $\begin{cases} x=A(t-\sin t), \\ y=A(1-\cos t) \end{cases}$ 在 $0<t<2\pi$ 的范围内确定了一个可微的函数 $y=y(x)$，并讨论 $y(x)$ 的单调性。

2. 证明旋轮线满足微分方程 $\left(1+{y_x'}^2\right)y=2A.$

证明. (1)

$$x'(t)=A(1-\cos t)>0, \quad \forall t\in(0,2\pi).$$

所以由例1知，$x(t)$ 有可微的反函数，于是 $y=y(t)=y(t(x))$ 可微。从而由链索法则，

$$1=t_t'=(t(x(t)))_t'=t'(x)x'(t),$$

$$y_x'=y_t'\cdot t_x'=\frac{y'(t)}{x'(t)}=\frac{\sin t}{(1-\cos t)},$$

当 $t\in[0,\pi]$，即 $0\le x\le A\pi$ 时，函数 $y(x)$ 严格增；当 $t\in[\pi,2\pi]$，即 $A\pi\le x\le 2A\pi$ 时，函数 $y(x)$ 严格减。

(2) 直接代入验证。 □

注：这是数学和力学史上的一条著名曲线，它有很多名字：旋轮线，最速降线、摆线。

二、高阶导数

【高阶导数的计算法则】

1. 线性：$(\lambda u+\mu v)^{(n)}=\lambda u^{(n)}+\mu v^{(n)}$.

2. Leibniz 公式：$(uv)^{(n)}=\sum\limits_{k=0}^{n}u^{(k)}v^{(n-k)}$，$n$ 次求导由u,v 分别承担其中一部分，C_n^k 是不同的分担方法种类数（n 次中哪些次由 u 承担，余下的由 v 承担）。

3. $f(ax+b)$ 的高阶导数：$\big(f(ax+b)\big)^{(n)}=f^{(n)}(ax+b)a^n$.

   注意：左右两边导数的含义是不同的，左边是对复合函数求导，右边是对 f 求导。

【高阶可微性】

1. 除法：当 $v(x)\neq 0$ 时，u/v 的高阶可微性与 $u.v$ 共同的高阶可微性一样。

   求商 $w=u/v$ 的高阶导数，可以转化为乘法计算：$v=uw$.

2. 复合函数：$v\circ u$ 的高阶可微性与 u,v 共同的高阶可微性一样。

3. 反函数：当 $u'(x)\neq 0$ 时，u^{-1} 的高阶可微性与 u 的高阶可微性一样。

   求反函数 $x=u^{-1}(y)$ 的高阶导数，可以把 $y=u(x)=u(u^{-1}(y))$ 对 y 求导（用到链索法则）。

4. 参数方程 $x=x(t),y=y(t)$，如果 $x'(t)\neq 0$，则 $x(t)$ 有反函数 $t=t(x)$，$y=y(t(x))$ 的高阶可微性与 $x(t),y(t)$ 的共同高阶可微性一样。

它们都没有简单的高阶导数公式。如何证明这些高阶可微性？

【高阶可微性的意义和用途】

1. 加速度是运动（位移对时间的函数）的二阶导数，力决定了加速度。

2. 曲线的二阶导数与曲线的弯曲性质（函数的凹凸性、曲线弯曲的方向、曲线弯曲的程度——曲率）有关。

3. 泰勒公式用高阶导数表示函数的近似形式。

例 7. 设 $y=u(x),z=v(y)$ 都是二阶可导，求复合函数得 $z=v(u(x))$ 的二阶导数。

解法1.

$$(v(u(x)))' = v'(u(x))u'(x),$$

再求导，

$$\begin{aligned}(v(u(x)))'' &= \big(v'(u(x))u'(x)\big)' = (v'(u(x)))'u'(x) + v'(u(x))u''(x) \\ &= v''(u(x))\big(u'(x)\big)^2 + +v'(u(x))u''(x).\end{aligned}$$

□

解法2. 使用 Leibniz 的符号。由链索法则

$$\frac{\mathrm{d}z}{\mathrm{d}x} = \frac{\mathrm{d}z}{\mathrm{d}y}\frac{\mathrm{d}y}{\mathrm{d}x}.$$

这视为对函数 z 的操作，左边是把 z 当作以 x 为自变量时的导数，右边是把它当作以中间变量 y 为自变量的函数。于是

$$\frac{\mathrm{d}}{\mathrm{d}x} = \frac{\mathrm{d}y}{\mathrm{d}x}\cdot\frac{\mathrm{d}}{\mathrm{d}y}.$$

从而

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

□

注：(1) 一阶导数有链索法则

$$\frac{\mathrm{d}z}{\mathrm{d}x} = \frac{\mathrm{d}z}{\mathrm{d}y}\frac{\mathrm{d}y}{\mathrm{d}x}, \tag{*}$$

它说明了变量 z 在坐标变换 $y = u(x)$ 下，导数是如何变化的。

虽然导数会变化，但是变量 z 的一阶微分 $\mathrm{d}z$ 是不变的，即它与坐标变换 $y = u(x)$ 无关：

$$\mathrm{d}_x z = \frac{\mathrm{d}z}{\mathrm{d}x}\mathrm{d}x = \frac{\mathrm{d}z}{\mathrm{d}y}\frac{\mathrm{d}y}{\mathrm{d}x}\mathrm{d}x = \frac{\mathrm{d}z}{\mathrm{d}y}\mathrm{d}y = \mathrm{d}_y z.$$

这也说明了微分是比导数更本质的概念。

(2) 小学生大概能用分数化简的办法看懂公式(*)，然而这方法在二阶导数时就失效了。本例中的计算表明，二阶微分

$$\mathrm{d}_x^2 z = \frac{\mathrm{d}^2z}{\mathrm{d}x^2}(\mathrm{d}x)^2$$

与坐标系有关：

$$\begin{aligned}
\mathrm{d}_x^2 z &= \frac{\mathrm{d}^2 z}{\mathrm{d}x^2}(\mathrm{d}x)^2 = \frac{\mathrm{d}^2 z}{\mathrm{d}y^2}\cdot\left(\frac{\mathrm{d}y}{\mathrm{d}x}\right)^2(\mathrm{d}x)^2 + \frac{\mathrm{d}z}{\mathrm{d}y}\cdot\frac{\mathrm{d}^2 y}{\mathrm{d}x^2}(\mathrm{d}x)^2\\
&= \frac{\mathrm{d}^2 z}{\mathrm{d}y^2}\cdot\left(\frac{\mathrm{d}y}{\mathrm{d}x}\mathrm{d}x\right)^2 + \frac{\mathrm{d}z}{\mathrm{d}y}\cdot\frac{\mathrm{d}^2 y}{\mathrm{d}x^2}(\mathrm{d}x)^2\\
&= \frac{\mathrm{d}^2 z}{\mathrm{d}y^2}\cdot(\mathrm{d}y)^2 + \frac{\mathrm{d}z}{\mathrm{d}y}\cdot\mathrm{d}_x^2 y\\
&= \mathrm{d}_y^2 z + \frac{\mathrm{d}z}{\mathrm{d}y}\cdot\mathrm{d}_x^2 y,
\end{aligned}$$

如果坐标变换 $y=u(x)$ 是一次函数，则 $\mathrm{d}_x^2 z=\mathrm{d}_y^2 z$，即两种坐标系下 z 有相同的二阶微分。但是坐标变换 $y=u(x)$ 通常是非线性的，此时附加项 $\frac{\mathrm{d}z}{\mathrm{d}y}\cdot\mathrm{d}_x^2 y$ 是坐标变换导致的。

(3) 如果视 $z=v(y)$ 为空间坐标系变换，$x=t$ 为时间，则

$$\ddot{z} = \frac{\mathrm{d}^2 z}{\mathrm{d}y^2}\cdot\dot{y}^2 + \frac{\mathrm{d}z}{\mathrm{d}y}\cdot\ddot{y}$$

说明了 y,z 两种坐标系下运动的加速度 $\ddot{y}$ 和 $\ddot{z}$ 之间的关系。

例 8. 设 $x(t),y(t)$ 二阶可微，$x'(t)\neq 0,\forall t$. 试用 $x(t),y(t)$ 的导数与二阶导数表示函数 $y=y(t(x))$ 的二阶导数。

解. 由链索法则，

$$1=\frac{\mathrm{d}x}{\mathrm{d}x}=\frac{\mathrm{d}x}{\mathrm{d}t}\frac{\mathrm{d}t}{\mathrm{d}x},\qquad \frac{\mathrm{d}y}{\mathrm{d}x}=\frac{\mathrm{d}y}{\mathrm{d}t}\frac{\mathrm{d}t}{\mathrm{d}x},$$

从而

$$\frac{\mathrm{d}t}{\mathrm{d}x}=\frac{1}{\frac{\mathrm{d}x}{\mathrm{d}t}},\qquad \frac{\mathrm{d}}{\mathrm{d}x}=\frac{1}{\frac{\mathrm{d}x}{\mathrm{d}t}}\frac{\mathrm{d}}{\mathrm{d}t}.$$

因此

$$\begin{aligned}
\frac{\mathrm{d}^2 y}{\mathrm{d}x^2} &= \frac{\mathrm{d}}{\mathrm{d}x}\frac{\mathrm{d}}{\mathrm{d}x}y = \frac{1}{\frac{\mathrm{d}x}{\mathrm{d}t}}\frac{\mathrm{d}}{\mathrm{d}t}\left(\frac{1}{\frac{\mathrm{d}x}{\mathrm{d}t}}\frac{\mathrm{d}}{\mathrm{d}t}y\right)\\
&= \frac{1}{\frac{\mathrm{d}x}{\mathrm{d}t}}\frac{-\frac{\mathrm{d}^2 x}{\mathrm{d}t^2}}{\left(\frac{\mathrm{d}x}{\mathrm{d}t}\right)^2}\frac{\mathrm{d}y}{\mathrm{d}t} + \left(\frac{1}{\frac{\mathrm{d}x}{\mathrm{d}t}}\right)^2\frac{\mathrm{d}^2 y}{\mathrm{d}t^2}\\
&= \frac{x'(t)y''(t)-x''(t)y'(t)}{x'(t)^3}.
\end{aligned}$$

□

常见错误：一阶导数的链索法则(*)以及

$$\frac{\mathrm{d}y}{\mathrm{d}x}=\frac{\frac{\mathrm{d}y}{\mathrm{d}t}}{\frac{\mathrm{d}x}{\mathrm{d}t}}$$

会给一些学生加深了“导数就是做比值”的印象，毕竟老师说“导数就是函数值变化量与自变量变化量的比值的极限”吗。于是他们认为

$$\frac{\mathrm{d}^2y}{\mathrm{d}x^2}=\frac{\frac{\mathrm{d}^2y}{\mathrm{d}t^2}}{\frac{\mathrm{d}^2x}{\mathrm{d}t^2}}.$$

但是本例和上例中的结论说明，这种形式化的理解是错误的。这也是 Leibniz 符号体系的陷阱。

注：本例中取 $y=t$，则可得反函数 $t=t(x)=y(x)$ 的二阶导数

$$t''(x)=\frac{\mathrm{d}^2y}{\mathrm{d}x^2}=\frac{-x''(t)}{x'(t)^3}.$$

例 9. 设 f 在区间 (a,b) 内满足 $f''>0$。证明

1. 对任意 $x_0\in(a,b)$，除点 $(x_0,f(x_0))$ 外，曲线 $y=f(x)$ 严格位于它在 $(x_0,f(x_0))$ 处的切线的上方。

2. 对任意 $x_1,x_2\in(a,b),x_1\neq x_2$ 以及任意 $t\in(0,1)$，都有

$$f((1-t)x_1+tx_2)<(1-t)f(x_1)+tf(x_2). \qquad (**)$$

满足 (**) 的函数 f 称为**严格凸函数**，并称曲线 $y=f(x)$ 为**严格下凸的**。满足与(**)相反不等号的函数称为**严格凹函数**，称其图像为**严格上凸的**。

证明. (1) 令 $g(x)=f(x)-f(x_0)-f'(x_0)(x-x_0)$. 则 $g'(x)=f'(x)-f'(x_0)$，$g''(x)=f''(x_0)>0$.

因此由例1结论，g' 在区间 (a,b) 内严格增。从而对任意 $u\in(a,x_0)$ 和任意 $v\in(x_0,b)$，$g'(u)<g'(x_0)=0<g'(v)$. 因此 g 在区间 $(a,x_0]$ 上严格减，在区间 $[x_0,b)$ 上严格增。从而

$$g(x)>g(x_0)=0,\quad \forall x\in(a,b)\backslash\{x_0\}.$$

即

$$f(x)>f(x_0)+f'(x_0)(x-x_0),\quad \forall x\in(a,b)\backslash\{x_0\}.$$

除点 $(x_0,f(x_0))$ 外，曲线 $y=f(x)$ 严格位于它在 $(x_0,f(x_0))$ 处的切线的上方。

(2) 令 $h(t)=(1-t)f(x_1)+tf(x_2)-f((1-t)x_1+tx_2)$. 则

$$h''(t)=-f''((1-t)x_1+tx_2)(x_2-x_1)^2<0.$$

所以 h' 严格减。

$h(0)=h(1)=0$.

假设存在 $t_1\in(0,1)$ 使得 $h(t_1)\le 0$.

则由例1结论知存在 $t_2\in[0,t_1]$ 以及 $t_3\in[t_1,1]$ 使得 $h'(t_2)\le 0\le h'(t_3)$。

由于 h' 严格减，所以 $t_2 = t_3 = t_1$. 从而 $h'(t_1) = 0$ 并且对任意 $u \in (0, t_1)$ 以及任意 $v \in (t_1, 1)$ 都有 $h'(u) > h'(t_1) = 0 > h'(v)$.

所以 h 在区间 $[0, t_1]$ 上严格增，在区间 $[t_1, 1]$ 上严格减，因此

$$h(t) > \min\{h(0), h(1)\} = 0, \quad \forall t \in (0, 1).$$

这与假设矛盾。因此对于任意 $t \in (0, 1)$，总有 $h(t) > 0$. 从而

$$f((1-t)x_1 + tx_2) < (1-t)f(x_1) + tf(x_2), \quad \forall t \in (0, 1).$$

□

注：如果使用 Rolle 定理，(2) 中的论证可以被简化。

例 10. 1. 证明旋轮线 $\begin{cases} x = A(t - \sin t), \\ y = A(1 - \cos t) \end{cases}$ 在 $0 < t < 2\pi$ 范围内确定了一个 $\mathscr{C}^\infty$ 函数 $y = y(x)$.

2. 求 $y''(x)$.

3. 证明旋轮线位于它的每条切线的下方（切点除外），为严格上凸的。

证法1. (1) 由例6知，$x = x(t)$ 有 $\mathscr{C}^\infty$ 反函数，从而 $y = y(t(x))$ 是 $\mathscr{C}^\infty$ 函数。

(2)

$$y'_x = \frac{y'_t}{x'_t} = \frac{\sin t}{1 - \cos t}.$$

从而

$$y''_x = \frac{\left(\frac{\sin t}{1-\cos t}\right)'_t}{x'_t} = \frac{\cos t(1 - \cos t) - (\sin t)(\sin t)}{A(1 - \cos t)^3} = -\frac{1}{A(1 - \cos t)^2}.$$

(3) 由例9结论即得。 □

证法2. (1)和(3)同证法1。

(2) 事实上，由例6中的

$$y(1 + {y'_x}^2) = 2A,$$

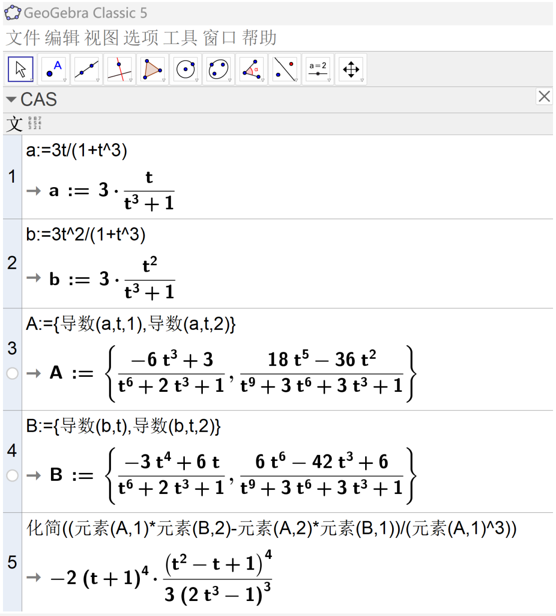
两边求导可得

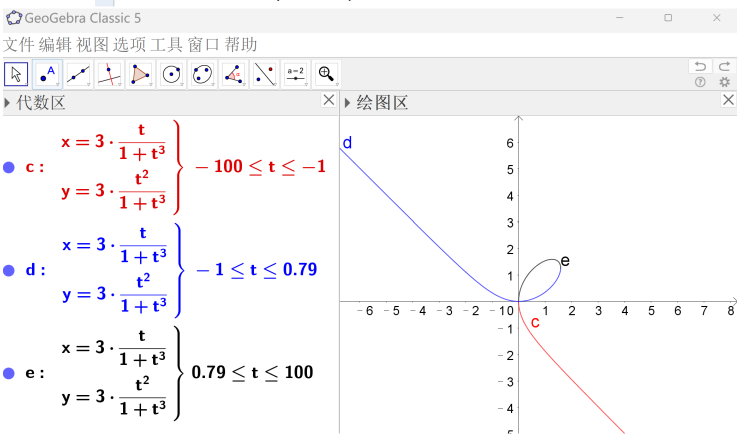
$$y'_x(1 + {y'_x}^2) + 2yy'_xy''_x = 0,$$

从而

$$y''_x = -\frac{1 + {y'_x}^2}{2y} = -\frac{2A}{2y^2} = -\frac{1}{A(1 - \cos t)^2}.$$

□





例 11 (笛卡尔叶形线). 讨论曲线 $x^3+y^3=3xy$ 的凹凸性。

解. 令 $t=\frac{y}{x}$，则
$$x^3(1+t^3)=3x^2t,$$
解得
$$x=\frac{3t}{1+t^3},\quad y=\frac{3t^2}{1+t^3},\quad t\neq -1.$$
这种对曲线进行参数化的办法来自欧拉。

原点在曲线上，它对应于 $t=0$.

用 GeoGebra 算得（见下图）
$$y''_x=\frac{x'(t)y''(t)-x'(t)y''(t)}{(x'_t)^3}=\frac{-2(t^3+1)^4}{3(2t^3-1)^3}.$$

所以当 $t \in (-\infty,-1)$ 时，曲线下凸；当 $t \in (-1, \frac{1}{\sqrt[3]{2}})$ 时，曲线下凸；当 $t \in (\frac{1}{\sqrt[3]{2}}, +\infty)$ 时曲线上凸。 □

注：我们在学习一元微积分时，应尽量避免隐函数求导，因为我们并不知道为什么一个方程能够确定一个隐函数以及隐函数为什么可导——这些是多元微积分中的隐函数定理保证的。数学的目的不仅仅是计算，它告诉我们为什么可以计算。所以在上面这个例子中，我们采取了参数化方法，把方程 $x^3+y^3-3xy=0$ 转化为参数方程，从而可以使用一元微积分的方法。把代数方程进行参数化，可以参见欧拉的名著《无穷分析引论》第一卷第三章。著名数学家 André Weil 说：今天的学生从欧拉的〈无穷分析引论〉中所能得到的益处, 是现代的任何一本数学教科书都比不上的。高斯说: “学习欧拉的著作，乃是认识数学的最好工具。”

例 12. 设 f 为 $\mathscr{C}^\infty$ 函数，求 $g(x)=\frac{f(x)}{x}$ 的 n 阶导函数的表达式。

解. 由商的高阶可微性定理知 g 也是 $\mathscr{C}^\infty$ 函数。用高阶导数 Leibniz 公式，得到

$$\begin{aligned}
g^{(n)}(x) &= \left(f(x)\cdot\frac{1}{x}\right)^{(n)} \\
&= \sum_{k=0}^{n} C_n^k f^{(k)}(x)\left(\frac{1}{x}\right)^{(n-k)} \\
&= \sum_{k=0}^{n} C_n^k f^{(k)}(x)(-1)(-2)\cdots(-n+k)\frac{1}{x^{n-k+1}} \\
&= n!\sum_{k=0}^{n}\frac{f^{(k)}(x)}{k!}\frac{(-1)^{n-k}}{x^{n-k+1}}.
\end{aligned}$$

□

例 13. 证明函数

$$f(x)=\begin{cases}\mathrm{e}^{-\frac{1}{x^2}}\sin\frac{1}{x}, & x\neq 0,\\ 0, & x=0\end{cases}$$

是 $\mathscr{C}^\infty$ 函数。

证明. 当 $x\neq 0$ 时，$f(x)$ 是初等函数，从而是 $\mathscr{C}^\infty$ 函数。所以只需要证明 f 在 $x=0$ 处任意阶可导。

$$f'(x)=\mathrm{e}^{-\frac{1}{x^2}}\left(\frac{2}{x^3}\sin\frac{1}{x}-\frac{1}{x^2}\cos\frac{1}{x}\right),$$

我们用数学归纳法证明，对任意正整数 n，对 $x\neq 0$，

$$f^{(n)}(x)=\begin{cases}\mathrm{e}^{-\frac{1}{x^2}}\left[P\left(\frac{1}{x}\right)\sin\frac{1}{x}+Q\left(\frac{1}{x}\right)\cos\frac{1}{x}\right], & x\neq 0;\\ 0, & x=0,\end{cases}$$

其中 $P(\cdot), Q(\cdot)$ 是多项式。

$$\begin{aligned}
f^{(n+1)}(x) &= \left(\mathrm{e}^{-\frac{1}{x^2}}\left[P\left(\frac{1}{x}\right)\sin\frac{1}{x} + Q\left(\frac{1}{x}\right)\cos\frac{1}{x}\right]\right)' \\
&= \mathrm{e}^{-\frac{1}{x^2}}\frac{2}{x^3}\left[P\left(\frac{1}{x}\right)\sin\frac{1}{x} + Q\left(\frac{1}{x}\right)\cos\frac{1}{x}\right] \\
&\quad + \mathrm{e}^{-\frac{1}{x^2}}\left[P'\left(\frac{1}{x}\right)\left(-\frac{1}{x^2}\right)\sin\frac{1}{x} + P\left(\frac{1}{x}\right)\left(-\frac{1}{x^2}\right)\cos\frac{1}{x}\right] \\
&\quad + \mathrm{e}^{-\frac{1}{x^2}}\left[Q'\left(\frac{1}{x}\right)\left(-\frac{1}{x^2}\right)\cos\frac{1}{x} + Q\left(\frac{1}{x}\right)\frac{1}{x^2}\sin\frac{1}{x}\right] \\
&= \mathrm{e}^{-\frac{1}{x^2}}\left[\tilde{P}\left(\frac{1}{x}\right)\sin\frac{1}{x} + \tilde{Q}\left(\frac{1}{x}\right)\cos\frac{1}{x}\right]
\end{aligned}$$

其中

$$\begin{aligned}
\tilde{P}(t) &= 2t^3P(t) - t^2P'(t) + t^2Q(t), \\
\tilde{Q}(t) &= 2t^3Q(t) - t^2P(t) - t^2Q'(t)
\end{aligned}$$

都是多项式。

$$\begin{aligned}
\lim_{x\to 0}\frac{f^{(n)}(x)}{x} &= \lim_{x\to 0}\frac{\mathrm{e}^{-\frac{1}{x^2}}\left[P\left(\frac{1}{x}\right)\sin\frac{1}{x} + Q\left(\frac{1}{x}\right)\cos\frac{1}{x}\right]}{x} \\
&= \lim_{x\to 0}\mathrm{e}^{-\frac{1}{x^2}}P_1\left(\frac{1}{x}\right)\sin\frac{1}{x} + \lim_{x\to 0}\mathrm{e}^{-\frac{1}{x^2}}Q_1\left(\frac{1}{x}\right)\cos\frac{1}{x} \\
&= 0,
\end{aligned}$$

所以 $f^{(n+1)}(0) = 0$. 于是对任意正整数 n， $f^{(n)}(0) = 0$.

因此 f 是 $\mathscr{C}^\infty$ 函数。 $\square$